Classroom Research Working Paper Series

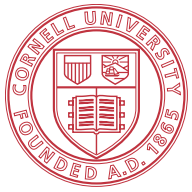
Volume 4

Featuring projects by the 2014-2015 Graduate Research and Teaching Fellows and Teagle Fellows:

Steffen Blings (Government)
Jae B. Cho (City and Regional Planning)
John Foo (Biomedical Engineering)
Molly Katz (English Language and Literature)
Nicholas Mason (Ecology and Evolutionary Biology)
Sarah Maxey (Government)
Kaylin Myers (English Language and Literature)
Joonsuk Park (Computer Science)
Vera Rinaldi (Comparative Biomedical Sciences)
Kristie Schlauraff (English Language and Literature)
Lindsay Springer (Food Science and Technology)
Meghnaa Tallapragada (Communications)

Center for Teaching Excellence
Sponsored by the Teagle Foundation and the
Center for the Integration of Research, Teaching, and Learning (CIRTL)

Editors: Kimberly Williams, Ph.D. and Runjini Raman



CLASSROOM RESEARCH WORKING PAPER SERIES

Volume 4

2014–2015 Graduate Research and Teaching/Teagle Fellow Contributors

Steffen Blings (Government)
Jae B. Cho (City and Regional Planning)
John Foo (Biomedical Engineering)
Molly Katz (English Language and Literature)
Nicholas Mason (Ecology and Evolutionary Biology)
Sarah Maxey (Government)
Kaylin Myers (English Language and Literature)
Joonsuk Park (Computer Science)
Vera Rinaldi (Comparative Biomedical Sciences)
Kristie Schlauraff (English Language and Literature)
Lindsay Springer (Food Science and Technology)
Meghnaa Tallapragada (Communications)

Editors: Kimberly Williams, Ph.D., and Runjini Raman

Cornell University
Center for Teaching Excellence

Sponsored by the Teagle Foundation and CU-CIRTL

Published: 2016
Ithaca, NY

CONTENTS

PART I: NEW PEDOGOGICAL FRONTIERS

PART II: STUDENT PERCEPTIONS ON THE LEARNING EXPERIENCE

PART III: TECHNOLOGY IN THE CLASSROOM

Center for Teaching Excellence
420 Computing and Communications Center
cornellcte@cornell.edu
www.cte.cornell.edu
607-255-3990

Methodological Paradigms and Assessment of Student Difficulties in Learning Advanced Quantitative Methods

Jae B. Cho

Graduate Research and Teaching Fellow 2014-2015

ABSTRACT

This study assesses student difficulties in learning advanced quantitative methods and provides reasoning as to why some students who are more qualitatively oriented experience more difficulty, as measured by grades, compared to those that are more quantitatively focused. Students enrolled in an advanced quantitative methods course offered in the City and Regional Planning Department at Cornell University were asked to rate different subjects based on their difficulty and to answer a series of questions regarding their methodological paradigms. Students were also observed within the classroom, and their assignments were used as data to draw significant conclusions. Results suggest that on average, quantitatively-oriented students perceived quantitative methods to be more accessible, yet with regards to actual performance, it was not disciplinary background but rather prior exposure that resulted in better grades. Qualitative results suggest that students find the math that is required to learn quantitative methods to be difficult, and that such methods should be taught within the context of real world problems. Overall, the results suggest that planning programs should include more qualitative methods courses within their curriculums in order to properly train students to address problems that our society faces today.

INTRODUCTION

Background

Planning considers itself to be one of the most interdisciplinary fields in the social sciences, utilizing both qualitative and quantitative methods in research and application. This is mainly due to the fact that planning as a discipline is very pragmatic and "problem-oriented": proposed solutions to current problems of our social and build environment take precedence over the methods used in research. Since there is rarely a single right or wrong way to address the complex problems different societies face, both qualitative and quantitative methods are seen as adequate ways to address research questions, so long as the methods are effective in the research context. In this sense, within the planning discipline, qualitative and quantitative literacy regards utilizing the related methods

appropriately such that the most relevant methods are used within the context of the research question at hand.

Due to the interdisciplinary nature of the planning field, graduate planning programs across the U.S. attract students from a wide variety of disciplines, ranging across urban studies, economics, political science, sociology, anthropology, and public policy, to name just a few. Students entering graduate school in planning generally see themselves as able to use the knowledge they expect to gain along with their ability to harness their good intentions to positively affect the world (Baum, 1997). Thus, students, much like planning academics and planning practitioners, yearn to use their knowledge in application rather than become hardcore theorists. This is especially true due to the fact that planning programs often serve relatively older students that are reentering school after some years of planning practice in the field (Kaufman & Simons, 1995). These students' interests and needs are largely driven by their current or previous jobs.

Both quantitative and qualitative research methods courses offered in graduate planning programs attract a very diverse body of students from various disciplines with different methodological paradigms and levels of competency. Especially when it comes to quantitative methods, students' previous educational experiences (i.e. whether they graduated from undergraduate programs that are more quantitatively or qualitatively aligned) and usage of such methods in the field determines to a substantive degree their overall proficiency in quantitative literacy. Yet even so, many students that have little background in quantitative methods and their usage still are either required or select to enroll in quantitative methods courses in order to learn these methods for future use. Thus there is a need to make methods courses accessible to students with differing levels of proficiency and research experience and to reconcile students' paradigmatic roots to enable students to better utilize quantitative methods effectively in their research and in application.

Research Setting

This study is based on an advanced quantitative methods course offered to students in the Urban and Regional Studies, City and Regional Planning, and Regional Science programs at Cornell University. The course focuses on quantitative methods in explaining and analyzing the geographical distribution of economic activity as well as the interaction between actors and places within and across geographies. Topics introduced in the semester-long course include advanced input-output analysis (including inter-regional and multi-regional models), structural path analysis, computable general equilibrium, and agent-based modeling, and students enrolled in the course are exposed to various computer packages as well as fundamental economic theory.

While the course is traditionally offered to graduate students enrolled in the planning and regional sciences programs, advanced undergraduate students enrolled in the urban and regional studies program who wish to develop more involved analytical skills are allowed to take the course with the consent of faculty advisors. For master's students the course is an elective course included as one of the classes needed to be taken to fulfill the requirement for advanced methods. However, Ph.D. students in both the planning or regional science programs at Cornell do not have a set core or elective curriculum to choose their courses from, and instead are given the liberty to choose courses from within the department or across the campus that best fit their interests, with the consent of their graduate committee. The vast majority of students taking the course have taken a previous course on input-output analysis, which is also offered as an elective for both undergraduate and graduate students. Therefore it is safe to say that students enrolled in the course are enrolled based on their motivation to develop their analytical skills, which distinguishes the students from those enrolled in core courses as a requirement.

The undergraduate program in Urban and Regional Studies has an interdisciplinary focus

and includes requirements for students to take courses in topics ranging across urban design, history, society, politics, land use, regional development, globalization, and planning methods. The City and Regional Planning graduate program is also very interdisciplinary in nature, with concentrations including economic development planning, land use, environmental planning, and international studies. Students enrolled in the graduate planning program come with a background in a variety of disciplines, including urban studies, planning, political science, history, sociology, and economics. The Regional Science program is more quantitatively oriented in its nature, including a focus among regional economic analysis, urban and regional economics, location theory, environmental and energy systems, transportation, and international spatial problems. Most Regional Science graduate students have undergraduate degrees in quantitatively-aligned disciplines such as economics or engineering.

The students enrolled in the course included a wide range of students from diverse backgrounds, which made the study better geared toward analyzing the philosophical grounding of research beliefs and how students leverage their previous exposure to both qualitative and quantitative research paradigms with course offerings. Moreover, the overall high motivation of students enrolled in the course assured to some degree that there would not be vast discrepancies in the effort put in by the students to complete the requirements of the course.

Objectives

The main purpose of this study is to assess the difficulties that students had in learning advanced quantitative methods, given their previous orientation towards qualitative or quantitative reasoning. In addition, the study attempts to identify whether students with a more qualitative orientation toward reasoning have more difficulties in learning quantitative methods. The overarching goal of the study is to suggest ways in which advanced quantitative methods can be made more accessible to the broad audience of students that take planning-methods courses, without sacrificing methodological depth in the topics covered.

LITERATURE REVIEW

Traditionally, the social sciences have utilized both quantitative and qualitative methods – as well as mixed methodologies – in conducting research. In recent years, the growth of qualitative research has sparked debate regarding the effectiveness of qualitative versus quantitative methods in research. Termed as the "paradigm wars," this debate indicates the dichotomy that exists within current social science research paradigms. Leech and Goodwin (2008) analyze the growth of qualitative methods research in doctorate education programs, and find that overall there is a large variability in the types of courses offered as well as the types of dissertations that are allowed. The findings signal a general messiness within the social science disciplines, and that students can become confused with regards to which methods they should use and how to learn such methods.

Students studying social science disciplines are demographically diverse and come from a variety of backgrounds. Furthermore, their academic interests also vary significantly. However, since most social science disciplines require some amount of competency in both quantitative and qualitative methods, many studies have shown that students – especially those that are less quantitatively literate – often have difficulties in learning quantitative methods through formal coursework. For example, Buchler (2009) observes that some political science students struggle in quantitative methods courses due to less exposure to formal training in mathematics. Because learning quantitative methods or mathematics requires a completely different form of education than other types of theory courses, students that have not been exposed to formal quantitative methods in their education previously are not used to this new learning style. Due to this lack of exposure, the authors find that students wrongly blame

the difficulty with math on other factors, such as the quality of teaching on the part of the instructor or the impenetrability of the subject matter. One of the main differences noted by the authors is that learning quantitative methods, especially math, is a linear process. The common response of students taking quantitative methods classes is to simply move on when a particular concept is not understood. This common response is offered as one explanation why students who are averse to math do not to well in these types of courses. Thus, the authors conclude that checking to make sure students understand content before moving on is critical to student success. Fundamentally, the authors stress that it is not simply teaching students to use technique A or B in a given situation, but explaining why we use technique A over B.

Becker and Greene (2005) identify similar problems with economics students. They suggest that studying the work of Nobel Laureates could provide for a way to effectively teach quantitative methods because often their work is well explained and linked to relevant questions in the real world. First noting that many undergraduate students are able to follow "cookbook" steps for estimation and hypothesis testing, they assert that even so, understanding of underlying principles of statistics is very low. The authors' main point is to emphasize that knowing "how to do" statistics and econometrics is very different from "understanding," and that intuitive examples and explanations are critical in bridging this gap. The main takeaway from this paper is that in a quantitative methods course, simply teaching students the nuts and bolts of methods is easy, but getting them to actually understand what they are doing is a completely different matter, and that student learning-outcomes are significantly enhanced by using intuitive real-life examples.

Murtonen and Lehtinen (2003) studied education and sociology students to find out whether they have different perceptions of difficulty related to quantitative and qualitative methods. Students were asked to rate different topics on the basis of their difficulty, and it was found that statistics and quantitative methods were perceived as more difficult than other domains, such as qualitative methods. Overall, it seems that students tend to polarize the academic subjects as "easy" or "hard"; language, major, and qualitative subjects are characterized as "easy" and mathematical, statistical, and quantitative subjects are characterized as "hard."

Five main categories of reasons for difficulties were established using student responses to open-ended questions: (1) superficial teaching, (2) linking theory with practice, (3) unfamiliarity with, and difficulty of, concepts and content, (4) creating an integrated picture of research in order to fully understand it, and (5) negative attitudes toward these studies. The students who gave high ratings for the difficulty of statistical and quantitative subjects cited teaching most frequently as the reason, while those students who did not have many problems in statistical and quantitative subjects mentioned negative attitudes as the main reason for difficulties. The implications are that instructors should pay more attention to the development of deep and holistic understanding of research as an integrated whole, and that students need to be examined on their prior knowledge and teaching must be adjusted accordingly. Many students noted how methods courses were "hurried," which begs the question of whether methods courses are too packed. Students also noted that it was difficult to link theory with practice, which instructors should take as advice to try and link material and course arrangements with real-life research questions.

Within the planning education literature, the emerging theme regarding quantitative methods education is the importance of application of methods to actual planning problems. This stems from the applied nature of the planning discipline itself, as well as the fact that many graduate planning programs award professional degrees that prepare students for actual jobs in the planning field. Davidson (1986) identifies five goals for

effective teaching of quantitative methods. The five goals are to 1) develop the students' "number sense," 2) develop the students' ability to formulate and substantiate well-thought-out quantitative arguments, 3) develop an ability to use a range of quantitative/statistical techniques, and 4) develop a critical perspective on quantitative technique. In order to do this, the author prescribes five principles for course design: 1) make activities in which planners engage with quantitative information as visibly as possible, 2) expand coverage in topic areas to go beyond the traditional statistical topics to more applied work, 3) resist the "right answer" syndrome, 4) resist the "back-of-the-book" syndrome, and 5) recognize the true complexity of planning problems. The general point made by the author is that quantitative reasoning requires not only acquiring knowledge, but applying that knowledge to your own research and practice, which requires extensive application of concepts and critical thinking on the part of students.

Prosperi (1986) raises questions regarding what is needed in order to effectively teach planning methods to students. The author asks two questions critical to planning pedagogy, namely 1) What is planning, and 2) Who are the planners? The conclusion drawn from these two questions is that students need to know a lot more about the contexts of planning practice than they currently do, and, once they learn more, they need to develop the set of methods (both analytical and otherwise) that these practitioner contexts demand. The author then asks two other relevant questions: 3) What is familiarity and competence in methods? and 4) How can microcomputers (technology, considering the article was published in 1986) be used to enhance teaching? The author fails to come to a cohesive conclusion regarding these questions, but stresses that competence and familiarity in methods requires both applied knowledge and use of various technologies in implementing this knowledge. Interestingly, the author stresses depth in methods courses, and concludes that a series of courses in one particular method is more useful than a flat model of introductory courses in many methods. This type of curriculum structuring is also advocated by Mahayni, Sanchez, and Kelly (1999). The main takeaway from this piece is that again, methodological competency comes from in-depth knowledge of methods resulting from a linear process of knowledge accumulation, and that planning pedagogy should be more aligned to practice. Baum (1997) also stresses that planning education should bridge the gap between academics and practitioners through a partnership between the two groups, and emphasize a holistic approach to planning scholarship, which includes both agendas.

METHODS

Participants

Data was collected for the 12 students enrolled in the course "Methods of Regional Science and Planning" offered within the City and Regional Planning department at Cornell University for the 2015 Spring semester. The group was comprised of one undergraduate student, eight master's students, and three Ph.D. students. The variation in the age, race, nationality, level of education completed, and methodological orientation was significant among the students, reflecting the interdisciplinary nature of the planning program and diversity of students enrolled within.

Because the enrollment for the course was not limited to either graduate or undergraduate students, the level of methodological rigor in which students had been exposed to prior to the course also varied significantly across students. Neither the Urban and Regional Studies undergraduate program nor graduate programs in Planning or Regional Science have a rigid methods curriculum in which the type of methods courses students must enroll in is mandatory, and thus even students that were in comparatively more advanced stages of their education had in some instances less background in the material covered in the course compared to other students that had fewer years of formal higher education.

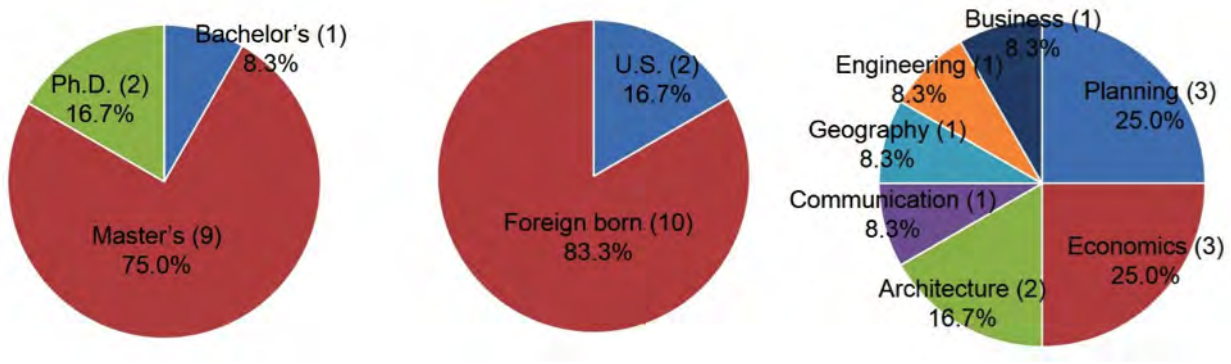


Figure 1. Demographic Characteristics of Students.
(a) student program, (b) nationality, (c) undergraduate background

Data Collection and Methodology

The main focus of this study was to investigate students' interpretation of quantitative methodology based on their previous exposure to different research paradigms, and how these students intended to utilize and apply these methods to their research or professional careers in the future. The intent was to use the insights gained from student feedback and classroom observation to suggest better ways to teach quantitative methods to students with diverse philosophical backgrounds. Thus the focus of data collection was to get students' own views, without directing them too much in terms of answers to specific questions or problems.

The first part of data collection consisted of a survey that was administered to the students roughly halfway into the semester. The survey consisted of three components and was administered voluntarily (although all students chose to respond) during a supplementary lab session of the course. The first component asked students to place seven academic subjects within a dimensional field defined by a coordinate system with two dimensions: easy-difficult and concrete-abstract, each ranging from -5 to +5 with the origin set to 0. The academic subjects included the Methods of Regional Science and Planning course itself, along with both quantitative methods such as math, statistics, computer programing and qualitative methods and also foreign languages (Murtonen & Lehtinen, 2003). In this way, we were able to compare student experience with specific topics covered in the course (that were mainly quantitative) with other areas including qualitative subjects.

The second component consisted of a series of Likert scale and questions regarding how each student viewed themselves within the quantitative/qualitative paradigm spectrum. Instead of asking directly whether the students considered themselves to be quantitatively or qualitatively aligned, the students answered a series of questions regarding their beliefs relating to logic, methods, epistemology, axiology, ontology, and the possibility of causal linkages and generalization. This portion was followed by open-ended questions related to how students viewed the material covered in class with regards to their methodological orientation, and whether they planned to actively use the methods introduced in class for their research or professional career. In addition, a series of questions relating to demographic characteristics, previous major, nationality, and self-reported grades were also asked to gather relevant data for statistical analysis.

In addition to the survey, students were also qualitatively observed during the lectures and the lab in order to get a deeper understanding of their learning dynamics. Observations included which types of students frequently asked questions, whether some students were consistently missing lectures or labs, students' performance in assignments, and whether students utilized office hours.

RESULTS

Quantitative Results

In the dimensional task, students placed different types of academic subjects within a plane, with the x-axis ranging from easy to difficult and the y-axis ranging from abstract to concrete. A total of seven subjects were included: 1) mathematics in general, 2) statistics in general, 3) foreign languages, 4) use of computer packages for analysis, 5) quantitative methods, 6) qualitative methods, and 7) methods covered in this course. Figure 2 shows the coordinate system and the means of student responses regarding the different academic subjects.

Surprisingly, overall the students regarded math to be the easiest, while a foreign language was reported to be the hardest subject. This may be because the course under study, "Methods of Regional Science and Planning," is an elective on advanced quantitative methods, resulting in self-selection of math-oriented students. Also noting that most of the students enrolled in the course were international students, the fact that foreign languages was reported to be the most difficult might represent difficulties faced by students learning quantitative methods in a language that they are not fully comfortable in. Generally, qualitative methods were regarded to be more difficult and abstract compared to quantitative methods. The fact that qualitative methods was reported to be the most abstract subject also suggests that self-selection of more quantitatively oriented students was present, which may have caused students to be less exposed to the qualitative aspects of methodologies.

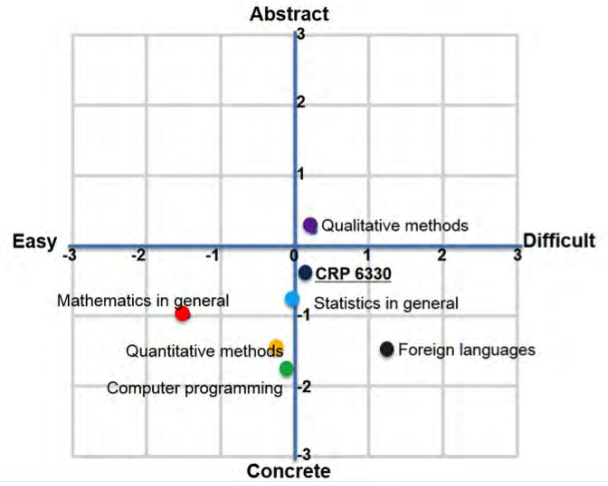


Figure 2. Student perceptions of difficulties in various academic subjects

In addition to their overall analysis, students were grouped based on their prior backgrounds into two groups, namely quantitative and qualitative groups. The quantitative group consisted of six students that were from more quantitatively oriented disciplines such as economics, business, and engineering. In the case of students that had planning or other backgrounds that were rather neutral in their orientation, their reported number of quantitative courses previously taken was used as a means to group students into the more quantitative group. Students that reported taking more than 10 quantitative methods courses during their undergraduate (or also master's studies for Ph.D. students) were included in the quantitative group. The qualitative group was also comprised of six students that came from disciplines such as geography, architecture, urban and regional studies, and communication. Again, in the case where students came from methodologically neutral disciplines, the number of quantitative methods courses previously taken was used as a criteria for selection into the group. Figures 3a and 3b graphically depict the means of the responses for students from each group, and Table 1 reports the differences between the two student groups.

The quantitative group reported math to be the easiest subject and statistics in general to be easy and concrete. Interestingly, the quantitative group reported Methods of Regional Science and Planning to be the most abstract, even more so than qualitative methods. This suggests that while they viewed general quantitative methods to be quite easy and concrete, the course material covered

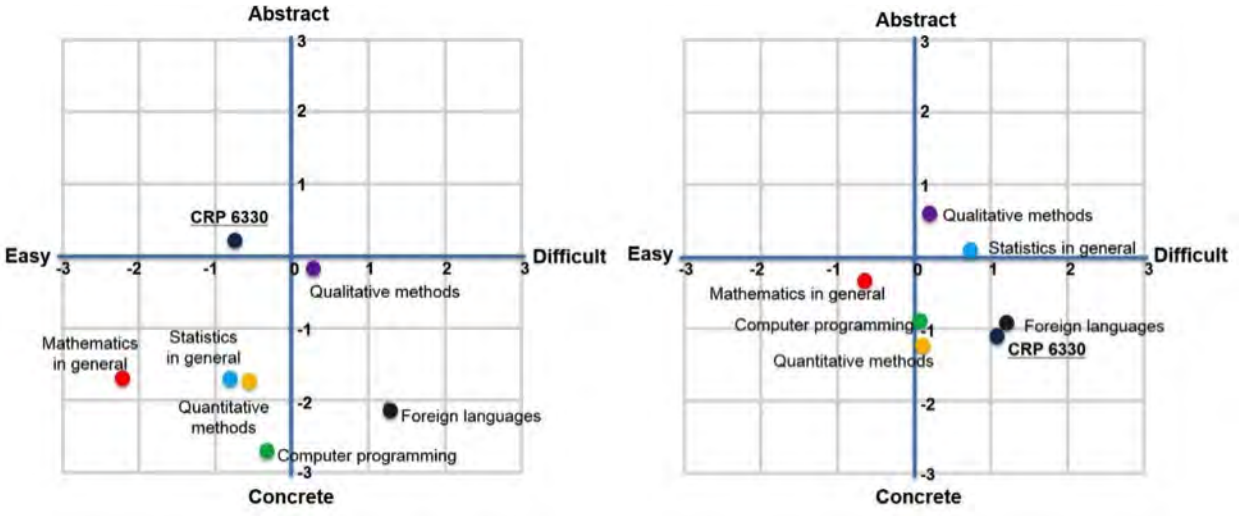


Figure 3. Student perceptions of difficulties in various academic subjects.

	Quantitative group		*Qualitative group*		*Difference*	
	x-coord.	*y-coord.*	*x-coord.*	*y-coord.*	*x-coord.*	*y-coord.*
Mathematics in general	-2.25	-1.67	-0.67	-0.33	-1.58	-1.33
Statistics in general	-0.83	-1.67	0.75	0.08	-1.58	-1.75
Foreign languages	1.25	-2.08	1.25	-0.92	0.00	-1.17
Use of computer packages for analysis	0.33	-2.67	0.08	-0.92	-0.42	-1.75
Quantitative methods	0.58	-1.75	0.08	-1.25	-0.67	-0.50
Qualitative methods	0.25	-0.17	0.25	0.58	0.00	-0.75
Methods covered in course	0.75	0.25	1.08	-1.08	-1.83	1.33

Table 1. Differences between the quantitative and qualitative groups in perceived difficulties

in the current course was more difficult and abstract. This may be due to particular teaching methods or possibly because the students did not have enough time to properly internalize the subject matter covered during the course of the semester. The qualitative group in general reported all subjects to be more difficult and abstract compared to the quantitative group, with the possible exception of qualitative methods and foreign languages. Both groups reported the same level of difficulty for qualitative methods, yet, surprisingly, the qualitative group found qualitative methods more abstract than the quantitative group. However, considering the relatively small sample size, the differences in perceived difficulties in qualitative methods seemed negligible. Overall, the largest differences between the two groups appeared for the math and statistics subjects, along with the methods covered in the current course.

The students were also asked a series of questions about their methodological beliefs regarding research in general. The questions were structured such that students' beliefs regarding generalization of research findings, causal linkages, the importance of values within research settings, and objectivity could be analyzed quantitatively. Figure 4 shows the results for the students as a whole, along with the average scores for the two groups.

The Likert scale was structured in such a

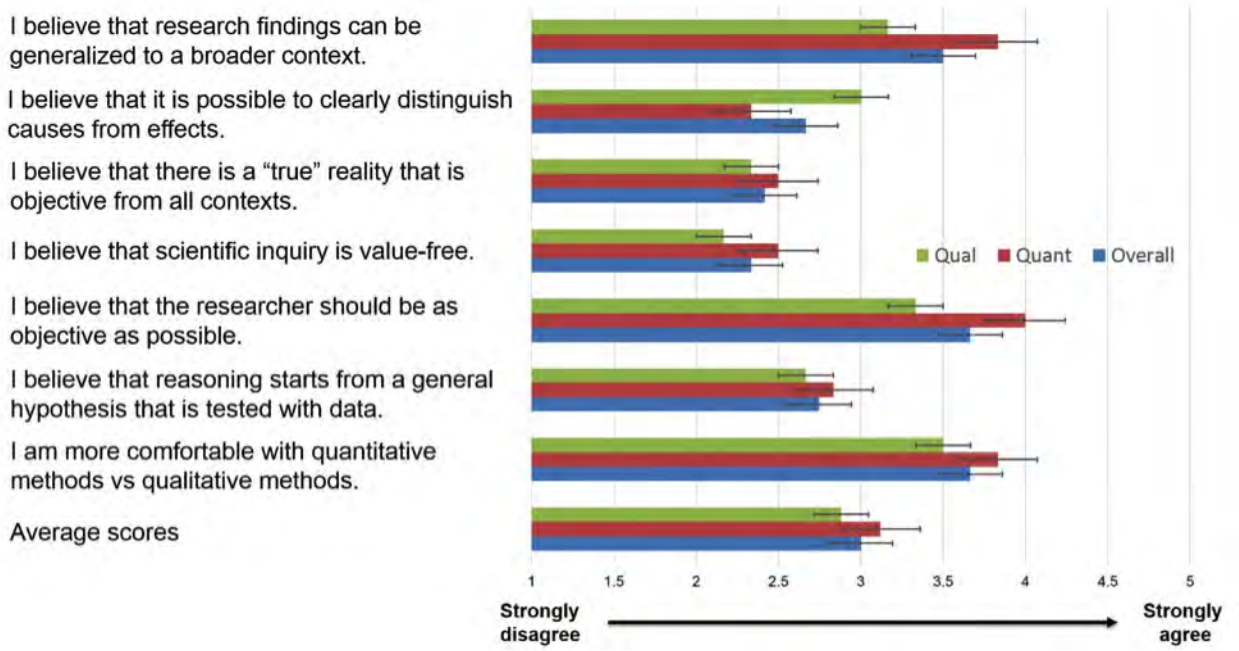


Figure 4. Students' Methodological Paradigms
(Bars indicate confidence intervals at the 95% level)

way that a higher score represented a more quantitative orientation. On average, the quantitative group scored higher relative to the qualitative group, as expected. However, the qualitative group on average scored higher than the quantitative group in the question regarding causal distinctions, and this difference was statistically significant at the 9% level. The question regarding researcher objectivity showed clearly that the quantitative group felt more objectively oriented, while the qualitative group viewed the researcher as more subjective and embedded within the research framework. The generalizability of research results was the same; the difference was also significant at the 95% level. While on average the difference in scores between the qualitative and quantitative group were not significant, the general tendency for quantitatively-oriented students to score higher along the Likert-scale was apparent across most questions.

Regardless of methodological paradigms, all students generally reported that they did not believe that there existed a "true" reality that is objective from all or that scientific reality is value-free. The two questions corresponding to these two concepts had the lowest score overall compared to others. On the contrary, students generally scored higher for the questions regarding generalizability and objectivity. Interestingly, one of the questions that had the highest overall score was that which asked whether students were more comfortable with quantitative versus qualitative methods. For this question, while the qualitative group scored lower on average, still the average score was much higher compared to other questions. Given that the current course was an elective advanced quantitative methods course, this could possibly be interpreted as the more quantitatively aligned students self-selecting into the course. However, this may also indicate that on average, relative to quantitative methods, qualitative methods are taught less within curriculums regardless of discipline and that students are more comfortable with quantitative methods simply because of prior exposure. This can be seen in the planning curriculum offered at Cornell,

where there is only one qualitative methods course offered compared to more than 10 quantitative methods courses. Overall, it can be seen that students' backgrounds in regards to both discipline and prior exposure impact their methodological paradigms.

A simple linear regression with the number of quantitative methods courses previously taken as the independent variable and final grades for the course as the dependent variable was run on the sample of students. Figure 5 shows the results of the regression.

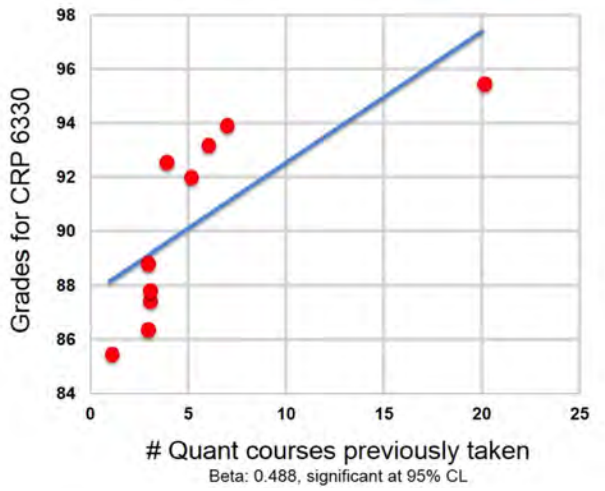


Figure 5. Simple Regression Results for Student Sample

The relationship between the number of quantitative methods courses previously taken and final grades was significant at the 95% confidence level, with a clear positive correlation. Due to limited sample size and thus limited degrees of freedom, more explanatory variables were not added to the regression. However, when comparing the average grades for the quantitative group versus the qualitative group, while the average for the quantitative group was higher at 90.5 compared to 89.63, this difference was insignificant at the 95% level. Again, due to limited sample size these results are not in any sense conclusive, yet they do suggest that students regardless of discipline do better in quantitative methods courses when they have more prior exposure to quantitative analysis. This was the case even for students from backgrounds such as architecture or communication (that are traditionally more qualitative in orientation) who had previously taken more quantitative methods courses. These students seemed to do better regardless of currently enrolled program (i.e. Bachelor's, Master's, or Ph.D.) or prior disciplinary background.

Qualitative Results

Utilizing the open-ended survey questions and observations of students within the classroom, four themes emerged from the qualitative analysis. The first was that abstract math and lack of perceived applicability to the real world made quantitative methods difficult to learn. Both the open-ended survey questions and observations of students seemed to support this finding. First, almost all of the students reported some form of perceived lack of applicability of quantitative methods to real world problems and difficulty in math.

-"If the instructor lacks the ability to simplify complex math, then learning becomes difficult."

-"I have no economics background. It is hard to understand concepts within quantitative models."

-"Math. Abstractness. Calculations [are difficult]."

-"The application of methods to real problems [is difficult.]"

-"I have trouble remembering multiple equations, so I need to constantly revise."

These qualitative findings contradict the quantitative analysis of students' perceived difficulties with different subject matter. In the quantitative analysis, students, regardless of qualitative or quantitative orientation, reported mathematics in general to be the easiest subject compared to others. However, the survey responses indicated the contrary, with many students experiencing difficulties with understanding complex mathematical equations regarding economic theory. This

could be seen in the classroom observations as well. Many more students asked questions when faced with tough problems related to math concepts, and students came to office hours more frequently when a problem set that needed more mathematical understanding was assigned. A possible explanation for the discrepancy between the quantitative and qualitative results could be that students view mathematics in general – such as calculus, including differential equations or linear algebra – to be different from the mathematics used in the class. Taking a look at the student responses to assignments where mathematical equations were extensively used, it could be seen that students understood the underlying mathematical principles, such as differentiation or matrix algebra, yet struggled to apply these concepts to economic analysis. This provides an explanation for why students viewed math in general as easy, yet still found it difficult to apply these concepts to the mathematical equations used in economic models.

Another theme that emerged was that students perceived qualitative methods to be more difficult that quantitative methods. However, the reasons behind this perceived difficulty varied among students. A sample of student responses are included below:

- "Sometimes the qualitative courses seem to be more difficult because of the language."
- "Qualitative courses require more personal knowledge of context, which means more effort. Quantitative classes are more objective."
- "Qualitative courses were more difficult than quantitative courses because qualitative skills require logical and critical thinking more. Also, I need to set my own point of view, and logic."
- "Qualitative courses could be more difficult compared to quantitative courses for me. How to observe correctly and how to make neutral and unbiased survey designs is difficult and fuzzy due to different situations."
- "I have not yet taken any qualitative methods courses (which I am open to taking in the future) mainly because I feel qualitative research is more time consuming, and the fact that there's no way to know the correctness of your results/ conclusions adds to the difficulty."

Some students noted that qualitative methods are more difficult because they require comprehensive knowledge of the research context, and thus are more time consuming. Some students erred in their understanding of qualitative methods in general, reporting difficulty in generalization of qualitative results and knowing the "correctness" of conclusions as being the factors that added to difficulties. Qualitative analysis is not about generalization or "knowing" the correctness of conclusions, and in this sense it can be seen that many students were not properly trained in quantitative analysis. This seemed to be the case even for students that came from more qualitatively oriented backgrounds with comparatively more experience and exposure to such methods. This suggests that the current ways in which different disciplines teach qualitative methods is not well received by students in a cohesive manner, and that students have difficulty in reconciling quantitative versus qualitative paradigms when conducting research.

Resonating with the quantitative results, qualitative analysis suggests that disciplinary background dictates the perceived usefulness of qualitative or quantitative analysis. With regards to the question of whether quantitative or qualitative methods seem more useful, some sample responses of students are included below:

- "Both are useful. There are different domains of applications and it's always better to combine them together."
- "Quantitative. We live in the era of "Big Data" and the value of qualitative analysis has been expended to a substantial degree."
- "Although quantitative methods have many

restrictions and assumptions, I still believe they provide solutions to verify intuitive hypotheses and allow for us to find general results."

-"Qualitative. Because if we define useful as more widely used then I think qualitative methods can be used in more areas."

Students that came from more qualitatively oriented disciplines were more favorable to qualitative methods compared to students that were more quantitatively aligned. Interestingly here, prior exposure – as in the number of qualitative or quantitative courses taken previously – did not seem to affect students' perceived usefulness of either methodological paradigm. Rather it was the disciplinary setting itself that seemed to matter for students determining which types of methods were more useful. Some students that had taken many quantitative courses and very few qualitative courses still thought that qualitative methods were more useful in practical applications. Also considering that there is a high probability of self-selection of more quantitatively-oriented students enrolling in the course, the fact that many of the students reported either qualitative methods to be equally as important or even more important compared to quantitative methods suggests that students yearn for more qualitative methods to be taught within the planning curriculum.

One of the strongest themes that seemed to emerge from the data was that students yearn for quantitative methods to be taught within the context of real-world problems. Previous research on teaching quantitative methods overwhelmingly suggests this perceived lack of applicability of quantitative methods as an obstacle in proper teaching of quantitative analysis. During the course of the semester in this particular course students were frequently given examples of how to utilize the methods learned in class in real-world planning problems. Furthermore, most of the assignments had an applied component to them where students were given certain policy scenarios or planning problems, which they were asked to provide insight into using the methods learned in class. Yet even so, students seemed to view the course as being more theory-oriented, and many seemed less motivated in the classroom due to the perceived lack of applicability of the quantitative methods taught in the course. This is apparent in student comments on how the course could be improved:

-"Practical, concrete projects [would benefit the course.]"

-"I would be motivated more to learn quantitative methods if I was aware of how they would help in real life situations."

-"I need more concrete examples on how we can use the software and methods in planning practice."

-"If I am aware of how much substantive help I can get from learning these methods, it would motivate me more so that I would be encouraged to improve my learning."

Even though a substantial amount of class and lab time was dedicated to show real-world applications of the methods learned in the course, it seemed as if students were bored and uninterested with the underlying mathematical and quantitative reasoning that went into the theory. In general, students were less interested and seemed to be less motivated in class when the instructor or the TA taught fundamental concepts critical in understanding how certain methods work, and more into the class when examples of applications and current research papers utilizing learned methods in real-world settings were covered. This could be observed in different ways. There was a relative lack of student engagement in lab sessions that covered background material that was more related to pure math. In addition, classroom attendance was higher when the material covered was more related to real planning problems, such as in sessions where the instructor covered his own work conducted in collaboration with different governments.

CONCLUSIONS

This study provides evidence supporting the long-standing notion that students perceive quantitative methods as being inherently difficult and abstract. Furthermore, the results suggest that students' prior exposure to quantitative and/or qualitative methods dictates to a large degree their competency in understanding more advanced methods. Overall, students seemed to yearn for more qualitative methods courses to be taught within the planning curriculum, and when quantitative methods were taught, they wanted the coursework to be complemented with real-world applications.

Students that came from more qualitatively-oriented disciplines found quantitative subjects to be more difficult and scored lower in terms of final grades. This difficulty stems from the observation that more qualitatively-oriented students have had less exposure to mathematics and other quantitative methods, making grasping more advanced concepts difficult. Even so, prior exposure to certain methodological paradigms did not seem to affect the perceived usefulness of either methodology, which is somewhat promising. As a discipline that prides itself in its interdisciplinary nature and utilization of different methods across a wide range of societies' problems, planning curriculums should include more qualitative methods courses for students. This is especially important since students seemed to struggle in reconciling the different philosophical beliefs that quantitative and qualitative methods advocate. Students often mistook quantitative beliefs such as generalizability and objectivity to be applicable to qualitative methods as well. As planning programs across the U.S. strive to train well-rounded professionals capable of providing insights into real world problems, a more balanced methodological core is needed to train students to cope with a diverse set of issues relevant to planning academics and practitioners.

Because the planning field is very diverse in both its composition and the problems it addresses, instructors are challenged to meet the needs of a wide range of students with different proficiencies and are also required to teach a wide range of different methods. As a result, teaching quantitative methods to planning students is a remarkably difficult task. One possible solution could be to teach quantitative methods courses through modules, where a linear progression takes place from more rudimentary applications all the way up to cutting-edge methods (Mahayni et al., 1999). Yet because of the variety of methods that a planning curriculum needs to cover, most of the planning curriculums in the U.S. are based on a flat model rather than a hierarchical one. In this regard, planning departments should thoroughly assess which types of quantitative methods are actually utilized by practitioners (Kaufman & Simons, 1995) and boldly eliminate those that are less utilized for those that are more needed in the field.

However, the diversity of students coming into planning programs also poses an opportunity unique to the planning discipline. As opposed to other fields that are predominantly quantitative or qualitative in nature, the planning field attracts students that are interested in a wide variety of methods. This can be seen in the qualitative results, where students reported substantial interest in learning qualitative methods even though their previous exposure to such methods was minimal at best. Again, a more balanced curriculum that addresses the needs of a wide variety of students is needed in order to properly train students to become successful practitioners or academics. Furthermore, the diversity of problems that planners are faced with regarding our society suggests that not just the methods, but even more so their applications, are important. More time should be allocated within the classroom to provide examples of applications, and students should be trained to not only understand the material, but to utilize methods for solving actual problems. Doing this would not only better the learning outcomes of all types of students, but

would also motivate students to learn methods – whether quantitative or qualitative – in a more comprehensive, in-depth fashion.

REFERENCES

Baum, H. S. (1997). Social science, social work, and surgery: Teaching what students need to practice planning. *Journal of the American Planning Association*, 63(2), 179-188.

Becker, W. E., & Greene, W. H. (2005). Using the Nobel Laureates in economics to teach quantitative methods. *The Journal of Economic Education*, 36(3), 261-277.

Buchler, J. (2009). Teaching Quantitative Methodology to the Math Averse. *PS: Political Science & Politics*, 42(03), 527-530.

Davidson, J. M. (1986). Quantitative reasoning in the planning curriculum. *Journal of Planning Education and Research*, 6(1), 30-36.

Kaufman, S., & Simons, R. (1995). Quantitative and research methods in planning: Are schools teaching what practitioners practice? *Journal of Planning Education and Research*, 15(1), 17-33.

Leech, N. L., & Goodwin, L. D. (2008). Building a methodological foundation: Doctoral-level methods courses in colleges of education. *Research in the Schools*.

Mahayni, R. G., Sanchez, T. W., & Kelly, E. D. (1999). Teaching planning methods through modules. *Journal of Planning Education and Research*, 18(4), 353-360.

Murtonen, M., & Lehtinen, E. (2003). Difficulties experienced by education and sociology students in quantitative methods courses. *Studies in Higher Education*, 28(2), 171-185.

Prosperi, D. C. (1986). Teaching planning methods: Some further questions and directions. *Journal of Planning Education and Research*, 6(1), 40-41.

Student Adaptation to the Modular Use of the Flipped Classroom in an Introductory Biomedical Engineering Course

John C. Foo[1,2*], Iwijn De Vlaminck[1], Kimberly Williams[2]

** Graduate Research and Teaching Fellow 2014-2015*

ABSTRACT

The Flipped Classroom model moves direct instruction outside of the classroom allowing more time for student-centered in-class activities. Guides to Flipping the Classroom suggest investing time at the start of the course to promote the Flipped Classroom model to students. However, this may not be the best use of time if the instructor only plans to use the Flipped Classroom model to teach a small portion of the class. One reason for not flipping the entire course may be that only certain portions of the course would benefit from the Flipped Classroom model. Or, perhaps flipping a small portion of the class serves as a precursor to flipping the entire course in subsequent offerings. To observe how students react to a partial flipping of a course, students in an Introductory Biomedical Engineering course were taught using the Flipped Classroom model only for the Biotransport module (2 class periods out of 19 for the term) in the middle of the semester. After the module, students were asked to complete a survey about their perceptions of the module. Classroom observations were recorded and student engagement data were obtained. The survey results indicate that the majority of the students found the transition to the Flipped Classroom easy or somewhat easy to adapt to and the majority were engaged in the course content both in and out of the classroom. These findings suggest that instructors may be able to use the Flipped Classroom model in a modular fashion.

INTRODUCTION

The Flipped Classroom is a pedagogical model that moves direct instruction outside of the classroom to allow more time in class for student-centered activities such as group work or peer-instruction. Bishop and Verleger define the Flipped Classroom as "an educational technique that consists of two parts: interactive group learning activities inside the classroom, and direct computer-based individual instruction outside the classroom."[1]

The Flipped Classroom model began gaining popularity among K-12 education circles in the late 2000s when online lectures by two high school chemistry teachers, Bergmann and Sams, started spreading beyond their school [2]. Since then, the Flipped Classroom model has been adopted by some educators in higher education as a possible means of improving student engagement and learning. In a recent scoping review, O'Flaherty and Phillips found

that there is already a significant amount of evidence, albeit indirect, suggesting that the flipped classroom model does improve student academic performance in higher education [3].

In spring 2015, we decided to use the Flipped Classroom model on the Biotransport module of our introductory biomedical engineering class, since the concepts and equations covered in this module are challenging for students to understand and apply. We believed that shifting the introduction of the concepts and the derivation of equations outside the class would allow us to better use class time by letting our students practice applying these concepts and equations, as well as addressing any student misconceptions.

However, given that a number of guides on flipping the classroom suggest spending a significant amount of time on the first day of class answering questions about the flipped classroom and convincing students of it merits, we became concerned about the feasibility of implementing it for only a small portion of the course [2], [4]. Considering that the Biotransport module comprised only two 75-minute long class periods, spending even 15 minutes to sell the idea and address student concerns seemed like a huge time investment.

A literature search on the flipped classroom in higher education mostly turned up studies on courses that have been either partially or completely flipped for the entire semester, or partially flipped for a few classes [5]–[11]. All these studies show that students generally responded positively to the Flipped Classroom model, but only one (a complete flip for the entire semester) explicitly mentioned an attempt to convince students of the merits of the Flipped Classroom – through the use of a "comprehensive syllabus" [5]. The students in this study were also provided a guide that emphasized what student were responsible for in and out of class. In addition, two other studies, both also featuring courses that have been completely flipped for the entire semester, recommend providing students with "some structure and guidelines" and "well-structured guidance" [6], [7]. For the studies in which the instructor completely flipped only a small portion of the course, students still responded positively, although it was not clear if any of the instructors first spent time selling the idea to students [12], [13]. Perhaps telling the students to watch the video lectures before class was all the instruction they needed prior to class?

The purpose of this study is to find out how easily students would adapt to the Flipped Classroom model if it was implemented without first convincing them of its benefits, as well as what students would have liked to have known prior to its implementation. This information would be useful to educators who want to completely flip only a small portion of their class, for reasons ranging from wanting to try the Flipped Classroom model, or deciding that only specific sections of the course would benefit from it. For this study, the flipped Biotransport Module comprised video lectures and online quizzes outside class, and group-based problem solving inside class. This study was granted exemption status by Cornell's Institutional Review Board.

METHODS

All first-year students in the college of engineering are required to take an "Intro-To" class, of which ENGRI 1310: Introduction To Biomedical Engineering is one of the nineteen options available. However, the class is also open to any student at Cornell University. The course was divided into 5 modules of varying lengths. Two modules were taught by guest instructors, one of which was the Biotransport module that was taught by the first-author.

In this Introductory Biomedical Engineering Course, the typical class period involved the instructor introducing and explaining new concepts to the students. The lessons typically include two to three multiple-choice "clicker" questions, either interspersed throughout the lecture, or saved for the very end of the lecture. To answer these "clicker" questions, students would use their personal response systems that have been preregistered with the class to

submit their answers. The instructor would then display a histogram of how the class answered, before querying the students for explanations. Students were awarded points for participating, regardless of the veracity of their answers.

The Flipped Biotransport Module

The Flipped Biotransport Module comprised two class periods, each 75 minutes long, during the 9th week of a 16-week semester. The first author was the instructor for the Biotransport module.

Video lectures (Figure 1) were made using a Wacom Intuos Pro Tablet and the Ink2Go screencast software, and uploaded onto YouTube unlisted. This meant that only people with the link are able to view the video.

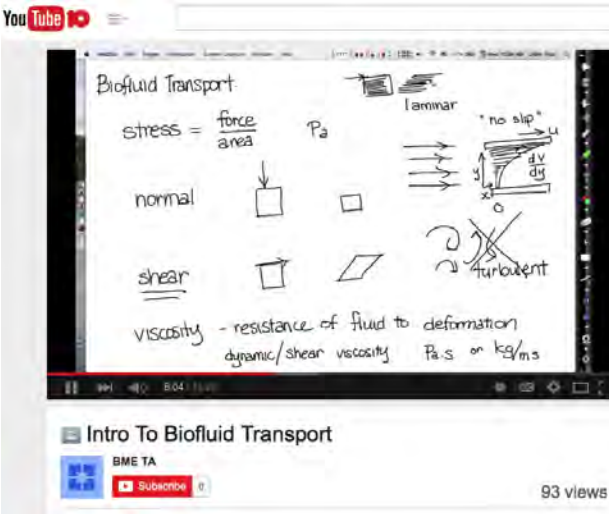


Figure 1: A screenshot of the Intro to Biofluid Transport video lecture.

Student responsibilities before class were to view a series of video lectures, which come up to about 30 minutes per class. After watching the videos, students had to complete an online quiz, which comprises 3 questions, one of which is a "muddiest point" question.

During class, students reviewed specific quiz responses as a class to address misconceptions. They then proceeded to apply the video lecture material through solving practice problems in groups.

Data Collection

To answer the research question, the following qualitative and quantitative data were obtained and used.

1. YouTube Viewership Data

The number of unique views, total views, and average view duration of each video were counted from the day the video was uploaded to the day of the corresponding class period, i.e. 3 for first class period, 2 for the second class period.

2. Online Survey

Students were specifically told that the online survey was anonymous and voluntary. The survey began with Bishop and Verleger's definition of the Flipped Classroom to ensure that participants understood exactly what the term refers to in the survey (see Appendix A).

For the Likert-like questions in the online survey, a neutral option was deliberately put in place to reduce agreeable bias.

To analyze the open-ended questions in the online survey, the responses were reviewed by the first author for general themes. The first author then went through all the responses again, and looked specifically for evidence of each theme in each response and labeled them accordingly. The number of times a theme was brought up was tabulated.

3. Classroom Observations

The instructor recorded his observations and reflections after each class period. The instructor also invited a peer to observe the first class period.

RESULTS

Results from Online Survey

Of the 44 students in the class, 33 students participated in the survey, giving a response rate of 75%.

Demographic Information

Of the survey participants, all were freshmen save one, who was a sophomore. The gender breakdown was 73% female, and 27%

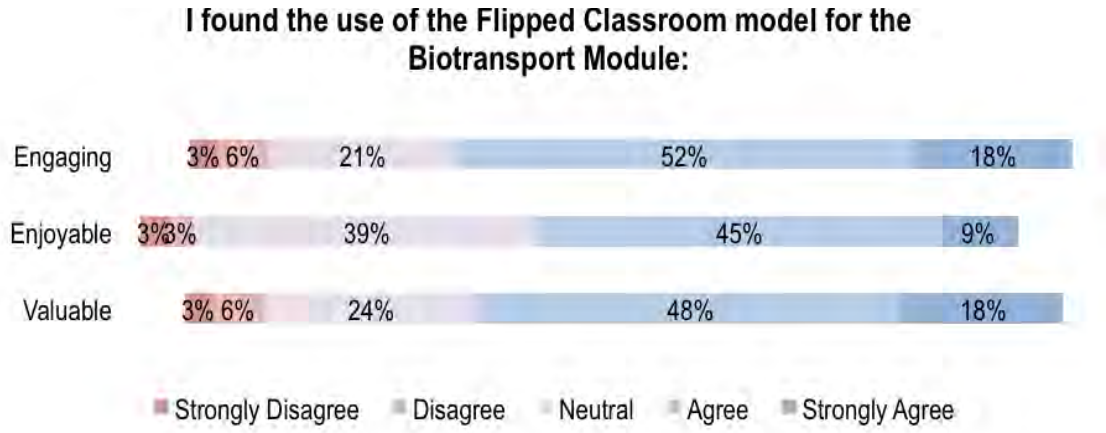


Figure 2: The survey respondents generally perceived the use of the Flipped Classroom model for the Biotransport Module positively.

male. In terms of college affiliation, 88% of the participants belonged to the College of Engineering, 6% to the College of Arts and Sciences, while the remaining participants had not yet declared a major. As for the racial composition of the class, 59% identified as Caucasian, 19% as Black or African American, 19% as Hispanic, 15% as Asian, and 4% as Native American. Participants who declared multiple racial identities were counted in all the races they identified with. When participants were asked what grade they expected in the class, 45% expected an A, 45% expected a B, while 10% expected a C.

Prior Experience with the Flipped Classroom Model

The majority of survey participants had experienced the Flipped Classroom model in some capacity prior to the Biotransport module – 33% first experienced it in high school, 37% first experienced it in college prior to the Biotransport module. For 30% of the participants, the Biotransport module was their first experience with the Flipped Classroom model. None of the participants reported that they had experienced the Flipped Classroom model in middle school.

Student Perceptions about the use of the Flipped Classroom Model for the Biotransport Module

The quantitative survey data on student perceptions are Likert-type data and have been presented in diverging stacked bar charts to highlight the spread of the positive and negative values, with the neutral point as the baseline [14].

The survey respondents generally perceived the use of the Flipped Classroom model for the Biotransport Module positively (Figure 2). The majority of the participants agreed or strongly agreed that the Flipped Classroom was engaging (70%), enjoyable (54%), and valuable (66%). Less than 10% of participants disagreed or strongly disagreed that the Flipped Classroom was engaging (9%), enjoyable (6%), and valuable (9%).

Of the four aspects of the Biotransport Module, the majority of survey respondents either agreed or strongly agreed that video lectures (91%) helped them learn the material, followed by the instructor addressing responses from the online quiz in class (79%), and solving problems in class (72%) (Figure 3). Only 33% of respondents agreed or strongly agreed that completing the online quizzes helped them learn the material, while 24% disagreed or strongly disagreed, with the remaining 42% having a neutral opinion.

Regarding the transition to the Flipped Classroom model, 51% found the transition smooth or somewhat smooth, and 63% found adapting to the Flipped Classroom model easy or somewhat easy (Figure 4). A sizable

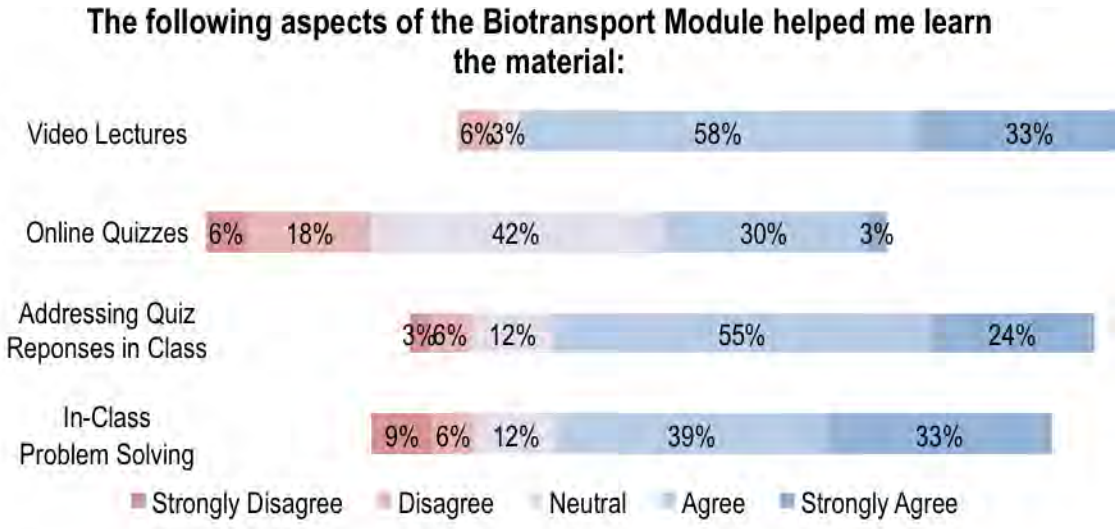


Figure 3: Video lectures and addressing quiz responses in class were the most helpful aspects of the Biotransport module in helping students learn the material..

minority of respondents found the transition to the Flipped Classroom model disruptive or somewhat disruptive (24%), and found adapting difficult or somewhat difficult (18%).

Student Suggestions to make the Biotransport Module more conducive for learning

The top three themes that emerged from student suggestions to make the Biotransport Module more conducive for learning are as follows:

a. Create more structure for in-class activities (9x).

One student found the class "kind of all over the place which makes it really hard to take notes and organize them later." Yet another student found the group problem solving "chaotic," preferring instead to "solve the problems independently and then review the correct answer." Another student suggested "multiple choice questions instead of free responses," which may suggest their preference for the order associated with clicker questions used in other modules of the class.

b. Provide more practice problems (7x).

This feedback relates less to the Flipped Classroom model than the class in its entirety, as this has been a common request throughout the semester. In fact, one student expressed that she enjoyed the Flipped Biotransport module because she got to do more practice problems.

c. Recap information from the video lectures in class (5x).

While a number students asked for "brief reviews" or "mini recaps", one respondent went as far as suggesting the instructor "redo the main derivations in class." In contrast to the requests for recapping the information from the video, a couple of respondents suggested "less time addressing quiz in class" and "not too much direct repeating of the videos in class, more moving on from the basic material."

Student Suggestions to make the transition to the Flipped Module smoother

a. Set clearer expectations for the Flipped Module (5x).

One student suggested that the instructor should "notify students so they don't think the videos and quiz are as impromptu and supplemental as the random readings we sometimes get…" Another said, "Because we were not notified that this was flipped classroom, I did not pay as much attention as I should have to the videos." A third student

who had a hard time adapting to the flipped classroom commented that the instructor "sort of just posted videos online with no pretext whatsoever. I had no idea what this was about until just now, so it was extremely awful. We jumped into it with no preparation."

It was interesting that students did not seem as concerned about the benefits of the flipped classroom, or why they were subjected to it. They instead seemed more focused on knowing what was expected of them, which may be due to the fact that the survey was not set up to garner such responses.

b. Flip other modules in the class (4x).
Four students suggested flipping the other modules in the class as well, although they did not list reasons for their suggestion.

c. Allow more time to complete pre-class activities (3x).
Students wanted "more time to watch the videos and answer the quiz." This appears to be due to the shorter interval between the Tuesday and Thursday classes as compared to the interval between the Thursday and Tuesday classes.

Results from YouTube Viewership Data
Viewership data obtained from YouTube suggest that most students viewed the video lectures, as the number of unique views for each video was close to the class enrollment of 44 students. The average view duration for each video is only about 60% of the video duration. Possible reasons for the low percentage include students re-watching specific sections of the video, as suggested by the total views outnumbering the unique views, or that students only watched portions of the video that they needed to answer the online quizzes. Because the videos were posted on YouTube and were unlisted on search engines, only students with the links were able to watch the videos, but there was no way to track which students watched the videos and how much of the videos they watched.

Results from Student Work
The online quizzes had an average completion rate of 98%.

Results from Classroom Observations
Despite being instructed to work in groups of 4 – 5, a number of individuals worked alone or in

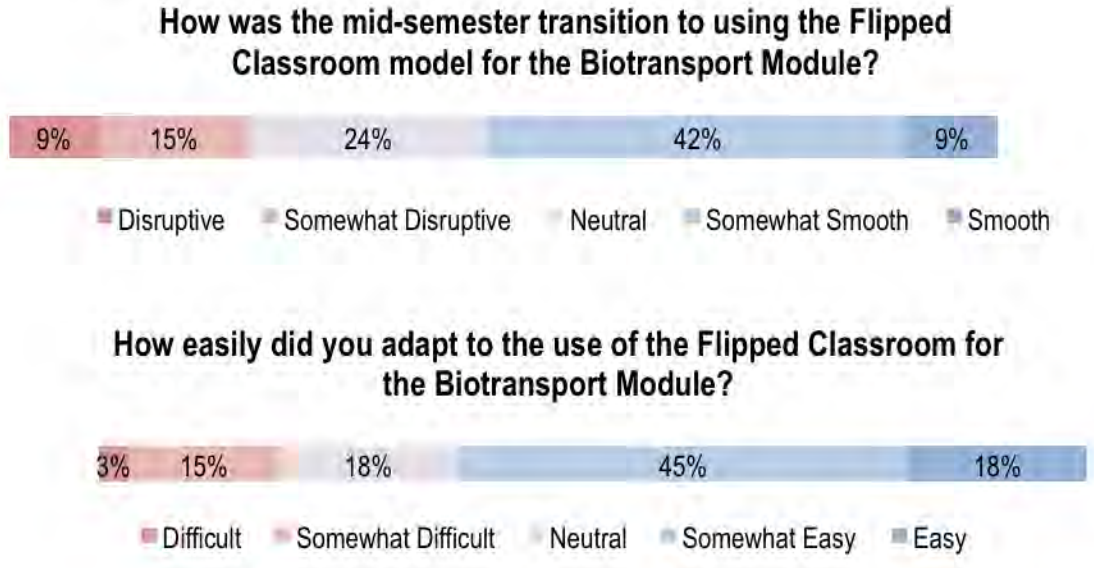


Figure 4: The majority of student respondents found the transition to the Flipped Classroom model smooth or somewhat smooth. An even larger majority adapted easily or somewhat easily to the use of the Flipped Classroom.

Lesson	Video	Unique Views	Total Views	Video Duration	Average View Duration	Percentage of Video Viewed
Biomass Transport	Intro to Biomass Transport	46	73	6:58	4:11	60%
	Fick's Law of Diffusion	48	66	12:36	7:11	57%
	Conservation of Mass	43	56	8:29	4:52	57%
Biofluid Transport	Intro to Biofluid Transport	41	57	15:21	8:38	56%
	Derivation of Poiseuille's Law	37	52	13:51	8:40	63%

Table 1: Although the number of unique views of each video is close to the total number of students, the average view duration for each video is only about 60% of the video duration.

pairs. If the question was open-ended, students tended to immediately begin discussing the problem; if the question involved mathematical calculations, students tended to work alone before consulting their group. Students used their cell phones to research answers for open-ended questions but also to access social media, specifically Facebook. Students worked at very different paces, with some students struggling, while others were done quickly. This made it challenging to pace the class. A few students did not work on the in-class questions and only paid attention when the answers were being addressed.

Instructor Reflections on Flipping the Classroom

The instructor found that creating video lectures was manageable with assistance. Using the tablet and screencasting software to create the video lectures was straightforward after getting guidance from Cornell's Academic Technologies Center. The instructor also found that scripting the video lecture greatly reduced the need for video editing.

The instructor was able to quickly and easily address misconceptions. Viewing the quiz responses the night before class alerted the instructor to common misconceptions, and gave him time to formulate a coherent response.

The instructor found structuring the in-class activity very challenging. It was difficult to get everyone working in groups, a problem that was exacerbated by the lecture hall setup. Some students who quickly solved the practice questions got bored while students who were struggling would consult the instructor instead of their peers.

Fielding open-ended questions was also immensely challenging, as it was impossible to anticipate the wide variety of student responses.

Having every student bring a device that can connect to the internet may allow for the entire class to conduct research together. Students

Lesson	Number Completed	Total	Percentage complaged
Biomass Transport	43	44	98%
Biofluid Transport	43	44	98%

Table 2: Almost all the students completed the online quizzes. The student that missed the Biomass Transport quiz is different from the student that missed the Biofluid Transport quiz.

can work in groups to ensure that at least one group member has access to the internet. Polling students at the start of class may be necessary to identify students who require assistance in obtaining a device.

DISCUSSION

Based on the Youtube Viewership data and the online quiz completion rates, most of the students seem to complete the necessary out-of-class activities.

While the majority of the freshmen in the class found the transition to the Flipped Classroom module easy or somewhat easy to adapt to, 18% found it difficult or somewhat difficult. While learning about the purported benefits of the Flipped Classroom may help these particular students adapt better, they seemed more concerned about the instructor "giving warning" so they did not go in to class "with no preparation." These ideas were also echoed by students who did adapt easily to the flipped classroom. This group of students appeared more concerned about what was expected of them and how they can do well in the class.

While the instructor did not detect any concerns regarding the effectiveness or use of the Flipped Classroom model per se, a handful of students expressed interest in how the material covered in the video lectures fit into the big picture. Therefore, highlighting the value of the information in the videos may further increase their contribution to student learning.

This study sought to address whether it is possible to implement the flipped classroom on a limited scale without selling the idea to students. This data suggest that students will easily adapt to the modular use of the flipped classroom when the instructor clearly informs them of what is expected of them and what their responsibilities are, not unlike informing students of the rules of an activity.

Limitations of this study were that the first author, who is also the instructor, is a graduate student teaching assistant who had no prior experience teaching this course and no prior experience implementing the Flipped Classroom model. He did, however, receive a lot of guidance from the professor of the course.

The content of the Biotransport Module was developed specially for this class and there were no prior lectures to copy from. The author's experience thus more accurately represents the experience of a new faculty member developing the course content for a course they have been assigned rather than that of an experienced faculty member wanting to try the Flipped Classroom for the first time.

No attempt was made to evaluate the efficacy of the Flipped Classroom model as this was new to this class. No comparison with previous years were possible and there was only one section in the class.

CONCLUSIONS

The majority of the freshmen in this class found the transition to the Flipped Classroom module easy or somewhat easy to adapt to. Setting clearer expectations for the Flipped Module and creating more structure for the in-class activities may further smoothen the transition. This suggests that educators may be able to implement the Flipped Classroom in a modular fashion, either for appropriate topics or as a step to eventually flipping the entire course.

ACKNOWLEDGEMENTS

We would like to thank the students and co-instructors in BME 1310, Bailey Cooper, Meghnaa Tallapragada, the instructor's peers in ALS 6016, and Benjamin Moss from the Cornell Academic Technologies Center.

REFERENCES

[1] J. L. Bishop and M. A. Verleger, "The flipped classroom: A survey of the research," in *ASEE National Conference Proceedings, Atlanta, GA*, 2013.

[2] J. Bergmann and A. Sams, *Flip your classroom: Reach every student in every class every day*. International Society for Technology in Education, 2012.

[3] J. O'Flaherty and C. Phillips, "The use of flipped classrooms in higher education: A scoping review," Internet High. Educ., vol. 25, pp. 85–95, Apr. 2015.

[4] "Coursera's Flipped Classroom Field Guide." [Online]. Available: https://docs.google.com/document/d/1arP1QAkSyVcxKYXgTJWCrJf02NdephTVGQltsw-S1fQ/edit?pli=1#bookmark=id.bslf3vgya1li. [Accessed: 02-Mar-2015].

[5] J. E. McLaughlin, M. T. Roth, D. M. Glatt, N. Gharkholonarehe, C. A. Davidson, L. M. Griffin, D. A. Esserman, and R. J. Mumper, "The Flipped Classroom: A Course Redesign to Foster Learning and Engagement in a Health Professions School," *Acad. Med.*, vol. 89, no. 2, pp. 236–243, Feb. 2014.

[6] G. S. Mason, T. R. Shuman, and K. E. Cook, "Comparing the Effectiveness of an Inverted Classroom to a Traditional Classroom in an Upper-Division Engineering Course," I*EEE Trans. Educ.*, vol. 56, no. 4, pp. 430–435, Nov. 2013.

[7] M. K. Kim, S. M. Kim, O. Khera, and J. Getman, "The experience of three flipped classrooms in an urban university: an exploration of design principles," *Internet High. Educ.*, vol. 22, pp. 37–50, Jul. 2014.

[8] S. G. Wilson, "The Flipped Class A Method to Address the Challenges of an Undergraduate Statistics Course," *Teach. Psychol.*, p. 0098628313487461, Jan. 2013.

[9] N. Warter-Perez and J. Dong, "Flipping the classroom: How to embed inquiry and design projects into a digital engineering lecture," in *Proceedings of the 2012 ASEE PSW Section Conference*, 2012.

[10] T. Stelzer, D. T. Brookes, G. Gladding, and J. P. Mestre, "Impact of multimedia learning modules on an introductory course on electricity and magnetism," *Am. J. Phys.*, vol. 78, no. 7, pp. 755–759, Jul. 2010.

[11] M. Moravec, A. Williams, N. Aguilar-Roca, and D. K. O'Dowd, "Learn before Lecture: A Strategy That Improves Learning Outcomes in a Large Introductory Biology Class," *CBE-Life Sci. Educ.*, vol. 9, no. 4, pp. 473–481, Dec. 2010.

[12] R. Pierce and J. Fox, "Vodcasts and Active-Learning Exercises in a 'Flipped Classroom' Model of a Renal Pharmacotherapy Module," *Am. J. Pharm. Educ.*, vol. 76, no. 10, Dec. 2012.

[13] T. P. Young, C. J. Bailey, M. Guptill, A. W. Thorp, and T. L. Thomas, "The Flipped Classroom: A Modality for Mixed Asynchronous and Synchronous Learning in a Residency Program," *West. J. Emerg. Med.*, vol. 15, no. 7, pp. 938–944, Nov. 2014.

[14] N. B. Robbins and R. M. Heiberger, "Plotting Likert and other rating scales," in *Proceedings of the 2011 Joint Statistical Meeting*, 2011.

Qualitative Assessment of Cognitive and Social Benefits Among Underrepresented First-year Biology Students in a Field Course: A Case Study of Experiential Learning in the Galápagos

Nicholas A. Mason[1,2*], Rebecca Brunner[2,3], M. Cissy Ballen[1], Kimberly Williams[4], Jeff McCaffrey[5], Irby Lovette[1,2]

** Graduate Research and Teaching Fellow 2014-2015*

ABSTRACT

Student attrition is a persistent challenge in life sciences education, particularly among underrepresented minorities, first-generation students, and women. Many undergraduate students are unsatisfied with the traditional introductory course; experiential learning opportunities diversify curricula by immersing students into non-traditional academic environments. However, most experiential learning opportunities are only available to upperclassmen. Here, we present a qualitative analysis of an experiential learning opportunity offered exclusively to underrepresented demographics, including a 10-day field component in the Galápagos Islands. We evaluated social and cognitive gains via ethnographic observations, students' natural history journals, and self-reported survey results. Students consistently reported strong cognitive gains in their understanding of basic evolutionary concepts, which was corroborated by journal entries and ethnographic observations. Most students also benefited socially from this experiential learning opportunity, although we observed higher variation in social gains. We suggest immersive field courses and other forms of experiential learning can help life science curricula engage and retain students of underrepresented demographics during their first two years of college and often strengthens students' social networks, self-efficacy, and volition.

INTRODUCTION

Student attrition in science, technology, engineering, and mathematics (STEM) is a perennial concern among postsecondary undergraduate institutions (Chen & Soldner, 2013). Within STEM fields, student attrition, which occurs when students switch to a non-STEM major or leave secondary education altogether (Tinto 2006), is higher among underrepresented demographics, such as ethnic and racial minorities (Matsui, Liu, & Kane, 2003; Dirks & Cunningham, 2006; Allen-Ramdial & Campbell, 2014), women (Bebbington, 2002;

Clark Blickenstaff, 2006; Griffith, 2010), and first-generation college students (Lam, Srivatsan, Doverspike, Vesalo, & Mawasha, 2005; Tate et al., 2015). STEM attrition occurs at higher rates among underrepresented students with weaker academic records (Mendez, Buskirk, & Lohr, 2008; Kokkelenberg & Sinha, 2010; Whalen & Shelley, 2010) and decreased self-efficacy, motivation, and confidence (Burtner, 2005; Beasley & Fischer, 2012). During the first two years of college, students from underrepresented demographics frequently cite negative experiences with lecture-based introductory courses, which lack personal interactions with instructors and peers (Seymour & Hewitt, 2000; Beasley & Fischer, 2012). Undergraduates also have little exposure to the breadth of STEM courses in their first two years (Bettinger, 2010), and poor performance in introductory STEM courses compared to non-STEM classes draws even more students away from sciences and mathematics (Rask, 2010). Thus, the challenges facing minority student retention in the STEM fields are multidimensional and complex.

One potential strategy for mitigating STEM attrition is to expose undergraduates to diverse learning opportunities that may be more enriching than traditional introductory courses. Active learning is a broad pedagogical framework that includes various instructional strategies that contrast with traditional, instructor-centered approaches to teaching (Bonwell & Eison, 1991). A common form of active learning involves experiential learning, which encourages learning through personal interactions with the material at hand and subsequent reflection (Kolb, 1984; Kolb & Kolb, 2005). A prominent goal of experiential learning is to move students beyond simply recalling factual knowledge and toward 'higher-level thinking and learning' (sensu Bloom's taxonomy; Airasian et al., 2001), which involves applying, analyzing, and synthesizing information for a more holistic understanding of a concept (Krathwohl, 2002). In the biological sciences, experiential learning opportunities often take the form of research internships or field courses. Research internships have well-documented positive effects on students' comprehension and volition in the sciences (Hunter, Laursen, & Seymour, 2007; Russell, Hancock, & McCullough, 2007), particularly among underrepresented minorities (Lopatto, 2007). However, the cognitive and social benefits of other active learning opportunities, such as field courses, are less well characterized—especially among underrepresented demographics

Within the academic disciplines of ecology and evolutionary biology, field courses constitute a prominent form of experiential learning (Zervanos & McLaughlin, 2003; Smith, 2004; McLaughlin & Johnson, 2006). Field courses foster discovery, exploration, and immersion into the course material, conferring benefits beyond traditional classroom environments (Eisner, 1982; McLaughlin, 2005; Sanders, 2007). Various studies have documented positive effects of field experiences on students' knowledge and attitude toward biology (Lisowski & Disinger, 1991; Magntorn & Helldén, 2005; Boyle et al., 2007; Easton & Gilburn, 2012; Prokop, Tuncer, & Kvasniák, 2007; Scott et al., 2012), as well as other fields, such as geography (Marvell, 2008) and geology (Rathburn & Weinberg, 2011). However, field courses are disappearing from many curricula due to issues concerning risk management, funding, and balancing other faculty obligations (Barker, Slingsby, & Tilling, 2002; Smith, 2004; Lock, 2010; Scott, Boyd, Scott, & Derek, 2015). Therefore, further assessments regarding the social and cognitive outcomes of field courses in STEM disciplines—particularly for underrepresented students—are necessary to optimize and evaluate postsecondary curricula amid ongoing financial, social, and educational concerns.

Here, we examine the cognitive and social benefits of a 10-day field course in the Galápagos, offered as an experiential learning opportunity paired with coursework over the course of a semester at Cornell University. This curriculum was exclusively offered to students from underrepresented demographics in the

biological sciences—minorities, first-generation college students, and women—offering an opportunity to deepen our understanding of the academic and social benefits that arise from experiential learning in the life sciences for these student groups. Through ethnographic observations of student behavior, assessments of student journals, and analyses of self-reported social and cognitive gains via post-curriculum surveys, we consider both social and intellectual gains of experiential learning in the Galápagos.

MATERIALS AND METHODS

Galápagos Curriculum

The experiential learning opportunity considered in this case study is a major component of a themed curriculum that includes eight credits of coursework at Cornell University, where a typical freshman course-load comprised 14-18 credits/semester. Our "Galápagos Curriculum" (GC) is offered via application to first-year students enrolled in the Biology Scholars Program (BSP), an academic community and support group that supported the participation of underrepresented demographics in biology (Summers & Hrabowski, 2006; Tsui, 2007). Since courses with a field component are often unintentionally restrictive in that they involve high costs associated with travel and equipment, it is important to note that the GC is heavily subsidized, which enables this opportunity to be financially accessible to students from different socioeconomic backgrounds. Twelve students have participated each year (except for 2012, the inaugural year, which had 8 students), bringing the total to 44 by the time of this study.

Students were selected via a written application and interviews with instructors. Accepted students concurrently enrolled in two courses: a core biological sciences course entitled Evolutionary Biology and Diversity (BIOEE 1780) and a separate, writing-focused seminar entitled The Enchanted Isles: Human Observation and Impact in the Galápagos (WRIT 1430). The structure of the GC divides the semester into three qualitatively different time periods: pre-trip, trip, and post-trip (Zervanos & McLaughlin, 2003). Prior to the trip, these courses taught background information about the ecology, evolutionary biology, human history, geography, and literature related to the Galápagos. Students then traveled to the Galápagos for eight days over spring break, where they participated in observation-based discussions, wrote and sketched in field journals, and completed short assignments related to evening lectures. After the trip, these courses offered an opportunity for extended reflection and synthesis, in addition to the continuation of coursework related to evolution and conservation.

While an in-depth description of each course's syllabus and structure is outside the purview of this study, we provide a brief overview here. BIOEE 1780 was a large lecture-based course that covered basic concepts in evolutionary biology and introduced students to major biodiversity groups across the tree of life. It was a required course for all Biology majors at Cornell University and was usually taken during the freshman or sophomore year. In addition to taking BIOEE 1780, GC students also enrolled in a seminar-format Writing in the Majors (WIM) section of that course, which was thematic and limited to GC students. The BIOEE 1780 course and the WIM section combined for a total of five credits. Through class discussions of primary literature, homework assignments, and in-class activities, the WIM section reinforced BIOEE 1780 lecture material; GC students learn background information on evolutionary patterns and processes via previous and ongoing research on Galápagos wildlife.

Taken at the same time, WRIT 1430 was a first-year writing seminar for GC students that focused on human perspectives in the Galápagos, including the history, culture, literature, and conservation of the archipelago. The first-year writing seminar met twice a week and included six graded essays and various informal, low-stakes writing assignments used to assess and design future assignments based on the writing abilities of GC students. Course material focuses on the perspectives and

experiences of different human groups on the Galápagos, such as early Spanish explorers, English and American whalers, evolutionary biologists, prominent authors such as Melville and Vonnegut, and current Galápagos residents. Taken together, the GC immersed students in an interdisciplinary study of the natural history and anthropology of one location over time.

While in the Galápagos, the GC facilitates experiential learning through immersive encounters with nature through hiking and snorkeling, expert-guided commentaries and discussions, and prompted (and unprompted) personal reflections in provided field journals. Unlike most other places in the world, humans have been historically absent from the archipelago. Therefore, animals are generally unafraid of humans and are comfortable with very close observations and interactions. Between island excursions, snorkeling, and boat rides, students live, sleep, eat and travel across the archipelago aboard a ship on which they and their instructors are the only passengers. While on the boat, students discuss primary literature relevant to the Galápagos and journal to reflect on their personal and shared experiences. Many of the students speak Spanish and engage in discussions with the ship crew members, most of whom live in the Galápagos full-time and have extended families there.

Assessment of Experiential Learning Gains

We assessed the cognitive and social gains of students in the GC in three ways: (1) observation of student behavior and interactions within an ethnographic framework; (2) evaluation of student entries in provided travel journals; (3) analysis of self-reported cognitive and social gains via an anonymous survey. Quantitative and qualitative approaches granted insight into different aspects of GC students' academic and personal experiences. We solicited student approval to undertake this project, anonymized student work, and performed all research in accordance with the Cornell University Institutional Review Board (IRB Protocol #1410005010).

Ethnography is a well-established framework in anthropology that involves descriptive observations of extended personal and social contact between researchers and their subjects (Clifford & Marcus, 1986; Willis & Trondman, 2000). Ethnographers assume a dualistic role of observer and facilitator, thereby seeking to describe and contextualize the idiosyncrasies and possible generalities that constitute an irreproducible human experience (Hammersley & Atkinson, 2007). In this case study, NAM observed and interacted with students as the instructor of WRIT 1430, including the 10-day excursion to the Galápagos. Here, we present their findings and observations within an ethnographical framework that roughly follows a chronological series of events throughout the semester.

Reflection is an integral component of experiential learning (Kolb, 1984), while observation is an essential skill in the life sciences (Magntorn & Helldén, 2005). Because natural history observations and reflections are often facilitated through the practice of keeping a field journal (Herman, 1986), we required students to maintain field notebooks throughout the Galápagos field component of the curriculum. We required students to carry their notebooks at all times, which were used to record natural history observations while in the field and personal reflections and responses to more formal writing prompts while on the boat, at least once per day. Following completion of the Galápagos field trip, student journals were handed in to NAM for grading and evaluation of cognitive and social effects. For this component of the study, we only considered journals from the 2015 cohort.

We solicited anonymous, self-reported information regarding the impacts of the Galápagos field experience through a post-trip survey of multiple cohorts of students. Using a five-option, Likert-scale survey (1 = strongly disagree, 5 = strongly agree), we asked students to indicate how the experiential learning opportunity impacted their cognitive (Supplementary Table 1) and social (Supplementary Table 2) gains. We also asked

students open-ended questions to identify the greatest challenges and rewards associated with the Galápagos field course component (Supplementary Table 3). Once we received survey responses, we treated each of the Likert responses as interval data, calculating the mean and standard deviation associated with each question as a measure of central tendency. We also combined all social and cognitive responses into two respective groups and performed a two-tailed, nonparametric Mann-Whitney-Wilcoxon test (Bauer, 1972) and a Fligner-Killeen test (Conover, Johnson, & Johnson, 1981) to determine if Likert-scale responses differed between social and cognitive gains with regards to median scores and variance, respectively.

RESULTS

Ethnographic Observations

The following observations were made by NAM as instructor of WRIT 1430 during the spring 2015 semester. At the beginning of the semester, GC students were eager about the course and visiting the Galápagos. However, students' incoming knowledge of the archipelago and its inhabitants was limited; most students' previous exposure to the islands was through brief excerpts in high school textbooks. For example, students generally knew very little about the geography of the archipelago; some students were uncertain whether the islands are in the Pacific or Atlantic Ocean or which country owned the Galápagos. Students were also unaware that people live on some of the islands, in towns such as Puerto Ayora or Puerto Villamil. Many students initially conceptualized the Galápagos as lush, tropical landscapes teeming with life. Prior to the spring break trip to the Galápagos, WRIT 1430 familiarized students with human experiences in the Galápagos and encouraged students to critically consider their own preconceptions of the archipelago. Students' conceptualizations changed quickly after considering various accounts of the islands, which transformed in their minds from a tropical paradise into a more foreboding setting of jagged lava rocks and little fresh water. Students were surprised to learn that the Galápagos have been a sporadic hotbed of human activity over the past few centuries. Relevant to both the natural and human history of the islands, we devoted two weeks reading and discussing Charles Darwin's career and the formation of his ideas on evolution and biodiversity. Students realized that the Galápagos and its inhabitants played a different role in Darwin's career than has been aggrandized by the superficial and abridged treatments in textbooks and many high school curricula.

In the week before the trip to the Galápagos, students had one-on-one interviews with the instructor of WRIT 1430. During these brief (approximately 15 minute) interviews, students expressed their feelings regarding the upcoming trip, including any sources of anxiety, excitement, or trepidation. All of the students were eager about the upcoming trip, although some were noticeably more excited than others. Some students were anxious about their first international travel experience, while others harbored fears of the open ocean or close encounters with organisms that they perceived as dangerous. Despite different backgrounds with respect to the amount of travelling and exposure to nature, students were generally optimistic and excited about the field trip to the Galápagos.

Upon departing for the Galápagos, the social dynamics among students and instructors changed substantially. While travelling to the Galápagos, some students choose to listen to music or interact with electronic devices on their own; other students conversed with peers and instructors about a wide variety of academic and non-academic topics. Even during the early stages of the 10-day trip, the prolonged contact between peers and instructors outside the classroom manifested social connections among the participants. Although the students had seen each other at least four days per week up until this point in the semester, the travel experience amplified the social dynamics already present among the participants.

Boarding the aircraft to the Galápagos facilitated a more concrete understanding of geography and human impacts in the Galápagos. Prior to departure, flight attendants fumigated the cabin to mitigate the possibility of introducing foreign species and pathogens into the Galápagos. Students were startled and puzzled at first, but the experience reinforced the prominent role that humans play as agents of dispersal of invasive species. We observed students discussing the geography of Central and South America while paging though an in-flight magazine; one student was surprised to learn that Ecuador was almost directly south of New York. Although every student had looked at a world map before, the students seemed to gain an improved sense of Western Hemisphere geography through travel. Similarly, even though students were provided with flight itineraries well in advance, many were still dismayed by the length of time it takes to reach the Galápagos from the mainland, an experiential measure of the archipelago's remote location. It was much easier for students to conceptualize distance through experience (i.e., air travel) than through lectures or readings.

Once the aircraft had landed on Baltra Island in the Galápagos, students immediately began to recognize organisms that we had discussed in class, such as land iguanas and the Opuntia cacti that dotted the arid, rocky landscape. As we made our way through customs and travelled to our ship, students became visibly (and audibly) excited about each new organism that they saw and their upcoming adventures. Prior to boarding our ship, we asked students to spread out along the coastline and take approximately 30 minutes to reflect on their initial impressions of the Galápagos and whatever was happening in their immediate vicinity. With all the excitement and anticipation of travelling, it was often difficult for students to focus on what they were experiencing and had experienced up until that moment. As instructors, we found that periodic reflection in isolation helped "center" students and encourage them to make mindful observations of their personal and shared experiences.

Although the hourly schedule during the Galápagos field course varies based on the trip itinerary, each day begins at dawn, which conflicted with the sleeping schedule of almost every undergraduate on the trip. One student claimed that the only other sunrise they had seen that school year was after pulling an all-nighter to study for exams. In fact, early wake-up calls and conflicts with "internal clocks" were among the most common complaints from students during the trip. Once awake, however, students began to appreciate dawn as a time of peak animal activity and comfortable temperatures amid the strong equatorial sun. We typically visited islands during the early morning and late afternoon to capitalize on the agreeable conditions for viewing wildlife and exploring the terrestrial ecosystems. Late mornings and early afternoons were typically reserved for snorkeling or on-boat academic activities, as heat from the equatorial sun peaks in the middle of the day. Students entered the course with varying levels of experience and comfort with swimming and physical activity. Although we encourage students to push themselves and participate as much as possible, water activities are generally optional and students occasionally choose to stay on the boat.

As a field course destination, the Galápagos provided an excellent environment for experiential learning while still allowing instructors enough certainty to plan with detail. Since travel itineraries are strictly regimented in the Galápagos and many organisms are easily observable due to restricted habitat ranges, fearless animal dispositions, and an open landscape, GC instructors can successfully predict much of the phenomena participants will observe each day, and we therefore organized paper discussions and class activities accordingly. For example, every GC participant observed marine iguanas from only one meter away because this species does not react to the presence of humans and often bask in predictable places on certain islands. We therefore assigned primary literature related

to these iguanas to be read directly before our encounters with them, which substantially enhanced student interest and understanding.

Compared to most field course settings, the Galápagos provides a much more controlled and logistically straightforward introduction to the wildlife encounters and international travel, ideal for an exclusively first-year class. However, some animal encounters are unpredictable, even in the Galápagos; students on the 2015 trip, for instance, were disappointed that we didn't encounter more dolphins. Moreover, a field course visiting the Galápagos during an El Niño year will have a very different impression compared to students that visit between El Niño events. While certain elements of the GC were impossible to predict and varied substantially from year to year, students learned that stochasticity, seasonality, and cyclical climatic effects are important components of any field course and nature itself.

The cast of humans in the Galápagos also changes from year to year. This last year, we were fortunate to interact with a group of college-age Ecuadorian student volunteers that were monitoring sea turtle nests. While we watched the students excavate a nest that should have finished hatching but still included many eggs, the GC students that spoke Spanish conversed with the volunteers. The students learned first-hand that a combination of larval flies and a fungal pathogen had attacked this particular brood of sea turtles. This serendipitous interaction with volunteers and aspiring wildlife biologists also exposed the GC students, who were almost entirely pre-medicine or pre-veterinarian, to additional career options in the life sciences. It was far more engaging for students to interact with authentic, wildlife biologists who were counting and developmentally staging partially hatched sea turtles compared to learning about career options in the life sciences through an in-class lecture.

Opportunities for experiential learning abounded while in the Galápagos—too many to comprehensively discuss here. After previously learning about sexual selection in their large lecture hall on campus, GC students observed blue-footed boobies perform their mating display in person. The immense size of giant tortoises and the observable size and color variation among different populations of lava lizards provided first-hand examples of island gigantism and biogeography. Underwater observations of lateral tail movements by marine iguanas and the powerful thrusts of flightless cormorants' hind feet demonstrated adaptations to predominantly aquatic lifestyles. Lava gulls that scavenged the carcass of a sea lion pup reminded students of the eternal struggle for survival and the process of natural selection. These immersive experiences reinforced course material and provided lasting examples of evolutionary processes in nature.

The students were also excited to experience the landscapes and historical sites that inspired written works about the Galápagos. Students revisited descriptions of the islands from prominent writers, such as Herman Melville and Charles Darwin, in light of their own writing and personal experiences. Students reflected on the similarity between their trip to the Galápagos and Darwin's famous voyage aboard the Beagle; Darwin was only a few years older than the GC students during his visit to the Galápagos, and the Beagle was a similar size to our boat in the Galápagos. We encouraged students to think critically about the similarities and differences between their experience and other human perspectives and histories in the Galápagos. By incorporating concepts and materials from literature and the humanities, we strove to create a more holistic, interdisciplinary academic experience. We created a video log that cataloged our daily activities and illustrated many of the encounters between students, instructors, wildlife, and historical sites in the Galápagos (Video S1).

In addition to gaining a deeper understanding of basic evolutionary concepts and human perspectives in the Galápagos, our time in the Galápagos offered students the opportunity to interact closely with their instructors. As first-year students, most of the participants

had not formed strong bonds with any of their introductory-level course instructors. However, after protracted interactions with instructors in academic and non-academic settings, students began to feel more comfortable conversing with both professors and graduate student instructors—perhaps for the first time appreciating the multidimensionality of instructors as human beings, rather than solely authoritative figures in the classroom. Once initial social boundaries had been broken down, students freely conversed with instructors about their personal and professional lives. These interactions were not unidirectional; the instructors also learned more about undergraduate social and academic life. We felt that this mutual exchange of information increased appreciation for undergraduate, graduate, and faculty life and helped foster social networks that have persisted well beyond the Galápagos field experience.

Upon return from the Galápagos, the classroom dynamic transformed dramatically. After the field trip, students conversed freely with each other and the instructor prior to the beginning of class; group activities were more fluid and inclusive. However, the students did feel that the climax of the semester had already passed; it was occasionally difficult to persuade students to continue reflecting and thinking about the Galápagos. Regardless, from our perspective as instructors, the field experience provided a strong platform to bolster core class concepts and strengthen social networks, both with peers and instructors.

Natural History Journals

At first, students were generally hesitant to engage in written reflection and required persistent prompts from instructors. Students also varied in their level of journaling; some students wrote minimal entries, while others were much more prolific (see Figure 1, Figure S1, Figure S2). Natural history entries toward the beginning of the trip consisted predominantly of relayed information from instructors rather than personal observations. Similarly, personal reflections began as mostly chronological records of daily activities rather than connections between their personal and shared experiences, core concepts of the course, or connections to ideas or concepts from other courses or background knowledge. Midway through the trip, we encouraged students to deepen their reflections and to use their notebooks to synthesize thoughts and forge cognitive connections beyond simply recording their daily activities in a diary. This intervention helped clarify what we are expecting as instructors, and, from our perspective, improved the overall quality and depth of the entries.

As a final journal entry, we asked students to reflect on the role that their field notebooks played in their Galápagos experience. Student responses to field notebooks were positive overall. Most students felt that keeping personal observations helped them to appreciate the experience and were excited to have a personal account to remember it by. One student mentioned that their field notebook "helped them stay present" during their time in the Galápagos (Figure S3). Another student remarked that their field notebook "forced [them] to pay attention to details I may have otherwise overlooked" (Figure S4). In contrast, other students had reservations about their notebooks being graded by instructors. One student said, "[The journal took] me out of the moment… but I am glad I have the knowledge recorded to look back on in the future" (Figure S5). Although there is varying opinion, overall the field notebooks provided a means for reflection that facilitated a more immersive experience for the majority of GC students during their time in the Galápagos.

Survey Responses

We received survey responses from 38 out of 44 students (86.4%; Table S1). The distribution of Likert-scale scores, as well as the mean and standard deviation, are reported in Figure 2. Assuming an interval scale, the average responses indicated positive effects for both cognitive (mean = 4.68 ± standard deviation = 0.62) and social gains (4.39 ± 0.9). We partitioned all question responses as either social or cognitive and used a Mann-Whitney

test with the Likert scale data as ordinal to identify a lower median value for responses to social (n = 304) compared to cognitive (n = 342) questions (W = 59641.5, P = 3.9e-05). Using a Flinger-Killeen nonparametric test of homogeneity of variances, we also found that responses to social questions were more variable than cognitive questions (2 = 20.57, df = 1, P = 5.75e-06).

Student responses to the open-ended question about the most challenging and most rewarding aspect of the Galápagos curriculum varied (Table S1). Many students said that the intense workload and the lack of sleep was the most challenging aspect of the course (Table S1). Seasickness, a particularly prevalent obstacle in this case study, and stomach illnesses were among the most common challenges for many students (Table S1). Certain students were critical of our role as "eco-tourists", rather than biologists, in the Galápagos and even thought "the trip [failed] to provide students with a practical portrayal of what it means to be a biologist." Overall, however, students were more forthcoming when asked to describe the most rewarding aspects of their time in the Galápagos. Multiple students cited close encounters with wildlife as the most rewarding aspect of the field component of the GC, in addition to a greater appreciation for nature (Table S1). Improved social connections with peers and instructors were also frequently mentioned as a positive aspect of the course (Table S1).

DISCUSSION

We documented evidence of cognitive and social gains associated with an experiential learning opportunity in the Galápagos for first-year students from underrepresented demographics in STEM fields. One distinct aspect of this particular case study is that the experiential learning opportunity was offered exclusively to first-year students of underrepresented demographics in the STEM fields. Because STEM attrition is even more pronounced among women, underrepresented minorities, and first-generation college students (Chen & Soldner, 2013; Allen-Ramdial & Campbell, 2014), providing diverse, affordable experiential learning opportunities specifically for these demographics may be an effective strategy to increase diversity in STEM disciplines and reduce attrition rates during the first two years of college education.

Students' feelings of self-efficacy or volition within the STEM disciplines are often shaped by an individual's academic and cultural background (Seymour & Hewitt 2000). Women, first-generation college students, and members of underrepresented ethnic groups often experience stereotype threats on campus that can exacerbate a sense of academic alienation and isolation (Beasly & Fischer 2012). Furthermore, a recent survey of underrepresented minorities interested in STEM disciplines revealed that the extent of previous exposure to science and the presence of individuals of a similar racial or ethnic background currently working in one's chosen field are important factors in determining whether minority students are interested in pursuing a career in that particular field after college (Mead, Clarke, Forcino, & Graves Jr. 2015). Therefore, experiential learning opportunities in ecology and evolutionary biology that actively recruit or are only available to members of underrepresented demographics—such as the GC considered in this study—may help promote solidarity and a sense of belonging among underrepresented students compared to lecture-based courses or traditional field courses, which are often prohibitively expensive or only available to upperclassmen.

Students from the Galápagos Curriculum reported strong cognitive gains in their understanding of basic concepts in evolutionary biology and ecology, which corroborate previous findings on the efficacy of field courses for teaching content (Lisowski & Disinger, 1991; Magntorn & Helldén, 2005; Boyle et al., 2007; Prokop, Tuncer, & Kvasničák, 2007; Easton & Gilburn, 2012; Scott et al., 2012). Previous studies have found that certain underrepresented demographics, particularly

African Americans, tend to view evolution more negatively and have a weaker understanding of general evolutionary concepts, which may be related to a higher prevalence of religiosity among ethnic or racial minority groups (Bailey, Han, Wright, & Graves, 2011; Rissler, Duncan, & Caruso, 2014; Mead et al., 2015). Thus, experiential learning opportunities that emphasize evolutionary concepts through immersion beyond traditional classroom settings may improve scientific literacy and interest in evolution among underrepresented ethnic or racial groups.

Our case study suggests that social gains experienced by students that partake in a travel-based field course are positive, albeit weaker and more variable than cognitive gains. The social benefits of an experiential learning opportunity for a specific individual are likely influenced by a large host of interactive factors. For example, an individual's ability to form positive relationships with their peers and instructors will undoubtedly impact their self-reported social experience. Student and instructor dynamics vary from year to year, such that different individuals may feel more or less comfortable in any given academic group. Group dynamics will also be influenced by the unique set of previous experiences and beliefs of each student. Together, these factors contribute to the variation among self-reported social gains.

In order to facilitate self-reflection and the formulation of synthetic ideas and connections, we assigned field notebooks to our students and required them to periodically reflect on their experience. Based on ethnographic observations, interactions with students, and assessment of field notebooks, we perceived an overall positive effect of field notebooks as facilitators of experiential learning. Students' reflective entries in field notebooks increased promoted connections of course material and personal experiences to their other academic interests and future aspirations. Among geography curricula, field journals have enhanced students' ability to critically reflect on their own learning experience and communicate their thoughts in a field course (McGuinness & Simm, 2005; Dummer, Cook, Parker, Barrett, & Hull, 2008). Therefore, reflective diaries and field journals appear to be an effective tool to facilitate experiential learning in the field across multiple STEM disciplines. However, we noticed that students initially treated the journal as more of a travel diary or activity log rather than a platform for deeper reflection. We recommend that instructors lay out clear guidelines and provide examples of entries from previous years so that student entries meet the instructors' expectations.

As a case study, the Galápagos Curriculum discussed here involves a few nuances that are worth considering, especially in comparison to more traditional field course opportunities at other institutions. First, we were fortunate to be able to underwrite the cost of participation for students. Not every institution of higher education can offer heavily subsidized field courses due to scarcity of funding in the current economy. The Galápagos Islands provide excellent conditions for experiential learning in evolutionary biology: the approachability of endemic wildlife and the breadth (yet manageable time period) of human history in the archipelago creates an excellent opportunity for total immersion. However, the benefits documented here are also translatable to smaller-scale experiential learning opportunities, including shorter field courses or even day trips to local parks or other natural attractions (McLaughlin & Johnson, 2006; Prokop et al., 2007).

While high attrition rates among historically underprivileged groups continue to plague the STEM disciplines, a central objective among colleges and universities has been to promote and maintain diversity in STEM fields. Experiential learning offers a powerful framework to diversify undergraduate curricula and improve student undergraduate performance and volition (Kolb 1984; Freeman et al., 2014). Field courses are a longstanding form of experiential learning in ecology and evolutionary biology; however, many field

courses are disappearing from course offerings, are too expensive, or are not available to first or second-year students. Our study demonstrates that immersive field experiences confer important social and cognitive benefits to first-year undergraduate STEM students from underrepresented demographics. We believe that higher education should support affordable, diverse learning opportunities to better educate students and strengthen social networks among individuals from different academic, socioeconomic, and ethnic backgrounds to promote and sustain diversity in the sciences.

ACKNOWLEDGEMENTS

We would like to thank Lina Arcila and Robin Schwenke who were co-instructors during previous renditions of this field course. Special thanks to Fausto Rodriguez for local expertise and guidance in the Galápagos. We are grateful towards Amita Verma of the Cornell Institutional Review Board with guidance in securing IRB approval for our study. Members of the ALS 6016 course provided constructive feedback on earlier versions of this manuscript.

REFERENCES

Allen-Ramdial SAA, Campbell, AG (2014). Reimagining the Pipeline: Advancing STEM Diversity, Persistence, and Success. Bioscience 64, 612–618.

Airasian PW, Cruikshank KA, Mayer RE, Pintrich PR, Raths J, Wittrock MC, Anderson, LW, Krathwohl DR (2001). A Taxonomy for Learning, Teaching and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives, Boston, MA: Addison Wesley Longman.

Bailey G, Han J, Wright D (2011). Religiously expressed fatalism and the perceived need for science and scientific process to empower agency. Int J Sci Soc 2, 55–87.

Barker S, Slingsby D, Tilling S (2002). Teaching biology outside the classroom: is it heading for extinction? A report on biology fieldwork in the 14–19 curriculum. FSC Occasional Publication 72. Preston Montford, Shropshire: Field Studies Council.

Bauer, DF (1972). Constructing confidence sets using rank statistics. J Am Stat Assoc 67, 687–690.

Beasley MA, Fischer MJ (2012). Why they leave: the impact of stereotype threat on the attrition of women and minorities from science, math and engineering majors. Soc Psychol Educ 15, 427–448.

Bebbington D (2002). Women in science, engineering and technology: A review of the issues. High Educ Quart 56, 360–375.

Bettinger E (2010). To be or not to be: major choices in budding scientists. In American Universities in a Global Market (ed. Charles T. Clotfelter), Chicago, IL: University of Chicago Press.

Bonwell CC, Eison JA (1991). Active Learning: Creating Excitement in the Classroom. ASHEERIC Higher Education Report No. 1, George Washington University, Washington, DC.

Boyle A, Maguire S, Martin A, Milsom C, Nash R, Rawlinson S, et al (2007). Fieldwork is good: the student perception and the affective domain. J Geo High Educ 31, 299–317.

Burtner J (2005). The use of discriminant analysis to investigate the influence of nonč cognitive factors on engineering school persistence. J Eng Educ 94, 335–338.

Chen X, Soldner M (2013). STEM attrition: College students' paths into and out of STEM fields. (NCES 2014-001). Washington, DC: National Center for Educational Statistics.

Clark Blickenstaff, J (2006). Women and science careers: leaky pipeline or gender filter? Gend Educ 17, 369–386.

Clifford J, Marcus G (1986). Writing Culture: The Poetics and Politics of Ethnography. Berkeley, CA: University of California Press.

Conover WJ, Johnson ME, Johnson MM (1981). A comparative study of tests for homogeneity of variances, with applications to the outer continental shelf bidding data. Technometrics 23, 351–361.

Dirks C, Cunningham M (2006) Enhancing diversity in science: is teaching science process skills the answer? Cell Biol Educ 5, 218–226.

Dummer TJB, Cook IG, Parker SL, Barrett GA, Hull AP (2008). Promoting and assessing "deep learning" in geography fieldwork: an evaluation of reflective field diaries. J Geo High Educ 32, 459–479.

Easton E, Gilburn A (2012). The field course effect: gains in cognitive learning in undergraduate biology students following a field course. J Biol Educ 46, 29–35.

Eisner T (1982). For love of nature: exploration and discovery at biological field stations. Bioscience 32, 321–326.

Fenske RH, Porter JD, DuBrock CP (2000). Tracking financial aid and persistence of women, minority, and needy students in science, engineering, and mathematics. Res

High Educ 41, 67–94.

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

Griffith AL (2010). Persistence of women and minorities in STEM field majors: Is it the school that matters? Econ Educ Rev 29, 911–922.

Hammersley M, Atkinson P (2007).

Ethnography: Principles in practice. London, England: Routledge.

Herman SG (1986). The Naturalist's Field Journal. Vermilion.

Hunter AB, Laursen SL, Seymour E (2007). Becoming a scientist: The role of undergraduate research in students' cognitive, personal, and professional development. Sci Educ 91, 36–74.

Kokkelenberg EC, Sinha E (2010). Economics of Education Review. Econ Educ Rev 29, 935–946.

Kolb AY, Kolb DA (2005). Learning styles and learning spaces: Enhancing experiential learning in higher education. Acad Manag Learn Educ 4, 193–212.

Kolb DA (1984). Experiential learning: Experience as the source of learning and development, Upper Saddle River, NJ: Prentice-Hall Inc.

Krathwohl DR (2002). A revision of Bloom's taxonomy: an overview. Theor Pract 41, 212–218.

Lam PC, Srivatsan T, Doverspike D, Vesalo J, Mawasha PR (2005). A ten year assessment of the pre-engineering program for under-represented, low income and/or first generation college students at the University of Akron. J STEM Educ 6, 4.

Lisowski M, Disinger JF (1991). The Effect of Field-Based Instruction on Student Understandings of Ecological Concepts. J Environ Educ 23, 19–23.

Lock R (2010). Biology fieldwork in schools and colleges in the UK: an analysis of empirical research from 1963 to 2009. J Biol Educ 44, 58–64.

Lopatto D (2007). Undergraduate research experiences support science career decisions and active learning. Cell Biol Educ 6, 297–306.

Magntorn O, Helldén G (2005). Studentč teachers' ability to read nature: reflections on their own learning in ecology. Int J Sci Educ 27, 1229–1254.

Marvell A (2008). Student-led presentations in situ: the challenges to presenting on the edge of a volcano. J Geo High Educ 32, 321–335.

Mathison S (1988). Why triangulate? Educ Res 17, 13–17.

Matsui J, Liu R, Kane CM (2003). Evaluating a science diversity program at UC Berkeley: more questions than answers. Cell Biol Educ 2, 117–121.

McGuinness M, Simm D (2005). Going global? Long-haul fieldwork in undergraduate geography. J Geo High Educ 29, 241–253.

McLaughlin J (2005). Classrooms without walls: a banana plantation, the random fallen tree, and a turtle nest. Int Educ 14, 52¬–54.

McLaughlin JS, Johnson DK (2006). Assessing the field course experientiallLearning model: transforming collegiate short-term study abroad experiences into rich learning environments. Frontiers 13, 65–85.

Mead LS (2015). Factors influencing minority student decisions to consider a career in evolutionary biology. Evol Educ Outreach 8:6, 1–11.

Mendez G, Buskirk TD, Lohr S (2008). Factors associated with persistence in science and engineering majors: An exploratory study using classification trees and random forests. J Eng Educ 97, 57–70.

Prokop P, Tuncer G, Kvasni Glyphčák R (2007).

Short-term effects of field programme on students' knowledge and attitude toward biology: a slovak experience. J Sci Educ Tech 16, 247–255.

Rask K (2010). Attrition in STEM fields at a liberal arts college: The importance of grades and pre-collegiate preferences. Econ Educ Rev 29, 892–900.

Rathburn SL, Weinberg AE (2011). Undergraduate student satisfaction and achievement at the GetWET Observatory: A fluid learning experience at Colorado State University. J Geo Educ 59, 47–55.

Rissler LJ, Duncan SI, Caruso NM (2014). The relative importance of religion and education on university students' views of evolution in the Deep South and state science standards across the United States, Evol Educ Outreach 7, 1–17.

Russell SH, Hancock MP, McCullough J (2007). The pipeline: benefits of undergraduate research experiences. Science 316, 548–549.

Sanders DL (2007). Making public the private life of plants: the contribution of informal learning environments. Int J Sci Educ 29, 1209–1228.

Scott GW, Boyd M, Scott L, Derek C (2015). Barriers to biological fieldwork: what really prevents teaching out of doors? J Biol Educ 49, 165–178.

Scott GW, Goulder R, Wheeler P, Scott LJ, Tobin ML, Marsham S (2012). The value of fieldwork in life and environmental sciences in the context of higher education: a case study in learning about biodiversity. J Sci Educ Tech 21, 11–21.

Seymour E, Hewitt NM (2000). Talking about leaving: Why undergraduates leave the sciences, Boulder, CO: Westview Press.

Smith D (2004). Issues and trends in higher education biology fieldwork. J Biol Educ 39, 6–10.

Summers MF, Hrabowski FA III (2006). Preparing minority scientists and engineers. Science 311, 1870–1871.

Tate KA, Caperton W, Kaiser D, Pruitt NT, White H, Hall E (2015). An exploration of first-generation college students' career development beliefs and experiences. J Career Dev 42, 294–310.

Tinto V (2006). Research and practice of student retention: what next? J Coll Stud Ret 8, 1–19.

Tsui L (2007). Effective strategies to increase diversity in STEM fields: A review of the research literature. J Negro Educ 76, 555–581.

Whalen DF, Shelley MC (2010). Academic Success for STEM and Non-STEM Majors. J Stem Educ 11,45–60.

Willis P, Trondman M (2000). Manifesto for Ethnography. Ethnography 1, 5–16.

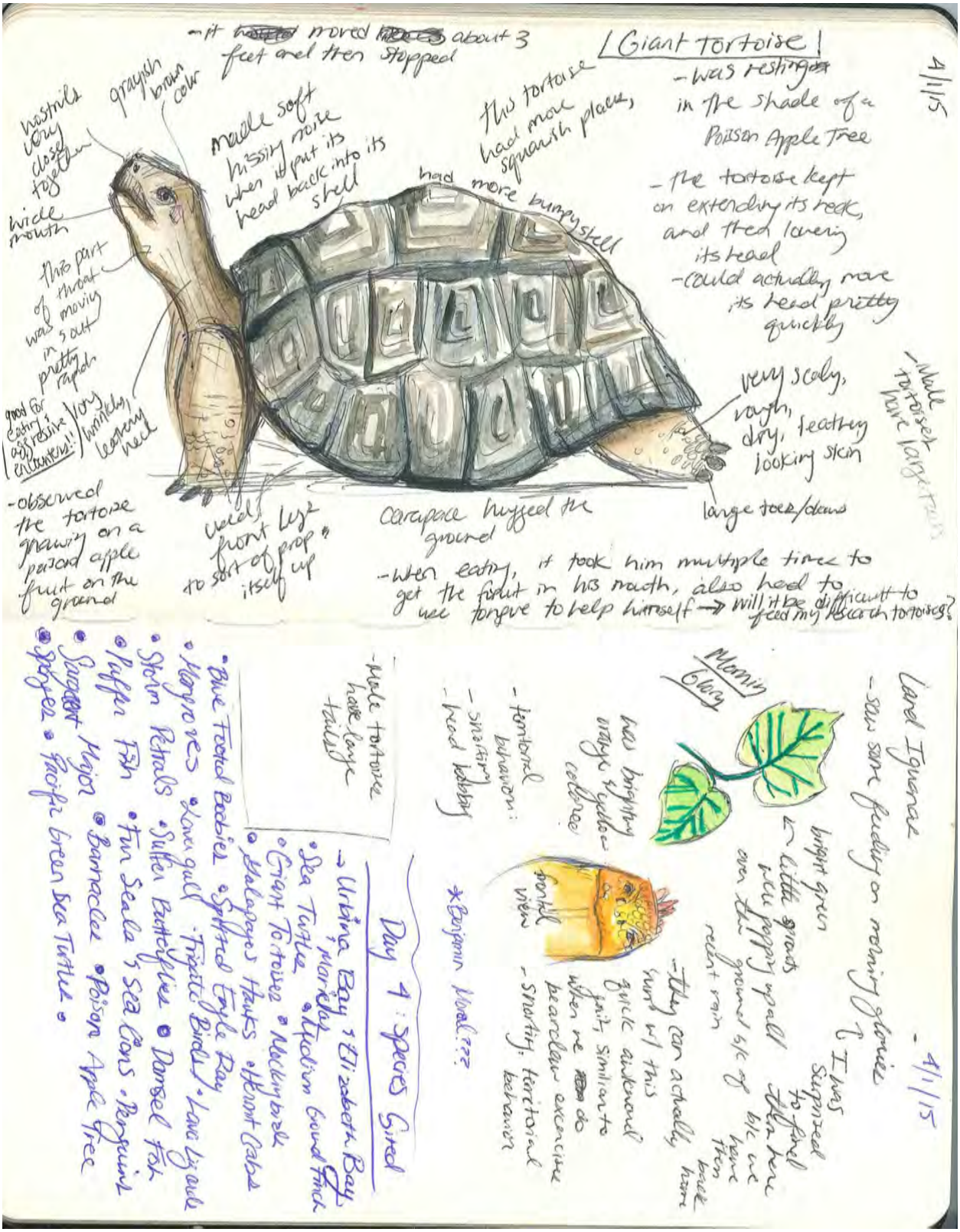
Zervanos SM, McLaughlin JS (2003). Teaching biodiversity & evolution through travel course experiences. Am Biol Teach 65, 683–688.

APPENDIX FIGURE LEGENDS

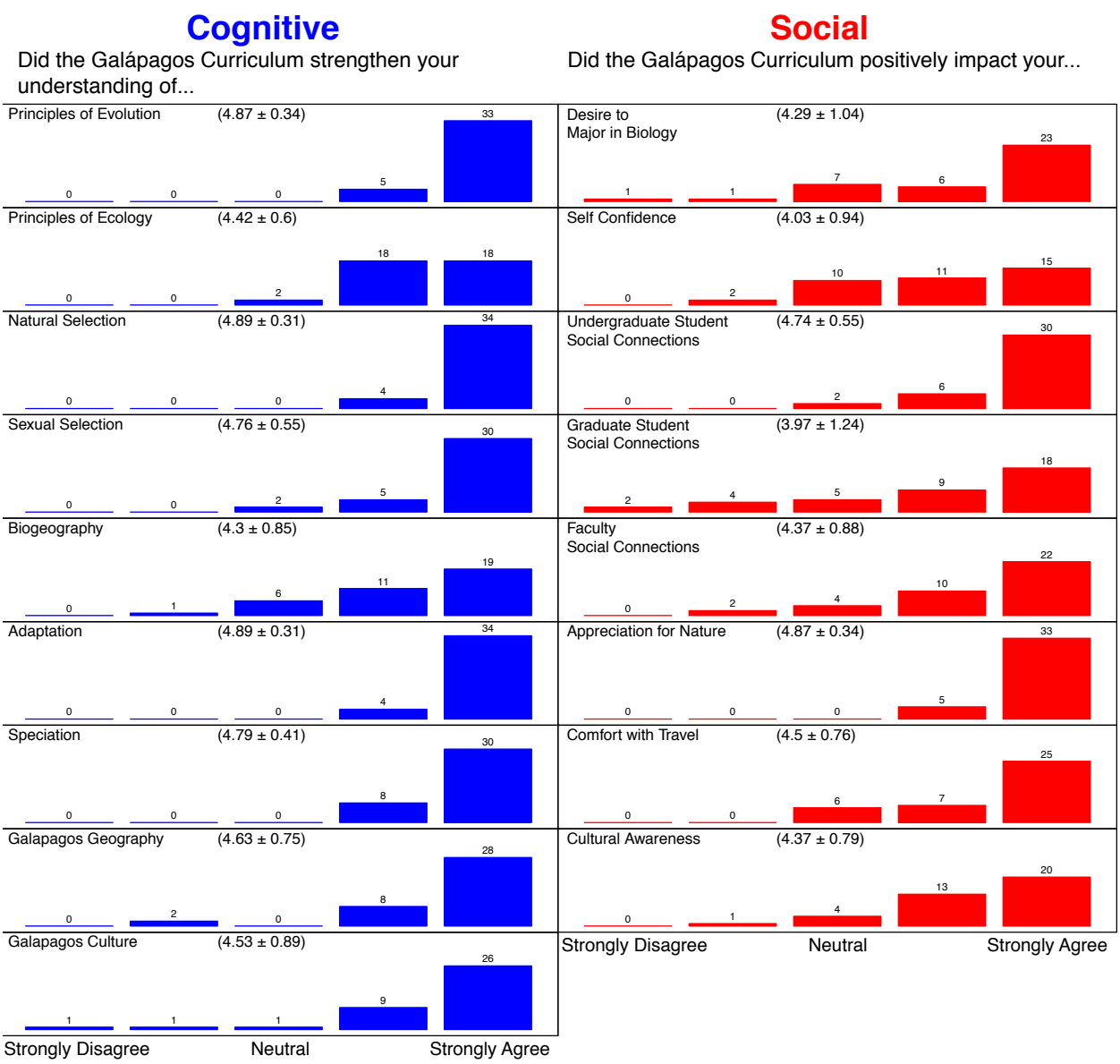
Appendix Figure 1: Exemplary student entry in a field journal showing detailed observations of a close encounter with a giant tortoise and a land iguana.

Appendix Figure 2: Post-trip survey responses from 38 (out of 44) participants. Participants responded to questions using a Likert-scale, in which 1 corresponded to strongly disagree, 3 corresponded to neutral, and 5 corresponded to strongly agree. The number of responses for each Likert score is indicated above each bar. The mean and the standard deviation for each question are also displayed.

Appendix Figure 1



Appendix Figure 2



The Effects of Peer- and Self-assessment on the Assessors

Joonsuk Park

Graduate Research and Teaching Fellow 2014-2015

ABSTRACT

Recently, there has been a growing interest in peer- and self-assessment (PSA) in the research community, especially with the development of massive open online courses (MOOCs). One prevalent theme in the literature is the consideration of PSA as a partial or full replacement for traditional assessments performed by the instructor. And since the traditional role of the students is as the assess in the assessment process, existing works on PSA typically focus on devising methods to make the grades more reliable and beneficial for the assessees.

What has been missing in the picture is the assessor: How are those conducting peer- and self-assessment impacted by the process? This question has become relevant from an educational perspective because in PSA, the students take on the role of the assessor as well.

We present PSA as an active learning exercise for the assessors and examine its impact. In order to do this, we incorporated PSA into a university-level Introduction to Natural Language Processing (NLP) course consisting of more than 100 students and analyzed student surveys and exam results of peer-, self-, and no-assessment groups. The final exam performance suggests that PSA is helpful for learning, which is consistent with the student survey results. Also, students generally enjoyed conducting PSA.

INTRODUCTION

Assessing the quality of student work, whether summative or formative, has long been a core element in most classes. It serves as feedback for both students and instructors, with the following potential uses, among others: Students can distinguish concepts they understand well from those with which they are less familiar and allot their time accordingly. Also, course instructors can gain a better estimate of the level of understanding their students have, which can be valuable when adjusting the content or the structure of the current or future classes to better accommodate the students' needs.

Conducting assessments, however, can be labor intensive. This is especially the case for Massive Open Online Courses (MOOCs), in which a handful of staff[1] have to support tens of thousands of students. The need for a scalable means of performing assessments has led to active research on peer- and self-assessment (PSA), the act of providing grades and feedback

[1] "staff" refers to the instructors and teaching assistants.

on peers' and one's own work [16, 17]. Because existing works on PSA typically focuses on how PSA would partially or fully replace traditional assessment methods, they tend to examine the benefits for the assessees and methods to achieve reliable grades [9, 12].

While an immediate goal of assessments is evaluating the quality of student work, performing assessments has the additional benefits of promoting learning for the assessor. Though several researchers have reported the benefits of PSA for the assessor, this is still an understudied aspect of PSA with a limited number of studies based on rigorous experiments [15, 7, 3, 14].

A more systematic study is necessary to better understand the impact of PSA on the assessors. This will allow the instructors to decide the time and resources that should be reserved for PSA in their courses. Also, positive findings will motivate the students to perform assessments more carefully, which can benefit both the assessees and the assessors [18].

In this paper, we detail our effort to analyze the effects of PSA on the assessors. Specifically, we compare the exam performance of more than 100 students after they conduct peer-, self-, and no-assessment in a semester long course on natural language processing (NLP). In addition, we analyze student surveys on PSA. The final exam performance suggests that PSA is helpful for learning, which is consistent with the survey in which students reported that they found PSA beneficial. Also, students generally enjoyed conducting PSA.

RELATED WORK

PSA and peer-assessment without self-assessment have been studied in many domains, including more open-ended tasks like writing [4] and oral presentation [8], as well as more objective tasks like computer programming [14] and algorithms [3]. In this project, PSA was conducted on problem sets consisting of conceptual and computational questions requiring knowledge in linguistics and probability theory. The questions are quite objective in that the questions mostly had definite answers, though multiple approaches may be possible depending on the question.

Different researchers are concerned with different aspects of PSA. For instance, Bauer et al. study how students feel about online vs in-person assessment [1]. With the advent of MOOCs, where the size of staff is significantly smaller than that of the students, research efforts have focused on ways to acquire reliable grades from peers [9, 12]. Also, many have concentrated on the impact of PSA on the assessees [10]. Some have reported that the assessors benefit from PSA as well [18]. However, not many systematic studies have been conducted to analyze how the assessors are affected as this project does.

The work by Chinn is one of the few works that look at PSA from the perspective of the assessors [3]. Chinn reports that the quality of peer-assessment and exam performance are correlated, though the direction of causality cannot be determined. In this project, we do not take the quality of assessments into consideration. This is because we are interested in the impact of PSA regardless of the assessors' levels of understanding, which is one of the factors influencing the quality of assessments.

By viewing PSA from the perspective of the assessors, we are effectively treating PSA as an active learning exercise for the assessors. Novel active learning strategies are often developed for classroom settings, where the goal is to make the traditional lecture style classes more active and engaging for the students [2, 11, 13]. Our implementation of PSA, however, takes place outside the classroom and thus is complementary to the active learning activities performed in classrooms.

METHODOLOGY

Overview

This work is based on experiments and observations made in a university-level

Introduction to Natural Language Processing course. There were over 100 students enrolled, the majority of them junior and senior level undergraduate and master's students, mostly majoring in Computer Science.

The experiment consisted of two cycles of problem set, PSA, and survey, followed by the final exam. The problem sets and the final exam covered topics as shown in Table 1. Note that the final exam also covered topics that were not covered in the problem sets, such as those covered in programming projects. These, however, are not mentioned in the table, as they are not part of the experiment.

Of the 103 students who completed the course, 94 students completed PSA for both Problem Set 1 and 2, and their final exam performance is used for the quantitative analysis in Section 4. Also, 67 students completed the survey after performing PSA for Problem Set 1, and 71 after performing PSA for Problem Set 2.

Topic*	Problem Set 1	Problem Set 2	Final Exam
LM	O		o
POS	O		o
WSD	O		o
CD		O	o
(P)CFG		O	o
Parsing		O	o

Table 1: Topics Covered in Problem Sets and Exams

**LM : Language Models, POS : Part-of-Speech Tagging, WSD : Word Sense Disambiguation, CD : Computational Discourse, (P)CFG : (Probabilistic) Context-Free Grammar, Parsing : Syntactic Parsing.*

Experiment Components

Problem Set

The students individually completed problem sets, each consisting of problems covering three topics that later appeared in the exam. The questions were designed to be as objective and clear as possible, involving concepts mostly from linguistics and probability. However, because the subject involves language and real-world applications which are inherently subjective and open-ended, the questions still had room for variations in approach.

The grade was broken down with an intention to motivate the students to perform PSA with more care: The peer- and self-assessments students conducted were graded by the staff to create an extrinsic motivation for students to get good grades. The peer-assessment results accounted for the peers' grades and provided intrinsic motivation for peers to be fair assessors. Numerically, 10% of the grade is based on the quality of the PSA, judged by the staff, and the remaining 90% on the solutions to the problem set: 81% is from the staff's grade, and 9% is from the average of peer-assessment grades.

Peer- and Self-assessment

The students performed double-blinded PSA on problem set submissions using a solution manual and a brief grading rubric provided for them soon after the problem set deadline. For each question being assessed, the students were to give a numeric score and written feedback.

In order to compare the effects of peer-, self-, and no-assessment on the exam performances, we needed to have students perform different types of assessment on the problem set submissions. One approach we considered was to split students into three groups, each with a designated type of assessment. This approach, however, couples assessment types and individuals too tightly: even if a statistically significant difference is observed in the final exam performance among the groups, it would not be distinguishable whether it is caused by the effectiveness of the assessment methods or the abilities of the students who happened to be assigned to the groups. Also, the students in the peer-assessment group would have to do much more work than those in other groups. Similarly, as the difficulty of the topics could vary, we did not want to assign specific assessment types for each topic.

To minimize the impact of the variations in

student abilities and topic difficulties on the experiment results, we set up PSA in a way that decouples individuals and topics from the assessment types: each student was assigned to a group with a designated topic to peer-assess (six peers' work) and another topic to self-assess, leaving the remaining topic for no-assessment. For instance, students in PSA Group 1 for Problem Set 1 peer-assessed LM questions for six peers and self-graded POS questions, and did no-assessment for WSD questions.

While the assignment to the groups was done randomly, the number of students in each group was kept as uniform as possible. This means that for each topic, there were roughly the same number of students performing each type of assessment. Under this setup, the exam result for each student-topic pair is associated with a single assessment type, and the final exam results can be clustered by the assessment type. This way, each student contributes to all three assessment types evenly, and so does each topic, minimizing the impact of the variations in student abilities and topic difficulties on the experiment results. In analysis, each score is divided by the best student score instead of the maximum possible score to partially control for the variations in topic difficulties.

Survey.
After completing PSA, the students filled out a post PSA survey. As it was largely voluntary, roughly 30% of students neglected to participate in the survey, as mentioned in the beginning of this section.

Most questions were 5-level Likert items, where a statement is presented and the respondent selects from one of five choices ranging from Strongly Disagree to Strongly Agree. Here are the question prompts presented :

1. Grader X's grades are fair and Grader X's feedback is helpful for each of the six peer assessors that assessed the given respondent's problem set.

2. My grades are fair and my feedback should be helpful.

3. Performing PSA helped me learn the materials better

4. Knowing my peers' level of understanding gave me more confidence

5. Performing PSA was beneficial overall.

6. I self-assessed every question, regardless of the PSA assignment.

Items 1 and 2 were presented to compare the peer- and self-reported quality of the assessments, items 3 - 5 were presented to measure the benefits of PSA, and item 6 was presented to determine how PSA was executed in real life.[2]

In addition to these Likert items, there were short questions asking for reasons for their choices on Likert items, as well as any comments about their experience and opinions about PSA.

Exam
After completing two cycles consisting of a problem set, PSA, and survey, students took the final exam. For each student, the grade with respect to each topic was recorded separately.

As discussed, each student's final exam performance on each topic is associated with an assessment type. We group the scores based on the assessment type to form three clusters of scores corresponding to peer-, self- and no-assessment, where each score is uniquely associated with a student and a topic. Then, the average scores of these score clusters are compared to measure the effectiveness of the assessment types.

Standard tests of statistical significance such as the (paired) t-test are not applicable in this case because the dependence among the scores

[2] We were concerned that students may self-grade every topic, not just assigned ones, because that would break the experiment setup in which scores are clustered by the assessment method. See Peer- vs Self-assessment in Section 4 for more discussion.

is a bit idiosyncratic: given two assessment types under comparison, you can pair scores by the student, but the scores themselves are on different topics. Thus, we chose to take the generally applicable bootstrapping approach, where 10,000 datasets are created via sampling with replacement and the percentage of the time the same or greater difference is observed is used as the estimate for the p-value [6].

RESULT 1 : EFFICACY OF PSA

The averages of the scores of each assessment type are presented in Table 2. Also, note that indented quotes that appear throughout the result sections are exemplar quotes from the student surveys. These quotes serve as anecdotal evidence for the qualitative analysis portion of the study.

Assessment Type	Peer	Self	No
Mean Score	74.6%*	74.8%*	72.8%
Standard Deviation	11.4%	10.6%	12.3%

Table 2: Comparison of Final Exam Performance

** The performance gain over no-assessment is statistically significant ($p<0.05$).*

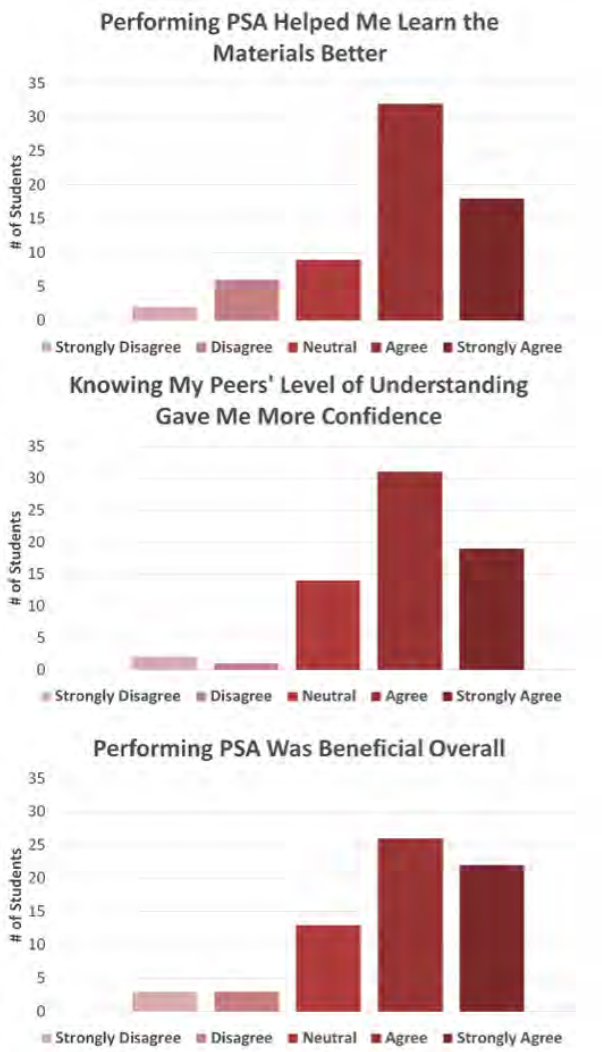


Figure 1: Post-PSA Survey Results

Peer- and Self- vs No-assessment

Peer- and self-assessment scores are higher than no-assessment scores. The difference is small but statistically significant. This is consistent with our hypothesis and student survey as shown in Figure 1. There are several benefits to PSA:

1. PSA makes students review the topics soon after the problem sets are due:

> "[...] I enjoyed the peer grading experience because in most classes, after an assignment is due, most people just forget it since "the work is done."

As depicted by Ebbinghaus's forgetting curve, frequent reviewing is crucial for retaining newly acquired knowledge, though the frequency can be lowered over time [5]. Performing PSA soon after the problem set deadline prevents the students form missing early review opportunities.

2. PSA exposes students to approaches different from the ones they are familiar with, which may or may not be correct:

> "I actually got more out of grading peers work than my own because it allowed me to see other ways of approaching the same

problems."

"It was a nice experience getting to know how others think. Also, it was a nice opportunity to see common mistakes."

The students have to reason through various approaches to evaluate their correctness. Apart from helping students gain multiple perspectives on a given problem, such critical thinking should also be helpful in learning more generally.

3. Students have additional intrinsic motivation to carefully study the concepts because they want to be fair to their peers:

"Personally, I understood the concepts even better than before because I studied it more in depth just so that I could grade accurately."

Everyone is driven by unique motivations. Also, existing motivations, such as grades, are mostly extrinsic in nature. Providing an additional source for intrinsic motivation in this way can fulfill the needs of certain students.

4. PSA satisfies students' curiosity:

"It's always interesting to see others' answers."

Students gain access to other students' work through PSA. Satiating the curiosity, along with other benefits, may keep them interested in the subject.

Peer- vs Self-assessment

What is less intuitive about the average exam scores from the experiment is that peer-assessment is no more beneficial than self-assessment. Note that benefit 2 and 3 above are effects of peer-assessment, specifically. Even then, why is there no significant difference in the exam scores between peer- and self-assessment groups?

One reason is that peer-assessment may not always be more beneficial than self-assessment. This is covered in more detail as we discuss the third section, When PSA Fails to Be Beneficial.

Also, students may have done self-assessment for the topic assigned for peer-assessment. The average response to survey item 6, "I self-assessed every question, regardless of the PSA assignment," was 3.03. In this case, 1 means that the given student self-assessed neither of the 2 remaining topics, and 5 means that they self-assessed both of the topics. Since 3.03 sits right in the middle of the scale, we can see that the students self-assessed 1 of the 2 remaining topics on average. Given the statistically significant difference in the average exam scores between no-assessment and the rest, we conjecture that the 1 additional self-assessment was performed on the topic the given student was assigned to peer-assess. This sounds reasonable as well because this is the topic that requires less additional work—after all, the student has already assessed 6 peers' work on that topic and is familiar with the correct solution. But it is also possible that the additional, voluntary self-assessment was done evenly across different topics, in which case the performance gain over no-assessment would be bigger in real life than what is shown in Table 2.

When PSA Fails to Be Beneficial

Just like any learning exercise, PSA is not too helpful when the assessor understands the concepts well already:

"I don't think it added much to the process given that I had already spent so long on the problems and for the most part understood the approaches."

The average final exam scores of the students who marked (strongly) disagreed for survey item 5, "Performing PSA Was Beneficial Overall," was significantly higher than that of those who (strongly) agreed with the statement, 81.83% and 70.2%, respectively. This suggests that students who think they know the concepts too well to benefit from an additional assignment like PSA indeed know the concepts well, instead of wrongly thinking that they do. However, it is still unclear whether they really did not benefit from PSA.

Also, when the assessee has a strong command over the materials and their solutions resemble those in the solution manual closely, there is little for the assessor to gain from PSA:

> "[...] An assignment with many mistakes will force us to think what goes wrong and look for place for partial credit if any. However, with perfect homework, one can learn nothing from peer grading."

In an extreme case, the assessee's solution could be identical to the sample solution. In such cases, though, the assessor has not much to lose either, except for perhaps an opportunity cost, since it will take little to no time to assess the student work.

Before we discuss the third case in which PSA is not helpful for the assessor, consider the following comment:

> "[...] On the last homework I found it useful to repeatedly go through the process of seeing how they arrived at answers. On this homework however, some of the peer solutions (while still with the correct answer) may not have solved it the same way I did. Thinking about multiple ways to solve a problem can in some ways make it more confusing, as I have a specific way most comfortable for me to correctly solve such problems."

This brings up an important point: To benefit from PSA, the student assessors need to have a sufficient level of command over the relevant concepts. They need to have the capacity to critically analyze and determine the correctness of a peer's approach that is different from the sample solution. Also, they need to be ready to absorb multiple correct approaches without getting them confused.

The level of sufficiency depends on how open-ended the question is. The more open-ended the question is, the more likely it is for the peers to take approaches that are divergent from the sample solution provided for them. Thus, should the instructor prepare questions with little to no room for alternative approaches? In such cases, however, there may be little benefit for the assessors because it is more likely to be in the aforementioned cases. Thus, the degree of subjectivity has to be carefully determined to balance the amount of benefit that can be gained from PSA and the level of understanding required to reap the benefits.

In our case, most students found PSA helpful, but some were ready for more challenges:

> "It's a lot more useful for assignments where there is more subjective grading criteria, such as essays. I'd rather peer grade the projects than the homeworks, honestly, since there's a lot more room for diverse solutions in the former."

RESULT 2 : STUDENT SATISFACTION

> "This is my first time to do things like PeerGrade. This is really helpful!"

> "Surprisingly, I enjoyed peer grading."

> "Hopefully we do this again!"

Students are generally happy about PSA, perhaps because they found PSA to be helpful, as discussed in the previous section. But there are other factors that contributed to the generally positive sentiment toward PSA.

Workload

When incorporating PSA into our class, we were very cautious about keeping the workload reasonable because it requires work in addition to all other assignments and responsibilities asked of the students. The students seem to appreciate it:

> "I'm glad that you guys gave us a reasonable chunk of peer grading since it didn't take too long."

Is it necessarily good that students are happy about the workload? What additional benefits would a more demanding implementation of PSA bring? These are interesting and practical questions that cannot be answered with our results. But what we can claim with our experience is that with a manageable workload,

PSA can be enjoyable to the students, even though they know that it requires extra work.

Perceived Assessment Quality

Table 3 summarizes how the students thought about the quality of the assessments they provided and received. Students are generally satisfied with the numeric grades and written feedback they received from the peers, which may partially explain the positive sentiment toward PSA. While this is not generalizable to all assessors, some assessors are quite meticulous, in a way that staff typically cannot afford to be due to limited resources:

> "I liked the way peers have explained the flaws. They have pointed out minute details of my homework. It definitely helps me to do better in my next home work. Overall, they have done a good job!"

Even if a student is assigned less meticulous peer-assessors for a given problem, he or she still receives feedback from multiple people, including a staff member. Thus, the feedback can collectively cover more aspects of the student work than it would in a setting without PSA.

Also, some peers tend to be more lenient than the staff. Again, this depends on the assessor, but perhaps some are lazy, do not understand the concept well enough to take off points and argue a case for it, or just want to be lenient to people in the same boat as they are:

> "[...] I did find that overall instructors were a bit harsher than peers in grading"

Judged by	*PSA for PSet 1*		*PSA for PSet 2*	
	Grade	*Feedback*	*Grade*	*Feedback*
Assessor	4.65(.57)	4.56(.59)	4.72(.70)	4.71(.69)
Assessee	4.61(.82)	4.48(.85)	4.72(.71)	4.72(.66)

Table 3: Average Survey Response Scores on the Quality of Peer-assessment (SD in parentheses)

The responses were given on a scale from 1 (Strongly Disagree) to 5 (Strongly Agree). Each value on the "Self" row is the average of self-reported quality scores of the assessments provided to the peers (e.g. "My feedback should be helpful" and "My grades are fair"). Each value on the "Peers" row is the average of quality scores of the assessments received from the peers (e.g. "Grader X's feedback is helpful" and "Grader X's grades are fair").

THREATS TO VALIDITY

This work details the experience and results from a study done in an NLP course with primarily senior-level students. Two of the implications are as follows:

First, the good balance between objectivity and the open-ended nature of the questions may make NLP a good fit for PSA, which in turn may have led to the promising results observed in this study. For instance, in language modeling, a question may ask the students to assign a probability to a sentence. Roughly speaking, students are to estimate the probability of seeing the given sentence in an unseen document, based on the number of occurrences of the substrings of the sentence in a provided corpus. Depending on the choice of preprocessing and smoothing, the answer may vary. However, given a complete description of an assessee's approach, there is little ambiguity in the steps and result. How effective would PSA be in other subjects? While this cannot be directly answered, the discussion at the end of Section 4 remains relevant.

Second, the maturity of the students might have had a significant role in making the overall experience beneficial and pleasant. It takes knowledge and character to evaluate others' work and provide constructive feedback. Can we expect to see similar maturity in younger students? To answer this question, additional experiments with students in earlier stages of their education is necessary.

CONCLUSIONS

In this study, we presented PSA as an active learning exercise for the assessors and studied its impact. After incorporating PSA into an Introduction to Natural Language Processing

course consisting of more than 100 students, we analyzed student surveys and exam results of a peer-, self-, and no-assessment group. The majority of the students found peer- and self-assessment to be helpful according to the survey, which is confirmed by the exam results. The final exam performance suggests that PSA is helpful for learning, which is consistent with the survey in which students reported that they found PSA beneficial. Also, students generally enjoyed conducting PSA.

This work compliments the vast literature on PSA that focuses on how the assessees benefit from PSA. It also provides additional motivations for students to perform more careful assessments for peers and themselves. For future work, we would like to explore different design decisions in implementing PSA in classes to maximize the benefit we see in this work.

REFERENCES

[1] C. Bauer, K. Figl, M. Derntl, P. P. Beran, and S. Kabicher. The student view on online peer reviews. In *Proceedings of the 14th Annual ACM SIGCSE Conference on Innovation and Technology in Computer Science Education*, ITiCSE '09, pages 26–30, New York, NY, USA, 2009. ACM.

[2] C. C. Bonwell and J. A. Eison. Active learning : creating excitement in the classroom. *ASHEERIC Higher Education Report No. 1*, 1991.

[3] D. Chinn. Peer assessment in the algorithms course. *In Proceedings of the 10th Annual SIGCSE Conference on Innovation and Technology in Computer Science Education,* ITiCSE '05, pages 69–73, New York, NY, USA, 2005. ACM.

[4] A. DiPardo and S. W. Freedman. Peer response groups in the writing classroom: Theoretic foundations and new directions. *Review of Educational Research*, 58(2):pp. 119–149, 1988.

[5] H. Ebbinghaus. Uber das gedachtnis leipzig, dunker. Translated by Ruger HA and Bussenius C*E (1913) Teachers College, Columbia University.[WWW document] http:// psychclassics. yorku. ca/Ebbinghaus/index. htm (last accessed August 16 2004)*, 1885.

[6] B. Efron and R. Tibshirani. An Introduction to the Bootstrap, 1993.

[7] S. Fallows and B. Chandramohan. Multiple approaches to assessment: Reflections on use of tutor, peer and self-assessment. *Teaching in Higher Education*, 6(2):229–246, 2001.

[8] D. Grez, Valcke, and Berings. Peer assessment of oral presentation skills. *Procedia - Social and Behavioral Sciences*, 2(2):1776 – 1780, 2010. Innovation and Creativity in Education.

[9] C. Kulkarni, K. P. Wei, H. Le, D. Chia, K. Papadopoulos, J. Cheng, D. Koller, and S. R. Klemmer. Peer and self assessment in massive online classes. *ACM Trans. Comput.-Hum. Interact.*, 20(6):33:1–33:31, Dec. 2013.

[10] N.-F. Liu and D. Carless. Peer feedback: the learning element of peer assessment. *Teaching in Higher Education,* 11(3):279–290, 2006.

[11] M. Prince. Does active learning work? a review of the research. *J. Engr. Education*, pages 223–231, 2004.

[12] K. Raman and T. Joachims. Bayesian ordinal peer grading. In *NIPS 2014 Workshop on Human-Propelled Machine Learning,* 2014.

[13] B. Simon, R. Anderson, C. Hoyer, and J. Su. Preliminary experiences with a tablet pc based system to support active learning in computer science courses. *SIGCSE Bull., 36(3):213–217*, June 2004.

[14] J. Sitthiworachart and M. Joy. Effective peer assessment for learning computer programming. In *Proceedings of the 9th Annual SIGCSE Conference on Innovation and Technology in Computer Science Education*, ITiCSE '04, pages 122–126, New York, NY, USA, 2004. ACM.

[15] D. Sluijsmans, F. Dochy, and G. Moerkerke. Creating a learning environment by using self-, peer- and co-assessment. *Learning Environments Research*, 1(3):293–319, 1998.

[16] K. Topping. Peer assessment between students in colleges and universities. *Review of Educational Research*, 68(3):pp. 249–276, 1998.

[17] K. Topping. Self and peer assessment in school and university: Reliability, validity and utility. In M. Segers, F. Dochy, and E. Cascallar, editors, *Optimising New Modes of Assessment: In Search of Qualities and Standards*, volume 1 of *Innovation and Change in Professional Education*, pages 55–87. Springer Netherlands, 2003.

[18] I. van den Berg, W. Admiraal, and A. Pilot. Design principles and outcomes of peer assessment in higher education. *Studies in Higher Education*, 31(3):341–356, 2006.

Teaching Students to Engage With Evidence: An Evaluation of Structured Writing and Classroom Discussion Strategies

Steffen Blings and Sarah R. Maxey

Graduate Research and Teaching Fellows 2014-2015

Note: A published version of this research is available in the Journal of Political Science Education. The full citation is

Steffen Blings & Sarah Maxey (2016): Teaching Students to Engage with Evidence: An Evaluation of Structured Writing and Classroom Discussion Strategies, *Journal of Political Science Education*, DOI: 10.1080/15512169.2016.1168303

To link to this article: *http://dx.doi.org/10.1080/15512169.2016.1168303*

ABSTRACT:

In their transition to college, students often struggle to identify and make connections between the main arguments, evidence, and empirical findings of articles from academic journals commonly assigned on political science syllabi. Which active learning techniques are most effective for teaching students to recognize and evaluate social science evidence?

To address this question, we conducted an experiment with students from two first-year writing seminars in political science. Students were randomly assigned to either an in-class writing activity or a group discussion, both of which required them to summarize the article's use of evidence. We found limited evidence that group discussions are more effective for teaching students to engage with evidence. The effects of discussions may be linked to the classroom environment, as students who participated in the group discussion with a familiar instructor were more likely to correctly identify the article's evidence.

INTRODUCTION

A central challenge in students' transition to college is their ability to critically evaluate and construct arguments. Our experience in the classroom suggests that students struggle to identify and make connections between the main arguments, evidence, and empirical findings of articles from academic journals commonly assigned on political

The authors would like to thank Matthew Evangelista, Adam Levine, Runjini Raman, Kim Williams, participants in the PREPARE colloquium and Classroom Research and Teaching Symposium at Cornell University, and students in their first-year writing seminars for their valuable comments and contributions to the paper. Any remaining errors are our own.

science syllabi. Teaching students to critically evaluate information forms the core of liberal arts education and is critical not only for the scientific study of politics, but also for students as citizens, and for those hoping to enter careers that require analyzing data and arguments. Active learning strategies provide tools to increase students' comprehension and critical thinking, but little is known about which active learning strategies work best for teaching different skill sets. We ask: Which active learning techniques are most effective for teaching students to recognize and evaluate social science evidence?

In this article, we present results from an experiment conducted with students from two first-year writing seminars in political science to investigate two of the most common and easily implemented active learning strategies: in-class group discussion and structured writing exercises. In particular, the paper analyzes the relative influence of discussion and writing on students' abilities to identify and evaluate the evidence presented in Heaney and Rojas' (2011) "Partisan Dynamics of Contention," which is typical of the scholarship students are likely to encounter in social science journals. We also consider factors such as previous exposure to the material and comfort in the classroom that may condition the effectiveness of different strategies.

We begin by embedding our study in existing scholarship on teaching evidence and active learning exercises. The following section describes the experimental setup and indicators for student comprehension. The next outlines the results of the experiment and the final two sections consider the effect of classroom context and the implications of our findings.

EXISTING RESEARCH

Teaching students how to evaluate information and become critical thinkers is central to higher education in general and to the study of political science in particular. As Atwater (1991) explains, critical thinking requires students to identify and evaluate the strength of arguments and to construct arguments of their own. Students' abilities to critically engage with causal claims and evaluate authors' evidence are especially important as technology increases the number of sources and density of information available. Despite the relevance of critical thinking to the goals of higher education and the creation of informed citizens, studies consistently show that students struggle to understand how evidence contributes to the broader argument (Fitzgerald and Baird 2011). As a result, students have difficultly integrating evidence into their own arguments (Çavdar and Doe 2012), both in the classroom and as a part of writing assignments.

Active learning techniques offer one promising way to improve students' critical thinking skills (Barr and Tagg 1995; Bean 2011; Bonwell and Sutherland 1996; Burch 2000; Lantis, Kuzma, and Boehrer 2000; Meyers and Jones 1993; Powner and Allendoerfer 2008). Higher education's shift towards active learning focuses on the classroom as an environment that encourages students to construct knowledge themselves (Barr and Tagg 1995, 15). By focusing on the student as a participant in the learning process rather than as a passive recipient of information, active learning has been shown to improve students' understanding of content, critical thinking, and writing (Bromley 2013). Integrating active learning strategies into political science classrooms and requiring students to engage with evidence first-hand thus has the potential to improve students' understanding of the data common in academic journals.

To date, much of the scholarship on active learning in political science has focused on designing in-class simulations.[1] As Archer and Miller (2011, 430) explain, active learning is a "natural fit for political science" because "[t]he subject matter lends itself to discussion

[1] See, for example Crossley-Frolick (2010); Dougherty (2003); Frederking (2005); Hensley (1993); Hoffman (2009); Kelle (2008); Loggins (2009); Mariani (2007); Shellman and Turan (2006); Sherman and Waismel-Manor (2004); Wedig (2010).

and debate, theories and decision-making can be evaluated in light of current events, and institutions such as Congress and the United Nations lend themselves easily to simulations." Through simulations, students can put theories of international relations into practice and better appreciate the complexity of negotiating between multiple actors at different levels of analysis. This approach both helps students understand how conflicts develop and appeals to multiple learning styles (Wedig 2010, 548). However, simulations require significant investments of time and resources (Crossley-Frolick 2010, 185), as well as instructor oversight that may not be feasible in all courses. Relying on simulations as the predominant form of active learning in political science classrooms thus limits the reach of active learning strategies. Alternatively, group discussions and structured writing activities are accessible and easily implemented across a variety of classroom contexts. By focusing on the effectiveness of these strategies, we demonstrate that active learning techniques can be integrated into political science instruction even when preparation time is limited or courses cannot commit entire class periods to activities.

While in-class discussions and structured writing activities can be integrated into a wide range of courses, evidence on the effectiveness of these strategies is mixed. Small group discussions increase students' critical thinking skills and learning outcomes (Hamann et al. 2012; Pollock et al. 2011). Opportunities for peer discussion within introductory lecture courses are shown to enhance students' understanding of the concept (Smith et al. 2009) and combining discussion with instructor explanation further increases these positive benefits (Smith et al. 2011). Discussions enhance understanding by encouraging students to actively participate in the class and to verbalize ideas, which leads to a better grasp of the logic underlying essential concepts (Huang 2005, 496). However, the benefits of group discussions may be limited to students who prefer social and verbal learning (Bromley 2013, 820) and are comfortable participating in the conversations (Dallimore et al. 2004).

Where the classroom environment makes group discussion less feasible, writing-to-learn offers an alternative approach. A branch of active learning, writing-to-learn is based on the logic that forcing students to think about a topic in their own words increases their understanding of the underlying concept (Fry and Villagomez 2012, 170). Writing is widely recognized as a beneficial tool for making thinking visible (Reynolds et al. 2012, 19) and developing critical thinking skills (Çavdar and Doe 2012). When writing activities were implemented in an introductory science course, Linton et al. (2014, 474) found that "students who write about a concept perform better on subsequent writing-based assessments" and the authors advocate for the use of individual writing exercises whenever possible. However, other studies find that writing does not improve student achievement (Armstrong et al. 2008; Fry and Villagomez 2012) and suggest that not only are writing exercises ineffective, they are also one of students' least favored activities (Bromley 2013, 821).

Whether and under what conditions active learning strategies are effective thus remains an open question. In particular, while existing scholarship has focused on the effect of writing activities and discussions on students' overall performance, little is known about whether these strategies are well-suited for teaching students to critically engage with evidence. Additionally, with notable exceptions including Bromley (2013) and Powner and Allendoerfer (2008), existing scholarship rarely compares the effectiveness of different active learning strategies in the political science context. Studies also center predominantly on the use of active learning techniques in large introductory lectures and do not consider the effectiveness of these strategies in seminar environments where they are more easily implemented. Focusing on lectures and introductory courses may underestimate the effectiveness of active learning techniques and students' ability to engage with the evidence in social science

journals. Instead, we aim to illustrate the full potential of active learning by testing strategies in the context where they are expected to have a particularly high impact—writing-intensive first-year seminars (Kilgo, Ezell Sheets, and Pascarella 2014). To more closely examine how instructors can improve students' comprehension of arguments and evidence in political science seminars, the following section outlines an experiment that tests the relative effectiveness of in-class writing exercises and group discussion.

USING AN EXPERIMENT TO EVALUATE ACTIVE LEARNING STRATEGIES

Experimental Design

To test the relative effectiveness of writing activities and group discussion, we conducted an experiment with students in our two first-year writing seminars (FWS) taught at a large research university in the northeast United States. The seminars focused on humanitarian interventions and the interactions between political parties and social movements, respectively. For the experiment, the courses were combined and students were assigned the same reading: Heaney and Rojas' (2011) account of how the 2008 presidential election demobilized the anti-war movement in the United States. In this article the authors argue that the anti-war movement demobilized after Obama was elected in 2008 because Democrats—initially a large percentage of protest participants—perceived the election as a policy success and stopped participating. As a result, the movement both shrunk and radicalized. The argument is supported by evidence collected from surveys of movement participants, interviews, and ethnographic observations from protests. The experiment focused on the Heaney and Rojas article for two reasons. First, its thematic content combined the FWS's topics, addressing both how protests affect military interventions and how party politics influence movements. Second, it represents scholarly work with a complex but clearly presented argument that incorporates multiple sources and types of evidence. The article is thus ideally suited for the purpose of this study because it requires students to engage carefully with the evidence and the argument, but its clear organization makes that task as easy as possible.

We conducted the experiment during a regular class session. On that day, 29 students from the two FWS met together in one room. Students were told in advance that the class would take place in a different room, but were not informed that there would be a survey until the beginning of class on the day of the experiment. While participation in the experiment was optional, this timing ensured that students did not prepare for the class meeting more or less than their normal routine. Thus, while limiting the number of participants, embedding the experiment in two existing seminars minimized changes to students' normal learning environments and helped ensure external validity. Student comments confirmed the importance of withholding this information, as one participant explained that, "I think beforehand you should tell the class about the activities so they would make sure that they read the piece more vigilantly and paid great attention to detail."

The experiment took place in two phases. The first phase measured whether writing or discussion activities created a significant difference in students' understanding of the argument and how it was supported by the authors' evidence. The second phase tested whether students could incorporate and evaluate evidence in their own summaries of the article. It also allowed us to investigate whether the potential effects of the different teaching strategies endured beyond the immediate active learning activity.

For the first phase we randomly assigned the students to one of two groups. The first group moved to another room and were handed prompts that asked students to write short responses to the following two questions: "What kind of evidence do the authors use to support their claims?" and "Do you find Heaney and Rojas' evidence compelling? Why

or why not?" This writing exercise enabled students to think through the argument while consulting the article and to reformulate their understanding of the argument and evidence in their own words. In answering these questions, a few students relied on figures and phrases directly from the article. For example, in a phrase taken from the article's abstract, one student wrote that, "Heaney and Rojas use survey results from 5,398 demonstrators at anti-war protests, interviews with movement leaders, and ethnographic observation." However, the majority of students provided original and comprehensive accounts of the article. One student lauded the authors' use of "oodles of evidence," while another wrote that, "Heaney and Rojas used, among other things, statistical analysis to support their claims. In their analysis, they polled members of the Democratic Party, and other third party movements to reveal how partisanship and the antiwar culture are related [...] I found this evidence compelling, but not conclusive." Because the writing activity required students to examine the text and address the questions in their own words, rather than relying on interactions with their peers, we expected it to help students retain insights about the connection between argument and evidence.

While the first group completed the structured writing activity, the second group remained in the original classroom and participated in an instructor-led discussion that addressed the same two questions as the prompts for the structured writing activity. The instructor both asked the prompt questions verbally and posted them on a slide that remained visible throughout the discussion. We were careful to keep the discussion focused on these two questions and students were encouraged to respond to each other, rather than to the instructor. While the discussion was student-driven, the instructor prompted students to continue searching for additional types of evidence until they had captured all of the approaches used in the article. This set-up is akin to a standard classroom discussion in which students gain knowledge by engaging with the material on their own terms, but are guided and kept on topic by prompts provided by the instructor. In particular, we designed the prompts to direct attention to the evidence without foreshadowing the questions used to measure students' comprehension in the evaluation surveys (see below). For example, in the class discussion the instructor asked, "You've mentioned surveys and statistical analysis, anything else?" but did not provide an overview or explanation of Heaney and Rojas' evidence. Students in this second group were therefore exposed to the exact same prompts and information as those in the writing group. Thus, any difference in learning outcomes between the two treatment groups was purely the product of the different methods of instruction: one group had the chance to structure their thinking through a writing activity, while the other group built knowledge and exchanged views through the discussion format.

After 15 minutes, both the discussion and writing group were given a questionnaire that used multiple choice and open-ended questions to measure students' preparation prior to class, comprehension of Heaney and Rojas' (2011) argument, and ability to identify the article's evidence (see the appendix for the full text of the questionnaire). To guarantee anonymity, the only background information we collected was the FWS in which students were enrolled. The questionnaires also asked for students' consent to use their responses in this study. One student chose to opt out and our two groups had 15 (group discussion treatment) and 13 (writing exercise treatment) participants, respectively. Thus, the number of students in the discussion group was large enough to create a productive exchange of knowledge, but small enough to allow for the participation of each student. This class size is representative of seminar-style courses and the typical teaching situation for group discussions.

Beyond creating two different treatment groups, random assignment also created a further source of variation. About half of the students in the discussion group had to engage

	Concept	Indicators
Phase 1	Preparation	"To what extent do you agree with the following statement: I read the Heaney and Rojas article very closely." "To what extent do you agree with the following statement: Before class I had a solid understanding of Heaney and Rojas' argument."
	Evidence	"Which of the following describes the data that Heaney and Rojas use to support their claims."
	Argument	"Which of the following best summarizes Heaney and Rojas' main argument?" "Which of the following best describes the structure of Heaney and Rojas' article?"
Phase 2	Evidence	Does the student's written summary mention a specific type of evidence (i.e., surveys, interviews, etc.)? Does the student's written summary mention the purpose for which the evidence is used?
	Argument	Number of correct components mentioned in student's written summary (0-5).
	Critical Engagement	Does the student mention evidence in his/her answer? Does the student mention a strength of the article? Does the student mention why the evidence is strong? Does the student mention a weakness of the article?

Table 1: Key Concepts and their Measurements

with an unfamiliar instructor, while the other half were led by their usual instructor. Since the experiment took place in the latter half of the semester, students and their instructors had the opportunity to get to know each other and build a productive working relationship over several weeks. If this direct, personal interaction matters for learning outcomes, we should observe differences not only between the treatment groups, but also within the discussion group. Accordingly, we exploit the variation in familiarity with the instructor by analyzing differences between these two sub-groups.

For the second phase of the experiment, we brought the two groups back together and conducted a 20-minute classroom discussion guided by a series of broader questions. These questions were relevant to our seminars but did not directly address the article's evidence or argument.[2] Students could take notes during this discussion period, but few made use of this opportunity and they were not allowed to refer to notes or any other materials while completing the final evaluation. For this final evaluation, we asked students to write a short summary of Heaney and Rojas (2011) and to evaluate the strengths and weaknesses of the article's argument without referring back to the text. The follow-up evaluation served two purposes. The first was to investigate students' abilities to incorporate evidence into their evaluations of the article. Second, the follow-up examined whether the potential treatment effects endured beyond the immediate

[2] The questions were: "What motivated people to join the protests?" "What role did politics play in individuals' decisions to participate in the protests?" "What effect did the protests have on American foreign policy and the Iraq War?"

aftermath of the active learning activities and whether participation in the full group discussion led to the diffusion of knowledge across the two groups. We coded students' responses based on their ability to reference the argument's five key components[3] and created indicators of students' comprehension, outlined in Table 1 and discussed in greater detail below.

Indicators of Student Learning

Following the first phase of the experiment, the initial evaluation was designed to measure students' level of preparation prior to class, comprehension of the article's evidence, and understanding of the main argument. First, we relied on straightforward measures of preparation by asking students to self-report how thoroughly they read and understood the article. We did not collect any identifying data and students thus had no incentive to exaggerate their efforts. The low level of reported preparation, discussed in more detail in the following section, further indicates that students answered these questions honestly. Second, we measured students' comprehension of Heaney and Rojas' (2011) empirical evidence with a multiple-choice question. This question asked students to select "which of the following describes the data that Heaney and Rojas use to support their claims" from four possible options, three of which were correct.[4] This item is well suited to measure students' comprehension of the evidence because it captures multiple levels of learning. Students with full comprehension identified the three correct answers only, students with a basic understanding missed one option—either by selecting the incorrect component or by failing to select a correct component—and students who struggled with the evidence missed multiple options. From this question, we created an overall score of students' comprehension by subtracting their total number of incorrect answers from the total number of correct answers. The resulting indicator makes students' levels of comprehension both distinct and easily comparable.

A final pair of items measured participants' understanding of the main argument. They asked students' to select the accurate summary of the article's argument and to identify the correct structure of the article. These items were designed to be difficult enough to capture variation in student comprehension while providing clearly comparable answers.

Following the combined group discussion, we conducted a second evaluation to determine whether students could independently incorporate evidence into their summaries of the argument and evaluate its strength and weaknesses. This evaluation consisted of two prompts: 1) Write a short summary of the authors' main argument, and 2) Evaluate the strength of Heaney and Rojas' argument. What's something the authors do well? What's a weakness of the article? From the first question, we coded students' summaries to create three variables. The first to related to students' understanding of evidence and measured whether the response included any discussion of evidence. The second indicated whether students who mentioned evidence could explain how it contributed to the main argument. For the third measure, which gauged students' comprehension of the argument, we coded the number of core components (see footnote 3) from Heaney and Rojas' argument that students mentioned in their responses.

From the second question, we coded students' responses to create binary measures of whether they addressed each part of the question—i.e., did they mention evidence, a strength, and a weakness—and whether they explained

---

[3] The five key components were: 1) The anti-war movement demobilized, 2) Democratic electoral success was the cause of the demobilization, 3) Democrats stopped participating in the movement, 4) Because Democrats perceived this as policy success, and 5) As a result the peace movement shrank and radicalized.

[4] The correct options included: surveys of movement participants, interviews with movement leaders, and ethnographic observations of the protests. The incorrect option was: interviews with party leaders.

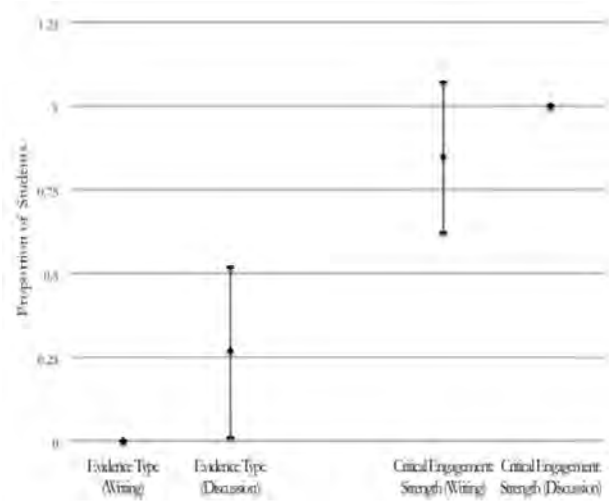


Figure 1: Effects of Treatment: Mentioning *(Strengths of) Evidence*[5]

why the evidence was strong or weak. Taken together, these measures capture students' levels of critical engagement with the evidence and the article's argument.

When combined, the survey items and open-ended responses provide a comprehensive picture of students' abilities to understand Heaney and Rojas' (2011) argument and evidence. Triangulating between closed survey questions and open-ended essay prompts limits measurement error by ensuring that our results are not the artifact of a specific way of measuring students' performance. This combination of indicators also enables us to explore whether writing exercises or discussion groups are better suited to helping students identify and evaluate evidence. The next section presents the results of the experiment and compares the effectiveness of these two strategies.

[5] All figures report differences in the writing and discussion groups based on the differences in proportions of students. While the figures report the mathematically accurate confidence intervals, theoretically values above one are not feasible for measuring proportions.

RESULTS

Group Discussion Is More Effective

The results show that in key categories, students assigned to the group discussion performed better than students in the writing group. Most notably, as seen in Figure 1,[6] participants in the group discussion were more likely to mention specific types of evidence in their written summaries of the article. Additionally, Figure 1 also shows that students in the discussion group were more likely to mention why the authors' evidence was strong in their open-ended evaluations of the article.[7] In fact, all students in the discussion group explicitly outlined why they thought the authors' evidence was strong. These findings suggest that exposure to other students' ideas and the instructor's prompting to identify all sources of evidence may help students internalize the relationship between evidence and argument more effectively than having students search for and write about evidence on their own.

Although group discussion positively encouraged students to engage with evidence, previous findings show that the benefits of discussion can vary by classroom environment (Dallimore et al. 2004; Bromley 2013). In the context of this experiment, the two relevant factors that could potentially condition the effect of group discussion were students' familiarity with the instructor and students' previous experience with the content of the material. As a result of random assignment, half of the students participating in the group discussion were led by their normal course instructor, while the other half were led by a

[6] All figures report differences in the writing and discussion groups based on the differences in proportions of students. While the figures report the mathematically accurate confidence intervals, theoretically values above one are not feasible for measuring proportions.

[7] The difference in the control and treatment group does not meet conventional standards of significance for this variable (p=0.1236); however, given the small sample size (N=28), the result is of substantive interest.

	Humanitarian Intervention		Movement Parties	
	Writing	Discussion	Writing	Discussion
Components of Argument	2.0 (0)	3.14 (0.34)	2.43 (0.48)	1.71 (0.68)
Difference	1.14*** (0.37)		-0.71 (0.83)	
Types of Evidence	0 (0)	0.14 (0.14)	0 (0)	0.43 (0.2)
Difference	0.14 (0.16)		0.43** (0.2)	
*Point estimates are means; standard errors reported in parentheses; ***p<0.01; **p<0.05; *p<0.10 (two-tailed tests)*				

Table 2: Effects of Treatment by Class

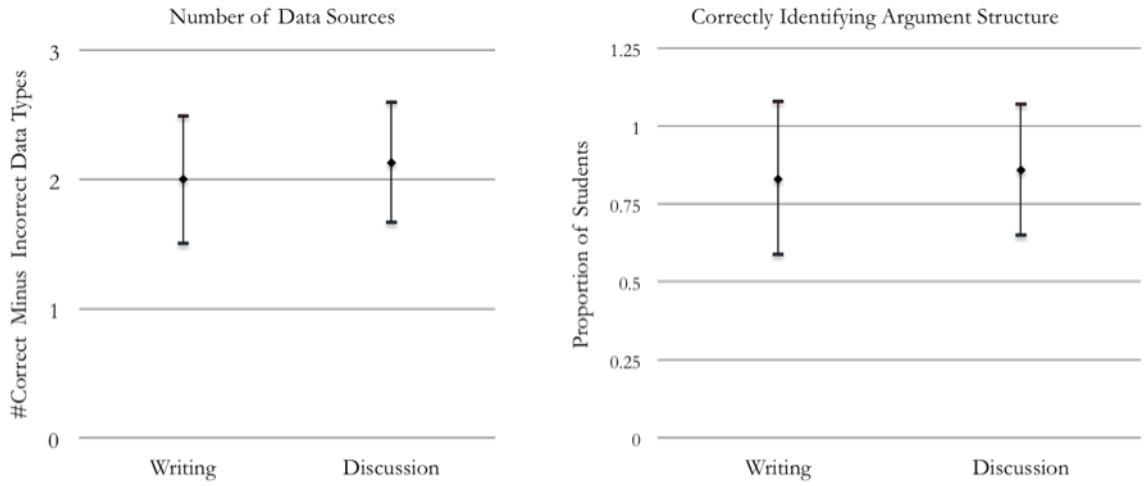


Figure 2: Effects of Treatment: Evidence and Argument Structure

new instructor. Students working with a new instructor were not familiar with her teaching style and may have been less comfortable or willing to actively participate, curtailing the positive benefits of discussion. Comparison of means tests (see Table 2) support this explanation: students in the discussion group with their own instructor (i.e., those in the course on humanitarian intervention) were significantly more likely to identify the correct components of the argument in their essays than the other discussion participants.

However, as Table 2 also shows, when it came to mentioning specific types of evidence in their own summaries, students in the discussion group who were unfamiliar with the instructor (i.e., those in the social movements class) were more likely to correctly list the article's sources of evidence. While not directly tested here, this difference may stem from these students' increased exposure to the article's measures of individual-level political behavior, which the discussion successfully activated. In their previous coursework, students in the course on humanitarian interventions focused on cases of intervention, elite rhetoric, and implications for foreign policy. By contrast, students in the course on political parties' interactions with social movements had received detailed instruction on individual-level political behavior and read articles with data similar to that presented by Heaney and Rojas.

Therefore, while group discussion appears to have significant benefits for student comprehension, the magnitude of its effect may depend on a combination of students' level of comfort in the classroom and familiarity with the material. Group discussion may thus

be particularly effective for teaching critical engagement with evidence in small seminars where students have the opportunity to interact with each other and with the instructor on a regular basis. Additionally, the positive benefits of discussion may be amplified when the course focuses on a single topic in detail and reading assignments consistently draw on and present evidence in similar ways.

Overall Student Competency

Despite the positive effect of group discussion on participants' abilities to independently identify and evaluate evidence, we find few differences in students' basic understandings of the article. As Figure 2 shows, in the first questionnaire, students in both the control and treatment group were equally likely to identify the article's data sources and the structure of the argument.

In both phases of the experiment, students demonstrated their ability to successfully engage with the evidence presented in Heaney and Rojas (2011). Regardless of their assigned treatment group, in the initial evaluation all students correctly identified the article's main argument. Almost all students also correctly characterized the structure of the argument (86%) and were able to distinguish at least two pieces of evidence used by the authors (79%). In the second phase of the experiment, which asked students to write their own summaries and evaluations of the article, the vast majority of students (27 out of 28) mentioned evidence in their discussion of the argument's strengths. The content of student responses also highlights their ability to use evidence to evaluate the strength of an argument. Students praised Heaney and Rojas for "presenting their data in a way that the reader can understand" and making it "clear what their argument is and what they are basing their conclusions off of." In addition to recognizing the benefits of the authors' evidence, students also identified some weaknesses, for example that "when using statistical analysis to make an argument, there will always be gaps. Correlation does not always prove causation and I would have liked the authors to address other possible reasons why anti-war sentiment declined." The vast majority of students were also able to identify the empirical expectations of the argument (25 of 28) and all students could offer a basic

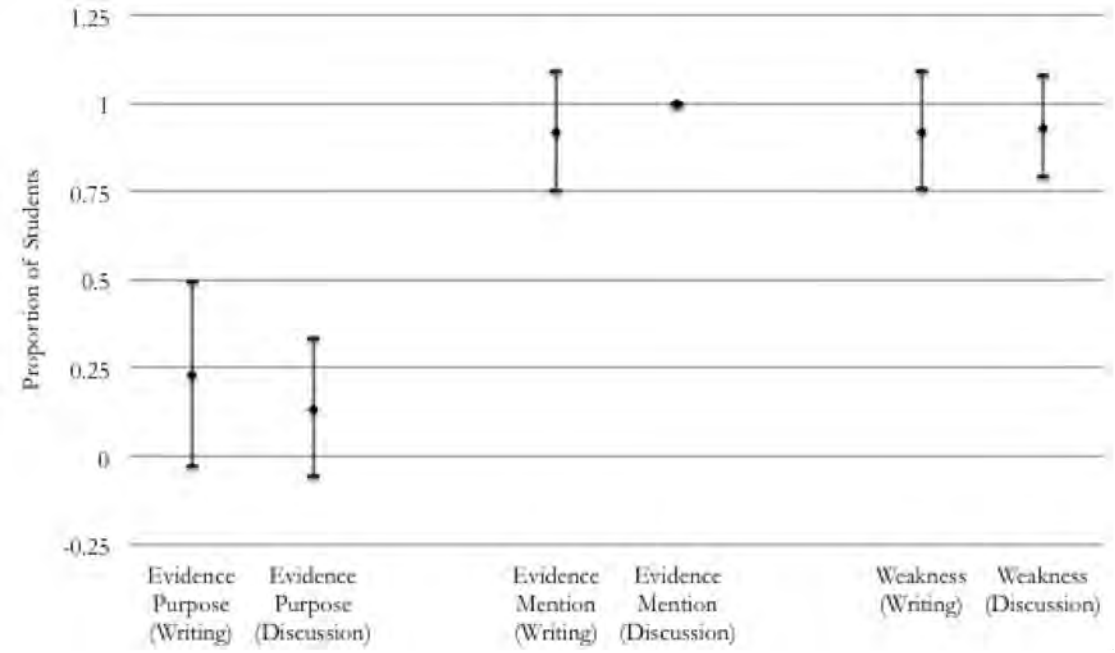


Figure 3: Effects of Treatment on Critical Engagement: Mentioning the Weakness/Purpose of Evidence, Questionnaire 2

explanation of how Heaney and Rojas used figures to backup their claims. These findings suggest that when engaging with academic articles, students are able to both identify the basic relationship between the argument and evidence and to independently recognize the central role that evidence plays in making a strong argument. Therefore, students are able to work with statistical evidence and the presence of quantitative analysis should not dissuade instructors from assigning material from academic journals in undergraduate political science courses.

In the second questionnaire, administered after a combined group discussion, treatment assignment did not create any noticeable difference in students' ability to recognize the purpose of the authors' evidence, incorporate a discussion of evidence into their own summaries of the argument, or identify and evaluate the article's weaknesses (see Figure 3). In other words, even after participating in a discussion that focused on different aspects of the article, students remained capable of identifying evidence and using it to critically evaluate the article's claims. They were, however, less capable of directly identifying the purpose of the evidence. Thus, students' overall competency and any effects of writing or discussion on their ability to work with evidence remain active for at least the duration of the class period.

Levels of Student Preparation

Finally, students' abilities to understand Heaney and Rojas' (2011) main argument and use of evidence did not depend on their preparation prior to class. While few students (7 out of 28) reported having read the article closely prior to class or that they had a solid understanding of the argument (10 out of 28), these variables did not significantly affect survey responses. Thus, even when students have not done the necessary background work and are not allotted time to read the article carefully in class, briefly revisiting the text through writing or discussion activities can help them understand the main argument and supporting evidence. This evidence suggests that active learning approaches may provide an alternative or supplement to pre-class preparation exercises (Trudeau 2005).

WRITING AND DISCUSSION ACROSS CONTEXTS

The experiment detailed above examined how students respond to group discussion and structured writing activities in the context of political science seminars. These seminars are ideal for implementing a wide range of active learning techniques because the class sizes are small, instructors are familiar with students' individual needs and abilities, and instructors can monitor participation to ensure that the relevant learning objectives are met. The results of this study thus speak most directly to the relative effectiveness of structured writing and group discussion activities in similar seminar-style classes. However, our findings also suggest that the benefits of these activities may be conditional on the classroom environment and students' familiarity with the discussion leader. In this section, we consider the potential benefits of and obstacles to implementing group discussions and structured writing in two alternative classroom environments—online courses and large lecture courses—and identify this as an important area for future research.

First, the interaction between active learning strategies and students' in-class experience speaks to the ongoing debate about the feasibility of teaching social science concepts through online courses (Keohane 2013; King and Sen 2013). Online courses have become increasingly prominent, with 22 of U.S. News and World Report's top 25 U.S. universities offering free classes online (Shah 2014). Understanding the extent to which classroom environments moderate the effects of active learning is thus particularly important. While both strategies encourage students to critically engage with the text, structured writing activities and group discussion differ in their emphasis on guidance from the instructor and interactions with other students. Structured writing activities ask students to revisit the text and think closely about their understanding

of the article's main argument in order to restate the claims in their own words. This strategy relies on writing prompts to guide students through a reexamination of the text and does not require any direct interactions with instructors or other students. If effective, structured writing activities could thus be easily implemented in online courses. For example, students could be assigned short writing prompts related to the article's evidence and required to post their responses on discussion forums.

Alternatively, group discussion relies on direct guidance from the instructor and interactions with other students to build comprehension. To the extent that comfort and familiarity with the instructor contribute to the effectiveness of group discussions—as our results suggest is the case—the benefits from this approach may not be easily translated to online environments. In online environments that allow students to interact with one another and the instructor in real time, recreating the familiarity and exchange of ideas that promote discussion may be possible. However, if the online setting prevents instructors from guiding discussion and creating a comfortable environment in which students can share ideas, it presents an obstacle to using class discussion to help students engage with evidence. Future research should more closely investigate the extent to which learning during discussion stems from face-to-face interactions in order to evaluate the utility of online courses for teaching critical thinking skills.

Second, the structured writing and group discussion activities implemented in this study also differ in the ease with which they can be adapted to larger classes or lecture environments. As with online courses, structured writing could be conducted with classes of any size—students simply need to be instructed to respond to writing prompts with access to the original text in class. Student responses could then be handed in to the instructor to ensure that they invested effort in writing comprehensive answers to the question. On the other hand, while it is feasible and recommended for students to discuss concepts with their neighbors during lectures (Smith et al. 2009), our findings suggest that care should be taken to ensure that students become familiar with each other and are guided through the discussion by an instructor. To create a comfortable and familiar environment for discussions, students could be assigned to the same groups over the course of the semester. Similarly, to guide students' discussions towards comprehensive accounts of academic articles, the instructor and teaching assistants should consider rotating among the groups over the course of the discussion.

Thus, while this study offers suggestive evidence that group discussion is the most effective strategy for teaching students to engage with evidence, this strategy also travels less easily between different classroom environments than the structure writing activities. To bridge this gap, future research should test different strategies for increasing students' familiarity with the instructor and guiding students through discussion activities. This research will be key for maximizing the positive benefits of group discussion across contexts.

CONCLUSION AND IMPLICATIONS

Combined, the results indicate that participation in group discussion better prepared students to integrate evidence into their evaluations of Heaney and Rojas (2011). Group discussions may be more effective because they encourage students to exchange ideas and correct misunderstandings over the course of the discussion, as well as allow the instructor to continue asking for additional responses until the group creates a comprehensive account of the argument. These positive effects are augmented when students are familiar with both the content of the material and the instructor, making group discussions of evidence an excellent fit for small seminars.

These findings are also in line with existing studies that find few benefits to in-class writing exercises (Armstrong et al. 2008; Fry

and Villagomez 2012; Bromley 2013). Thus, as an approach to active learning in general and the comprehension of evidence in particular, discussion appears to be the more effective pedagogical tool. However, this is not to say that writing activities are not useful for assignments that focus on personal opinion or experiences, or as an out-of-class follow-up to reinforce the content of group discussions. Future research should consider how discussion and writing activities can be combined with other approaches to maximize students' critical engagement with the material.

While discussion most effectively increases comprehension, we find that students are generally capable of critically engaging with academic articles and quantitative evidence. With the opportunity to revisit and reflect on the text in class, even students who had not carefully completed the assigned reading before class were able to grasp and evaluate the main argument. These results are particularly impressive given that the overwhelming majority of the students involved were in their first year of college. Therefore, articles from academic journals, including those that use statistical analysis, can be effectively used in undergraduate seminars when coupled with active learning strategies that guide students through the evidence.

Finally, at a time when liberal arts institutions search for ways to improve the feasibility of teaching online (Young 2015), this study's findings suggest that the benefits of active learning strategies may not easily translate outside of the classroom environment. While the study does not provide a direct test of how active learning strategies operate online or in large lecture environments, evidence that students were more likely to reap the benefits of discussion when they were familiar with the instructor highlights an aspect of the in-class experience that may be difficult to translate beyond smaller seminars. Future research should examine this finding more closely and compare the effectiveness of both structure writing and group discussion across different environments.

REFERENCES

Archer, Candace C, and Melissa K. Miller. 2011. "Prioritizing Active Learning: An Exploration of Gateway Courses in Political Science." *PS: Political Science and Politics* 44 (2): 429– 434.

Armstrong, Noris A., Carolyn S. Wallance, and S. Chang. 2008. "Learning from Writing in College Biology." *Research in Science Education* 38 (4): 483–499.

Atwater, Thomas. 1991. "Critical Thinking in Basic U.S. Government Classes." *PS: Political Science and Politics* 24 (2): 209–211.

Barr, Robert B., and John Tagg. 1995. "From Teaching to Learning: A New Paradigm for Undergraduate Education" *Change*. 27 (6): 12-25.

Bean, John C. 2011. Engaging Ideas: The Professor's Guide to Integrating Writing, Critical Thinking, and Active Learning in the Classroom. 2nd ed. San Francisco: Jossey-Bass.

Bonwell, Charles C., and Tracey E. Sutherland. 1996. "The Active Learning Continuum: Choosing Activities to Engage Students in the Classroom." *New Directions for Teaching and Learning.* 67 (1): 3-16.

Bromley, Pam. 2013. "Active Learning Strategies for Diverse Learning Styles: Simulations Are Only One Method." *PS: Political Science and Politic*s 46 (4): 819–822.

Burch, Kurt. 2000. "A Primer on Problem-Based Learning for International Relations Courses." *International Studies Perspectives.* 1 (1): 31-44.

Çavdar, Gamze, and Sue Doe. 2012. "Learning through Writing: Teaching Critical Thinking Skills in Writing Assignments." *PS: Political Science and Politics* 45 (2): 298–306.

Crossley-Frolick, Katy A. 2010. "Beyond Model UN: Simulating Multi-Level, Multi-Actor Diplomacy Using the Millennium Development Goals." *International Studies Perspectives.* 11 (1): 184-201.

Dallimore, E.J., J.H. Hertenstein, and M.B. Platt. 2004. "Classroom Participation and Discussion Effectiveness: Student-Generated Strategies." *Communication Education* 53 (1): 103– 115.

Dougherty, Beth. 2003. "Byzantine Politics: Using Simulations to Make Sense of the Middle East." *PS: Political Science and Politics.* 36 (2): 239-244.

Fitzgerald, Jennifer, and Vanessa A. Baird. 2011. "Taking a Step Back: Teaching Critical Thinking by Distinguishing Appropriate Types of Evidence." *PS: Political Science and Politics* 44 (3): 619–624.

Frederking, Brian. 2005. "Stimulations and Student Learning." *Journal of Political Science Education.* 1 (3): 385-93.

Fry, Sara W., and Amanda Villagomez. 2012. "Writing to Learn: Benefits and Limitations." *College Teaching* 60 (4): 170–175.

Hamann, Kerstin, Philip H. Pollock, and Bruce M. Wilson. 2012. "Assessing Student Perceptions of the Benefits of Discussions in Small-Group, Large-Class, and Online Learning Contexts." *College Teaching.* 60 (2): 65-75.

Heaney, Michael T., and Fabio Rojas. 2011. "The Partisan Dynamics of Contention: Demobilization of the Antiwar Movement in the United States, 2007-2009." *Mobilization* 16 (1): 45–64.

Hensley, Thomas R. 1993. "Come to the Edge: Role Playing Activities in a Constitutional Law Class." *PS: Political Science and Politics.* 26 (1): 64-68.

Hoffman, Donna R. 2009. "Representation and the Rules of the Game: An Electoral Simulation." *PS: Political Science and Politics.* 42 (3): 531-35

Huang, L. 2005. "Fine-Tuning the Craft of Teaching by Discussion." *Business Communication Quarterly* 68 (4): 492–500.

Kelle, Alexander. 2008. "Experiential Learning in an Arms Control Simulation." *PS: Political Science and Politics.* 41 (2): 379-85.

Keohane, Nannerl O. 2013. "Higher Education in the Twenty-First Century: Innovation, Adaptation, Preservation." *PS: Political Science and Politics* 46 (1): 102-105.

Kilgo, Cindy A., Jessican K. Ezell Sheets, and Ernest T. Pascarella. 2015. "The Link Between High-Impact Practices and Student Learning: Some Longitudinal Evidence." *High Education*. 69 (4): 509-525.

King, Gary, and Maya Sen. 2013. "The Troubled Future of Colleges and Universities." *PS: Political Science and Politics* 46 (1): 83-89.

Lantis, Jeffrey S., Lynn M. Kuzma, and John Boehrer. 2000. The New International Studies Classroom: Active Teaching, Active Learning. Boulder, CO: Lynne Rienner Publishers, Inc.

Linton, Debra L., Wiline M. Pangle, Kevin H. Wyatt, Karli N. Powell, and Rachel E. Sherwood. 2014. "Identifying Key Features of Effective Active Learning: The Effects of Writing and Peer Discussion." *CBE Life Science Education* 13 (3): 469–477.

Loggins, Julie A. 2009. "Simulating the Foreign Policy Decision-Making Process in the Undergraduate Classroom." *PS: Political Science and Politics.* 42 (2): 401-07

Mariani, Mack D. 2007. "Connecting Students to Politics through a Multi-Class Campaign Simulation." *PS: Political Science and Politics.* 40 (4): 789-94.

Meyers, Chet and T.B. Jones. 1993. Promoting Active Learning: Strategies for the College Classroom. San Francisco: Jossey-Bass.

Pollock, Philip H., Kerstin Hamann, and Bruce M. Wilson. 2011. "Learning Through Discussions: Comparing the Benefits of Small-Group and Large-Class Settings." *Journal of Political Science Education.* 7 (1): 48-64.

Powner, Leanne C., and Michelle G. Allendoerfer. 2008. "Evaluating Hypotheses about Active Learning." *International Studies Perspectives.* 9 (1): 75-89.

Reynolds, Julie A., Christopher Thaiss, Wendy Katkin, and Robert J. Thompson Jr. 2012. "Writing-to-Learn in Undergraduate Science Education: A Community-Based, Conceptually Driven Approach." *CBE Life Science Education* 11 (1): 17–25.

Shah, Dhawal. "MOOCs in 2014: Breaking Down the Numbers." Published December 26, 2014. https://www.edsurge.com/n/2014-12-26-moocs-in-2014-breaking-down-the-numbers.

Shellman, Stephen M., and Kürsad Turan. 2006. "Do Simulations Enhance Student Learning? An Empirical Evaluation of an IR Simulation." *Journal of Political Science Education*. 2 (1): 19-32.

Sherman, Daniel J., and Israel Waismel-Manor. 2004. "Get It in Writing: Using Politics to Teach Writing and Writing to Teach Politics." *PS: Political Science and Politics* 36 (4): 755–757.

Smith, M.K., W.B. Wood, K. Krauter, and J.K. Knight. 2011. "Combining Peer Discussion with Instructor Explanation Increases Student Learning from In-Class Concept Questions." *CBE Life Science Education* 10 (1): 55–63.

Smith, M.K., W.B. Wood, W.K. Adams, C. Wirman, J.K. Knight, N. Guild, and T.T. Su. 2009. "Why Peer Discussion Improves Student Performance on In-Class Concept Questions." *Science* 323 (5910): 122–124.

Trudeau, Robert H. 2005. "Get Them to Read, Get Them to Talk: Using Discussion Forums to Enhance Student Learning." *Journal of Political Science Education*. 1 (3): 289-322.

Wedig, Timothy. 2010. "Getting the Most from Classroom Simulations: Strategies for Maximizing Learning Outcomes." *PS: Political Science and Politics.* 43 (3): 547-555.

Young, Jeffrey. "New Consortium's Mission: Improve Liberal-Arts Teaching Online." Chronicle of Higher Education. Published May 12, 2015. http://chronicle.com/blogs/wiredcampus/new-consortiums-mission-improve-liberal-arts-teaching-online/56621.

APPENDIX A

Questionnaire One: Multiple Choice and Open-Ended Responses

We're going to ask you a few questions about how you prepared for class today and about the content of the Heaney and Rojas article, "The Partisan Dynamics of Contention." Please answer honestly and carefully.

1. Circle the First-Year Writing Seminar in which you are enrolled:

"From Social Movements to Political Parties"

"Humanitarian Intervention"

2. To what extent do you agree with the following statement: I read the Heaney and Rojas article very closely.

Leave blank if you did not have a chance to read the article before class.

a) Strongly agree

b) Agree

c) Neither agree nor disagree

d) Disagree

e) Strongly disagree

3. To what extent do you agree with the following statement: Before class I had a solid understanding of Heaney and Rojas' argument.

a) Strongly agree

b) Agree

c) Neither agree nor disagree

d) Disagree

e) Strongly disagree

4. Did you take any notes on the article before coming to class? This could include written or typed notes, comments in the margins of the article, or highlighting or underlining important phrases in the article. Circle your answer below.

Yes No

5. Which of the following describes the data that Heaney and Rojas use to support their claims? Circle all that apply.

a) Surveys of movement participants

b) Interviews with party leaders

c) Interviews with movement leaders

d) Ethnographic observations of the protests

6. True or False: The authors expect the end of the Bush presidency to increase antiwar activism. Circle your answer below.

True False

7. Which of the following best summarizes Heaney and Rojas' main argument?

a) Democrats are more likely to protest wars than Republicans.

b) The anti-war movement demobilized after the U.S. achieved policy success.

c) The anti-war movement demobilized after Obama was elected president because Democrats were no longer motivated to participate.

d) The Iraq war led to large-scale protests around the world.

8. How does the following figure contribute to Heaney and Rojas' (2011, 52) main argument? Explain your answer in a few sentences below.

9. Which of the following best describes the structure of Heaney and Rojas' article?

a) Evidence that Democrats stopped participating → Evidence that Democrats' attitudes changed → There was an active antiwar movement in response to the Iraq War → The movement declined after Obama was elected because Democrats stopped participating → Partisanship shapes movement participation and endurance

b) There was an active antiwar movement in response to the Iraq War → The movement declined after Obama was elected because Democrats stopped participating → Evidence that Democrats stopped participating → Evidence that Democrats' attitudes changed → Partisanship shapes movement participation and endurance

c) There was an active antiwar movement in

response to the Iraq War → The movement declined after Obama was elected because Democrats stopped participating → Partisanship shapes movement participation and endurance → Evidence that Democrats stopped participating → Evidence that Democrats' attitudes changed

APPENDIX B

Questionnaire Two: Student Summaries and Evaluations

1. In today's class we discussed the article "The Partisan Dynamics of Contention" by Heaney and Rojas. Based on what you remember from the article and from our class discussion, but without looking back at the article itself, write a short summary of the authors' main argument in the space below.

2. In the space below, evaluate the strength of Heaney and Rojas' argument. What's something that the authors do well? What's a weakness of the article?

3. What did you think about today's class? What's one thing that you think went well? What's one thing you think could be improved?

Intellectual Safety: Does Your Personality Type Contribute to Whether or Not You Take Intellectual Risks in Classes?

Meghnaa Tallapragada[1*], Dr. Kimberly Williams[2], & Dr. Dawn E. Schrader[3]

** Graduate Research and Teaching Fellow 2014-2015*

ABSTRACT

While scholarship has recognized that physical, emotional, and intellectual safety together constitute classroom safety, there has been a dearth in investigation in regards to intellectual safety. This research study aims to contribute to the discussion on intellectual safety and explores how personality contributes to one's personal feelings of intellectual safety in an educational setting. One hundred and ninety six students (N=196) attending a Northeastern private university participated in an online survey and results reveal that there is little association between student personalities and their feelings of intellectual safety. Certain aspects of professors, peers, class structure, and materials however seem to contribute to when and how students feel intellectually safe/unsafe in classrooms.

INTRODUCTION

There is no denying that schools are required to provide a safe environment for their students, staff, and teachers. There has been considerable amount of media discussion, government initiatives, and research emphasis on creating more physically and emotionally safe schools (Harris, 2015; US Department of Homeland Security: School safety). Physical safety implies providing a safe space from any physical violence or any physical danger, and emotional safety suggests a safe space where students feel confident in talking about issues such as teasing, cruelties, bullying, or other such potentially traumatizing experiences to their teachers and having them respond in appropriate ways (Merrow, 2004; Bucher & Manning, 2005). In order to create safe school spaces, Bucher and Manning (2005) recommend a holistic approach to safety where schools invest not only in security scanning technologies, but also strive to create a nurturing school environment that embraces various student identities and engages in continuous preventative programs.

In addition to creating physically and emotionally safe spaces, Merrow (2004) adds that schools also need to provide for intellectual safety, where students feel safe to risk and expand their intellectual horizons. Relative to physical and emotional safety, there is little discussion on how to provide more intellectually safe classrooms (Kohn, 2004; Schrader, 2004; Bucher & Manning, 2005; US Department of Education: School safety, 2015).

This research study aims to contribute to the discussion on intellectual safety in the classroom and explores the potential connections to students' personalities among students at a private Northeastern university.

LITERATURE REVIEW

Intellectual Safety

Scholars who have examined intellectual safety have associated it with students feeling comfortable enough to say that they do not understand or follow the material being covered by a teacher who listens to them and cares about them (Merrow, 2004; Bucher & Manning, 2005; Call, 2002). Schrader (2004) defines intellectual safety in the classroom as an "environment in which the professor is open and caring, demonstrates respect, and embraces the uniqueness of the students and their perspectives and does so in a classroom format where all are invited to participate actively, engage in personal disclosure while trusting the confidentiality of such openness, and where the professor maintains a sense of control and direction to facilitate learning" (p. 95-96). Schrader's definition emerged from an analysis of student responses using five themes that Schrader and Call (2002) found were dominant from the definitions students reported on intellectual safety. The five themes are:

Self. Students described their "emotions, ability, confidence, effort, as well as their perception that they were encouraged or discouraged to talk by the professor or by the class" in response to positive or negative classroom experiences (Schrader, 2004, p.94).

Professor. Students often described their experiences in relation to their professor's attitudes, beliefs, opinions and behaviors inside and outside classrooms.

Class structure. Students discussed the format of lecture or discussion, class size, the classroom space arrangement, and classroom environment.

Course materials. Students expressed their like/dislike for the course material and their presentation in the forms of Microsoft Powerpoint, handouts, etc.

Peers. Students wrote about the influence of peers and their opinions on their modes of communication, expression, and participation in classrooms.

Along with these five factors, Schrader (2004) also identified two specific concepts of "epistemic fit" and "epistemic stretch," which explain why there cannot be one type of intellectually safe classroom for everyone. Schrader synthesizes Perry (1970), Belenky and colleagues (1986), and Magolda (1992, 1999, 2001) to propose the concept of epistemic fit. According to Schrader (2004), if the student and teacher have an epistemological fit, then the student is more likely to feel intellectually safe. For example, if a student's epistemic perspective is one where the teacher is seen as an authority who delivers knowledge, knows it all, and is the one in control, then a classroom where the teacher lectures and does not seek student participation can be perceived by the student as a safe classroom. Any type of question or discussion from the teacher can be perceived as feeling intellectually unsafe for this student. On the other hand, if the student's epistemic perspective is one where the teacher is seen as a guide, who pushes students to present their opinions, challenges and critiques their positions, and engages in participation, then the student will perceive a classroom where there are stimulating interactions between teacher and students to be intellectually safe. An intellectually unsafe classroom for this student would be one in

which the teacher would lecture and stifle any student discussions or differences in opinion. These examples illustrate the need for an epistemic fit between the teachers and students to establish an intellectually safe environment.

Schrader (2004) emphasizes that only when there is a collective construction of classroom norms where students can begin to express their opinions, reflect on their underlying assumptions, and engage in a dialogue to discuss their perspectives, can students begin to engage in the much required process of "epistemic stretch." Stretching in this context suggests the process wherein students are able to "express and examine their own assumptions about the nature of knowledge and their involvement in it" (p. 93, Schrader, 2004). In other words, students are engaging in stretching or growing their epistemologies. The experience of engaging in epistemic stretch does not usually result in a "cushy," warm, and comfortable feeling, and instead renders a much needed challenging environment in order to transform students into critical, independent, and intelligent thinkers. This discomfort, however, could be associated with feeling intellectually unsafe in a classroom. To avoid this, Schrader (2004) emphasizes that students need to feel that a "safety net" has been established before proceeding to stretching their epistemologies. Of course, epistemic stretch can occur only when there is an epistemic fit achieved within the classroom between the teacher and students.

Schrader (2004) has provided epistemic fit and epistemic stretch as factors that contribute toward feeling intellectually safe in a classroom. This study aims to explore if student personalities also have any association with these feelings of intellectual safety in classrooms.

Student Personalities

Scholarship in education has recognized the need to assess student personalities in order to estimate their influence on teacher evaluations (McCann & Gardner, 2014), teaching environments (Caspi et al, 2006), and assessment preferences (Furnham et al, 2008). More recently, Pawlowska and colleagues (2014) found that classrooms could no longer operate on a "one size fit all" mode because they found that students' performance and level of satisfaction depended on the different students' personality types. While student personalities have become an integral factor in education scholarship, their influence on intellectual safety is yet to be explored. This study aims to address this missing link.

The Myers Briggs Type Indicator (MBTI) was used to assess student personalities in this study. Katherine Briggs and her daughter Isabel Myers developed Carl Jung's (1971) initial ideas of personality to devise the MBTI, which is a widely used self-assessment survey to characterize individuals along four pairs of personality dichotomies (Harrington & Loffredo, 2010). The four pairs include: Introversion (I) – Extroversion (E); Sensing (S) – iNtuition (N); Thinking (T) – Feeling (F); Judging (J) – Perceiving (P). According to the official MBTI website and Kuipers and colleagues (2009), the four bipolar dichotomies are based on four essential questions: does one choose to focus on their outer world (E) or their inner world (I), does one prefer to take in or perceive information as is (S) or keen on perceiving its interpretation (N), during decision-making, does one focus on the logic or the thinking judgment (T) or the people involved or the feeling judgment (F), and does one like to have things decided i.e., a judging attitude (J) or remain open to new possibilities or perceiving attitude (P). The MBTI instrument places each individual in one of the16 possible personality types. According to the website, the MBTI instrument has been used by organizations interested in better understanding its members, individuals trying to decide upon suitable careers and interested in personal growth, legal teams aiming to better understand their jury, health care professionals devising better communication strategies with their patients, couples in counseling, and also by teachers striving to better understand themselves and their students.

Using the MBTI instrument to assess students' personalities and the questions developed by Call (2002) and Schrader and Call (2002), this study explores the following research question (RQ): How are students' personality types associated with their perceptions of an intellectually safe environment?

METHODS

An online survey was hosted on a Northeastern private university's secure online survey website. The instrument was approved by the university's Institutional Review Board to ensure ethical research conduct.

Measurement

An online survey was designed to measure intellectual safety, personality, and student characteristics such as their year of enrollment, major, college, nationality, and gender.

Intellectual safety

The survey began with a reliable set of questions developed by Schrader and Call (2002) and subsequently used by Schrader (2004) to measure intellectual safety. The list of open-ended questions is listed in the Appendix A. According to Schrader (personal communication, March 3, 2015), the first set of questions were designed to provoke an emotional response leading to more "truthful" or "accurate" recall, and the next set of questions was asked to first define intellectual safety and then to prompt participants to describe elements of safe and unsafe classroom environments. In order to be able to compare results of this study to these previous studies, the same questions were used and their order was maintained.

Personality

The online survey included an external link to measure participants' MBTI personalities (external link: http://www.humanmetrics.com/cgi-win/jtypes2.asp). Participants were asked to click on the link, which opened in a separate tab, to take the MBTI test. Participants then returned to the survey page to enter their MBTI personality type, before they could submit the survey.

Student characteristics

All student characteristics were measured using nominal measures. Participants could indicate their year of enrollment (freshman, sophomore, senior, graduate student, staff, other), major and college (based on the university's academics list), nationality (US citizen, international citizen, other), and gender (male, female, other). All these questions that included the "other" or "yes" option were asked to further explain their response through an open-ended question.

Recruitment

Students enrolled at a Northeastern private university were invited to participate in the study through their university's secure online survey website, which hosted the university's secure research management system. The study was publicized in multiple undergraduate classes at the university. Students were offered extra credit for participating in the study. No identifying information, such as their name or date of birth was collected, however students were asked in the survey to enter their website ID number in order to ensure that they would get credit for participating in the study. The researchers had no access to connecting these ID numbers to specific students. This information was maintained confidential by the university's secure online survey website. This ensured anonymity for the students so that they would feel comfortable to freely divulge their experiences around intellectual safety, personality, and their other personal characteristics. In this paper, pseudonyms are used to further protect the anonymity of the students.

Analysis

Besides traditionally analyzing qualitative data using a thematic approach, a software called "Voyant" was used. Voyant creates word clouds that depict the most and least frequently used terms and provides an option to analyze their contexts. SPSS was used to run the preliminary descriptive analysis.

RESULTS

Demographics

One hundred and ninety six students participated in the study, but only 163 students completed the entire survey. Figures 1-4 display the distribution within the sample of gender, nationality, year of enrollment, and major respectively.

As shown in Figure 1, 124 students identified themselves as female, 45 as male, and 2 as other. Figure 2 shows that there were 22 students who reported being international students and 147 who reported saying they were not international students. As for the year of enrollment shown in Figure 3, there were 19 freshmen, 47 sophomores, 53 juniors, 35 seniors, and 15 graduate students. Figure 4 shows that 70 students indicated being Communication majors; 27 as Information Science; 16 Biology and Society; 9 as Computer Science; 6 as Applied Economics and Management; 3 each as Animal Science, Biological Sciences, Computer and Electrical Engineering, Economics, and Nutritional Science; 2 each as Biological Engineering, Food Science, Hotel, Industrial and Labors Relations, and Natural Resources; and 1 each as Agricultural Sciences, Environmental Engineering, Global and Public Health Services, Fashion Design Management, Fine Arts, Interdisciplinary Studies, Plant Sciences, Psychology, and Science and Technology Studies. While the sample seems quite diverse, the students participating in the study were mostly women, non-international students, and not Communication or Information Science majors.

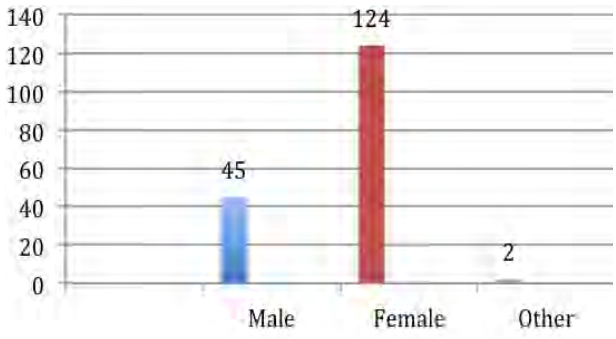


Figure 1: Gender

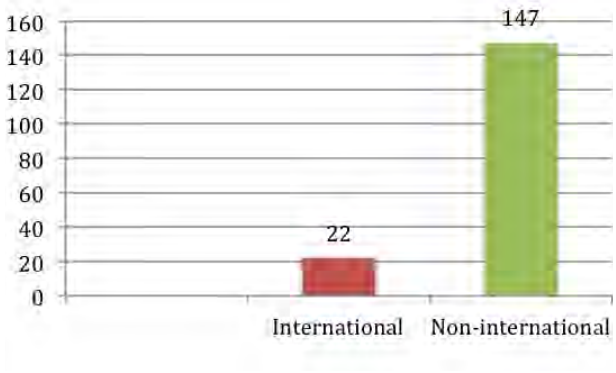


Figure 2: Nationality

Physical and Emotional Safety

Although these two aspects of classroom safety were not the focus of this study, there were a couple of instances when students reported feeling physically or emotionally unsafe in the classroom. For example, consider the following two quotes from two different students describing their unsafe experiences:

Ashlee, a freshman: "I often participate in class and try my hardest. After class one day, some guy in my class, [came] up to me and start[ed] making comments and touching me. He mean[t] it in a friendly way, and I guess after getting to know him better I understand this now. But it really bothered me, when he said he was 'just admiring [my] figure' and other inappropriate comments."

Nicole, a junior: "There was one instance where she [the professor] mentioned the term 'faggot' during class and tried to play it off as being humorous, but it was very offensive and was not even relevant to the discussion."

Ashlee's comment highlights the aspect of physical safety that is rarely discussed in classroom safety scholarship. Oftentimes articles discussing physical safety focus mainly on addressing and preventing violence and rarely broach issues pertaining to sexual harassment. This comment serves as a reminder to investigate often neglected issues

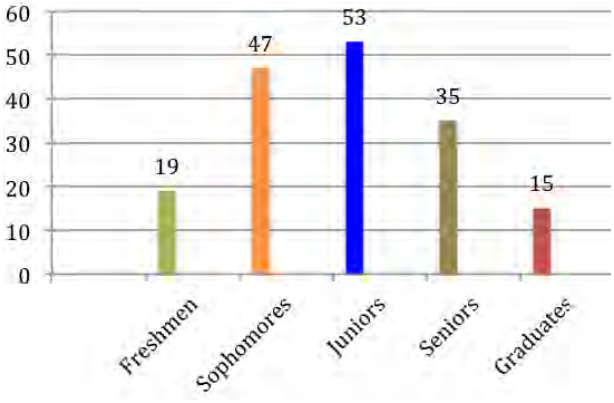


Figure 3: Year of Enrollment

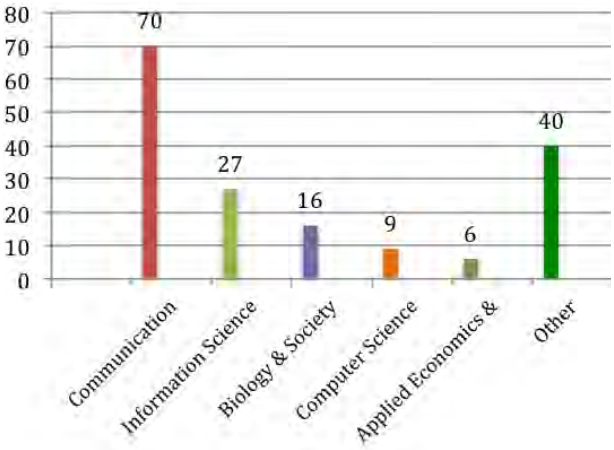


Figure 4: Major

of sexual harassment and other aspects of physical safety, not just violence in classrooms.

Nicole's comment portrays the potentially negative effects that the use of some emotionally charged terms and issues can have on classroom safety. The same student expressed that in classes where there was "intentionally biased speech, hate talk, closed-mindedness, intolerance, [and] disrespect," they felt unsafe. It is important to state that the student did not isolate this one instance to frame this professor as being "bigoted," but instead felt this was a pattern with that professor. The student also mentioned how these instances with this professor were upsetting even though the comments were not personally offensive to them, but for their peers who were gay or in the closet. The student went on to state that, "This kind of [offensive] language used in a professional setting could re-affirm destructive norms surrounding homosexuality and [is] incredibly inappropriate." It is striking how the student identifies that these discussions were not even relevant to the class topic; it begs the question as to why professors feel the need to express and at times humor such personal opinions with their students. Research on humor in classrooms and their influence on emotional safety could potentially add to this discussion.

Intellectual Safety

Schrader defines (2004) intellectual safety as an "environment in which the professor is open and caring, demonstrates respect, and embraces the uniqueness of the students and their perspectives and does so in a classroom format where all are invited to participate actively, engage in personal disclosure while trusting the confidentiality of such openness, and where the professor maintains a sense of control and direction to facilitate learning" (p. 95-96). The data revealed similar definitions. However it also produced five themes around intellectually safe classrooms (self, professor, structure, material, and peers). Some of these themes were more dominant than the others.

Themes

The theme of "self" was present through most responses with students describing how they experienced classes and their thoughts and emotions regarding these classes. . This is not too surprising given the nature of questions asked. It seemed more pertinent to explore the remaining four themes. Rarely did students express only one of the themes in their responses, and often constituted a combination of these themes. For example, consider the following (for a list of questions, please refer to the Appendix A):

Francis, a graduate student: "The people in the classroom. It starts with the instructor. The instructor has to really ensure that they guide the class in making sure everyone feels respected, and they have to be a role model for that. At the same time, students have to really

put an effort in to make sure they are not just attacking someone because they questioned something they said. Respect is mutual, but it has to be nourished. Once you lose trust and respect in the classroom, it is not safe. You have to also make sure you feel like you can approach the instructor. That's hard, I know, because sometimes personalities clash. It is when the instructor doesn't recognize how they are being exclusionary that things become [a] real [problem]. You can disagree with the ideas someone is saying without being a [****] about it. But that requires effort, and most people just don't care."

You can see in the quote above, that for Francis, it is about the people in the classroom – both the professor and their peers. Students seem to emphasize the role of the professor more than that of their peers. Following this logic, the following four major themes that emerged are discussed in the order they were found to be the most dominant to the least:

1. Professors. Most of the students reported their professors to be a significant factor in developing and nurturing intellectual safety in classrooms. They seemed to consider professors to be most in control of creating a safe atmosphere in class. Students Michael, Missy, and Anthony help explain this:

Michael, a senior: "The most important piece is the professor. How s/he controls the class and acts is relayed onto the students who reflect and build upon it."

Missy, a sophomore: "The teacher sets the tone of the classroom. If the teacher gives a level of respect to the students, then they will hopefully give it back. If the teacher is honest, the students will follow suit."

Anthony, a junior: "I think the expectations of the environment should be set before hand, such that it should be expected that people can express themselves without [judgment] from others. I believe that teachers are crucial in creating a safe intellectual environment, as they can steer conversations in the right direction if discussions get off-topic or stray from the original topic."

Along with respect, as Anthony described, students felt safe with a professor who was more tolerant and less judgmental about their contributions in class. Consider the following responses from students:

Paula, a junior: "[Some of the elements that I think help create a safe intellectual environment are] having a teacher that is open and doesn't make you feel stupid. A teacher [who] isn't condescending but instead welcomes questions and actually considers student's opinions [makes for a safe intellectual environment]."

Meghan, a junior: "An understanding Professor who doesn't make students feel dumb is a big part of creating a safe intellectual environment. Also, giving students the support they need and allowing them to feel secure about themselves and their knowledge. Providing a comfortable learning environment [where] students won't be ridiculed."

Stephanie, an international student and a freshman: "An element that might create an unsafe or threatening intellectual environment is where the instructor/professor is very critical of the students. In one of my classes, my instructor was the type of person that has very exactly lined out what is right and what is wrong (in terms of our opinions and interpretations of text). So he would always direct the discussion toward what he felt was right and would hint that we [were] wrong if we said anything other than what he had in mind as correct. I think this kind of environment can make the students feel discouraged to participate and critically think."

Paula, Meghan, and Stephanie described what many other students also expressed in their responses; it is the rigidity or the need for more flexibility around discussing differing perspectives that seems to be an important factor in classroom discussion.

2. Peers. As Francis described "people" in the classroom matter a lot in developing and maintaining intellectual safety, and peers

comprise most of the "people" in a classroom. For Arthur, peers seem to matter more than the instructor:

Arthur, a sophomore: "I think this idea has a lot more to do with your peers than it has to do with the instructor. I think a safe intellectual environment is one where a person does not have to be worried about being embarrassed to actively participate."

Students seemed to be most bothered about being misunderstood or judged by their peers for their thoughts and perspectives. Consider the following comments made by students Ashlee and Catherine:

Ashlee, a freshman (who also discussed physical safety above): "Students should not be assuming or frustrated ([or] at least [should] not convey that [openly]) when other students are [unable to understand some materials]. It is hard to understand [how materials that are easy for you might be harder for some others.] [Everyone's] brain works at different paces and in different ways."

Catherine, a senior: "[Not having] negative or hurtful commentary from classmates about a student's beliefs or thoughts on a subject [and instead having] an environment that is open-minded and accepting of everyone [can help create an intellectually safe environment]."

Students also commented on how peers could add to an intellectually safe classroom environment:

Adrian, a junior: "I know that with all of my close friends, I can say anything that's on my mind with no negative consequence, so having a close student body would probably help create a safe intellectual environment, but that's probably wishful thinking."

3. Structure. In order to create a safer classroom, some students responded by saying they would appreciate having the instructor set some ground rules for the structure of the class. Ariana and Diana illustrate this:

Ariana, a senior: "In the classroom, I think that there should be a policy (it can be informal) that students are welcome to contribute to discussion and voice dissent, if they feel something could be done differently. Also, it should be made clear that there is support for students who feel uncomfortable or need [accommodations]."

Diana, a senior: "I think if teachers say, either during the first day of class or toward the beginning of the semester, that students' thoughts will be heard and they should not feel uncomfortable voicing their opinion in class, [it can create an intellectually safe classroom environment]."

Along with people in the classroom, students also seem to consider class structure as a factor in feeling safe or unsafe in classrooms. As described by Erica, Heather, and Ray, class size seems to be a component of this structure:

Erica, a sophomore: "It is hard to have an intellectually safe environment in a big class. In a class of over 40 (give or take), there [are] simply too many people for every person in the room to have some form of bond. When you hear words of a random stranger that would make a good topic of conversation later in the day, there is little incentive to keep quiet."

Heather, a sophomore: "I think that smaller class sizes are one of the most effective ways to ensure a safe intellectual environment. Not only is it less intimidating to share ideas, opinions etc. to a small group of people, but it also makes it easier to facilitate a more community-type (versus competitive) environment which fosters a feeling that your classmates want to know your ideas and respect your intellectual contributions. Additionally, a community-type environment fosters the element of friendship between classmates, allowing students to feel that their classmates will respect their ideas, opinions, etc."

Ray, a junior: "I think a small classroom helps to create a safe intellectual environment because it fosters participation and relationship building. I also think that settings ground

rules at the beginning of a class can help, so that the students [know] what to expect, and what boundaries they can and cannot cross. Ultimately, the most important part of a safe intellectual environment is [to] refrain from using personal [judgment] against people whose ideas may differ from your own"

It seems like the need for smaller class sizes seems to correlate with the students' need to feel more connected with their professor and peers.

4. Material. This was the least discussed theme in the responses of this study. Materials did not seem to affect feeling intellectually safe in classrooms, but there were some instances that, as Joe and Jennifer put it, lacked sensitivity when discussing controversial issues that contributed toward unsafe intellectual classroom environments:

Joe, a senior: "Disparity between different social identity groups that aren't discussed, addressed, or resolved (without anyone overseeing the environment to ease tension and help protect disadvantaged minorities) [contribute toward an unsafe intellectual classroom environment]."

Jennifer, an international student and a senior: "Singling out based on differences (in race, ethnicity, religious preferences, sexual preferences, socioeconomic status, language, culture) [contributed toward a threatening or unsafe classroom environment]."

It seems that it is only when sensitive course materials are not taught/discussed in an appropriate manner that they contribute to an intellectually unsafe environment.

Epistemic fit and stretch

There seemed to be an underlying willingness among students to participate in discussions, which suggests that most of the students in the study were not expecting the professors to tell them what to think and were hoping to have their professors push and challenge their ideas. However it is disconcerting that instead some students who feel like they want to participate fear that they will be pointed out for their lack of understanding. For example, consider the following student responses:

Karen, a sophomore: "I think that making a student comfortable with the professor and the peers by allowing for wrong answers and creativity [creates an intellectually safe classroom environment]."

Brittany, a sophomore: "Having positive responses to what I say and being able to speak up and not feel that I am being judged [contribute toward an intellectually safe classroom environment]."

Karen and Brittany illustrated something that many other students also noted – a resistance to being proven wrong. This creates a potential obstacle for many professors who wish to execute an epistemic stretch. For students to "stretch," there needs to be a change in perspective, which means, for the most part, that they are indeed wrong in some perspectives. Some students seem to recognize this aspect. Consider Tyler, Jeff, and Chris:

Tyler, a junior: "Warm environment created by professors and TAs and constructive criticism, [and] explaining to students where they may have gone wrong instead of cutting them down in front of the class [makes for an intellectually safe classroom environment]."

Jeff, a sophomore: "An environment that encourages ideas to be shared, [and] if the ideas are wrong, the responses aren't negative, [and instead] are constructive [creates an intellectually safe classroom environment]."

Chris, a junior: "Having an open-minded classroom, [where students are] not [belittled for providing] the wrong answer [and are instead given] constructive criticism, [where] everyone [feels] welcome to learning and participating [is an intellectually safe classroom environment]."

Tyler suggests saving face by addressing these issues in private and not in front of the class. However for epistemic stretch to occur

for the entire class, it seems imperative that these issues are raised and discussed more openly. Jeff and Chris seem to suggest turning these discussions into opportunities to present constructive criticism without necessarily publicly embarrassing the student. This seems to be a more achievable and desirable outcome in order to provide for epistemic stretch. Carol seems to provide a concrete example to illustrate this:

Carol, a graduate student: "I think that instructors can do a lot to create a safe atmosphere by building a formal structure for participation. One of my professors made a point of calling on people by name in class to contribute. It was a big class, which could have made people nervous, but our professor took time on the first day to let us know that a) she would call on us to answer questions, and b) it would be fine to say 'I don't know.' or to make a guess and be wrong. By establishing a norm of active participation, the professor helped us all feel comfortable speaking up when she posed questions. The way that the instructor responds to a student's contribution can make a big difference, too. If a student asks a question, it's important that the question be acknowledged as legitimate. If a suggested answer is incorrect, it's helpful for the instructor to show that they appreciate the contribution and acknowledge any positive aspects of the answer before correcting it. Obviously not all answers will be right, but it can make a big difference for a professor to say, 'Here's what I'm hearing you say. I can see where you're coming from, and why you might think that. However, here's a piece you maybe didn't think of/take into account that might change your answer.' That sort of response can make students feel more comfortable, because they know that their contribution is valuable."

Carol described the necessary condition of establishing the "safety net" before professors proceed towards engaging in any type of epistemic stretching. Simply stated, students needs to first feel safe to intellectually explore in the classroom before they can start to take any intellectual risks in order to stretch or grow their epistemologies.

Personalities

Figure 5 illustrates the representation of personalities in this study. As shown, there were more students who were Extroverts than Introverts (NE=90, NI=74), more students with iNtuitive perception than Sensing perception (NN=90, NS=74), more students with Feeling-based judgment than Thinking-based judgment (NF=101, NT=63), and very slightly more of students with a Judging attitude than a Perceiving attitude (NJ=83, NP=81). Voyant showed that there were some factors more frequently reported than others based on their personality type.

Professors. Regardless of the personality, professors were a highly discussed factor in influencing a safe classroom environment. It seemed to matter for more Extroverts than Introverts, more students with iNtuitive perceptions than Sensing perceptions, more students with Thinking judgment than Feeling judgment, and also more students with a Judging attitude than a Perceiving attitude.

Peers. The influence of peers seemed to matter more for students with iNtuitive perceptions than Sensing perceptions, slightly more for Extroverts than Introverts, and for students with a Perceiving attitude than a Judging attitude. This factor seemed to matter almost equally for both students with Thinking judgment and Feeling judgment.

Structure. As for classroom structure, it seemed to matter more for students with Thinking

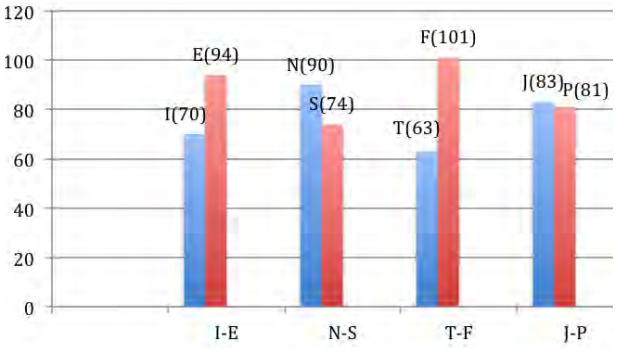


Figure 5: MBTI Personalities

judgment than Feeling judgment, and also for Extroverts than Introverts, and slightly more so for students with a Sensing perception than iNtuitive perceptions, and also very slightly for students with a Perceiving attitude than a Judging attitude.

Materials. In the very few times this was mentioned, it seemed to matter relatively greater for Thinkers than Feelers, and very slightly for students with a Judging attitude than a Perceiving attitude. It seemed to matter equally and fairly less to Extroverts and Introverts, and also for students with iNtuitive perceptions and Sensing perceptions.

While there were some subtle differences between these types, there were no stark trends to estimate that different personality types significantly defined intellectual safety differently. Refer to the word clouds developed for each personality type describing intellectually safe and unsafe classrooms in Appendix B. The word clouds depict how students in general seemed to focus on the same factors much to the same extent, regardless of their personality types. For the sample in this study, personality types do not have a significant association with feelings of intellectually safety in classrooms. Future research should extend sample size and increase diversity in sample types, and explore other factors that might be influencing intellectual safety in classrooms.

DISCUSSION

As Merrow (2004) mentioned, it seems imperative to address all three aspects of classroom safety – physical, emotional, and intellectual. Comparing quotes from students describing physical, emotional, and intellectual safety, there is an indication that feeling physically, emotionally, and intellectual safe/unsafe seem to be interconnected. Future research should investigate these interconnections and recommend ways that could help make classrooms safe spaces for learning.

As for intellectual safety, it seems like the behaviors of people in the classroom i.e., professors and peers are the most influential factors contributing to intellectual safety in classrooms among all the themes discussed by Schrader and Call (2002). Students in Schrader's (2004) study mentioned that they felt intellectually safe with professors who were: caring, compassionate, authoritative with an open and flexible character, and honest, and knew students by their name, spoke to them "on [an] equal level," disclosed their personal experiences, and maintained a good physical stance with eye contact and open body stance. In this study, students did not seem too concerned with physical stance, eye contact, body stance, or personal disclosure. Instead the emphasis seemed to be on professors being less judgmental and more tolerant and open toward students' opinions/perspectives. The same criteria seemed to apply to peers as well. This change in emphasis on intellectual safety indicates that this is a dynamic concept and needs to be measured continually with different samples to assess the factors influencing intellectual safety.

While research has shown that student personalities have a significant association with teacher evaluations (McCann & Gardner, 2014), teaching environments (Caspi et al, 2006), and assessment preferences (Furnham et al, 2008), this study suggests that there might not be an association with intellectual safety. Future research (in addition to extending the sample size, diversifying sample types, and using other methods) should also investigate other factors that could help us understand the influences for intellectual safety.

CONCLUSION

This study began with the question: Does student personality type contribute to intellectual safety? According to this study, despite differences in personality, students' definitions of what constitutes an intellectually safe or unsafe classroom seems to be more cohesive than disparate. Future research should further investigate the factors

influencing intellectual safety, and also the interdependencies of physical, emotional, and intellectual safety in classrooms.

ACKNOWLEDGEMENTS

I would like to thank Dr. Kimberly Williams, without whom this project would have remained a dream. Thank you for your continuous (scholarly and personal) support and guidance. I would also like to thank Dr. Dawn Schrader for her scholarship advice and guidance. I would like to especially thank the GRTFS for their feedback and support throughout this teaching as research experience: Jaebeum Cho, John Foo, Molly Katz, Nick Mason, Kaylin Myers, Jonathan Park, Vera Rinaldi, Kristie Schlauraff, and Lindsay Springer. I would like to extend my thanks to Runjini Raman for her feedback on this paper and its related poster. Finally, I would also like to express my gratitude to the CTE staff, especially Dr. Derina Samuel, for the unconditional support, encouragement, and nurturance of my passion for teaching.

REFERENCES

Belenky, M., Clinchy, B., Goldberger, N., & Tarule, N. (1986). *Women's ways of knowing.* New York: Basic Books.

Brotman, J. S. & Moore, F. M. (2008). Girls and science: A review of four themes in the science education literature. *Journal of Research in Science Teaching,* 45(9), 971-1002.

Bucher, K. T. & Manning, M. L. (2005). Creating safe schools. *The Clearing House, 79(1), Classroom Management for Middle and Secondary Schools*, 55-60.

Call, C. M. (2004). Intellectual safety and epistemological development in the college classroom. *Doctoral dissertation,* Cornell University, Ithaca, NY.

Caspi, A., Chajut, E., Saporta, K., & Beyth-Marom, R. (2006). The influence of personality on social participation in learning environments. *Learning and Individual Differences*, 16, 129–144.

Furnham, A., Christopher, A., Garwood, J., & Martin, N.G. (2008). Ability, demography, learning style, and personality trait correlates of student preference for assessment method. *Educational Psychology*, 28, 15–27.

Guzzetti, B. J. & Williams, W. O. (1996). Gender, text, and discussion: Examining intellectual safety in the science classroom. *Journal of Research in Science Teaching*, 33(1), 5-20.

Harrington, R. & Loffredo, D. A. (2010). MBTI personality type and other factors that relate to preference for online versus face-to-face interaction. *Internet and Higher Education*, 13, 89-95.

Harris, E. A. (2015). Suspension rules altered in New York city's revision of school discipline code. *NY Times: NY Region*. Retrieved on March 03, 2015 from http://www.nytimes.com/2015/02/14/nyregion/suspension-rules-altered-in-new-york-citys-revision-of-school-discipline-code.html.

Kohn, A. (2004). Safety from the inside out: Rethinking traditional approaches. *Educational Horizons*, 83(1), 33-41.

Kuipers, B. S., Higgs, M. J., Tolkacheva, N. V. & de Witte, M. C. (2009). The influence of Myers-Briggs type indicator profiles on team development processes. *Small Group Research*, 40(4), 436-464.

Magolda, M. B. (1992). *Knowing and reasoning in college: Gender-related patterns in students' intellectual development*. San Francisco: Jossey-Bass.

Magolda, M. B. (1999). *Creating contexts for learning and self-authorship. Nashville*, TN: Vanderbilt University Press.

Magolda, M. B. (2001). *Making their own way: Narratives to transform higher education to promote self development*. Sterling, VA: Stylus Publishing.

McCann, S. & Gardner, C. (2014). Student personality differences are related to their responses on instructor evaluation forms. *Assessment and Evaluation in Higher Education*, 39(4), 412-426.

Merrow, J. (2004). Safety and excellence. *Educational Horizons*, 83(1), 19-32.

Myers Briggs Type Indicator (MBTI). Retrieved on March 08, 2015 from http://www.myersbriggs.org.

Pawlowska, D. K., Westerman, J. W., Bergman, S. M. & Huelsman, T. J. (2014). Student personality, classroom environment, and student outcomes: A person-environment fit analysis. *Learning and Individual Differences*, 36, 180-193.

Perry, W. G. (1970). *Forms of intellectual and ethical development in the college years: A scheme*. New York: Holt, Rinehart and Winston.

Schrader, D. E., & Call, C. (2002, November). Intellectual safety, moral atmosphere and epistemology in college classrooms. Poster presented at the Association for Moral

Education annual meeting, Chicago, IL.

Schrader, D. E. (2004). Intellectual safety, moral atmosphere, and epistemology in college classroom. *Journal of Adult Development*, 11(2), 87-101.

Schrader, D. E. (2015). Personal communication on March 3, 2015.

US Department of Education: School Safety. Retrieved on March 03, 2015 from http://www.ed.gov/school-safety.

US Department of Homeland Security: School safety. Retrieved on March 03, 2015 from http://www.dhs.gov/school-safety.

APPENDIX A

Questions to measure intellectual safety:

1. Have you ever had a class (or an experience in a class) that impacted you in a negative way? Describe the class in as much detail as possible.

2. What were the emotional reactions you had to this particular class or experience? Be as detailed as possible.

3. Have you ever had a class or a classroom experience that impacted you in a positive way? Describe the class in as much detail as possible.

4. Again, what were the emotional reactions you had to this class or experience?

5. How would you define or explain the concept of "intellectual safety?"

6. What are some of the elements that you think help to create a safe intellectual environment?

7. What are some of the elements that might create an unsafe or threatening intellectual environment?

8. Any other additional comments?

APPENDIX B

Figure 6: Intellectually safe class – Introverts

Figure 7: Intellectually safe class – Extroverts

Figure 8: Intellectually safe class – iNtuition

Figure 9: Intellectually safe class - Sensing

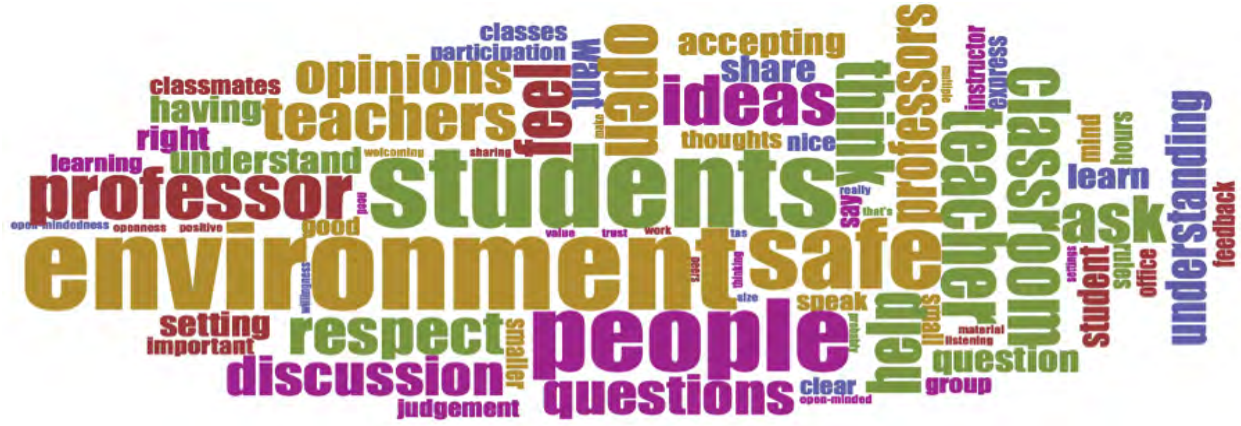


Figure 10: Intellectually safe class – Thinking

Figure 11: Intellectually safe class – Feeling

Figure 12: Intellectually safe class – Judging

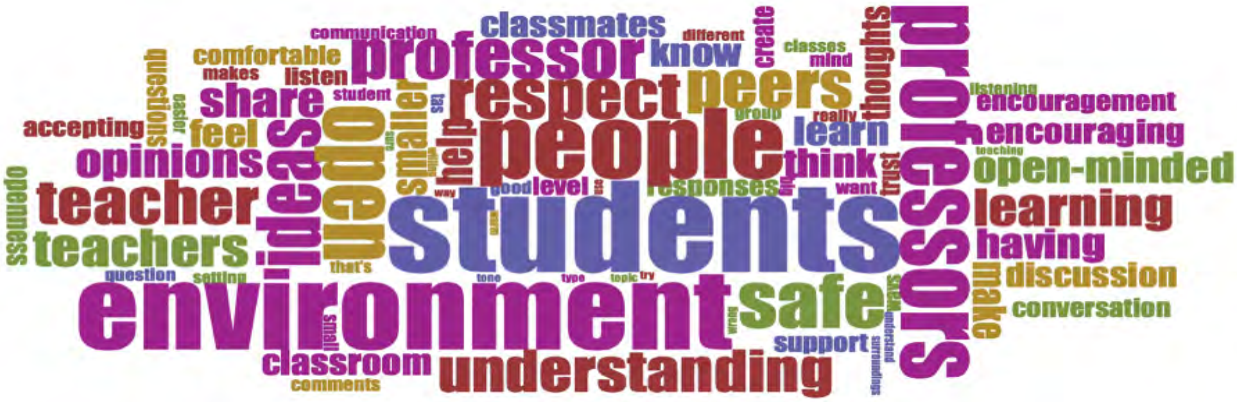


Figure 13: Intellectually safe class - Perceiving

Figure 14: Intellectually unsafe class – Introverts

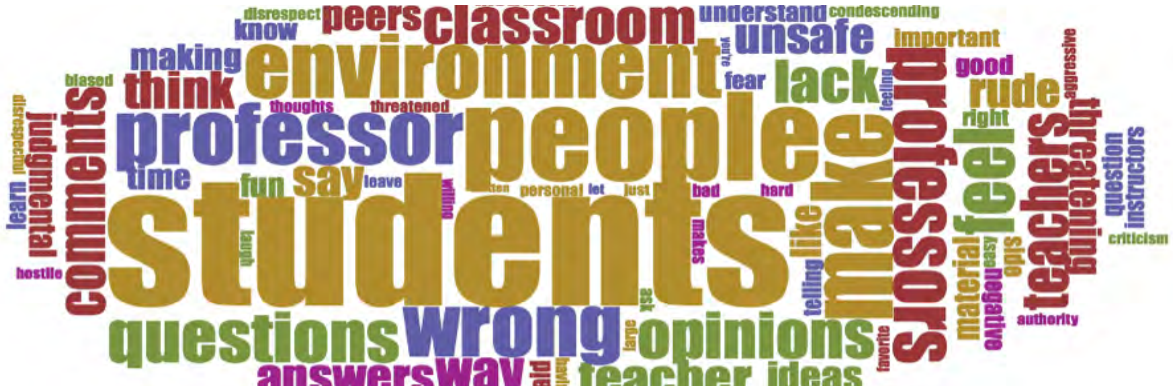


Figure 15: Intellectually unsafe class - Extroverts

Figure 16: Intellectually unsafe class – iNtuition

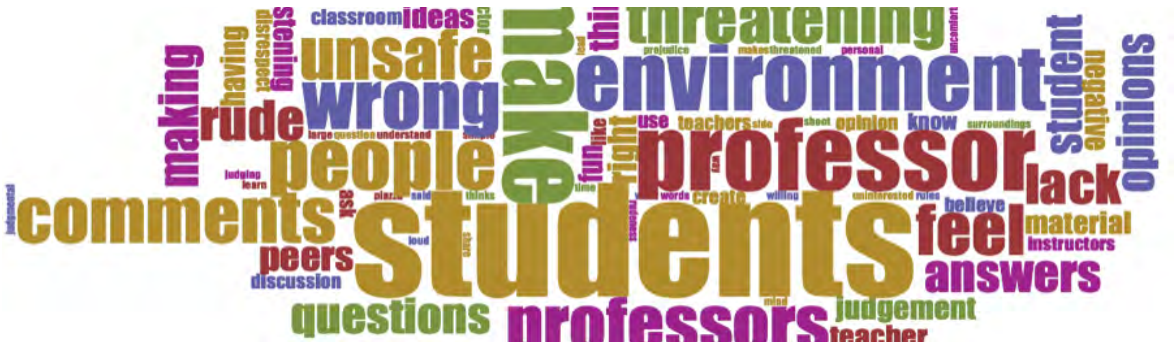


Figure 17: Intellectually unsafe class - Sensing

Figure 18: Intellectually unsafe class – Thinking

Figure 19: Intellectually unsafe class - Feeling

Figure 20: Intellectually unsafe class – Judging

Figure 21: Intellectually unsafe class - Perceiving

Thinking Aloud Through Writing Prompts: A Case Study

Molly Katz

Graduate Research and Teaching Fellow 2014-2015

ABSTRACT:

This paper documents the researcher's experiences interviewing three college freshmen about a writing assignment. The researcher conducted a "think aloud" with one student, asking her to "please read this writing prompt and think aloud as you figure out what this prompt is asking you to do. Please tell me everything that comes into your mind as you complete this task." The researcher also conducted a focus group with two other students, asking them to read the same prompt and to suggest how the researcher might improve the assignment. The findings reveal potential thoughts, actions, and reactions a given student might have to a writing assignment. By analyzing in detail these few students' complex thoughts and reactions to a writing prompt, the researcher hopes to help her fellow instructors: 1) imagine more vividly their students as readers 2) more readily consider the various possible reactions and thought processes of their students and 3) think through what their students might want that might be different from the implicit or explicit demands of their assignments. The paper also proposes two new pedagogical methods for instructors planning to distribute writing assignments to their students. In the course of this study, the researcher concluded that the research methods employed in order to gain insight into these students' reading processes and opinions had great potential as pedagogical tools. This paper explores the pedagogical benefits of these methodologies to the participants in this study and suggests potential strategies for reaping these benefits in the classroom.

BACKGROUND

A college freshman's mind is a black box—at least, it can seem that way to professors who are developing writing prompts for their students. In Spring 2014, I taught my fourth Freshman Writing Seminar (FWS) course at Cornell University. The course was ENGL1168 Cultural Studies: Fanfiction, and it was my second time teaching that specific course. I'd designed it myself from the bottom up: title, course description, topics, readings, writing assignments. The primary objective of the Freshman Writing Seminar program is to teach writing. FWS courses are meant to focus more on developing writing skills than on topic or content coverage, and the courses are writing intensive rather than reading intensive, with no more than 75 pages of reading assigned per week. According to Cornell's "Indispensible Reference for Writing Seminar Instructors" (Gottschalk, 2014): "Seminars should require at least six—and at most nine—formal essays on new topics, totaling 25–30 pages of polished prose." That means that in the Spring of 2014, I asked my 17 students to collectively produce a minimum of 425 pages of polished prose. Assuming it takes a student half an hour to write a page, that's over two hundred hours of work. I wanted to respect my students' labor

by making clear, comprehensible requests that serve a purpose that is evident both to me and to them. I knew my writing prompts were having a strong influence on the work my students were producing. Research has demonstrated that writing prompts have a significant impact on the quality and nature of students' academic work. As Pillai (2014) noted, "Since Kroll's (1979) seminal study, the rhetorical demands of university writing assignments have been identified as a factor that affects academic success. The studies reviewed above show that although some task prompts do not explicitly refer to a preferred rhetorical structure, the organization of information in student research papers plays a significant role in determining how these papers are assessed."

Each time my students turned in their essays to me, I evaluated their responses with anxiety. Did the prompt I had written clearly communicate what I expected my students to do? Did the prompt make clear how to go about meeting those expectations?

When grading papers for another professor, I found myself thinking often about the perils of a poor prompt. The professor for whom I was grading generally provided her students with several writing prompts from which to choose. We noticed, reviewing the essays, that students who had chosen one of the prompts seemed consistently to have performed more poorly. Was it, we wondered, because this prompt had attracted the weaker students, or was it that this prompt had encouraged thinking along lines that we found less sophisticated and desirable? When, in Spring 2014, three out of my seventeen students responded to my first writing assignment in a way that I felt did not reflect my goals, I decided it was time to learn more about how to write a good essay prompt.

DEFINING TERMS

When I asked my colleague, who also taught FWS courses, if she would send out an email to her students inviting them to participate in a study about how students interpret writing prompts, she expressed concerns about the eligibility of her students on the grounds that she doesn't "give writing prompts." What did she mean, I asked. All of her writing assignments, she told me, were the same. She asked students to develop arguments about the books they read according to a rigid yet also capacious formula: "X book does Y thing for Z reasons" (capacious because students were allowed a great deal of freedom to choose their topic; rigid because the basic structure of their arguments was predetermined.) For her, the phrase "writing prompt" did not describe this approach to writing assignments. In general, literature on writing prompts seems to tacitly agree with her assessment that she does not give writing prompts, though perhaps not for the same reasons she provided. Her prompt might be classified as an "open prompt" according to Reid and Kroll's (1994) landmark classification system. Most literature on writing prompts largely treats prompts as examination tools (e.g. Reid and Kroll, 1994; Allison and Gupta, 1997). Given this trend, I considered the idea that perhaps it would be better to refer simply to writing assignments and to excise all mention of writing prompts. In a recent study by Pillai (2014), however, "writing prompts" do refer to essay assignments, rather than essay examinations. I have chosen to join Pillai (2014) in using the term "writing prompts" because I believe it is the term that best suggests the complex negotiation process between student and instructor that occurs when an instructor gives out a writing assignment. The term "prompt" implies that the words the instructor has written down are merely a starting point, something to trigger the writer's thought process. The term therefore makes clear that the student must use a great deal of imagination and reasoning to arrive at a point where the task he or she is going to execute becomes clear.

A goal of this paper is to shift the focus toward how students can play a more active role in making meaning out of writing assignments. In a recent article, Rank and Pool (2014) noted that "Faculty members care deeply about student

writing…however, they do not often directly examine the part of the assessment process over which they maintain complete control and on which they rarely receive feedback: the formatting of assignments." [emphasis mine] Rank and Pool (2014), in an attempt to help faculty correct this trend, provide a typology of possible writing tasks, inspired by Bloom's Taxonomy (Bloom, Engelhart, Furst, Hill & Krathwohl, 1956). Their work pushes instructors to understand their own purposes, goals, and aims for a given assignment. Their work also suggests ways in which instructors can organize their prompts in order to make those aims clearer. However, while Rank and Pool are correct that instructors maintain "complete control" over the formatting of assignments, it is crucial to remember that students, in the act of interpreting an assignment, may "reformat" and revise that assignment according to their own schemata and expectations. Ackerman (1989) notes that "in many cases the assignment [or topic] given by an instructor and the assignment [or topic] taken by a student are not a reciprocal fit" (p. 96), further pointing out that "giving and responding to an assignment is an act of negotiation" (p. 96). The process of making meaning out of writing prompts is always a collaboration between students and instructors. This study will argue that it is important to work to make the collaborative nature of this task an explicit part of the prompt making process. As Hamp-Lyons (1991) points out, "Regrettably, we have not yet achieved a rule-of thumb for the degree of freedom and constraint that allows a writer to show her best self" (p. 53). This paper will recommend a way to circumvent this problem, allowing students to show their best selves without an instructor needing a "rule of thumb."

One of the limitations of a study like Rank and Pool's (2014) is that, although their advice is concrete and specific in many ways, it does not take into sufficient account the fact that instructors, even when well versed in the theory of what should make a prompt work, often fail to predict which prompts will be difficult for students in practice. In their article, "Why some questions don't work," Allison and Gupta (1997) noted that the instructors in their focus group struggled to identify the question that "didn't work." The instructors voted on four prompts, with instructions to vote for the prompt they thought was the most likely to have failed in practice. Seven out of ten of the instructors did not vote for the failed prompt. The study suggests that professors are not always able to anticipate which prompts will go awry. In the conclusion to their study, Allison and Gupta (1997) note: "It would be valuable to learn more about students' own expectations as they approach writing tasks, and we would advocate classroom studies in this area. Use could be made of practice tasks, perhaps involving past questions as group discussion topics or as essay planning or writing tasks to be explored in later interviews or group discussions. Another possibility involves think-aloud techniques, though Polio & Glew (1996) chose to avoid these, cautioning that they may alter the writing processes of respondents. I agree entirely with their assessment. I believe that Polio and Glew (1996) are correct to believe that think aloud techniques might alter the process of students, but I believe that much can still be gained from adopting this technique, and it formed a major component of my methodology.

METHODS IN THEORY AND PRACTICE

Part 1: The Think Aloud

For decades, Think Aloud Protocols have been used to gain insight into reading and writing processes (Bereiter and Bird 1985; Perl 1979; Lau, 2006; Olk, 2002; Pressley and Afflerbach, 1995; Sainsbury, 2003). These studies, often meticulously coded, have yielded understanding about the different ways in which individuals may approach a problem, while achieving the same 'answer' (Someren, Barnard & Sandberg, 1994). Think aloud protocols can be used both as a pedagogical technique, to make students more aware of their cognitive processes, and as a tool for researchers looking to conduct descriptive analysis (Jahandar, Khodabandehlou, Seyedi, &

Abadi, 2012)

Many think aloud protocols have focused on comprehension Koro-Ljungberg, Douglas, McNeill, Therriault & Malcolm (2013) note, "In general, during a typical TA (Think Aloud) study, researchers provide students with problems and ask them to verbalize what they are thinking while attempting to solve the problem." The primary advantage of Think Aloud Protocols is that they provide a 'real time' look at a problem solving process (Someren, et al 1994). My study followed in this tradition, and sought to reap similar benefits from using this methodology. Few have done similar work. To date, only one other study (Nelson, 1990) has employed "think aloud" to gain insight into how students interpret writing prompts. Nelson, however, looked at responses to prompts given to students in the middle of completing a course in which they were enrolled. Nelson's work focused on the tendency for students to make "interpretive errors" that were not really errors at all, but were rather based on a correct understanding of a different set of standards than the ones explicitly stated in the prompt. One student in Nelson's study memorably remarked, "this was an easy assignment," flatly contradicting the professor, who stated that he believed the assignment would be highly challenging. This student correctly determined what the teaching assistant, who graded the assignment, would require of him and wrote to those standards, which were quite low.

As Reid and Kroll (1991) observe, a wide range of variables across multiple categories ought to be taken into consideration when it comes to evaluating the efficacy of a writing assignment:

1. The writing situation (contextual variables),
2. The subject matter (content variables),
3. The wording of both the prompt and the instructions (linguistic variables),
4. The task(s) (task variables),
5. The rhetorical specifications (rhetorical variables), and
6. The scoring criteria (evaluation variables)

It is important to note my approach to several key variables across these six categories. I did not provide the students with a rubric (category 6), and I did not provide them with any "course content," beyond what occurred in the prompt (Category 1). I also shaped the students' reading process by asking them to read in specific ways (i.e. following the think aloud or focus group procedure) (Category 1). I therefore did not attempt to examine how students might interpret writing prompts under what one might call "normal classroom conditions," but rather looked at what kind of readers my research subjects would become when asked to think aloud through the prompt or to think about how they could improve the prompt.

The think aloud protocol reduced the opportunity for students to make errors based on inattention and incomplete reading. For those wishing to study careless errors and what might cause them, the think aloud would not be advisable. However, understanding what helps and hinders students when students are reading with full concentration, which the think aloud offers, is also valuable. It is possible to find ways to minimize, if not to eliminate entirely, rushed reading, just as it is possible to make certain that grading rubrics and grading practices help, rather than hinder, student comprehension of the prompt. I believe it is important to understand how students read when they are concentrating fully and when contextual variables are not interfering as they did in Nelson's study.

One other study shaped my methodology substantially. Pillai (2014) interviewed 24 students enrolled in the "foundation program," a course meant to give first year university students a solid grasp of the conventions of academic writing. Pillai (2014) interviewed students about how they chose the prompts they did, why, and what they thought of them. Pillai (2014), in examining student responses, drew a crucial distinction, inspired by Swales (1982), between "what is required of the task" and "how the student writers may be

expected to accomplish the task." They found in their interviews with students that students focused much more heavily on talking about the cognitive demands placed on them, much more easily divining from the prompt "what is required" over "how to accomplish it." I decided to focus specifically on determining what in the prompt *does* help students think about how to accomplish the task, and included several interview questions at the end of the "think aloud" that would explicitly induce the student to comment on what in the prompt helped her understand how she might go about completing the assignment.

The Prompt: I chose a variation on the first prompt I'd given my students in my Spring 2014 course on Fanfiction. I chose the prompt because several students had clearly struggled with it, and because despite these problems I believed the assignment had the potential to help students think about important issues. The prompt asked students to write about what they had done, actively, in order to get pleasure out of a piece of writing they'd enjoyed. I decided to revise the prompt before using in my study, but to use the same prompt in both the focus group and the think aloud, avoiding further revisions until the study was complete. My goal was not to see what had gone wrong, but rather to see how students would respond to a fresh effort at expressing the question. To see the prompt I presented to the three participants in my study, please turn to Appendix A.

Recruitment for the Think Aloud: This project was approved by Cornell University's Institutional Review Board. Students read and signed a consent form for participation in the study. I asked ten of my colleagues to send out a recruitment email to their classes, asking students to contact me about setting up an interview if they were interested. The recruitment email explained the basics of the study and the task I would be requiring, as well as informing them that they would be given a $5 gift card for their efforts. In order to be eligible, one needed to be a college freshman between the ages of 18 and 19. The compensation must have been too low in relation to the task I described, because none of the 150 potential research subjects contacted me. I report this unsuccessful recruitment procedure in the hopes that others attempting similar studies can learn from my mistakes and modify their procedures accordingly. I eventually did successfully recruit a subject, the sibling of one of my colleagues, who was a freshman at another university. I did not previously know this student. We arranged a time and I conducted a 30 minute Skype interview.

Procedure for the Think Aloud: Ten minutes of our interview were spent with the subject simply "thinking aloud" through the prompt. I asked the student to "Please read this writing prompt and think aloud as you figure out what this prompt is asking you to do. Please tell me everything that comes into your mind as you complete this task." I followed up with several free form questions, inspired by my observations of the student during the think aloud. I made sure to include among my follow up questions those that would invite comments about how the prompt was helping—or not helping—the student think about how she would accomplish the task. Questions included: "What, if anything, made this prompt difficult to understand?" and "What would your first step be, if you were to attempt to write an essay on this prompt?"

Results of the Think Aloud: I gained a vivid picture of this student as a reader through this procedure. At one point during the think aloud, the student told me, "I think so now I'm thinking about--I'm a skim reader, I think. So I don't develop like--or not develop, I don't absorb every word. I kind of look at the bigger picture of the paragraph. So I read through the paragraph and I'm like--ok, I understand the bigger picture. 'Cause in my mind I'm almost staging it again."

In keeping with this statement, the student reread phrases she didn't understand. She also, however, continued reading beyond the phrases she didn't understand. I noticed this pattern in her reading, but she remarked upon

it herself, as well. Part way through the think aloud, she said, "If you reread it and you still don't understand it then you should continue to read until you get to the bottom of the prompt. And now I understand."

Perhaps most striking to me was the student's use of the "learning outcomes" (see Appendix A). The student used the learning outcomes not just to figure out what the assignment was asking her to do but also to figure out how she was going to do it. As previously stated, the work of Pillai (2014) led me to take a particular interest in what about my prompt might help students develop a sense not just of what they were meant to do, but how they were expected to do it. I noted that the student used the learning outcomes to keep her priorities straight: *"Learning outcomes are really helpful…you [can] get very bogged up in the finesse of the writing… and not the structure of it … [or] the continuation of the prompt."* The student also commented, *"And so having these three learning outcomes I would almost use this as like a part of my outline,"* which indicates that she used learning outcomes to envision a writing process, rather than just a writing product.

Finally, the student read to simplify, looking for keywords: "I would look for about three words that are what I need to do." Seeing that she did this led me to think about how, in the future, I might develop writing assignments that would help a student like her pick out the three words that would be most helpful, rather than three that might mislead and confuse. I could, perhaps, indicate keywords with boldface or italics.

Pedagogical Value of the Think Aloud: The Think Aloud achieved at least three learning outcomes for this student. She:

1. Gained practice reading and interpreting a complex text

Although ideally a writing prompt would be simple, the prompt I gave this student was a complex text that required her to prioritize, search for key words, skip over difficult passages, return to those passages later to see if her understanding had changed, and, lastly, to form a "gestalt" sense of what the prompt was asking her to do out of a series of complex directions. Most instructors who teach writing would agree that learning to read a difficult text is a crucial skill, and a difficult one to develop. The think aloud helped this student develop this proficiency.

2. Learned to talk about reading experience as an active process

Reading is a task that requires effort—reading is something one does, not something that simply happens to one. However, some do not understand what they are doing to actively shape their interactions with a piece of writing.

3. Examined the implications/results of implementing that process

She seemed to begin to develop a sense of herself as a reader, and her habits, good and bad (i.e. "I'm a skim reader"), and how they might be helping her or hindering her in her efforts to understand this task.

Strikingly, the learning outcomes I'd hoped students would achieve from *writing the essay* I assigned in the Spring 2014 class were remarkably similar to those achieved simply by asking the student to think aloud through the assignment. The outcomes I outlined for the *assignment* were as follows:

1. Identify an experience of reading that was enjoyable to you
2. Learn to talk about your reading experience as an active process
3. Examine the implications/results of implementing one part of that process

The alignment of the outcomes of employing this research methodology with the outcomes I had desired to achieve as a teacher of writing show how valuable the think aloud process could be to instructors.

Part 2: The Focus Group

To date, no study of the efficacy of writing

assignments has employed a focus group consisting of students. Some studies such as Reid and Kroll's (1994), have used pilot testing to determine potential problems with assignments, but none have worked to document what could be gained from asking students to consider how they might like to see a prompt revised.

Recruitment for the Focus Group: I am hired to train new tutors for the Cornell University Knight Institute Writing Walk-in Service. My duties include teaching Cornell undergraduates how to help their fellow students revise their work at the sentence and structural level, how to ask questions and listen carefully, and how to help their peers achieve their goals. I invited my tutors in training to assist me with this study as an extra, entirely optional activity, which I told them would provide them with a novel experience that would, perhaps, complement their training. It is in this context that the focus group occurred. Two out of five of my trainees agreed. Although both of the students who participated in the group were college freshman, they were students who had already been selected for their strong written and oral communication skills, their interest in writing and the writing process, their ability to read carefully, listen well, and think flexibly and on their feet. Additionally, the training with which I had already provided them might have influenced their approach to the focus group. They had, after all, already learned how to help a writer clarify and improve their work. It is worth noting that a personal connection with the subjects and the ability to provide them with an intrinsic rather than an extrinsic motivation resulted in a significantly more effective recruitment process than the one I experienced during the think aloud part of the study.

Procedure for the Focus Group: The focus group lasted half an hour, from 6:00 to 6:30 pm. The focus group took place immediately following a tutor training session, over a pizza dinner (which I provided). I asked the two students to "read the prompt and help me improve it, focusing on what might be unclear." I did not attempt to withhold any of my own thoughts about the potential strengths and weaknesses of the prompt, informing them of what had previously gone wrong. They began by asking me what class the prompt was for, and I told them it was for a class on fanfiction. They asked me about the purpose of the prompt, and I told them that it was to help students think more clearly about what they were doing to get pleasure out of their reading. One student then said "so a correct execution of this prompt would look like…" and described a potential paper. I stated that this was, indeed, correct. We then had a discussion in which we all made various suggestions. "Would the prompt be better if….?" Each of us asked versions of this question, and we considered the pros and cons of various changes, examining what might be gained and lost through each alteration.

Results of the Focus Group: The students expressed appreciation for the level of specificity of the prompt, and also for the fact that the way it was written, it would encourage a different kind of writing and thinking than they were used to. "It would be hard to respond to this with just a five paragraph essay," one noted.

The students who participated in the focus group wanted a prompt that:

- Explicitly identified the essay's audience;
- Positioned them as authorities; and
- Identified the expected components of the essay

Although they expressed enthusiasm for the prompt as it was, they thought it would be easier to do well if it was written as a "scenario," in which the student imagined him or herself as a blogger, addressing his or her blog readers. Immediately after the focus group, I revised the prompt based on our conversation. That revised prompt is included in Appendix B of this paper.

Allison and Gupta (1997) cited Hamp-Lyons (1988) in their study, noting that Hamp-Lyons, "Proposes the four rhetorical categories of

topic, comment (the instructional V or VP), focus (topic-narrowers) and perspective (determining viewpoint to be taken) in analyzing a prompt. Her example (which she presents in tabular form) is 'Discuss (comment) the use of (focus) nuclear energy (topic) to benefit mankind (perspective)'" Allison and Gupta (1997) note that not all four of the rhetorical categories need to be present in all prompts, citing "perspective" as a category that might not always be needed. However, the students in the focus group valued a clear and explicit perspective extremely highly. I will, in the future, hesitate to omit this component of the prompt.

Pedagogical Value of the Focus Group: The benefits of the focus group were rich and varied, and included inspiring in students:

Investment: The participants in the group became stakeholders in the project. At one point one of the participants remarked, *"I'm worried that the specific audience would detract from the personal evaluation. I don't want them to be too focused on the other person."* They became generous readers, willing to consider what was already working about the prompt, rather than acting as critics.

Imagination: The participants in the group demonstrated their ability to envision what an essay in response to the prompt might look like: *"Are you saying the thesis is going to be 'these elements of the reading contributed...'"* The students were freshmen, imagining what other freshmen might produce in response to the prompt. Alone, I could not have as effectively imagined how writers at that level would respond to the question.

Lateral Thinking: The participants drew useful, vivid analogies as they worked to develop their own ways of articulating the task. One participate noted that *"It's like a Star Wars argument."* We were collaborating in order to make the prompt better, and that placed the students in a constructive, authoritative position.

CONCLUSION

Both the think aloud and the focus group offered insights that could influence those wishing to revise their approach to writing prompts. The student in the think aloud part of the study did not focus on the "action verbs" of the prompt itself to figure out what the prompt was asking her to do. Instead, she focused on the learning outcomes, which helped her understand that overall purpose of the assignment, and the role the prompt was asking her to take on. If this student's response is representative, it would be beneficial for instructors to put additional time and effort into developing learning outcomes, and to include them in every assignment, and even at the beginning of every class session, as they introduce work for the day.

The methodologies used for this study have the potential to be useful not just as *research* tools, to help instructors better tailor their prompts to their students' cognitive needs, but as *teaching* tools. One could incorporate these methods into one's classroom. Depending, of course, on the structure of the course in question, it might be possible to implement "focus groups" and "think alouds" as prep activities for difficult prompts. I personally intend to incorporate both think aloud and focus group activities into my classroom.

I suggest, as a means of implementing the think aloud protocol, that an instructor might consider asking students to record themselves reading aloud and thinking aloud as they attempt understand a writing assignment. They could then send in the audio file and any questions that might have arisen for them in the course of "thinking aloud." I would recommend that this assignment be graded on a pass/fail, complete/incomplete scale. An added benefit of adding such an assignment to a course is that it will force students to think in detail about the writing assignment at an earlier date than they might otherwise. I have often found that when I review a prompt in class, students tell me they understand and have no questions, but when they begin writing, they find they have

many questions. This assignment might also help students get started earlier on their writing assignments, as they will be required to start thinking about the assignment at an earlier date.

Focus groups could also be incorporated into most classrooms. An instructor could distribute a writing assignment in advance of assigning it formally, and ask students to get into small groups to talk about what they believe could be done to improve it. They would then report their findings to the class, and the instructor could take them into consideration and revise the prompt accordingly.

REFERENCES

Ackerman, J. (1989). *Students' Self-analyses and Judges' Perceptions: Where Do They Agree?*. Center for the Study of Writing.

Allison, D., & Gupta, A. F. (1997). Why Some Questions Don't Work: Evaluating Examination Prompts in an Educational Setting. *Language and Education*, 11(3), 147-162.

Bereiter, C., & Bird, M. (1985). Use of thinking aloud in identification and teaching of reading comprehension strategies. *Cognition and instruction*, 2(2), 131-156.

Bloom, B. S.; Engelhart, M. D.; Furst, E. J.; Hill, W. H.; Krathwohl, D. R. (1956). Taxonomy of educational objectives: The classification of educational goals. Handbook I: Cognitive domain. New York: David McKay Company.

Gottschalk, K (2015). Indispensable Reference for Teachers of First-Year Writing Seminars.

Hamp-Lyons, L. (1988). The product before: Task-related influences on the writer. *Academic Writing: Process and Product. ELT Docs*, 129.

Hamp-Lyons, L. (1991). *Assessing second language writing in academic contexts*. Ablex Publishing Corporation, Norwood, NJ

International Journal of Scientific & Engineering Research Volume 3, Issue 9, September-2012 1

Jahandar, S., Khodabandehlou, M., Seyedi, G., & Abadi, R. M. D. (2012). The think-aloud method in EFL reading comprehension. *International Journal of Scientific & Engineering Research Volume*, 3(9).

Koro-Ljungberg, M., Douglas, E. P., Therriault, D., Malcolm, Z., & McNeill, N. (2013). Reconceptualizing and decentering think-aloud methodology in qualitative research. *Qualitative Research*, 13(6), 735-753.

Kroll, B., & Reid, J. (1994). Guidelines for designing writing prompts: Clarifications, caveats, and cautions. *Journal of Second Language Writing*, 3(3), 231-255.

Pillai, A. D. (2014). *Malaysian Journal of ELT Research*, 10(2), 82-106.

Polio, C., & Glew, M. (1996). ESL writing assessment prompts: How students choose. *Journal of Second Language Writing*, 5(1), 35-49.

Rank, A., & Pool, H. (2014). Writing Better Writing Assignments. *PS: Political Science & Politics*, 47(03), 675-681.

Swales, J. (1982). Examining examination papers. *English Language Research Journal*, 3, 9- 25.

Van Someren, M. W., Barnard, Y. F., & Sandberg, J. A. (1994). *The think aloud method: A practical guide to modelling cognitive processes (Vol. 2)*. London: Academic Press.

Nelson, J. (1990). This was an easy assignment: Examining how students interpret academic writing tasks. *Research in the Teaching of English*, 362-396.

APPENDIX A:

The assignment, given to focus group and think aloud participants in May 2015:
Some people eat Oreos by eating the cream first, while others eat the cookie bite by bite. It's common knowledge that it's possible for two people who enjoy a food to enjoy that food through different processes. And yet we often don't talk about how our processes as we read writing we enjoy vary from person to person. In this assignment, I'm asking you to write about how you went about reading a piece of writing you enjoyed. Your first task is to choose something you read. You might choose:

- A specific novel
- A specific webcomic
- A specific blog
- A text message that made you smile
- A reading assignment for a class (!?!)

You should choose a specific piece of writing, NOT a type of writing.

After you've identified a piece of writing you've enjoyed, I would like you to identify on one element of your reading process as you read that thing. Your process is probably quite complex, with many features. I'm asking you to choose just one feature. Elements of your process might include:

- Whether you read every word or skimmed over some.
- How often you started and stopped.
- Whether you read the text in order or out of order.
- What expectations you had before you began reading
- What questions you asked yourself as you read
- What associations and connections you made in your mind as you read

Now, you're almost ready to write your essay. Your writing task is to assess whether the element of your process that you chose contributed to or detracted from your enjoyment of whatever it is that you enjoyed reading. Answering this question will form the core of your essay/blog post.

Learning Outcomes of this assignment: When you complete this assignment, you will have:

- Identified an experience of reading that was enjoyable to you
- Learned to talk about your reading experience as an active process
- Examined the implications/results of implementing one part of that process

APPENDIX B

The Assignment Revised According to Input from the Focus Group:

Some people eat Oreos by eating the cream first, while others eat the cookie bite by bite. It's common knowledge that it's possible for two people who enjoy a food to enjoy that food through different processes. And yet we often don't talk about how our reading processes vary. In this assignment, I'm asking you to write about how you went about reading a piece of writing you enjoyed.

The scenario: You maintain a popular blog. You made a post recommending a piece of writing. One of your loyal readers posted a comment thanking you for the recommendation but telling you that they read it and didn't like it. They suspect that the problem was not that the writing was bad, but that they didn't go about reading it in a way that would maximize their enjoyment. They politely request that you make another post shedding some light on your reading process.

Your first task is to choose something you read and enjoyed. You might choose:

- A specific novel, or series of novels
- A specific webcomic
- A specific blog
- A reading assignment for a class (!?!)

You should choose a specific piece of writing, NOT a type of writing. Remember, in this scenario, you've ALREADY COMPLETED a blog post recommending this piece of writing, so you don't need to spend any time talking about why you think it's good.

Your process is probably quite complex, with many features. You'll need to find a way to focus your essay. You might focus on one feature of the text that made a few different reading techniques necessary. Or you could focus in on just one element of your process and explore in some detail all of the positive effects that came from that. Elements of your process might include:

- Whether you read every word or skimmed over some of the text.
- Whether you reread the text
- How often you started and stopped.
- Whether you read the text in order or out of order.
- Expectations you had before you began reading
- Questions you asked yourself as you read
- Associations and connections you made in your mind as you read

Once you've chosen elements (or an element) of your process, you're almost ready to write your essay. Your writing task is to assess how those elements (or element) of your process that you chose contributed to enjoyment of whatever it is that you enjoyed reading. Answering this question will form the core of your essay.

So, for example, you might write an essay about all of the reasons why reading George R. R. Martin's Game of Thrones after watching the HBO series helped you enjoy it more (you already had a more vivid image of the world, you were able to look out for changes the show had made, you had a sense of 'getting ahead' of your friends who only knew what happened up to the ending of the latest episode, etc).

Learning Outcomes of this assignment: When you complete this assignment, you will have:

- Identified an experience of reading that was enjoyable to you
- Learned to talk about your reading experience as an active process
- Examined the implications/results of implementing one part of that process

The Digital Divide: Perspectives on Integrating Digital Media into First-Year Writing Seminars

Kaylin Myers

Graduate Research and Teaching Fellow 2014-2015

INTRODUCTION

We live in a world surrounded by screens, images, and videos. For undergraduate students especially, media technologies shape everyday life, including the acquisition of information, the development of an identity, and the construction of social relationships. The evolution of language, the publication of print materials, and the establishment of a larger community have always been a unique part of what it means to be human. And now, all three of these issues are wrapped up in digital media and social networking, which function as virtual spaces in which readers can become users, authors, and audience all at the same time. These days, because the Internet has created a network of millions, readers can recreate their favorite stories through fan fiction and develop them communally with other users. This vast collaborative mode of creation changes the way we think about authorship, ownership of intellectual property, and textuality itself.

At the beginning of a class I taught called "Female Monsters and Monstrous Females" at Cornell, I asked students to leave their laptops and cell phones in their bags. Because the course was a small First-Year Writing Seminar, I thought it would be more beneficial for students to be engaged with physical texts rather than being distracted by their computers. In many ways, I believe limiting technology use within the classroom is beneficial because it incites students to participate in discussions and engage with their peers. Inevitably, perhaps, I noticed several students texting on their phones under the table. I pulled a student aside after one class and reminded her of my laptop/phone policy for the class. She apologized, but quickly assured me that she had not been texting—instead, she had been live-tweeting some of our discussions in class. It was at this moment that I began thinking about integrating technology into the classroom. What could it do for students to use these tools to learn writing? How would integrating technology affect not only the way they write, but also the way they think about writing as a skill? After this course, I set out to investigate how integrating technology into my courses could influence student learning—and potentially how harnessing students' zeal for their devices could benefit their own writing.

The purpose of this pilot study is to examine the differences in how students and instructors perceive the incorporation of digital media and traditional writing in Humanities First Year Writing Seminars (hereafter referred to as FWS). My questions include: Do students believe that using media to enhance their

essays will improve their final grade and/or their understanding of content? Do instructors believe that using digital media in the classroom will create more engaged and motivated students? To help answer these questions, I performed a two-part study that I describe fully below.

There have been a variety of studies that investigate the use of digital media and digital storytelling within university classrooms, as well as how technology can be a powerful pedagogical tool for students. And yet, despite the pervasiveness of media technology in our own lives and our students' lives, comparatively few classrooms incorporate digital media with traditional writing assignments. Although much attention has been paid to issues of students' technological experience and access to technology (or lack thereof) in schools, less attention has been paid to helping students develop critical digital literacies. We know, for example, that the vast majority of students at Cornell University have access to and use computers and social media—either from their own devices or from devices available on campus. As instructors, we constantly notice that our students are glued to their cell phones and their computers. In my Graduate Research and Teaching Fellowship, we even spent an entire week discussing how technology changes our students' relationship to the material we teach.

And yet, students rarely get to have these comparative discussions about media and the texts they read in class. For them, digital media and technology are often separate from their work in humanities courses. In order to bridge this gap between technology and writing courses, it is first necessary to study how students and instructors perceive digital media before we can develop long-term methods of integrating the two—methods that allow students to incorporate technology that they use in their daily lives with literature and writing. I began this current study because I would like to integrate digital media and traditional forms of writing into my own classes, but I was initially unsure how to begin the process. The purpose of this study is therefore to examine the differences in how students and instructors perceive the incorporation of digital media and traditional writing in the Cornell FWS.

"WHAT'S AT STAKE? DEFINING 'DIGITAL MEDIA' AND ITS PARAMETERS"

One of the greatest challenges of studying media in the classroom is simply defining the phrase "digital media" for both instructors and students. Because there is a vast array of literature on digital media and web-based learning, definitions tend to vary wildly. Traditionally, digital media refers broadly to audio, video, and photo content that exists in computer-readable formats (Lake & May, 2012). The pace at which technology evolves and changes is swift—as Carroll (2014) suggests, "it makes a book on writing for digital media a bit like chasing one's tail." The growth of social media, the adoption of tablets and e-readers, interactive and responsive web design, and the proliferation of digital publishing are only a few ways that technology has changed literary studies and the humanities (Carroll, 2014). Even though it is difficult, as Carroll notes, to keep up with such a rapidly changing landscape, digital media provides educators with an opportunity to both expand their own pedagogical methodologies and improve their students' learning experience.

Mark Windschitl (1998) and others (Greenhow, 2008; Hartley & Bendixen, 2001; Roschelle & Pea, 1999) have outlined extensive research practices for incorporating digital media and technology into classrooms. Windschitl (1998) suggested that researchers focus on three subjects specifically: using technology for student inquiry, examining student communications via the internet, and using qualitative-based research methods to study web-based learning. He described using computers and web-based technology as a potential repository for information and new knowledge that could in turn help students more effectively understand lessons and communicate their understanding to both

their peers and their instructors. Jeffrey T. Grabill (2003), on the other hand, examined the association of computers and composition to focus on the direct relationship between different economic classes and their access to technology. Although Grabill observed that from 2000 to 2002 access to computers increased among all user groups, he also suggested that other gaps (like understanding how to use technology productively or understanding how to organize and sort large volumes of information) pose significant problems even as access increases.

This research remains valuable today, especially now that access to technology has skyrocketed in our society. It is now expected that every college freshmen should own or have access to a computer—that they should be able to access the internet at all times for research and other course purposes. Similarly, high school students now come to college with a wider array of technological skills. Wells & Lewis (2006) determined, for example, that since the mid-1990s, the percentage of public high schools connected to the Internet exploded from 35% to 100%. Public high school instructional classrooms with Internet access even grew to 94%, up from 14% a decade earlier. In 2008 outside of schools, more than two thirds of people in the United States have Internet connections at home, more than half of which are broadband (Horrigan, 2008).

As access to technology has expanded, both the nature of the web and student experience of the web has also shifted dramatically. Ten years ago, the use of the internet in classrooms was viewed as an educational resource akin to traditional classroom tools—for example, a website was a source of information comparative to a book, or even a means of displaying content like an overhead transparency. Websites and digital platforms were also controlled by a small number of providers, and users were limited in the ways they could add to or interact with such sites. In other words, individuals were not able to share knowledge, content, and posts as easily as they can now if they lacked knowledge of HTML. More recently, technology and digital media has shifted in a new direction—digital programs and platforms now facilitate a participatory, collaborative, and interactive method of learning (Kim & Bagaka, 2005).

In a similar vein, students who graduate from college are expected to know how to use basic technology in order to both communicate and express themselves well. According to the International Society for Technology in Education, today's graduating students face a job market that requires the integration of digital media and communications—a job market that is often immersed in a digital environment that emphasizes the capacity for innovation, leadership, interdisciplinary collaboration, and collaborative problem solving (ISTE, 2007; D'Aloisio, 2006). New standards emphasize the learner, his or her experiences and choices, and the cognitive, social, and cultural dimensions of how technology is used in various settings. For instance, student skills must often include creative and original multimedia work in project-based teams in which the problems, tasks, roles, and techniques are constantly changing (ISTE, 2007).

The question remains, however, how we as instructors can integrate technology successfully into the classroom. As a guideline for this pilot study, I have used D'Aloisio's 2006 report on a large study by McCorkle et al. (2001). McCorkle et al. ran a study in which they surveyed 765 students and 308 instructors about their perception of digital media in the classroom. Their research questions included: "How do technology preferences differ between students and instructors?" and "What are the student and instructor perceptions of technology use, technology support, and the effectiveness of digital tools? (D'Aloisio (2006) and McCorkle et al., 2001)." The major finding of these two studies appear to be that both students and instructors want to teach and learn from new technologies, despite issues of logistical failure, lack of training, and the potential for distraction. They also found that students appreciate and use digital tools more

than instructors. According to D'Aloisio and McCorkle et al., students are more actively motivated in class when they can connect each lesson or skill with a needed skill in their own lives. In other words, students will embrace new technologies or digital media if they understand how it will help with future goals. Instead of using technology for social value or for entertainment, they can learn to use new technologies as a skill set for the future. Communication-wise, professors have attempted various formats that tap into today's complex mix of communications technologies.

Like D'Aloisio and McCorkle et al., this present study is interested in examining both the disparities and similarities between student and instructor perceptions of technology. Today, students have a copia of devices and technologies at their fingertips. They also have more choices than ever about how and where to spend their time learning—whether that is online, in the classroom, at home, or in public. In his 2007 study, Lenhard et al. observed that students prefer non-traditional forms of communication like text messaging, instant messaging, and social networking to traditional methods of e-mail and in-person communication. These are the students who come to fill the First Year Seminar classrooms at Cornell University—students who are deeply immersed in social media and interactive digital media within their own lives. These are also students who could benefit from learning about writing across different genres. After all, composing an e-mail, a tweet, an essay, and a letter are all vastly different activities that we require of our students daily. I believe incorporating digital media into the classroom in order to teach students about different genres of writing would be immensely productive during the learning process.

It is my hope that the following study will give some insight into how Cornell students and instructors regard technology in the composition classroom. While many current studies focus solely on one perspective—either from the viewpoint of the student or the instructor—this pilot study examines and compares both perspectives in order to better understand the gaps that exist between these two groups. This study thus investigates a select group of Cornell undergraduates and graduate student instructors in order to document their digital fluencies and perceptions of digital media in the classroom. I began the study with two goals: to identify gaps in attitude and valuation in order to help instructors harness digital media in writing seminars, and to better understand how students approach writing as a field. Perhaps on a smaller scale, this study could also provide First-Year Writing instructors with relevant information to create courses that students are both interested in and motivated to take. I believe that writing can become more accessible because of this interdisciplinary approach, especially for those students who are less comfortable with writing as a field.

METHODS

This data was collected during the spring semester of 2015 from Cornell University undergraduates currently enrolled in a FWS. Even though sophomores, juniors, and seniors are allowed to take these courses, the majority of participants who completed this survey were first year students. The only participation criterion for this survey was that individuals were enrolled in a FWS. Students are required to take two First Year Writing Seminars during their tenure at Cornell University. A typical FWS course enrolls up to seventeen students from a range of departments—from English to Engineering to Agriculture and Life Sciences. The courses must include six to nine short essay assignments and no more than 70 pages of reading per week. One-third of these courses are taught by faculty, while the remaining two thirds are taught by graduate students with varying levels of teaching experience and interest in pedagogy. These courses give Cornell undergraduates a space to experiment with a range of genres, gain a sense of their own writing style, and improve their clarity, coherence, argumentation, and stylistic control.

The two parts of my study consist of

two separate surveys that ask Cornell undergraduates and FWS instructors about their experiences of writing seminars—both of these surveys can be found in the appendix. In my experience teaching writing seminars, I have found that there is a large gap between my own understanding and my students' understanding of technology. I therefore created both of these surveys to analyze the differences in experience and perception among undergraduate and their instructors in order to bridge this gap of experience and practice. I began by creating a survey for Cornell undergraduates that focuses on their experience with and perceptions of digital media in the classroom. It consists of nineteen questions and uses both multiple choice and short fill in the blanks. Sample questions include: "Do you prefer to write a traditional essay that does not incorporate digital media?"; "How many hours a day do you use social media?"; and "What are your top social media sites?"

First, I sent out an online version of this survey across the Humanities list servers. I asked instructors (both graduate students and faculty) to forward my survey to students that were currently enrolled in a FWS. The selection process for this survey was random and the responses are anonymous. In addition to sending out this survey via e-mail, I also had the opportunity to visit and observe a FWS in the Medieval Studies this semester after they implemented digital media into a writing assignment. The instructor of the course asked her students to write a digital essay using Scalar—a publishing platform that allows authors to assemble media from a variety of sources and juxtapose them with their own writing. One of the most fascinating aspects of Scalar as a platform is that it allows authors to create unique digital writing, including nested, recursive, and non-linear formats. The instructor of this Medieval Studies FWS asked her students to research one aspect of Chaucer's Wife of Bath, write a paper in Word using critical secondary sources, and upload their essay onto Scalar. She then asked students to juxtapose one piece of media with each one of their written pages—media could include but was not limited to images, videos, audio, and links to other websites. On the day that the final project was due, I attended this class in order to survey and observe her students. The instructor allowed me to briefly explain my study to her students and pass out paper copies of the online survey. The surveys were anonymous, and students had the option of leaving class five minutes early or staying to complete the survey. Every student stayed to complete the survey, and there were nine students in the course.

At the same time that I was surveying undergraduates, I also sent out an online survey to instructors across the humanities list servers. The recruitment for instructors was randomized and like the undergraduate survey, the results were anonymous. Sample questions from this survey include: "Have you ever incorporated digital media into a Freshmen Writing Seminar at Cornell?"; "If yes, in what capacity?"; and "Did you encounter any specific problems?" I therefore collected a broader base of survey data with the online surveys, as well as data from the specific FWS in the Medieval Studies Department. This FWS provided a unique opportunity to gauge how students perceived the integration of digital media into a final project before and after its completion.

RESULTS

Student Perceptions of Digital Media in the Classroom:

As previously described in the methods, there were between 16 and 18 student respondents for surveys. Students who completed the paper surveys occasionally skipped questions or left questions blank, making the tally for certain questions lower than the overall number of completed surveys. Of these 18 students, 17 of them were freshmen and 1 student was a junior. Students came from a variety of departments and majors, but the majority of participants were in the College of Agriculture and Life Sciences and in Engineering. There was only 1 participant who identified as a humanities student within the College of Arts

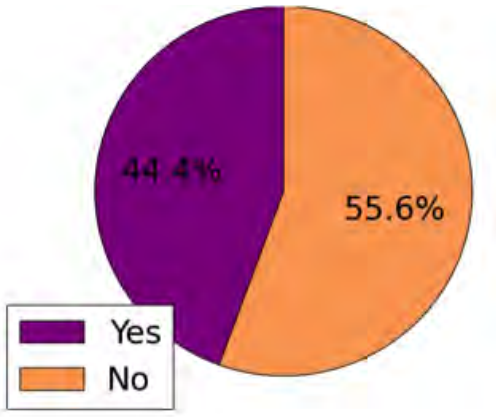


Figure 1: Students were split on whether they preferred traditional writing or writing with digital media (n=18)

and Sciences—this student reported that his/her major was Linguistics. Of these 18 students, 8 were male and 10 were female.

For instructors, between 16 and 20 instructors responded to each question on the survey. Like the undergraduate participants, several instructors did not answer every single question—thus, several questions have varying n values. Of these graduate student participants, 4 were third years, 7 were fourth years, 7 were fifth years, and 2 were sixth year students. The overwhelming majority of these participants were women—this study was comprised of 16 female and 4 male graduate students. While the majority of participants also taught within the English Department, there were also graduate student instructors from the Medieval Studies, Art History, and History Departments.

As Figure 1 shows, the undergraduate survey indicated that students were generally split on whether they preferred traditional writing to writing that incorporated digital media. When asked: "Do you prefer to write a traditional essay without using digital media?" 44.4 percent said yes, while 55.6 percent said no.

In Figures 2 and 3, there is a wide gap between instructor and student perception. I provided the following statement: "Incorporating digital media and writing is too time-consuming for students" and asked both students and instructors to indicate their preference on a Likert Scale. In the undergraduate responses, the overwhelming majority either agreed or strongly agreed with this statement (62.5% agreed, 25% strongly agreed, and only 12.5% disagreed). For instructors, 75% disagreed, 12.5% strongly disagreed, and 12.5% agreed. The difference between these figures is significant, suggesting that there is a disconnect between the way students and instructors understand how much time it takes to incorporate digital media into student writing.

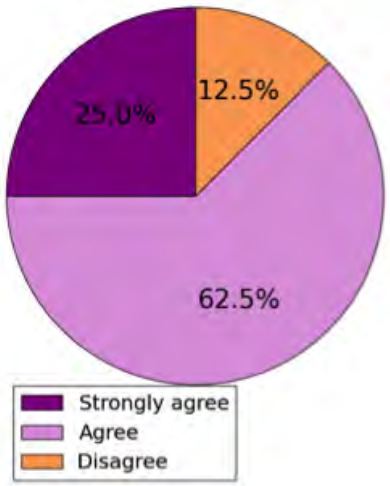


Figure 2: Student time consuming (n=16)

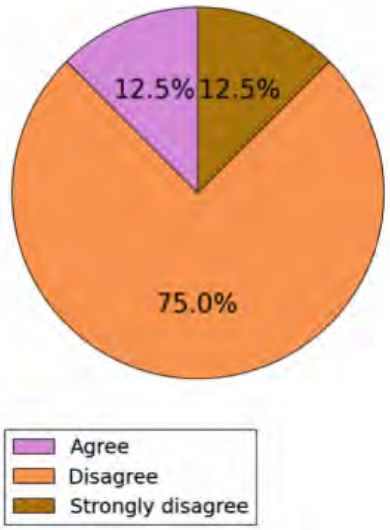


Figure 3: Instructor time consuming (n=16)

As opposed to Figures 2 and 3, Figures 4 and 5 show agreement between student and instructor perceptions. When provided the following statement with an accompanying Likert Scale, 64.7% of students agreed that incorporating digital media produces a more creative final product. In instructor responses, 50% agreed that digital media produces more creativity. Moreover, 23.5% of students and 20% of instructors strongly agreed in favor of increased creativity.

Observations from a Medieval Studies FWS: When I surveyed the Medieval Studies FWS directly after their assignment was due, I was able to observe their discussion about the integration of digital media and ask questions. The course was a small seminar consisting of 9 students, and the atmosphere was relaxed when the instructor asked students to give their opinions on the project they had just completed. Within their discussion, the class was firmly split down the middle as to whether they enjoyed the project or not. For those who voted positively, they seemed to like best the creative freedom of choosing images to support their own writing. However, they spoke mostly about digital media as corroborating or enhancing their critical analysis. One student commented, "I liked the visual aspect of the project—that I could talk about art and painting and add to the Wife of Bath." Another student who enjoyed the project reflected on how the assignment changed the way she viewed some of her other courses: "It made me realize that digital media is something important to learn—and also that almost all of my professors add videos and images to their lectures even if they aren't relevant." At this point in the conversation, other students chimed in to agree, and one added specifically: "Yeah, in classes it seems like professors use pictures and other technology to grab students attention rather than in specific lessons."

The students who voiced negative reactions were mostly focused on the amount of time the project required. One student said, "Finding pictures or videos was hard because I felt like it wasn't adding anything to my essay."

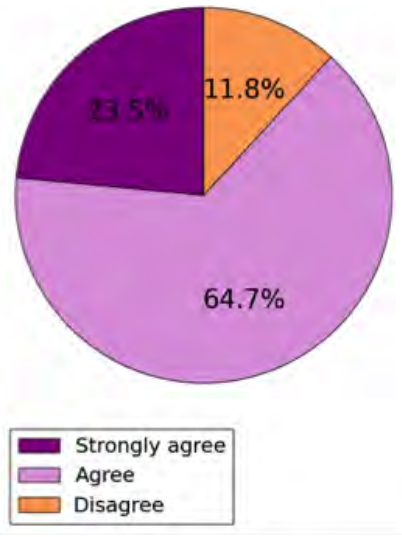


Figure 4: Student creative (n=17)

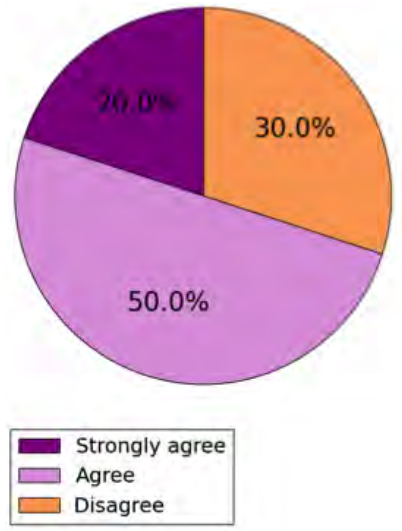


Figure 5: Instructor creative (n=20)

Following this student's comment, another jumped in to say, "Yeah, I agree. I was writing about something abstract so it was hard find pictures that weren't generic." 7 out of 9 students reported that they had problems using Scalar—these issues included glitches with the platform itself as well as difficulties learning and navigating a new digital tool. Almost all of the students said that it took a significant amount of time to "play around" with the platform before they were comfortable turning in their projects for a final grade. As the discussion wound down, I asked students

if they thought that using digital media in their essay made their work more creative. 5 of the 7 students thought that their work was more creative—the remaining 2 students said that it seemed as though they were trying to fulfill an assignment rather than adding true creativity to their writing. But one student also added that the end product was "probably more fun" to read than a normal essay.

Instructor Perceptions of Digital Media in the Classroom:

Out of the instructors who completed my survey, 19 out of 20 reported that they had incorporated digital media into a prior classroom. When asked to record different types of digital media that they use, instructors listed the following examples in descending frequency:

1. Powerpoint
2. Blackboard
3. YouTube
4. Google Docs
5. Audio files/content
6. Twitter
7. Blogs
8. Social media
9. Moodle
10. Scalar

Recent literature shows that instructors often do not consider integrating technology into their classrooms because of insufficient training and lack of time (Schrand, 2008; Vie, 2008). In my results, however, 50% of instructors reported that their use of digital media in their FWS was 'successful' and 40% reported 'very successful'. When asked whether or not they felt confident with using technology in the classroom, 70% answered positively (6 instructors felt 'very confident' and 8 felt 'confident'). Moreover, instead of viewing digital media as a large time commitment, instructors largely disagreed when asked whether incorporating technology is too time consuming (see Figure 3). This finding is contrary to student responses, who largely believed that incorporating digital media into their writing was a laborious time commitment. These findings nevertheless match up with D'Aloisio's (2006) results in that instructors largely perceive the integration of technology as positive—for D'Aloisio, 78% of instructors thought that student engagement improved as the use of digital tools increased and 87% thought that learning outcomes improved.

Perhaps the most interesting result of the instructor survey, however, were responses to the question, "Did you encounter any specific problems when you incorporated digital media into the classroom?" 16 out of 18 instructors that answered this question reported logistical problems when they used technology in the classroom. These problems included weak internet signals, projector malfunctions, and lack of reliable audio equipment. One instructor reported: "The internet on campus can be unreliable; I always have a back-up plan." Another simply responded, "There is limited tech capability in the classrooms." One instructor also added, "Most of the problems were logistical (layout of the room made it hard for some students to see, projector malfunctions, etc.)." Similarly, although 90% of instructors said that they were aware of resources at Cornell that will help them incorporate technology into their classroom, 70% of instructors reported that they had not taken advantage of these resources. Instructors seemed to take the burden of incorporating digital media on themselves rather than seeking out aid or training from the university.

Along with potential logistical issues, other instructors mentioned how digital media occasionally makes their students un-engaged or opens up the gates for distraction. A combination of logistical failure and dis-engagement prompted one instructor to say, "The hookups sometimes worked and sometimes didn't, so on several occasions I ended up asking students to take out their own devices and follow along…I don't like to do that, since students rapidly get distracted by e-mail, etc." Another instructor said, "If I

try to use YouTube clips too early in the class it will make them sleepy/less engaged/more passive, unless it involves small group work." Despite these issues, however, instructors on the whole seemed to agree that technology and digital media was worth the hassle. In general, it appeared that they were eager to learn and teach with a variety of technologies, just as students were generally in favor of learning with them.

CONCLUSIONS

In my limited findings, there seems to be a disconnect in the way that students view technology and the way they view learning writing. Although students use digital media frequently in their daily lives, the push to include it in writing courses is just beginning. Student surveys suggested that despite being willing to experiment with new digital media in the writing classroom, they do not think about technology as a tool for learning writing—rather, they find value in its potential for creativity or as a new skill to be learned in and of itself. Moreover, although students on the whole seemed to be in favor of using technology in a composition class, they also expect instructors to give extensive support for its use. This means that for complex digital platforms like Scalar, students need to be taught explicitly how to work and learn from these instructional technologies. These findings are congruent with the studies and reports from D'Aloisio and McCorkle et al.

While instructors largely said they had the skills to integrate technology, there seemed to be a substantial worry that technology is not sufficiently available or reliable in Humanities classrooms. Instructors also worried that the use of technology like laptops, iPads, and smartphones in the classroom will provide too much of a distraction for students. At the same time, however, they recognized the merit of digital media and technology—perhaps most of all, the majority saw its value in their students' creativity and their increased access to information. Like students, instructors tended to use digital media and new technology to enhance comprehension about specific content. For example, using images or audio files to help students understand a complex passage in a book. They were less likely to use digital media as a tool for learning writing specifically. This is perhaps where the disconnect stems from in student and instructor perceptions of technology. Because instructors use digital media for content rather than composition, students are less likely to see these tools as beneficial for the writing process.

Based on these limited surveys and observations, I have compiled some initial recommendations for integrating technology in a FWS. I hope to further corroborate these recommendations by adding to this present study—specifically, I will be surveying my next class entitled "Text(ing) in the Age of Digital Media" throughout the entire semester on their experiences with digital tools and platforms. Based on this study, however, I recommend that when integrating technology into the classroom, instructors explain to students in detail before each project begins how each digital platform could aid or affect their writing. By taking time to discuss the direct benefits for their writing, I believe that students will be more likely to engage with learning new tools and procedures. To get started with integrating technology into the classroom, I would also recommend (at least initially) that instructors choose a digital platform that does not have a steep learning curve. While digital platforms like Scalar are exciting tools for non-linear publication, they are often tricky for students to learn in a short amount of time without extensive support. In the recent past, scholars have produced an astonishing amount of research about technology in the classroom. This study continues such research by adding new voices to the chorus of scholars who examine the complex relationship between both learning and teaching and between instructor and student. While the number of participants in this pilot study is too small to draw broad conclusions, it does hint at approaches or techniques that can help instructors integrate technology more successfully.

REFERENCES

Carroll, B. (2014). *Writing and Editing for Digital Media*. New York: Routledge. xi.

D'Aloisio, A. (2006). Motivating students through awareness of the natural correlation between college learning and corporate work settings in *College Teaching*, 54, 225-230.

Day, J. M., & Kumar, M. (2010). Using SMS text messaging to create individualized and interactive experiences in large classes: A beer game example in *Decision Sciences Journal of Innovative Education*, 8, 129-136.

Grabill, Jeffrey T. (2003). On divides and interfaces: Access, class, and computers in *Computers and Composition*, 20(4), 455–472.

Greenhow, C. (2008). Commentary: Connecting formal and informal learning in the age of participatory media: A response to Bull et al. in *Contemporary Issues in Technology and Teacher Education*, 8(3).

Hartley, K., & Bendixen, L. D. (2001). Educational research in the Internet age: Examining the role of individual characteristics in *Educational Researcher*, 30(9), 22–26.

Horrigan, J. B. (2008). Home broadband adoption 2008. Washington, DC: Pew Charitable Trusts. Retrieved April 6, 2015 from http://www.pewinternet.org/files/old-media/Files/Reports/2008/PIP_Broadband_2008.pdf.

International Society for Technology in Education. (2008). *National educational technology standards for teachers* (2nd ed.). Eugene, OR: Author.

Kim, S. H., & Bagaka, J. (2005). The digital divide in students' usage of technology tools: A multilevel analysis of the role of teacher practices and classroom characteristics. *Contemporary Issues in Technology and Teacher Education*, 5, 316-29.

Lake, S & May, K.B. (2012). *Digital Media: Concepts and Applications*. Ohio: South Western Cengage Learning.

Lenhart, A., & Madden, M. (2007). Pew Internet Project data memo. Washington, DC: Pew Charitable Trusts. Retrieved April 15, 2015 from http://www.pewinternet.org/files/old-media/Files/Reports/2007/PIP_Teens_Social_Media_Final.pdf.pdf.

McCorkle, D. E., Alexander, J. F., & Reardon, J. (2001). Integrating business technology and marketing education: Enhancing the diffusion process through technology champions in *Journal of Marketing Education*, 23, 16-24.

Roschelle, J., & Pea, R. (1999). Trajectories from today's WWW to a powerful educational infrastructure in *Educational Researcher*, 28(5), 22–25.

Schrand, Tom. (2008). Tapping into Active Learning and Multiple Intelligences With Interactive Multimedia: A Low-Threshold Classroom Approach in *College Teaching*, 56, 78-84.

Vie, Stephanie. (2008). Digital Divide 2.0: "Generation M" and Online Social Networking Sites in the Composition Classroom in *Computers and Composition*, 25, 9-23.

Wells, J., & Lewis, L. (2006). Internet access in U.S. public schools and classrooms: 1994–2005 (NCES 2007–020). Washington, DC: U.S. Department of Education, National Center for Education Statistics. Retrieved April 2, 2015, from http://nces.ed.gov/pubs2007/2007020.pdf.

Windschitl, M. (1998). The WWW and classroom research: What path should we take? in *Educational Researcher*, 27(1), 28–33.

Using Technology in the First-Year Writing Seminar

Kristie A. Schlauraff

Graduate Research and Teaching Fellow 2014-2015

ABSTRACT

In an atmosphere where students and instructors alike are inundated with myriad technologies, how do we construct a functional definition of the term "technology" in the context of higher education? What potential does technology have as a means of enhancing student learning in First-Year Writing Seminars (FWS)? My study explored how FWS students and instructors in the Cornell University English Department define and experience "technology" in the classroom. I surveyed 73 students and 11 instructors in order to better understand how these groups define "technology," how they perceive technology being used in their FWS courses, what technologies they use on a daily basis for academic and non-academic purposes, and what technologies they believe might be productively used in FWS courses. In addition to surveys, I had a discussion with my FWS students where they had the opportunity to talk about this issue in greater depth. Based on my results, "technology" might be defined as a set of tools that can be used in FWS courses to distribute course materials and information, enable interactive and collaborative learning, and supplement content. While instructors were optimistic about the potential of technology in FWS courses and envisioned how it might enhance learning, many students shared the sense that "FWS is one of the few classes that does not need much technology in order to learn everything we need to." My results suggest FWS instructors should think carefully about why and how they incorporate technology; students need to understand how the technology will help them more effectively achieve the goals of the assignment.

INTRODUCTION

What is the role of technology in the First-Year Writing Seminar (FWS)? In recent years, questions about how technology might be used to enhance FWS courses have led educators to explore the potential of electronic devices from computers to e-readers, digital platforms and programs such as Blackboard and Netflix, and social media from wikis to Facebook. However, the diversity of these technologies raises questions about how we define the term "technology" itself, especially in an atmosphere where we are inundated with new innovations and breakthroughs every day. More importantly, the array of available technologies means that instructors seeking to integrate them into the classroom need to carefully consider what each technology has to offer and how it will enable students to more effectively achieve the goals set forth in the course objectives.

While new technologies have impacted learning and teaching in all disciplines, the digital age has been particularly important (and controversial) in relation to composition.

In a 1999 article in *The Chronicle of Higher Education* titled "Technology Transforms Writing and the Teaching of Writing," Wendy R. Leibowitz explains that computers have fundamentally changed the way students write. Although she points to some positive aspects of computers, Leibowitz offers a relatively negative perspective that underscores the article's subtitle: "Many professors try to combat the bad habits they fear their students pick up on computers." Computers have allowed for electronic submission, electronic comments and easy distribution. They open up a broader audience and allow students who may not be comfortable speaking in class the opportunity to participate online. However, computers have also led students to adopt more conversational tones in their writing, produce longer (but not necessarily better) essays, rely too strongly on spelling and grammar check programs, and submit writing that has not been revised. Computer culture in general encourages speed, Leibowitz argues, which produces not only bad writing practices, but bad reading practices. Essentially, the article presents technology as a double-edged sword that introduces new possibilities, but also creates new problems: "The perils are clearer […] The possibilities are exciting, but their effectiveness is largely unproved, say faculty members who teach writing" (Leibowitz, A67). In the sixteen years since the publication of Leibowitz's article, the "perils" of computers, and technology more broadly, continue to shape course design in the humanities. While some instructors embrace it, creating activities and assignments that specifically ask students to use various technologies, others view it as a distraction, taking steps like banning the use of laptops and other devices in the classroom.

What drives instructors to engage in or shun technology? One of my recent frustrations with higher education, which may have some bearing on this question, has been the commodification of courses. Students tend to adopt consumerist attitudes, "shopping" for classes and looking for instructors who entertain rather than educate (though admittedly, the two are not mutually exclusive). At a recent meeting of the Graduate and Professional Student Assembly (GPSA), where I serve as a Voting Representative for the Humanities, an undergraduate Student Assembly (SA) Representative asked the GPSA to support a resolution for the Implementation of Open Course Evaluations. Described as a kind of "Yelp for courses," these evaluations are intended to provide students with more information about classes before they enroll including the median grades, course syllabi, and student feedback. One set of evaluative questions the SA Representative proposed as a model, which has been instituted at peer institutions like Yale University, asks "Was this class easier or more difficult than the other classes you have taken at Cornell?" During our discussion of this resolution at GPSA, many of the humanities students expressed exasperation regarding this particular question because of the likelihood that it would impact the humanities more than the sciences.

Students majoring in science are unlikely to avoid a required course in their discipline because of the level of difficulty; however, they might choose to prioritize these courses and select easier classes to fulfill humanities requirements like the First-Year Writing Seminar. Alternatively, they might decide not to take elective courses in the humanities that are evaluated as being "more difficult than the other classes." These decisions can have dire consequences for graduate students and faculty members in the humanities. Courses can be cancelled if enrollment is too low, which can have ramifications for graduate students whose funding is dependent on teaching or for junior faculty members who are up for tenure. In the current job market, which has spurned articles in *The Chronicle of Higher Education* such as "Graduate School in the Humanities: Just Don't Go," graduate students are under great pressure to teach courses that attract undergraduates. While GPSA members said that they did not want to be compelled to teach "easier" classes in order to fill seats, they also feared that teaching more difficult courses that

could be cancelled for low enrollment might impact their chances on the job market.

Even without the kind of open course evaluations the SA is proposing to implement, graduate students in the English department at Cornell University work hard to design First-Year Writing Seminars with "sexy" titles and course descriptions that undergraduates will find alluring. In recent years, some of the FWS titles offered through the English department have included "Sluts, Spinsters and Drag Queens," "Spy Writers," "The Doctor is In," "Beyond *The Hunger Games*," and "Fools, Fops, and Idiots." In an atmosphere where students want to "shop" for classes, I would argue that the incorporation of technology and media becomes another way to make courses "sexy." Although catalog descriptions do not generally list the kinds of technology that will be used throughout the semester, many instructors suggest openness to technology through their inclusion of films, television series, podcasts, video games, and other non-textual materials. While my study focuses on the former, technologies used in the classroom, it is important to recognize that instructors must think about technology on many different levels when designing their courses. How will technology be used in the course as a whole, how will it be used in high-stakes assignments, and how will it be used on a daily basis?

In the study that follows, I address three central questions: first, how do students and instructors define "technology" that might be used to enhance student learning in higher education? Second, what technologies do students and instructors use in their daily lives for academic and non-academic purposes? Third, what potential do students and instructors envision technology having as a means of enhancing the teaching of writing and the understanding of course materials in FWS courses? In the Literature Review section I consider the expectations of technology as part of the English curriculum. Furthermore, I explore the ways in which instructors have attempted to integrate technology into the classroom, calling attention to both the benefits and drawbacks of such assignments. In the Methodology section, I describe my objectives in conducting this study. I also explain the design of my study and the nature of the data I collected from students and instructors. The Results and Conclusions sections that follow present a comprehensive overview of student and instructor responses to my survey questions and identify the central conclusions I draw from my data.

LITERATURE REVIEW

In a recent article publish in *Pedagogy*, Kerri Hauman, Stacy Kastner, and Alison Witte contend that "implementing technology in the classroom seems to be a shared goal of departments, programs, and universities," but point out the difficulties of actually achieving this objective (Hauman, Kastner & Witte 52). Specifically, they identify the lack of consistent techno-pedagogical instruction for graduate students as a major stumbling block. While some institutions have a required course, others have elective courses that instruct graduate students on how to incorporate technology in the classroom. Some have no courses at all, but offer departmental and/or non-departmental workshops or seminars. Although ten of the forty-one survey respondents indicated their institution had a required techno-pedagogy course, "five described their courses as driven by first-year composition (FYC) curricula, with the purpose of preparing graduate students for the first-year writing sequence" (Hauman, Kastner & Witte, 49). While the authors admit that first-year composition practicums could serve as viable techno-pedagogy courses, they argue that since not all respondents identified these courses as fulfilling that requirement, the possibility should be approached with caution. Again, inconsistency across institutions poses a problem. Compounding this difficulty is the fact that even the definition of techno-pedagogy itself is inconsistent. Ultimately, Hauman, Kastner and Witte make the following suggestions:

1. Techno-pedagogy should involve more than functional literacy.

2. Courses should offer both theoretical and practical instruction, addressing historical and contemporary issues.

3. Techno-pedagogy instruction should be supported by a core course in the program's curriculum. A "computers-and-writing specialist" should teach this course in a wired space.

The idea that departments have a responsibility to prepare graduate students to effectively use technology in the classroom suggests that techno-pedagogy has immense value in higher education. However, the pressure to incorporate technology into writing courses in particular seems to derive from a sense that these courses will become immaterial otherwise: "The threat of irrelevancy is upon those of us who occupy English departments if we do not have the time and resources to create curricula that prepare students to teach with and about technology" (Hauman, Kastner & Witte, 55). While teaching with technology certainly has advantages, it is important to remember that our primary role as instructors is not simply to ward off the "threat of irrelevancy" by employing technology without deeply considering how it enhances the teaching of writing. In order to provide a more comprehensive picture of the relationship between writing and techno-pedagogy, I turn to examples of how technology is currently being used to teach writing, and consider what aspects of these assignments make them successful or unsuccessful.

In an age saturated with digital technologies, many instructors take for granted that students will implicitly know how to use technology and, more importantly, that they will want to use technology. Discussing a wiki assignment that was overwhelmingly unpopular, Debra E. Allwardt cautions that just because students use things like Facebook and text messaging in their daily lives, doesn't mean they necessarily want to use similar technologies in the classroom. When an instructor asked students "how posting information on the wiki was different from posting information on Facebook or MySpace, which use essentially the same tasks of editing and saving," one student replied, "But we want to do that" (Allwardt, 602). What emerges from the various studies on using wikis, blogs, tablets, and other devices, is the idea that students require a lot more guidance on assignments involving technology than instructors imagine. In part, this has to do with the fact that when technology is being used to enhance student learning, it presents challenges that aren't inherent to daily use.

One popular tool that has been used in writing classrooms is the wiki, a platform that allows students to write collaboratively maintain a record of revisions and comments. In "A Wiki for Classroom Writing," Brian Morgan and Richard D. Smith describe the benefits of using wikis to teach writing. Wikis are easy to use, letting authors add and change text, images, and charts. Students can collaborate, but can do so at their own pace without having to be in the same place at the same time. The instructor has more oversight in this type of collaborative project, and can control access to the different parts of the wiki. Morgan and Smith provide some basic guidelines for using wikis, including providing students with a tutorial so they understand how to use the technology, and establishing a clear schedule of due dates. Throughout the article the authors reinforce the ease and success they have had with wikis, stating, "Our experience using the wiki has been entirely positive," "wikis are nearly bulletproof," and "Once you start [using wikis in the classroom], you will wonder how you ever lived without one" (Morgan & Smith, 81; 82).

In her article "Teaching Note Writing with Wikis: A Cautionary Tale of Technology in the Classroom," Allwardt presents a contrary argument. She describes an assignment using a wiki to collaboratively write a literature review as overwhelmingly negative; more than one student wrote "NO MORE WIKI" on the course evaluation (Allwardt, 602). Although Allwardt identifies the same advantages to using wikis as Morgan and Smith, namely the ability to collaborate and see all the changes made over time, she provides what is perhaps a less

idealistic, though more productive example. Three general critiques emerged from the student comments: (1) time management issues, (2) group coordination concerns, and (3) assignment parameters (Allwardt, 600). Many of the students were frustrated that group members didn't always reply in a timely manner and often remained inactive until just before the assignment was due. While the wiki supposedly eliminates the need for face-to-face meetings, students found this challenging in that not everyone was on the wiki at the same time, group members often failed to communicate effectively, and they were unsure what to do about peers who didn't participate or used the site inappropriately (Allwardt, 601). Finally, students wanted more guidance both in terms of the technology and the assignment itself. The fact that many of these issues seem like they could be resolved, or at least mitigated by the instructor, points to the idea that assignments that incorporate technology require careful planning and oversight if they are going to be effective. Although planning and oversight are integral to the success of any assignment regardless of whether technology is involved, it is apparent that every technology presents its own set of challenges that instructors need to account for when developing their activities. Having clear learning objectives and understanding how the technology can help rather than hinder students is key. In Allwardt's assignment, "the technology seemed to overshadow student learning," an outcome that does not benefit the student or the instructor (Allwardt, 602).

In addition to wikis, writing instructors have incorporated blogs, devices like iPads, and Facebook into their classrooms as a means of enhancing student learning. Charles Tryon uses blogs as a means of making real-world connections and providing his students with a broader audience for their work. In doing so, he increases the stakes of their argument and forces them to think about how others will perceive their writing. One of the first steps Tryon takes is to have his students look at models that they then analyze. This helps students understand this particular mode of writing in relation to other forms of writing. Rebecca M. Sullivan exposes another kind of technology use in her article on writing instruction using the iPad. Sullivan's students all received an iPad with specific applications installed (and uninstalled) that allowed them to access course content and other features. Having an identical set of devices for the whole class that were synced with one another allowed Sullivan to create collaborative assignments both in and out of the classroom. In their article, "Engaging Introductory Writing Students Through Facebook Assignments," Elyse D'nn Lovell and Betsy Palmer describe an instance where an instructor replaced an unsuccessful journaling assignment with a Facebook assignment. One of the benefits of this was that students were already habitually posting on Facebook, meaning it was a regular practice like journaling that they were already engaging in on a daily basis. Students felt the experience led to social bonding and enjoyed sharing their personal reflections about the course. However, Lovell and Palmer suggest that instructors be prepared to address different reactions to this kind of assignment. Some students were less than enthusiastic about using Facebook as part of the course. I would add that the examples of student writing Lovell and Palmer provide are not necessarily the kind of writing one would expect in a student journaling assignment. Rather than reflect on course material, students share personal responses to the course itself. For instance, one student writes "I paid money to be here, and learn a specific skillset, but I am forced to takes classes such as this one that have zero bearing on my future" (as cited in Lovell and Palmer, 27). Lovell and Palmer establish their goals for the assignment as hoping for an increased sense of connectedness among students and improved mechanics in their writing (Lovell and Palmer, 26). While these are not by any means unconventional objectives, I am not convinced that the Facebook assignment was the best way to achieve these goals.

All of these examples illustrate how technology

can be used to create non-traditional, interactive assignments that help students learn to write. However, they also reinforce the argument that instructors need some kind of training in techno-pedagogy if they are going to successfully integrate technology into their course design. Using technology haphazardly or just for the sake of using it can have a negative, rather than a positive, impact on student learning. My own study does not specifically consider how technology is being used to teach writing, but rather questions how students perceive the use of technology in the writing classroom. What do students consider "technology" and how do they see it impacting their own learning? Many of the articles call attention to the fact that computers, e-mail, and Facebook have become routine parts of life that students take for granted. How does this perspective change the perception of what counts as "innovative" uses of technology in the classroom? Do students still consider PowerPoint a form of technology, and if so how does it function differently than more contemporary technologies like wikis or blogs? What about platforms like Blackboard that primarily facilitate the distribution of materials and information? Do students need to be active users of the technology in order for it to help them learn? By understanding how students perceive the use of technology in writing classrooms in contradistinction to how instructors perceive it, my study aims to expose gaps between these two demographics. In doing so, I hope to gain a better understanding of how students and instructors define "technology," and how they envision it being used most productively in FWS courses.

METHODOLOGY

Objectives

The primary objective of my project is to gain a better understanding of which technologies are being used and how they are being used to enhance the teaching of writing in FWS courses offered through the English department at Cornell University. While current research on the use of technology often presents case studies, offering examples of how specific assignments do or do not work to enhance student learning, few provide a definition of what constitutes "technology" more broadly. These studies frequently attend to individual technologies, meaning researchers are unable to speak comparatively about the effectiveness of related technologies in achieving common goals. For instance, what is the difference between a journal assignment like Lovell and Palmer's that uses Facebook and a journal assignment like Tryon's that uses blogs? In other words, what are the strengths and weaknesses of these technologies as academic tools? Another shortcoming of such articles is that they typically address not only individual technologies, but rather individual assignments. Rather than considering how a range of technologies could be consistently employed in the classroom throughout the semester, these studies create the sense that assignments using technology are "special occasions" that occur only sporadically because they require a great deal of preparation and time. In part, I think this mentality stems from the unspoken idea that things like PowerPoint, which can be easily used on a daily basis, have become passé, and consequently, instructors who use them are using "old" technologies rather than engaging more innovative and cutting edge technologies. Instructors who want to remain "current" might feel pressure to keep up with the times by demonstrating an awareness of and ability to use "new" technologies. If technology is being used to make courses "sexier," the rarity of these projects might also be symptomatic of a tendency to include them primarily as a means of attracting undergraduates and not because instructors sincerely believe that they have greater educational value than more conventional assignments.

A main objective of my study is to construct a more functional definition of "technology" shared by instructors and students. Rather than treat available technologies as inherently productive academic tools, I aim to differentiate between those used daily by instructors and students for academic and non-academic

purposes. I do not intend to argue that technologies primarily used for non-academic purposes cannot be effectively incorporated into the classroom, but rather, I hope to address a point raised by several critics: just because students enjoy using certain technologies in their daily lives, does not mean they will necessarily enjoy using them in the classroom. Defining "technology" more effectively might help to close the gap between how instructors perceive the functionality of technology vs. and how students perceive its functionality. In doing so, we might be able to make the incorporation of technology into the classroom more productive for both instructors and students.

Methods

My study uses a series of surveys and observations in order to determine how instructors and students currently teaching and taking First-Year Writing Seminars in the English Department define "technology." Since all undergraduates attending Cornell University are required to take two First-Year Writing Seminars (students may be able to use AP credits to replace one semester of FWS), these courses tend to be quite diverse in terms of academic interests. Although there are some exceptions, most students enrolled in these classes are in their first year at Cornell. A profile for the Class of 2018 created by Cornell Institutional Research and Planning states that 51.1% of freshmen are women and 48.9% are men. Most of the students, 64.6%, attended public high schools. The average age of the students is 18. Providing a profile of FWS instructors in the English department is somewhat more challenging. During the Spring 2015 semester 57 FWS courses were offered through the English department, not including those that are cross-listed as English; these courses are taught by PhD students, MFA students, and faculty members in all stages of their careers.

My primary mode of data collection consists of two short 5-10 minute surveys: an electronic one for instructors created on Cornell's Qualtrics Survey Software and a paper one for students. With a few exceptions, these surveys are identical. Both groups were asked to "identify examples of what [they] consider to be 'technology' that might be used in the college classroom," explain how technology has been used in their FWS in relation to writing and reading comprehension, and fill in a chart that details which technologies they use in their daily lives for academic and non-academic purposes. Finally, they were asked which of the "technologies" included in the chart they believe could be used effectively in an FWS. Students were also asked which technologies their other instructors have used that their FWS instructor has not. I distributed the instructor survey through the English department list-serve and received 11 responses. I administered the paper survey in 5 First-Year Writing seminars and received 73 responses. In selecting which seminars to survey, I tried to choose a range of instructors who confessed varied levels of technology use in their courses. Both surveys were completely voluntary and anonymous. In addition to these surveys I also observed my own class to obtain qualitative data regarding their understanding of technology as well as to determine what kinds of suggestions they have for incorporating technology into the classroom.

RESULTS

My results are organized around the three central questions laid out in my introduction: how do students and instructors define the "technology" that might be used to enhance student learning in higher education? Which technologies do students and instructors use in their daily lives for academic and non-academic purposes? What potential do students and instructors envision technology having as a means of enhancing the teaching of writing and the understanding of course materials in FWS courses?

Both the student and instructor survey asked respondents to "Identify examples of what you consider to be 'technology' that might be used in the college classroom to improve student learning. Come up with as many examples as you can and list them below."

My intention in asking this question was to gain a better sense of what each group considered to be "technology" in the broadest sense of the term. However, when confronted with this question few individuals "thought outside the box," meaning their responses were not especially diverse and did not move beyond technologies already being used in FWS courses on a regular basis. For example, the most frequent responses from the 73 student respondents were computers/laptops, projectors, cell phones, tablets/iPads, iClickers, and PowerPoint. The most frequent responses from the 11 instructors were computers/laptops, projectors, Blackboard, movies/DVDs, Youtube, and audio recordings. These results largely correspond to technologies that students and instructors are currently using on a daily basis for academic purposes. One of the key differences between these lists is the fact that students seemed to focus on devices that could be used to present or respond to material (i.e. projectors, iClickers), suggesting that they think about technology as a tool to find and/or display information. Students were more likely to list visually oriented devices like projectors, videos (or related platforms like Youtube), and digital texts rather than auditory devices like podcasts, MP3s, or music. The most divergent student responses were "medical technologies" and "walkie talkies," which could potentially be used to enhance student learning in some courses, but would likely not be practical in an FWS. Instructors seemed to focus on alternative course materials such as movies, Youtube clips, and audio. This difference demonstrates that instructors tend to think about technology as something that can be integrated into FWS courses as both a mode of presenting information (i.e. projectors, Blackboard) as well as a form of content (i.e. movies/DVDs, audio recordings). One instructor listed a number of ideas for how technology could be incorporated into the classroom to enhance student learning, including "in-class writing via social media, interactive game-like programs," and "video game study." Another mentioned specific digital learning tools such as University of Virginia's "For Better for Verse." Although these answers are not necessarily representative, the fact that instructors were more detailed in their responses implies that they tend to evaluate technologies in terms of relevance to their discipline and be more aware of specific technologies that could be useful to them. This mode of thinking is not unexpected, but means that creating a definition of "technology" that could be applied across disciplines could be challenging. On the other hand, students taking courses in many different fields may not give much thought to why certain technologies are better suited to certain disciplines, providing a broader definition of "technology."

The next few questions on my survey addressed the question: how are instructors currently using technology in FWS courses? I asked students and instructors what technologies were currently being used in relation to teaching writing and understanding course materials, and how those technologies were being used. One important conclusion that emerged from these questions was the fact that neither students nor instructors view the teaching of writing and the understanding of course materials as separate processes. In other words, respondents commonly identified the same technologies and activities in relation to both. The overwhelming response on the part of students was that instructors are using technology to project PowerPoint presentations related to course material, display short readings, display examples of writing, and show movies, film clips, and Youtube videos. For instance, one student wrote, "She uses the projector a lot—she shows PowerPoints with English lessons or pieces of writing. Many times we're asked to read essay excerpts from the projector and discuss them." Another student responded that the instructor "projected examples of writing […] in order to teach us how to structure introductions, thesis statements, conclusions, etc. And to watch videos." Several students said they listened to audio or made their own recordings, but this was far less common. Another prominent technology was Google Docs, which is used for group editing, providing feedback, and

allowing for collaboration. The analysis of film clips was also common. Most students did not provide evaluative statements about whether they found these exercises useful, suggesting that they may simply view these activities as everyday occurrences in class. However, students who encountered innovative uses of technology, such as one student who's instructor used PowerPoint to "show the class where our homes are or what we define as home" recalled these activities more vividly, making claims like "This is a great way to connect the class and have students engage." These kinds of responses were rare.

Although many of the same technologies used to teach writing were used to help students understand course materials, there were some subtle differences. Student responses indicate that technology was used to provide them with access to readings, but also to allow them to make connections that move beyond the readings. For example, one student wrote, "Podcasts have been used to give examples of related works as well as background information on a reading. Some TED talks, videos, and a music video, have been used in a similar way to relate themes of readings to the real world. Movie and musical adaptations have been shown to visualize the readings in a new way." This response indicates that instructors may view technology as a means of helping students make connections between readings and current events, exposing real-world implications and creating stakes. One instructor claimed to use "Youtube videos of literary adaptations, [and] images or digital slides to jumpstart readings for the day." The idea of using media to "jumpstart readings" positions technology as what Mary Shelley might call a "spark of being" that brings the readings to life for the students in a way that other methods do not.

Another aim of my study was to determine which technologies students use daily for academic and non-academic purposes. As part of the student surveys, I provided a list of 23 "Technologies" and asked students to check boxes corresponding to the following four criteria: "I use this for non-academic purposes in my daily life"; "I use this for academic purposes in my daily life"; "My FWS instructor has used this technology and it enhanced my understanding of writing/course materials"; and "My FWS instructor used this technology but it DID NOT enhance my understanding of writing/course materials." When reviewing the responses I separated the list of technologies into three categories: Electronic Devices, Programs and Platforms, and Social Media. The first category, Electronic Devices, was the least polarized in terms of being used for non-academic and academic purposes. As the graph below demonstrates, most electronic devices are being used by students for both non-academic and academic purposes on a daily basis with the exception of iClickers and Cameras/Video Cameras.

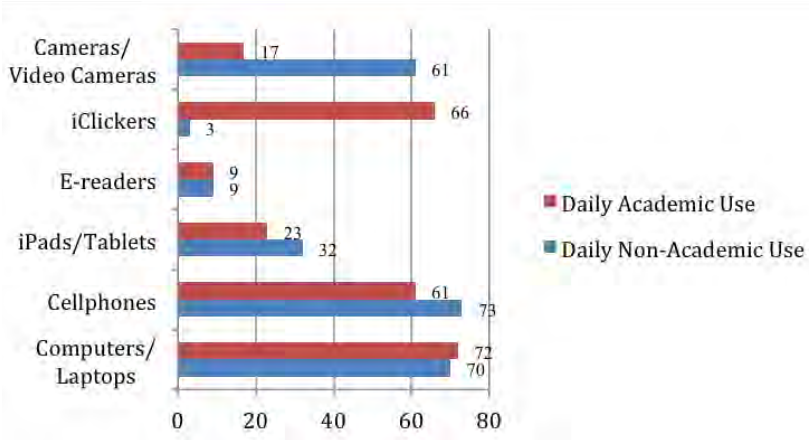


Figure 1: Electronic Devices (Student Reponses n=73)

In terms of use in FWS courses, 49 students identified computers/ laptops as being used and enhancing their understanding of writing/course materials and 3 identified computers/laptops as being used but not enhancing their understanding. 7 students identified iPads/Tablets as being used and enhancing their understanding of writing/course materials. 1 student identified e-readers as being used and enhancing their understanding of writing/course materials. 4

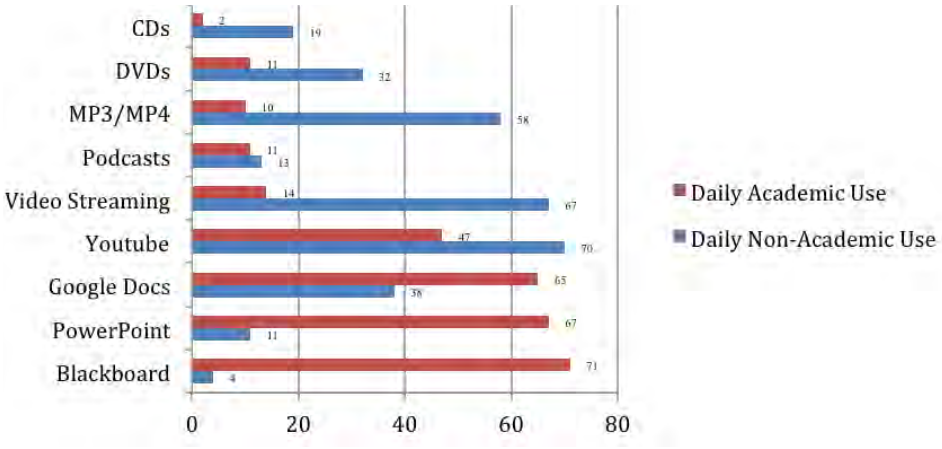


Figure 2: Programs and Platforms (Student Responses n=73)

students identified cameras/video cameras as being used and enhancing their understanding of writing/course materials. These results indicate that when electronic devices were used in FWS courses they were generally perceived as having a positive impact on student learning.

The second category, Programs and Platforms, was slightly more polarized in terms of non-academic and academic use on a daily basis.

The majority of the platforms and programs were used primarily for non-academic use. For instance, students tend to use video streaming, DVDs, CDs, and MP3/MP4 files for non-academic use. However, they tend to use Blackboard, PowerPoint, and Google Docs primarily for academic use. One of the platforms with the most overlap is Youtube, indicating that instructors are more likely to incorporate short clips into their courses rather than full-length movies. Student comments indicate that Youtube can be a useful way to jump start class discussion. I would argue that Youtube is especially popular because it is a format that students are very familiar with and feel able to comment on, it offers a wide range of clips that could relate to any subject, and it is a quick and engaging way of introducing a topic for discussion without much preamble. Additionally, this kind of activity may be useful in creating the connections to current events mentioned in student comments. The other platform with the most overlap is Podcasts, with a much smaller number of student users. Responses to the use of programs and platforms in FWS courses similarly suggested

Programs and Platforms	*My FWS instructor has used this technology and it enhanced my understanding of writing/ course materials.*	*My FWS instructor has used this technology and it DID NOT enhance my understanding of writing/course materials.*
CDs	1	0
DVDs	10	0
MP3/MP4	11	1
Podcasts	13	0
Video Streaming	16	0
YouTube	32	1
Google Docs	13	1
PowerPoint	31	1
Blackboard	53	1

Table 1. Programs and Platforms (Student Responses n=73)

that when these technologies are used they are perceived as having a positive impact. For instance, 53 students indicated that their FWS instructors used Blackboard and it enhanced their understanding of writing/course materials. Only 3 students said their instructors used Blackboard and it did not enhance their understanding. The results for other programs and platforms were similar.

The technologies that are most polarized in terms of academic and non-academic use, perhaps unsurprisingly fall into the category of Social Media. This is the only category in which certain technologies received 0 responses in terms of academic use.

Students indicated that they use Twitter, Instagram, Snapchat, and Tumblr exclusively for non-academic purposes. Although students responded that they use wikis for academic and non-academic purposes it is more likely that they are using sites like Wikipedia rather than generating their own wikis in collaborative projects like those described by Allwardt and Brian Morgan and Richard D. Smith. In conversations with my own students, they indicated that they consider these technologies part of their "private" lives and are not necessarily eager to use them in the "public" forum of the classroom. If instructors similarly categorize things like Facebook and Twitter as "private" this may account for the fact that they are not frequently used in FWS courses. Neither students nor instructors perceive the potential benefit of assignments involving social media as outweighing their discomfort with merging the private and public spheres of their lives.

The final question on student and instructor surveys asked the following question: "Which of the 'technologies' on this list do you feel might be used to enhance student learning in First-Year Writing Seminars? Identify each technology and briefly describe how it might be used." While students did make some recommendations, none of their ideas stood out as things that have not been explored before by those teaching First-Year Writing Seminars. Several students expressed that FWS is not the place for technology, or claimed that they did not like the use of technology. For example one respondent said, "Google Docs make for easy collaboration in the event a teacher wants to make a group project (which I do not enjoy). Otherwise, I am not a fan of technology in the classroom." Another student definitively claimed: "None is better for English classes." Most of the suggestions were aimed at making students lives easier rather than creatively teaching writing skills. For instance, one of the most popular suggestions was Google Docs because it allows for easy commenting and collaboration. A student responded, "I believe that computers, tablets, digital texts, Blackboard, Google Docs, PowerPoint, Youtube, and Wikis are useful for enhancing student learning. They can be used for making students' lives more convenient." In terms of social media, one student specifically requested that social media not be used in FWS courses because they had a bad experience in high school. Another student stated that Blackboard was a good tool because all other technologies are distracting. These responses suggest that students are not specifically looking for FWS instructors to use technology in more innovative ways. In fact, more than a few

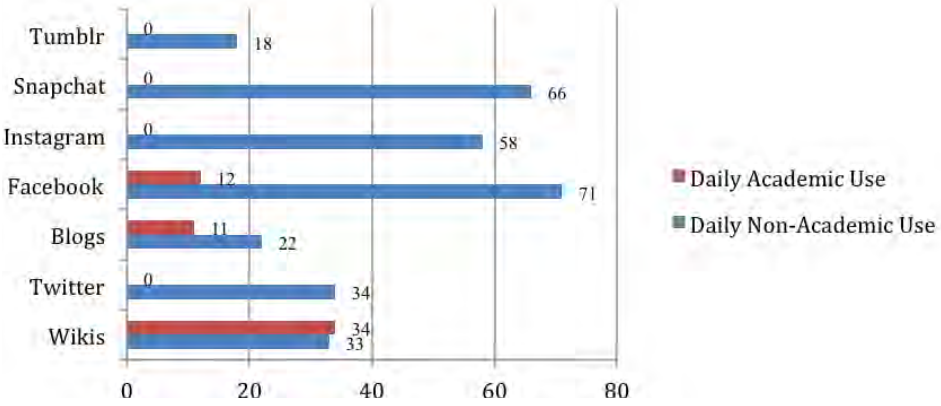


Figure 3: Social Media (Student Responses n=73)

students don't seem to think it has a place in the writing classroom at all.

My conversation with my own students corroborated this assumption. When I asked my students whether a course description including movies, television series, or other media would make them more or less likely to take the course, they said that it really depended on the topic. One said, "I don't think the media would be an extra incentive for me to take the course." Others suggested that if the same course was offered with or without media they would likely take the former, indicating that students might consider a course with media slightly more dynamic and therefore slightly more interesting. When I pushed them, they confessed that they automatically assume that FWS courses with media in the description are easier than those without media, which might make them more likely to choose a course with media. "We're forced to take an FWS," they explained, "so whatever one sounds easier is probably the one that I'm going to take." When I asked them if a course description that indicated they would have to complete a project using technology would make them more or less likely to take the class their responses were more emphatic. My students wanted to know what kinds of projects I was talking about, so I shared some of the assignments from the articles in my literature review. One of my students asked, "Who thought that would be a good idea?!" Another said "Those assignments just don't sound fun." Underlying all of their answers was a general sense that FWS courses are meant to be about writing. Students said, "I think writing should be taught, not technology."

When I posed the same question on the instructor survey, "Which of the 'technologies' on this list do you feel might be used to enhance student learning in First-Year Writing Seminars? Identify each technology and briefly describe how it might be used," the responses were much more positive and dynamic. The suggestions from instructors were more detailed and included both things they had tried in the classroom and things they thought might be productive for teaching writing. In general, instructors came up with more innovative ways of potentially using technology to enhance student learning. One responded, "I feel they all (conceivably) could be used to enhance student learning—it all depends on what subject or practice the instructor is trying to teach." Another instructor suggested that computers could be useful tools in the classroom because "asking students to bring in computers allows you more freedom to use other digital platforms. It also allows students to research topics in class." Confronted with the question of how technology could be used, instructors never stopped to consider that it might not be useful like the students did. Instead of seeing the potential problems of using technology, instructors envisioned the many possibilities. This disconnect might be indicative of a rift between these two groups in terms of how they think about technology in the context of higher education.

CONCLUSIONS

The major conclusions I have reached based on my study are as follows:

1. "Technology" might be defined as a set of tools that can be used to distribute course materials and information, enable interaction and collaborative learning, and supplement content.

2. Neither students nor instructors "thought outside the box" when asked to identify examples of technology that could be used in higher education. This could indicate that we have become complacent in thinking about technology's potential in academia. Students and instructors may feel as though the ways technologies are currently being used are sufficient.

3. Students and instructors primarily identified technologies and classroom activities that supplement more traditional content/teaching methods. For instance, many stated that technology was useful in displaying visual content that could spark class discussion.

4. Students understand certain technologies as "private" or "non-academic" and do not necessarily want to use them in the "public" or "academic" sphere. In fact, many students were resistant to incorporating technologies like social media into the classroom. Instructors should be aware of this disparity and design assignments accordingly.

5. While instructors view most technologies as having potential in the classroom, students have a more limited perspective. They do not consider technology as integral to enhancing learning in FWS classes and do not find courses that integrate technology inherently more attractive.

Based on my conclusions, FWS instructors should think carefully about why and how they choose to employ technology in the classroom. Students want to understand *why* they are being asked to use technology and *how* it is helping them achieve the goals of the assignment.

REFERENCES

Allwardt, D. E. (2011). Teaching Note Writing With Wikis: A Cautionary Tale of Technology in the Classroom. *Journal of Social Work Education*, 47(3), 597-605.

Hauman, K., Kastner, S., Witte, A. (2015). Writing Teachers for Twenty-First-Century Writers: A Gap in Graduate Education. *Pedagogy* 15(1), 45-57.

Leibowitz, W. R. (1999). Technology Transforms Writing and the Teaching of Writing: Many professors try to combat the bad habits they fear their students pick up on computers. *The Chronicle of Higher Education* 46(14), A67-A68.

Lovell, E., Palmer, B. (2013). Engaging Introductory Writing Students Through Facebook Assignments. *About Campus*, 25-28.

Morgan, B., Smith, R. (2008). A Wiki for Classroom Writing. *The Reading Teacher* 62(1), 80-82.

Sullivan, R. (2013). The Tablet Inscribed: Inclusive Writing Instruction With the iPad. *College Teaching* 61, 1-2.

Tryon, C. (2006). Writing and Citizenship: Using Blogs to Teach First-Year Composition. *Pedagogy* 6(1), 128-132.

Integrating Formative Assessment during the Laboratory Section of a Histology Course

Vera Rinaldi[1*], Nancy Lorr[1] and Kimberly Williams[2]

** Graduate Research and Teaching Fellow 2014-2015*

ABSTRACT

The use of active learning in undergraduate courses in sciences has been shown to improve students' examination scores and concept inventories. However, the suitability of a learning activity varies with students' learning style and personality (Chen, 2015). Therefore, it is important to obtain feedback about students' understanding, i.e. formative assessment, in order to adjust the teaching to fit the needs of each student (Ludvigsen et al., 2015). Technology-supported formative assessment provides feedback about student knowledge, facilitates classroom dialogues, and modifies student's learning strategies (Baleni, 2015; Ludvigsen et al., 2015). Here we evaluate the use of an interactive application (e.g Learning CatalyticsTM ("Transforming the Lecture Space (EdSurge News)," n.d.)) to integrate formative assessment to active learning strategies during the laboratory section of an upper-level Histology course. Our study was conducted during the 2015 BioAP4130-Histology course at Cornell and led to the conclusion that such applications can be helpful for identifying misunderstandings as well as forming a new communication venue between students and instructors. However there is room for improvements, with the weakest point being the image display interface and graphical resolution. We also surveyed students to evaluate their opinion about the traditional glass slide microscopy and the use of computer-based virtual slides. Interestingly, even though there is a preference for the virtual slides, this cohort of students is aware of its limitations. The take-home message is that student' diversity requires instructors to use different teaching tools: the newest teaching tools aren't necessarily better and new and old tools can be used to complement each other.

INTRODUCTION

Undergraduate courses often rely on students' grades to assess learning outcomes; therefore, it is not a surprise that many students are grade-motivated. Besides the obvious impact grades have on student motivations, lower grades often mean the student failed to understand or has a wrong interpretation of a foundation concept(Pulfrey et al., 2013). As instructors we want to prevent students from making faulty interpretations of topics; but if these are inevitable we want to be able to identify them early (Watkins and Mazur, 2013). Failing to promptly identify and correct misconceptions in foundation concepts may irreversibly impair the student's understanding of a topic (Crouch and Mazur, 2001). With advancements in technology, teachers of morphology-based courses have tools that

go beyond computer visualization. "Clickers," (Briggs and Keyek-Franssen, 2010) the Piazza web-based interface ("Piazza • Ask. Answer. Explore. Whenever.," n.d.), and internet-based voting applications (Mathiasen, 2015) are examples of teaching tools currently used to engage students and provide feedback about their state of knowledge. These tools are widely used; however there are no formal reports about their effectiveness when used in morphology-based biomedical courses (Karolcík et al., 2015; Selvig et al., 2015). We hypothesize that the integration of interactive teaching software with visualization tools (e.g. virtual slides) will allow instructors to identify misconceptions "on-the-go." Therefore, we simulated, to the best of our abilities, how an interactive histology laboratory session would feel.

Histology Course

Histology is a detail-oriented subject based on identification and description of the cellular organization found in healthy tissue (Mione et al., 2013). It is traditionally taught as a combination of descriptive lectures and practical laboratory sections (Bloodgood, 2012; Bloodgood and Ogilvie, 2006). The latter aims to develop the students' abilities to identify subtle morphological differences in characteristics of a tissue and integrate their knowledge of the functional aspects of an organ to their observation of a two dimensional microscopic image (Bloodgood and Ogilvie, 2006; Hortsch and Mangrulkar, 2015; Mione et al., 2013; Selvig et al., 2015). This is a complex and multistep process that is usually facilitated by interaction between learners and teachers.

The most common laboratory settings usually allow for students to interact with instructors and benefit from having multiple teaching assistants (TAs) that work with the students one-on-one or in small groups. The laboratory section tends to rely on the students' interaction with professors and TAs to clarify what the student is expected to identify and also clarify doubts the student may have from observing the histological sample (Bloodgood, 2012; Collier et al., 2012). Therefore students that are shy or that are not comfortable asking questions and/or seeking clarification may be marginalized and fail to achieve their full potential towards learning the subject. In order to provide alternatives to students with these personality traits some histology laboratories have available alternative learning resources (such as books, sample images, or an interactive atlas) (Khalil et al., 2013). Another alternative many instructors attempt is to offer unsolicited help while hovering around students during the learning activities. Due to the diverse nature in learning styles of each student, every classroom has its unique culture, and the instructor must adapt to it in order to facilitate the learning process (Twenge, 2009).

It is also important to recognize how virtual slides, each a high resolution digital image of the glass histological slide, improved the learning experiences by allowing multiple individuals to collectively analyze a histological structure (Braun and Kearns, 2008; Harris et al., 2001; Husmann et al., 2009). Before the advent of virtual slides, teachers had to interpret the students' questions by looking at a glass slide through the same microscope the student used. This approach did not allow for group discussion and required a lot of one-on-one interaction between teacher and student (Collier et al., 2012; Husmann et al., 2009). With increase in class size and didactic material to be covered at each laboratory session, many educational settings have moved away from using microscopes and started using only computer aided visualization tools, the most common of which is the virtual slide (Bloodgood and Ogilvie, 2006). Using virtual slides, multiple individuals are able to collectively analyze a histological structure. Even though virtual slides facilitate group learning, the optical microscope is still the most used method outside of the classroom, for instance in clinical and academic settings (Pratt, 2009). In training students for their professional careers, a working knowledge of the use of optical microscopy represents a skill necessary in many of their potential careers. Taking advantage of the 2015 BioAP4130-Histology

course setting at Cornell, where students have access to both virtual and physical microscopes, we surveyed the students' opinion on using the virtual microscope compared to the physical-optical microscope during their learning experience. The survey responses aided in the creation of a formative assessment activity as well as provided a better understanding of that particular classroom culture.

Objective

The purpose of this study is to evaluate formative assessment as a tool to identify misconceptions "on-the-go". The targeted audience was undergraduate students enrolled in morphology-based biology courses, in this study limited to Cornell's 2015 BioAP4130-Histology course. This histology course has a lecture and a laboratory component; the laboratory sections comprise more than two thirds of the assigned classroom schedule. The laboratory activity thrives on students' active learning and peer learning. As a constant effort to improve students' learning outcomes, the learning activities constantly incorporate the use of new teaching tools, currently using the virtual slide technology coupled with the traditional glass-slide microscopy and problem-based learning activities. However misconceptions still arise and are not detected until grading high-stakes assessments. Here we tested the effectiveness of an interactive technology-based teaching tool in early detection of student's misconception. Even though the technology is not tailored to morphology-based courses, we concluded that such tools have the potential to improve student engagement in learning activities, help identify otherwise unnoticed misconceptions and open a new venue of communication between students and teachers.

METHODOLOGY/EXPERIMENTAL PROCEDURE

In order to achieve an interactive laboratory session, small modules that incorporated histological images into an interactive learning software (e.g. Pearson's Learning CatalyticsTM) were generated and delivered as 15 to 30 minutes review sessions during the last half hour of the laboratory (Favero, 2011) to the students taking the 2015 BioAP4130-Histology course at Cornell. The learning software application provides an interactive interface and allow for customization of the learning activity. However it was not designed for the purpose of morphology-based teaching and we had to use still images (not the high resolution virtual slides) in these modules. Students and instructors were told about the experiment and asked to provide feedback about their experiences by answering voluntary and anonymous surveys. The impact on student's grades was not a focus of our study.

Context

The study was performed during the laboratory session of the Histology course BioAP/MS4130 taught in the spring semester of 2015 at Cornell University in Ithaca, NY. This is an upper-level undergraduate course offered by the Department of Biomedical Sciences at the College of Veterinary Medicine. It is a four-credit course, offered during the spring semester; the class meets twice a week (Monday and Wednesday) for 14 weeks. When the study was conducted, the course consisted of a 55 minutes lecture, followed by a two-hour laboratory section. The classroom size for this semester was 39 enrolled students. The course content covered all major organ systems. Both glass and virtual slides from a variety of vertebrate species were available. Every week students had access to the material corresponding to that week's topic, and printouts of the slide descriptions were made available during laboratory time and through blackboard. The slide descriptions included concise explanations of how to correlate observed structure to organ function. Even though students had access to most specimens through virtual slides, some samples were provided only as glass slides with the goal of encouraging students to use the optical microscope and develop their microscopy skills. During the laboratory sessions, students were encouraged to work in pairs or small groups, but a few elected to work independently (Braun and Kearns, 2008).

The staff for this laboratory semester consisted of one faculty member, one postdoctoral teaching assistant, one graduate teaching assistant and five undergraduate teaching assistants. The faculty member and the graduate teaching assistant, authors of this article, envisioned and implemented the interactive activities. During the laboratory, all staff (for now referred to as instructors) were encouraged to hover around the room waiting for students to call for help. Instructors had weekly meetings in which the slides and slide descriptions were thoroughly reviewed to ensure comparable and homogeneous knowledge of the topics. However, instructor's ability to guide students through a histological section was dependent on background, experience, and the student's preferences. Since help was provided only when solicited, students' preference was the major influence upon students-instructors' interaction. Whereas some students only trusted the faculty member's responses, others only felt comfortable posing questions to undergraduate TAs while others were not comfortable seeking help at all. Therefore, the students' diverse learning styles in the classroom are relevant and need to be accounted for during every activity (DiLullo et al., 2011; Twenge, 2009).

Assessments & Assignments

Three types of assessments were used: quizzes, written assignments, and exams. There were four quizzes, five graded written assignments (and four non-graded and voluntary), which together were worth 30% of the total grade. Three exams were given and together represented the remaining 70% of the grade. The low-risk written assignments and quizzes were designed to engage students with the subject matter and to prepare them for the exams. The majority of the written assignments were based on glass slides that the students needed to examine using the optical microscope.

Study Design

Making use of the laboratory setting available for the BioAp4130 Histology course at Cornell, four interactive review modules were generated. Each module was delivered on the second day of the weekly meetings (e.g. Wednesday) and contained 10 to 15 questions covering the material suggested for the week. These interactive reviews were only implemented during the second half of the semester. We speculated that since this was a voluntary experiment that required students' active participation and was the first time some students were utilizing this kind of interactive learning software, we would have a better participation rate if the students had time to build a relationships with instructors and with each other. We reasoned that group activities in which people are more familiar with each other tend to have higher engagement rates in interactive exercises than when the participants do not know each other (Freeman et al., 2014).

In order to assess how the students experienced these review sessions, we designed internet-based surveys specific for each module that were distributed soon after class ended utilizing the Cornell Qualtrics platform. We also generated two paper-based surveys: one to assess how these students perceived the overall interactive activity and another to obtain information on how the students perceived the different histological slides (e.g. virtual or glass slides). The combination of these surveys allowed us to determine the students' perception of each review activity, assess how they perceived the interactive activities, and establish a line of written communication. The web-based surveys also helped us improve the review modules based on students' self-reported experiences and expectations from the previous review session. We also used our own observations coupled with instructors' surveys to evaluate the activity from a non-student-centric perspective.

All of the students enrolled in the spring semester of 2015 BioAp4130 Histology course at Cornell were encouraged to participate in the study. Every intervention and survey request provided participants with information regarding the research goal and procedures of the study. Participation was not required, and students were allowed to withdraw at any

time without any penalty. The Institutional Review Board (IRB) at Cornell approved the research methods used. The Pearson's Learning Catalytics™ software was kindly made available by Pearson representatives, free of charge, for the purpose of this study.

RESULTS

Classroom Demographics

There were 39 students enrolled; 19 females, 20 males. From these, 28 volunteered (12 males and 16 females) to answer the first paper based survey, which provided us with further information about the students' career goals and expectations as well as how they related to the microscope and virtual slides. From this survey we learned that 57% of the responders envision using the traditional glass-slide microscopy throughout their professional career (Figure 1A). Anecdotally we know that more than half of the students that take this course seek a medical or veterinary degree; therefore we expected a higher percentage of students envisioning the use of the traditional glass-slide microscopes in the clinical setting. We expected them to know that, even though the use of slide scanners is becoming more prevalent, most clinical settings still require the use of the traditional microscope. To determine the precise number of students seeking a professional degree we did a second survey and found that 75% of the students are likely to pursue a MD/DVM degree (Figure 1B). This result suggests that students are not completely aware of what will be required from them during their training and future career in medicine. Another interesting result is that very few (two out of 24) students foresee going to the job market without further study. This is likely a reflection of their awareness of the highly competitive job-market. We also found that 39% of the students believe they benefit from having access to both methods to visualize histological slides. Therefore we conclude that it is important to provide both visualization tools in order to improve the number of students reaching their full potential towards understanding the topic.

Interactive Review Session

The goal of the review modules was to obtain information about the student's understanding of foundation concepts before high stakes assessments, thus identifying and remediating misconceptions "on the go." For that reason we made use of images from both glass slides as well as virtual slides but not necessarily the same images used for the laboratory activity of the week. We speculated that, by exposing students to unfamiliar images, we would discourage memorization of particular aspects of a slide (for example histological artifacts) and help guide their study to relevant histological characteristics of tissues (for example organ structure and cell morphology). We expected that with this approach we would be giving students the guidelines necessary to allocate study time to concepts matching the learning outcomes for the course. We used the Pearson's Learning CatalyticsTM software as a tool because it provides an extensive and comprehensive list of question types and uses an interactive interface (see figure 2 for two examples). It also allows the students to give feedback about their understanding of the material in real-time. With this tool we were able to identify which morphological aspects of a specific histological structure were leading students to misconceptions. Once we were able to spot the problem, addressing it was much easier. The four modules provided information about students' misconceptions and weak points from our explanations. As we became more aware of this classroom's culture, we became more proficient in predicting the topics they would likely struggle to understand.

Student's Self-Reported Experiences

The qualitative analyses of the surveys show a positive impact in students' self-reported understanding of the material. Although rated as helpful, four students out of 19 felt that the review modules were not well structured. This is probably a reflection of our (the instructors) inexperience using the software. However, when instructors suggested cancelling the last review module, students expressed dissatisfaction. We concluded that, although the

use of the interactive software was faulty and needs improvements, it is overall beneficial and useful. When asked about how to improve the review experiences the most common theme from both instructors' and students' survey was regarding image display and quality. We speculate that an ideal interactive tool for morphology-based teaching should incorporate the ability to increase image size and resolution or have it coupled with virtual slide visualization tools.

Observations from the Authors

During the review modules the students appeared interested and engaged. About one third used their own cell phones to access the interface. During the later sessions we noticed that students were more likely to ask questions by sending messages using the software rather than by voicing it. A couple of students from a group that rarely requested help during the regular laboratory activity, were noticeably more active and seemed comfortable at texting their questions to the instructors. Therefore we conclude that the interactive interface was able to engage a different cohort of students that might have otherwise been marginalized (DiLullo et al., 2011).

Although the process of creating the review modules per-se required a lot of up-front work, the positive response of the students made the effort worthwhile. Besides the student's engagement in the activity, we were able to visualize how certain definitions were causing students to misunderstand certain topics. For example, during one of the activities we learned that several students believed that the layers of the epidermis were to have a maximum depth (of only a couple of cell layers). While most students were able to list the layers that formed the epidermis, when asked to draw a line in the interface between dermis and epidermis on a sample image they were not familiar with, many failed (figure 3). This interactive exercise provided us with a prompt answer to what was happening and we found that when provided with this unfamiliar image several students resorted to the color differences that best resembled the slide they were familiar with, instead of accepting that one of the layers in the new image was much wider than what they expected. The fact that we can address any issue promptly seems much more effective than waiting for assignments to provide students with feedback.

Every review module provided us with different observations. For instance, during the first review module, we noticed students were not expecting to actively answer questions. Even though they were informed about the need to actively participate, they seemed to be expecting a passive review in which instructors lecture about the important topics. This probably reflects previous experiences in which instructors provide a list of important concepts without the need to actively participate. However there was a shift in students' behavior for the next modules, with the majority studying the material, and preparing for the review. The students clearly did not like to answer questions incorrectly even when their identity was not disclosed. Since most students were on the pre-vet, pre-med track our interpretation is that there is an intrinsic desire to excel and show their intelligence. Another general observation that supports our interpretation is that when questions were easy, students stopped participating, whereas if the questions were challenging, they would engage more in the activity. Therefore our observations might not represent how students would react in a less competitive environment.

DISCUSSION

Students' misconceptions vary between classrooms depending on a multitude of factors that are not always clear to the instructors. It is not always trivial for teachers to predict topics in which students will have difficulties, or why and how students' confounding interpretations arise (Fouché, 2015). In order to facilitate this process we used an interactive teaching tool that helped identify misconceptions "on-the-go." We used the software Learning CatalyticsTM from Pearson to design our activity, but envision achieving similar results with other freely or commercially available

teaching tools. We believe our results can be extrapolated to other morphology-based biomedical courses that require practice in interpretation of images, thus requiring students to identify differences between technical and biological variability.

This work constitutes an initial attempt to inspire instructors of morphology-based biology courses to experiment with new teaching technology that provides a real-time formative assessment. Our analysis suggests that students taking morphology-based courses will benefit from such tools and that the use of them facilitates designing active learning activities. Our conclusion is that the use of interactive software, such as the one used in this study, facilitates the identification of confounding concepts "on-the go" and helps instructors understand what is causing the wrong interpretation. However there is a gap in the development of teaching tools suited for morphology-based courses in that they are not able to incorporate high resolution images such as those generated from scanned slides (virtual slides). Although there is room for improvement, both students and instructors agreed that the use of the interactive software was beneficial. We believe that our study also advocates in favor of having varied stimuli in order to reach a diverse cohort of students. We hope that this study inspires instructors and software developers to invest in integrating interactive software for morphology-based courses.

REFERENCES

Baleni, Z.G., 2015. Online Formative Assessment in Higher Education: Its Pros and Cons. Electron. J. E-Learn. 13, 228–236.

Bloodgood, R.A., 2012. Active learning: A small group histology laboratory exercise in a whole class setting utilizing virtual slides and peer education. Anat. Sci. Educ. 5, 367–373. doi:10.1002/ase.1294

Bloodgood, R.A., Ogilvie, R.W., 2006. Trends in histology laboratory teaching in United States medical schools. Anat. Rec. B. New Anat. 289, 169–175. doi:10.1002/ar.b.20111

Braun, M.W., Kearns, K.D., 2008. Improved learning efficiency and increased student collaboration through use of virtual microscopy in the teaching of human pathology. Anat. Sci. Educ. 1, 240–246. doi:10.1002/ase.53

Briggs, C.L., Keyek-Franssen, D., 2010. Clickers and CATs: Using Learner Response Systems for Formative Assessments in the Classroom. Educ. Q. 33.

Chen, Y., 2015. Linking Learning Styles and Learning on Mobile Facebook. Int. Rev. Res. Open Distrib. Learn. 16, 94–114.

Collier, L., Dunham, S., Braun, M.W., O'Loughlin, V.D., 2012. Optical versus virtual: teaching assistant perceptions of the use of virtual microscopy in an undergraduate human anatomy course. Anat. Sci. Educ. 5, 10–19. doi:10.1002/ase.262

Crouch, C.H., Mazur, E., 2001. Peer Instruction: Ten Years of Experience and Results. Am. J. Phys. 69, 970–77.

DiLullo, C., McGee, P., Kriebel, R.M., 2011. Demystifying the Millennial student: a reassessment in measures of character and engagement in professional education. Anat. Sci. Educ. 4, 214–226. doi:10.1002/ase.240

Favero, T.G., 2011. Active review sessions can advance student learning. Adv. Physiol. Educ. 35, 247–248. doi:10.1152/advan.00040.2011

Fouché, J., 2015. Predicting Student Misconceptions in Science. Educ. Leadersh. 73, 60–65.

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. Proc. Natl. Acad. Sci. 111, 8410–8415. doi:10.1073/pnas.1319030111

Harris, T., Leaven, T., Heidger, P., Kreiter, C., Duncan, J., Dick, F., 2001. Comparison of a virtual microscope laboratory to a regular microscope laboratory for teaching histology. Anat. Rec. 265, 10–14.

Hortsch, M., Mangrulkar, R.S., 2015. When students struggle with gross anatomy and histology: A strategy for monitoring, reviewing, and promoting student academic success in an integrated preclinical medical curriculum. Anat. Sci. Educ. doi:10.1002/ase.1519

Husmann, P.R., O'Loughlin, V.D., Braun, M.W., 2009. Quantitative and qualitative changes in teaching histology by means of virtual microscopy in an introductory course in human anatomy. Anat. Sci. Educ. 2, 218–226. doi:10.1002/ase.105

Karolcík, Š., Cipková, E., Hrušecký, R., Veselský, M., 2015. The Comprehensive Evaluation of Electronic Learning Tools and Educational Software (CEELTES). Inform. Educ. 14, 243–264.

Khalil, M.K., Kirkley, D.L., Kibble, J.D., 2013. Development and evaluation of an interactive electronic laboratory manual for cooperative learning of medical histology. Anat. Sci. Educ. 6, 342–350. doi:10.1002/ase.1350

Ludvigsen, K., Krumsvik, R., Furnes, B., 2015. Creating formative feedback spaces in large lectures. Comput. Educ. 88, 48–63. doi:10.1016/j.compedu.2015.04.002

Mathiasen, H., 2015. Digital Voting Systems and Communication in Classroom Lectures--An Empirical Study Based around Physics Teaching at Bachelor Level at Two Danish Universities. J.

Interact. Media Educ. 2015.

Mione, S., Valcke, M., Cornelissen, M., 2013. Evaluation of virtual microscopy in medical histology teaching. Anat. Sci. Educ. 6, 307–315. doi:10.1002/ase.1353

Piazza • Ask. Answer. Explore. Whenever. [WWW Document], n.d. URL http://www.piazza.com (accessed 1.15.16).

Pratt, R.L., 2009. Are we throwing histology out with the microscope? A look at histology from the physician's perspective. Anat. Sci. Educ. 2, 205–209. doi:10.1002/ase.100

Pulfrey, C., Darnon, C., Butera, F., 2013. Autonomy and Task Performance: Explaining the Impact of Grades on Intrinsic Motivation. J. Educ. Psychol. 105, 39–57. doi:10.1037/a0029376

Selvig, D., Holaday, L.W., Purkiss, J., Hortsch, M., 2015. Correlating students' educational background, study habits, and resource usage with learning success in medical histology. Anat. Sci. Educ. 8, 1–11. doi:10.1002/ase.1449

Transforming the Lecture Space (EdSurge News) [WWW Document], n.d. . EdSurge. URL https://www.edsurge.com/news/2014-10-14-transforming-the-lecture-space (accessed 1.15.16).

Twenge, J.M., 2009. Generational changes and their impact in the classroom: teaching Generation Me. Med. Educ. 43, 398–405. doi:10.1111/j.1365-2923.2009.03310.x

Watkins, J., Mazur, E., 2013. Retaining Students in Science, Technology, Engineering, and Mathematics (STEM) Majors. J. Coll. Sci. Teach. 42, 36–41.

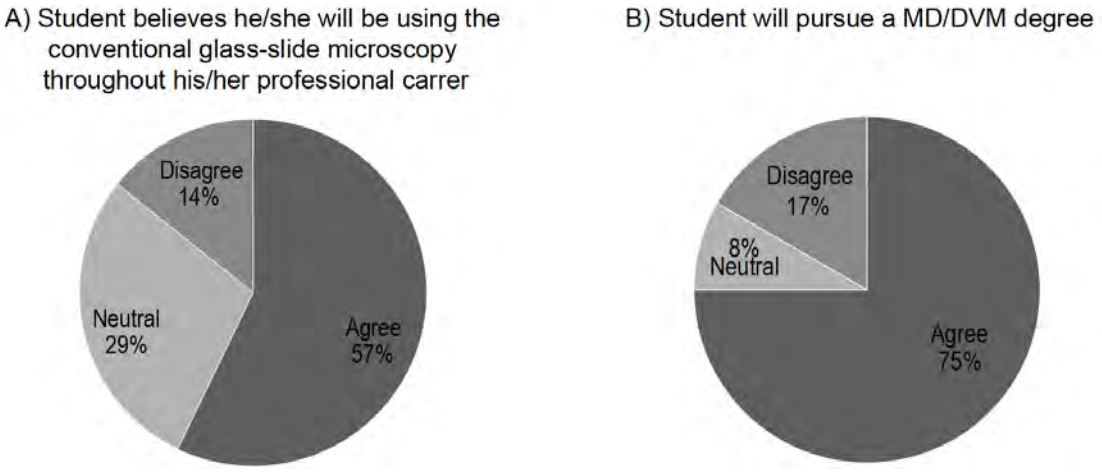
APPENDIX FIGURE LEGENDS

Appendix Figure 1: Pie chart summarizing results about future career. **A-** Students' response for the question about their expectations of using the traditional glass-slide microscopy throughout their career. **B-** Students' response about seeking a professional degree. Responses were divided in three groups based in if they are likely to (agree), are not going to (disagree) or are not sure about (neutral) using the microscope (A) or pursuing a DVM/MD degree (B).

Appendix Figure 2: Example of two different types of questions illustrating the use of the interactive tool during review sessions. **A-** Example of a "composite sketch" type question. In this type of question the instructor presents the task to the student followed by an image. The students respond to the task in their own device by drawing in the image provided what he/she believes is the answer. The blue lines in the image to the right represents a superposition of all the answers. **B-** Example of a "regions" type question. In this type of question the student is asked to identify a specific region in the image provided. The image in the right represents a superposition of all answers. Correct answers appear as green dots and wrong answers as red dots. This particular software allows the instructors to visualize individual answers as well.

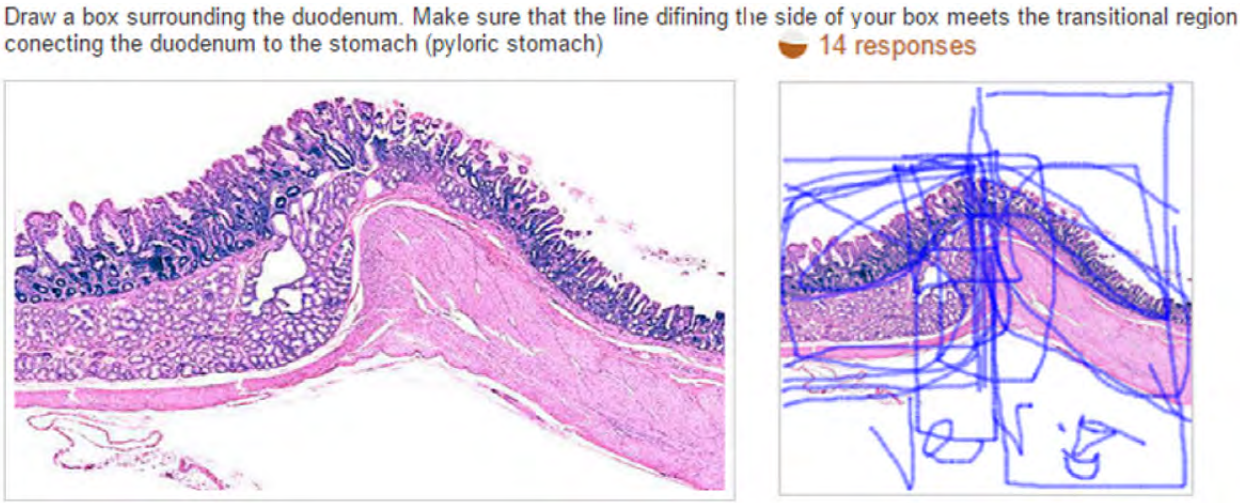
Appendix Figure 3: Students answers when asked to indicate the interface between dermis and epidermis on a sample image they were not familiar with. Students' answers are in blue; the image shows an overlay of all responses. To the left are their initial answers. To the right are their revised answers after discussing the concepts and clarifying misunderstandings in the topic.

APPENDIX FIGURE 1

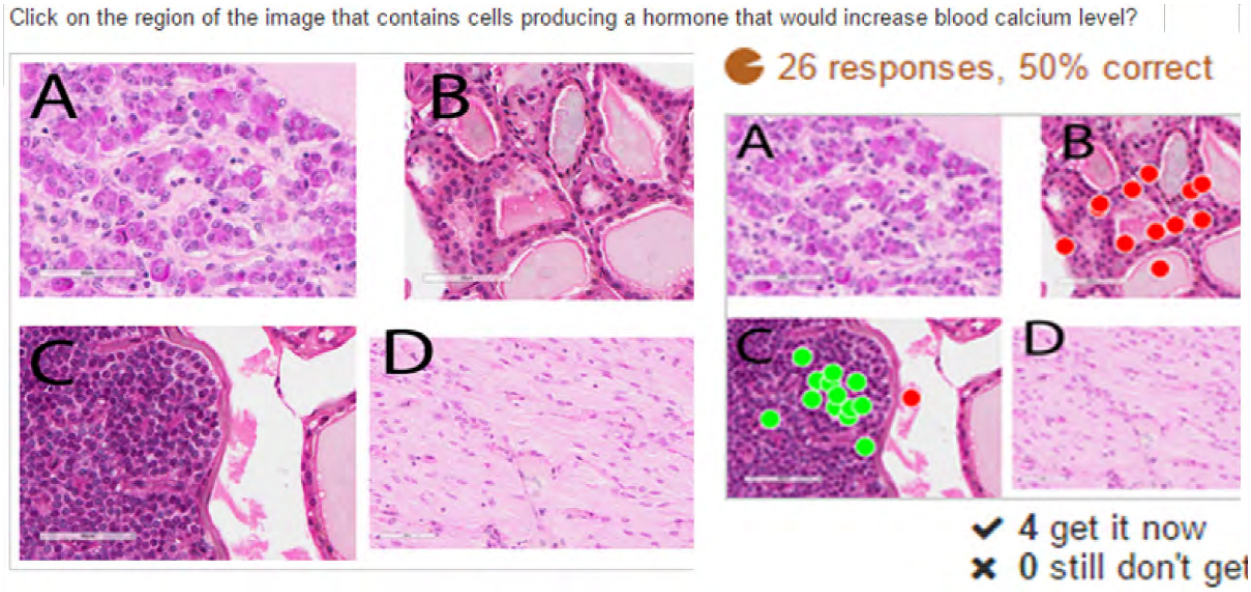


APPENDIX FIGURE 2

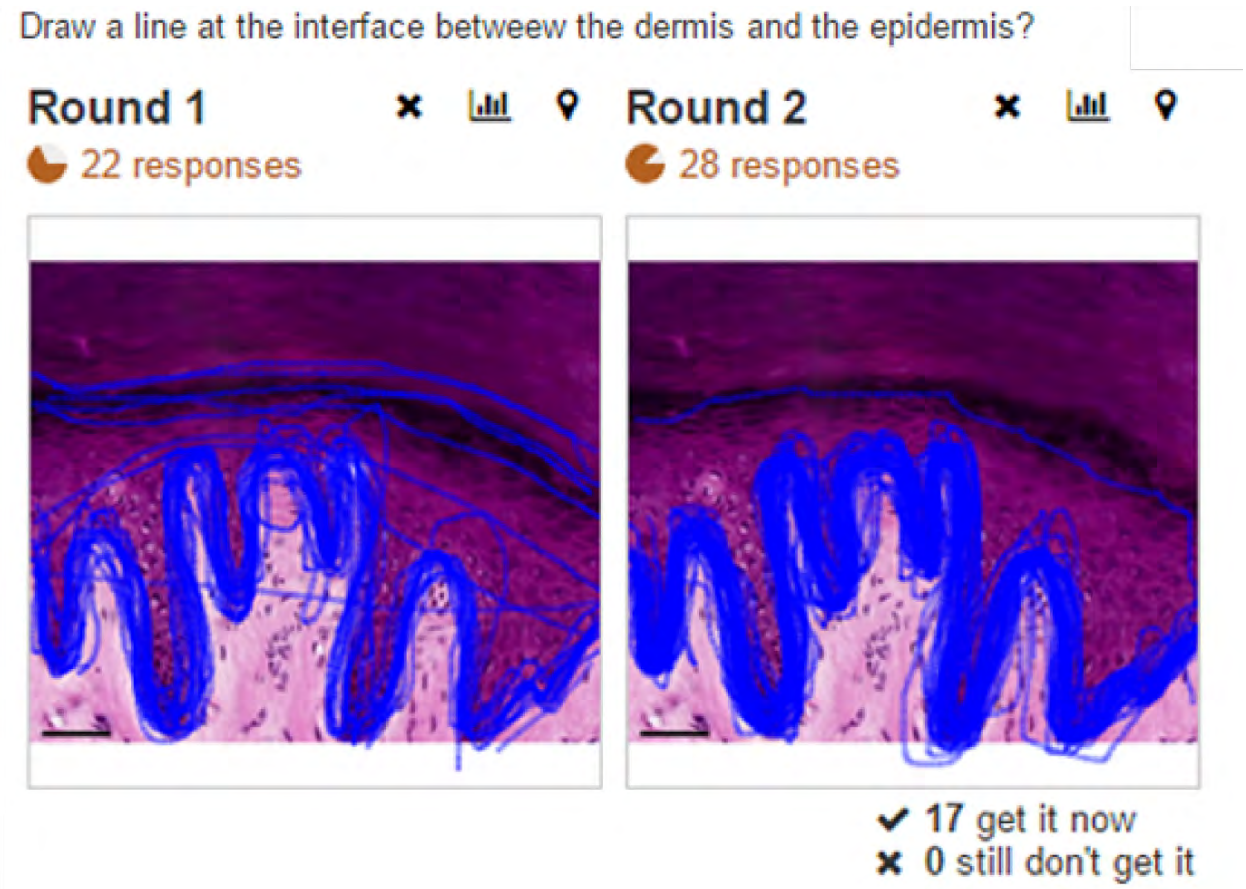
A)



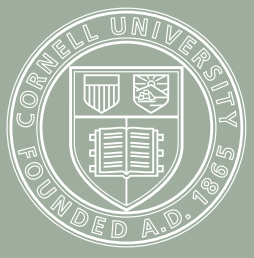
B)



APPENDIX FIGURE 3



This material is based upon work supported in part by the National Science Foundation under award number 1231286.



Center for Teaching Excellence
420 Computing and Communications Center
cornellcte@cornell.edu
www.cte.cornell.edu
607-255-3990