THE IMPACT OF A FLIPPED CLASSROOM MODEL OF LEARNING ON A LARGE UNDERGRADUATE STATISTICS CLASS

PERPETUA LYNNE NIELSEN
Brigham Young University
nielsen@stat.byu.edu

NATHAN WILLIAM BEAN
University of North Carolina at Chapel Hill
nbean@live.unc.edu

ROSS ALLEN ANDREW LARSEN
Brigham Young University
ross.larsen@byu.edu

ABSTRACT

We examine the impact of a flipped classroom model of learning on student performance and satisfaction in a large undergraduate introductory statistics class. Two professors each taught a lecture-section and a flipped-class section. Using MANCOVA, a linear combination of final exam scores, average quiz scores, and course ratings was compared for the two groups after controlling for the effects of students' previous achievement, gender, teacher, degree of learner autonomy, and attitudes about math and statistics. The results show significant improvement in the students' performance and course satisfaction with the flipped classroom. Overall, the results showed that the flipped classroom model can be used in large lecture classes with the help of undergraduate teaching assistants and the use of additional labs.

Keywords: *Statistics education research; Learner autonomy; Math anxiety; Statistics confidence*

1. INTRODUCTION

In the traditional method of instruction, students are first exposed to the lesson content in class and then they are given assignments outside of class to practice and reinforce what they were taught in class. In a flipped classroom, students learn the basics outside of class by reading a textbook, watching videos, or accessing online materials where formative assessments may take place. After studying the lesson content on their own, they come to class to practice and apply what they learned, under the supervision of the instructor or teaching assistants (TAs). During class, instructors or TAs provide student-centered learning activities based on the results of formative assessments students completed before class. The flipped classroom format allows students to be given more personalized feedback (Winquist & Carlson, 2014). Previous studies have shown the positive impact of restructuring teaching sequences on student performance and satisfaction in introductory statistics courses (Peterson, 2016; Strayer, 2012; Vidic & Clark, 2016; Wilson, 2013; Winquist & Carlson, 2014). These positive aspects of the flipped class provided the impetus for this quasi-experiment.

The hypothesis that was tested for this study is that students enrolled in the flipped classroom sections have higher final exam scores, have higher average individual quiz scores, and give higher course ratings than students in the traditional lecture after controlling for their American College Testing (ACT) Math

Statistics Education Research Journal, 17(1), 121-140, http://www.stat.auckland.ac.nz/serj

scores, high school grade point average (GPA), gender, instructor, level of learner autonomy, level of math anxiety, attitude regarding the usefulness of statistics, and confidence in learning statistics.

1.1. LITERATURE REVIEW

Definition The New Media Consortium (NMC) Horizon Report in the 2014 Higher Education Edition defines a flipped classroom as a "model of learning that rearranges how time is spent both in and out of class to shift the ownership of learning from the educators to the students" (p. 36). The teacher adapts instructional and collaborative approaches in class to suit the learning needs of students and to create a more efficient and enriching use of class time. According to NMC, "the flipped classroom model is part of a larger pedagogical movement that overlaps with blended learning, inquiry-based learning, and other instructional approaches and tools that are meant to be flexible, active, and more engaging for students" (p. 36).

Honeycutt and Garrett (2013) described a flipped classroom as one where instructors do not rely on lectures and focus on other ways to enhance learning by using active learning strategies that put students in the center of the learning experience. Several scholars have defined or described the "flip" in different ways. Peterson (2016) and Wilson (2013) described the flipped classroom as a learning environment in which the activities traditionally completed outside of class as homework are now completed in class during instruction time. The activities traditionally completed in class, transmission of knowledge, are now completed in a student's own time before class. Brame (2014) further amplified this description by adding that knowledge-level learning outcomes (remembering and understanding) are achieved by students outside of class and higher-level learning outcomes (analyzing, applying, and evaluating) are achieved during class time where they have the support of their peers and instructors.

Vanderbilt University's Center for Teaching proposed that a flipped classroom should provide the following key elements: (a) an opportunity for students to gain first exposure prior to class, (b) an incentive for students to prepare for class, (c) a mechanism to assess student understanding, and (d) in-class activities that focus on higher level cognitive activities (Brame, 2014). More recently, Abeysekera and Dawson (2015) defined the flipped classroom as a

> set of pedagogical approaches that (a) move most information-transmission teaching out of class, (b) use class time for learning activities that are active and social, and (c) require students to complete pre- and/or post-class activities to fully benefit from in-class work. (p. 3)

Comparative studies examining the efficacy of the flipped model of learning have been few in number (Hamdan, McKnight, McKnight, & Arfstrom, 2013). Bishop and Verleger (2013) examined 24 studies relating to the flipped classroom. They defined 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" (p. 5) and excluded studies that did not use videos as an outside classroom activity. Theirsurvey of the research showed that most studies only explored student perceptions and used single-group designs. Only six studies investigated student learning outcomes as measured by test scores and only one of these six examined student performance throughout a semester. The other 18 studies used informal assessment or subjective opinion surveys. Their survey also showed that student perception of the flipped classroom are mixed but are generally positive overall. Students preferred in-person lectures to video lectures but preferred interactive class activities over lectures. Bishop and Verleger recommended the use of objective learning outcomes and controlled experimental or quasi-experimental designs in future studies.

Implementation The flipped class model of learning can take many forms. O'Flaherty and Phillips (2015) conducted a scoping review of the use of flipped classrooms in higher education. A scoping review examines the extent and nature of research activity, determines whether a systematic review is necessary, and identifies potential research gaps within the existing literature. Using their key search descriptors, the authors identified 1,084 articles. After excluding K–12 studies and research made prior to

1995, they identified 28 studies from five countries (United States, Australia, United Kingdom, Taiwan, and Malaysia) as being relevant to their research topic.

O'Flaherty and Phillips found the following technologies were used in the pre-class asynchronous activities: pre-recorded lectures in the form of podcasts/vodcasts, screencasts, annotated notes, captured videos, pre-readings, automated tutoring systems and study guides, and interactive videos from an online repository (e.g., the Khan Academy's case-based presentations and simulations). The following technologies were used in the in-class face-to-face activities: smart-phone apps, tablets, and iClickerquizzing for real time formative assessments. The following activities were also used in class: problem-solving, case-based presentations, team-based discussions, panel discussions, expert-led discussions, role-plays and student presentations, debates, and think-pair-and-share activities.

Their review of the current literature showed that there was little empirical validation for the claims regarding enhanced class preparation for students, increased classroom interactivity, and improved academic performance. O'Flaherty and Phillips cite that

> there is future potential to explore individual differences in response to different types of course structures i.e., is there a particular demographic or personality that may predict reactions to a flipped class. Placing materials online is a good format to teach lower order cognitive skills but not higher order (Prunuske, Batzli, Howell, & Miller, 2012). (p. 93)

Their findings show that there is no single model for the flipped learning approach but the core features are "content in advance (generally the pre-recorded lecture), educator awareness of students understanding, and higher-order learning during class time" (p. 95). The authors recommend having a stronger link or feedback loop between the pre-class and the face-to-face sessions.

Jensen, Kummer, and Godoy (2015) used a quasi-experimental design to provide quantitative and controlled data to investigate the effectiveness of the flipped learning model. They concluded that

> the flipped classroom does not result in higher learning gains or better attitudes compared with the non-flipped classroom when both utilize an active-learning, constructivist approach and propose that learning gains in either condition are most likely the result of the active-learning style of instruction rather than the order in which the instructor participated in the learning process. (p. 1)

These peer-reviewed studies from 2000–2016 found that a flipped classroom setting was positive and beneficial to student learning, although some of the learning gains may not have been practically significant. Some studies have shown that a flipped classroom approach has a positive impact on student attitude and performance, retention and grades (Butt, 2014, Davies, Dean, & Ball, 2013; Peterson, 2016; Wilson, 2013). Other studies have shown that students were less satisfied with the flipped classroom method (Missildine, Fountain, & Gosselin, 2013; Strayer, 2012).

The studies that found the flipped classroom beneficial tended to include older students or students of a specific major (McLaughlin et al., 2013; Missildine et al., 2013; Pierce & Fox, 2012; Tune, Sturek & Basile, 2013). In several of the studies, the flipped classroom was used for only a short period of time during the semester (for a week or for a particular topic or set of topics), which may have had an impact on the results (Deslauriers, Schelew, & Wieman, 2011; Moravec, Williams, Aguilar-Roca, & O'Dowd, 2010; Vidic & Clark, 2016). A few studies were mostly qualitative in nature, comparing teacher and student perceptions of the effectiveness and efficiency of the flipped classroom with traditional lecture (Dove, 2013; Lage, Platt, & Treglia, 2000; Strayer, 2012). Most comparative studies came from observational studies comparing the flipped approach course with the ones taught by traditional methods in previous semesters (Crouch & Mazur, 2001; Mason, Shuman, & Cook, 2013; Moravec et al., 2010; Wilson, 2013; Winquist & Carlson, 2014). The more recent studies used quasi-experimental designs with small sample sizes without controlling for instructor-, semester-, or time of day-effect (Peterson, 2016). Bishop and Verleger (2013) in their survey of 24 flipped classroom studies also found that researchers did not clearly describe their in-class and out-of-class activities.

Use in undergraduate statistics classes In undergraduate statistics classes, four studies have been conducted regarding the flipped classroom: Several studies employed the flipped classroom idea of having students read the material first before attending class (Wilson, 2013; Winquist & Carlson, 2014).

Peterson (2016) and Vidic and Clark (2016), on the other hand, used prerecorded lectures as the preclass assignment. Additionally, the Peterson (2016) study had the students working on problem sets and quizzes written by the instructor during class. All studies found benefits to the flipped design in terms of higher exam scores (Peterson, 2016; Vidic & Clark, 2016; Wilson, 2013). Winquist and Carlson (2014) found higher retained knowledge 21 months later as measured by a standardized test.

Nevertheless, the literature is missing stronger causal claims which could be remedied by having a quasi-experimental design where controls are put into place (same instructor, same semester, and same time of day of instruction). Many of these studies used smaller sample sizes (in terms of sections taught and number of replication), necessitating additional research in this area.

Predictors of student success Previous studies have shown that student performance in statistics can be explained by math aptitude and gender. Lim and Morris (2009) included gender as one of learner variables influencing perfomance in a blended learning environment. Kintu, Zhu, and Kagambe (2017) also included gender in their study regarding student characteristics and learning outcomes in a blended learning enviroment. Although gender was not a significant predictor of student success in these studies, it needs to be included among our control variables because female students express more positive attitude for online learning (Lim & Morris, 2009) and female students in online learning environments are more engaged and can out-perform their fellow male students (Kintu, Zhu, & Kagambe, 2017). Vella, Turesky, and Hebert (2016) confirmed the latter claim when they found that female students had a significantly higher probability of success in blended courses. In addition to gender, other researchers like Wilson (2013) reported that "students' attitudes, motivation, math anxiety, and preparedness can negatively impact the student and instructor experience and have the potential to negatively impact student learning" (p. 193) in undergraduate introductory statistics courses. Kohli, Peng, and Mittal (2011) also found GPA, Math prerequisite grade, and statistics anxiety as significant predictors of student success in their undergraduate business statistics course. According to Bean and Metzner (1985) and Hawkins, Graham, Sudweeks, and Barbour (2013), prior academic success is also a predictor of student performance in online learning environments. Hence, Math ACT score and high school GPA were added to this study as indicators of prior academic performance.

An important characteristic of successful students in online learning environments is that of learner autonomy. Kerr, Bynearson, and Kerr (2006) reported a positive relationship between independent learning and online student success. Kintu et al. (2017) showed that self-regulation was a significant predictor of learner satisfaction and learner knowledge construction. Peterson (2016) attributes the advantage of flipped classrooms to the facilitation of "more efficient and autonomous interaction with the material (Snodin, 2013; Yang, 2012), allowing for greater collaboration in the classroom (Bernard et al., 2014; Johnson & Johnson, 2009)" (p. 10). Duarte (2013) also investigated the relationship between students' perception of their own learner autonomy and their willingness to learn in a blended learning environment. Research in the science of learning shows that learning motivation is also a significant predictor of student performance. Lim and Kim (2003) identified six elements of learning motivation—among these are perceived relevance or usefulness of subject matter, self-competence or confidence in achieving a certain task, and learner control. Thus, learner autonomy needs to be included as a control variable because it has the potential to impact student success in a blended learning environment. A preliminary search of the literature, however, revealed that there is no single consensual definition of what is meant by "autonomous learning" and its related constructs of "independent learning" and "self-directed learning" (Macaskill & Taylor, 2010). There is also a lack of psychometrically sound, brief measures of learning autonomy outside of language learning.

Two other characteristics of successful online students are self-efficacy, a person's confidence in learning a particular task, and perceived relevance of the subject matter (Bandura, 1986; Lim & Kim, 2003; Shea & Bidjerano, 2010). In the context of statistics education research, attitude towards statistics has been shown to be a good predictor of student achievement and course satisfaction (Finney & Schraw, 2003; Garfield & Ben-Zvi, 2007). The two particular constructs of interest in this study are confidence in learning statistics and opinion on the usefulness of statistics.

The lead author developed abbreviated scales measuring Learner Autonomy, Math Anxiety, Opinion about the Usefulness of Statistics, and Confidence in Learning Statistics. Psychometric properties of these scales have been evaluated. Math Anxiety was used instead of Statistics Anxiety because the survey measuring these constructs was given a week before school started and very few students had been exposed to statistics.

1.2. SUMMARY

The flipped class is a blended learning system which combines face-to-face instruction and computer mediated instruction (Graham, 2006). In addition to blending these two modes of instruction, the flipped class inverts the traditional sequence of teaching by requiring students to learn the basics of the current lesson and to complete formative assessments before class. During class, the instructor uses results of the pre-class assessment to plan learning activities. Students work in groups solving problems to practice and apply what they have learned prior to class or completing quizzes using instant response systems like iClickers in order to encourage class participation among students and to inform instructors of students' level of understanding.

The flipped class contains elements of both the cognitive and constructivist theories of learning. The cognitive theory describes how information is gained, processed, retained, and retrieved. Repeated exposure leads to the information being stored in long term memory which facilitates retrieval. In a flipped class setting, students get at least two exposures to the new information and are given more practice exercises in class. Cognitive theory predicts better student performance in flipped classes (Akyol & Garrison, 2011; Ginsburg & Opper, 1988). The constructivist theory of learning emphasizes the involvement of students in constructing their own knowledge. In a flipped class, students are expected to come to class with some prior knowledge and are given learning activities in class to reinforce and develop what they have previously learned on their own. Constructivist theory also predicts higher student performance in flipped classes (Bishop & Verleger, 2013; Dubinsky & McDonald, 2001; Huang, 2002).

This study aims to address problems associated with small sample sizes, use of observational studies from different semesters, lack of replication, and lack of control for instructor effect and other confounding variables that have been identified as problematic in previous research studies.

2. METHOD

2.1. PARTICIPANTS

The setting for this study is the Department of Statistics at a large private university in the western United States with about 33,000 undergraduate students. The school adopts a trimester schedule of 14 weeks each. With limited faculty and lecture rooms, the department offers an introductory statistics course that has an annual enrollment of at least 4,000 students. As a result, each fall and winter semester, one of the large lecture sections has more than 800 students and the other three lecture sections each have more than 200 students enrolled. The department wanted to explore the viability of using the flipped class approach to teach all students in a 200+ classroom setting where they only meet with the instructor once a week and attend labs twice per week taught by undergraduate TAs instead of the traditional model of meeting with the instructor three times a week and lab once per week.

The institutional review board for human subjects (IRB) at the university approved this study and participants gave their informed consent online. A total of 365 introductory statistics students enrolled in fall 2014 participated in the study: 229 students were in the control sections and 136 students were in the flipped class sections.

At this private university, the introductory statistics course fulfills two general education requirements. As a result, students in both the control and flipped class sections came from all levels of class standing (i.e., freshman, sophomore, etc.) and had a variety of majors. Students were not expected to have completed any mathematical courses beyond college algebra, the course's only prerequisite. Some upper-level undergraduate students also enrolled in the course to fulfill a prerequisite of graduate programs.

2.2. DESIGN OF STUDY

The Department of Statistics designed a quasi-experiment in fall 2014 to investigate the impact, if any, of a flipped classroom on students' final exam scores, average individual quiz scores, and overall course satisfaction ratings in an introductory statistics course with college algebra as a prerequisite. The course consisted of 38 lessons with 38 online quizzes associated with each lesson. These quizzes were created by a committee consisting of four faculty members teaching the course. Quiz questions were aligned with the learning outcomes of the lessons they were associated with. The course had three midterms and a comprehensive final exam developed by the same committee that created the quizzes. All of these exams are reviewed and revised each semester by this committee. Final exam scores were used as a measure of aptitude, individual quiz scores as a measure of effort, and end-of-semester course satisfaction ratings as a measure of attitude. Four large lecture sections of the statistics course were compared, two using the traditional three lecture-one lab per week format (the control group) and the other two using a flipped classroom model using a one class-two labs per week format (the treatment group). Lectures and labs were 50 minutes each.

In this study, we incorporated the following features of the flipped class that contributed to improved student performance and course satisfaction in past studies: making students accountable for completing assignments given prior to class (Crouch & Mazur, 2001; Wilson, 2013; Winquist & Carlson, 2014), acting on feedback from pre-class formative assessments (Crouch & Mazur, 2001; O'Flaherty & Phillips, 2015); using active learning strategies and testing in class (Dove, 2013; Jensen et al., 2015), and using peer instruction or group activities in class (Crouch & Mazur, 2001; Lage et al., 2000; Mason et al., 2013; Peterson, 2016; Strayer, 2012; Wilson, 2013).

Flipped class setup Students were given reading assignments for each lesson using an interactive online courseware provided by Carnegie Mellon's (https://community.oli.cmu.edu/) Open Learning Initiative before class and lab. This courseware contains learning activities (e.g., Learn By Doing, Did I Get This?, and CheckPoints) which are graded. For each assigned lesson, students in the flipped classes were required to complete two separate online quizzes: a group quiz and an individual quiz. Each week, students in the experimental group had one class meeting (every Wednesday) for 50 minutes with the instructor and two labs (every Monday and Friday) for 50 minutes each with an undergraduate teaching assistant. These class meetings were different from traditional lectures as they involved group quizzes and focused on application. On the first week of class, students were randomly assigned to groups of four or five in their labs. In the first 10 minutes of the lab, undergraduate teaching assistants discussed the results of and answered students' questions about the previous lesson's individual quiz and group quiz results. These quizzes contained 7–10 questions that focused on higher order thinking skills like understanding, applying, and analyzing. After the quiz feedback, students worked together in their assigned groups to complete a group quiz pertaining to the current lesson's reading assignment. Students submitted their group quizzes individually and these group quizzes comprised 4% of students' grades. Groups that finished early worked on the individual quiz associated with the current lesson without the help of their group members. Students were expected to complete these individual quizzes on their own and these quizzes were worth 8% of the student's grade. These quizzes had no time limit but had daily deadlines and could only be taken once. Results of these group and individual quizzes were used by the instructor and teaching assistants to prepare for class and lab activities. On the Wedneday meeting with the instructor, the first 10 minutes of class was spent discussing the most frequently missed questions of previous quizzes. Based on the results, the instructor decided either to review misunderstood topics or go on with a 10-minute overview of the current lesson topic. After the quiz feedback and/or mini-lecture, students worked in their assigned groups to answer quiz questions regarding the current lesson or reading assignment using their iClickers under the guidance of their instructor. Instructors gave instant feedback to these iClicker questions in class. Depending on the quiz results, the professor either moved on to the next question or conducted a mini-lecture on the questions or topics where a majority of the students missed the correct answer.

Traditional lecture setup Students in the control group, the traditional lecture in the department, met three times a week with their instructor and once a week with their assigned lab TA. These lectures and labs were all 50 minutes in length. Students in the flipped and lecture classes used the same online course materials and took the same individual quizzes and exams. Lecture students were given online practice quizzes instead of group quizzes but the content and coverage of the group and practice quizzes were similar. These practice quizzes were not graded and had no deadlines, and students could take them multiple times. Lecture students were given the same online reading assignments as the flipped class students but were not tested on these materials. In lectures, the instructor discussed the content of the reading assignments using PowerPoint slides developed by the department and provided additional examples and practice exercises. Both lecture and flipped class students had access to these PowerPoint slides which were posted on the school's learning management system. To encourage class participation, a few questions were posed throughout the lecture, and students responded to these questions using iClickers. These questions were mostly remembering and understanding questions that took less than a minute to answer. Instructors then gave instant feedback. Outside of lecture, students were asked to complete the online individual quizzes which comprised 15% of their course grade. In labs, the TAs discussed the most frequently missed individual quiz questions and reviewed topics discussed in lectures based on students' questions. In contrast to the flipped class students, lecture students used the time in labs to review the previous class material rather than to work on group quizzes pertaining to the current lesson.

Study protocol The course manager, who is also the principal investigator and first author, prepared the lab materials and trained and supervised the teaching assistants. During weekly TA meetings, all undergraduate teaching assistants received training to master the course material and improve teaching skills. The TAs who were assigned to the flipped class labs received additional weekly training from the course manager. The purpose of this training was to help the TAs understand how they should conduct each lab in accordance with the design of the flipped classroom model so that each lab experience would be similar across the different TAs.

The study was designed as a quasi-experiment because students could not be randomly assigned to the two groups of interest. However, when classes opened for registration, they were made to look similar. One week before class started, a pre-course survey was given to all students in the sections taught by the two instructors who participated in the course. This survey informed the students of the purpose of the study and asked for their consent to participate as agreed upon with IRB. Students who declined to participate enrolled in another lecture section of the course taught at the same time by a different instructor. The study sought to control for teacher effect by having two instructors use both teaching methods. Two faculty members each taught a flipped class and a lecture class. It also sought to control for time of day by offering both teaching methods in the morning (9:00 AM) and in the afternoon (2:00 PM).

On the first day of class, the two instructors discussed the details of the study and the different class format with students in the flipped class sections. Labs for the introductory statistics course started on the second week of class but students in the flipped class sections met their lab TA on the first week of class to discuss the lab format and be assigned to their groups. They were also given time to register on the required *e*-book website to access the course content, register their iClickers, and complete the pre-course survey if they had not done so or they added the class late.

2.3. VARIABLES OF INTEREST

Explanatory variable The explanatory variable in this study is the method of instruction, whether flipped classroom or traditional lecture, referred to as Group.

Outcome variables The three response variables used in the study were:

- Final exam percentage obtained from the school's Testing Center website and did not include any extra credit points
- Average Individual Quiz scores obtained from the school's learning management system
- Course Rating obtained from the University's Assessment office coming from one question regarding overall course rating. The actual question asked was, "Comparing this course with other university courses, please indicate an overall rating for this course: 1 = Exceptionally poor, 2 = Very poor, 3 = Poor, 4 = Somewhat poor, 5 = Somewhat good, 6 = Good, 7 = Very good, 8 = Exceptionally good

Control variables The following eight variables were used as control variables:

- Gender obtained from the school's Records Office
- Instructor
- ACT Math score obtained from the school's Records Office
- High School GPA obtained from the school's Records Office
- Opinion regarding the Usefulness of Statistics – this construct was measured by a pre-course survey given to all study participants
- Confidence in Learning Statistics – this construct was measured by a pre-course survey given to all study participants
- Learner Autonomy obtained from the pre-course survey given to all study participants
- Math Anxiety – an abbreviated measure of this construct was used in the pre-course survey given to all study participants

In fall 2014, a total of 766 students in the introductory statistics course completed a pre-course survey consisting of 17 statements to measure their attitude towards statistics and their level of math anxiety and learning autonomy. They were asked to choose which of the following options (with their point values in parentheses) reflected their current feelings: Strongly Disagree (1), Disagree (2), Somewhat Disagree (3), Somewhat Agree (4), Agree (5), and Strongly Agree (6). Exploratory factor analysis (EFA) extracted four factors which explained 48.43% of the common variance and yielded a Kaiser-Meyer-Olkin (KMO) sampling adequacy measure of 0.908 which is in the good range, according to Worthington and Whittaker (2006). Table 1 shows the loadings and Cronbach's alpha of the four factors that were extracted using Principal Axis Factoring and Promax oblique rotation.

2.4. STATISTICAL ANALYSES

After the data files from the different sources were merged and verified for accuracy, intra-class correlations (ICCs) were calculated to estimate Design Effects (DEFF), where DEFF = 1 + (average cluster size – 1) × ICC to determine whether clustering or nesting effects of students in different sections can be ignored. According to Muthen and Satorra (1995), if DEFF is less than 2, we can exclude the random effects of section in the analysis. To test for equivalence between the two groups prior to group assignment with regards to the covariates, two-sample *t*-tests comparisons were obtained for High School GPA, Math ACT score, Math Anxiety, Learner Autonomy, Confidence in Learning Statistics, and Opinion on the Usefulness of Statistics. In addition, because the dependent variables were strongly correlated, a multivariate analysis was employed. More specifically a multivariate covariance analysis or MANCOVA was used to tease out the effects of the explanatory and control variables on the three dependent variables of interest using SPSS, JMP, and Mplus.

Table 1. Loadings and Cronbach's alpha for the four factors extracted

Item description	Loading
Confidence in Learning Statistics (α = .879)	
Statistical reasoning will be easy to understand.	.919
Statistics will be easy to understand.	.781
Statistics is a difficult subject. (reverse scored)	.710
I will feel nervous and tense in this Statistics class. (reverse scored)	.626
I feel confident in my ability to learn Statistics.	.579
I will have no idea of what is going on in this course. (reverse scored)	.545
Math Anxiety (α = .787)	
I have taken as few Math courses as possible.	.763
Dealing with numbers makes me feel uncomfortable.	.730
I will make a lot of math errors in this course.	.488
Usefulness of Statistics (α = .770)	
The study of Statistics will be very useful in my daily life.	.802
The study of Statistics will be very useful in my work.	.768
Statistical thinking is not useful to most professionals. (reverse scored)	.653
I am looking forward to learning Statistics.	.505
Learner Autonomy (α = .455)	
I need external rewards in order to feel motivated. (reverse-scored)	.620
I do very well learning on my own.	.499
I can keep to a schedule.	.363
I would rather sit through an hour of lecture than study for an hour on my own. (reverse-scored)	.306

3. RESULTS

A total of 392 students consented to participate in the study and 365 students (93%) completed the course. Of the students who completed the course, more than a third (37%) were in the flipped class, less than half were female students (44%), and one of the instructors taught at least 71% of the students. The mean ACT Math scores for all participants was 26.62 with standard deviation 4.05. The mean High School GPA for all students in the study was 3.68 with standard deviation 0.42. A comparison of the mean ACT Math scores and mean High School GPA for the flipped and control sections did not result in any significant difference, using a significance level of 0.05. The proportion of male and female students in the flipped and control sections were also very close and showed no significant difference (see Table 2). There was also no significant difference in the mean Math Anxiety score, mean Learner Autonomy score, and mean rating of Confidence in Learning Statistics and Opinion on the Usefulness of Statistics between students in the two groups. We can safely conclude that students in the flipped and control groups were comparable in these aspects. Note that higher values indicate higher levels of the attribute; for example, higher values indicate higher levels of Math Anxiety or Higher Learner Autonomy.

Table 2. Comparison of participants by group

Variable	All participants	Flipped class	Control group	*p*-value[a]
Gender	*n* = 365	*n* = 136	*n* = 229	
Female	44.1%	45.6%	43.2%	.371
Male	55.9%	54.4%	56.8%	
Instructor	*n* = 365	*n* = 136	*n* = 229	
Instructor 1	71.2%	71.3%	71.2%	.538
Instructor 2	28.8%	28.7%	28.8%	
ACT Math score	*n* = 324	*n* = 118	*n* = 206	
Mean (SD)	26.62 (4.05)	26.41 (4.09)	26.74 (4.03)	.475

High School GPA	*n* = 332	*n* = 121	*n* = 211	
Mean (SD)	3.68 (.42)	3.67 (.45)	3.69 (.39)	.727
Math Anxiety	*n* = 256	*n* = 93	*n* = 163	
Mean (SD)	9.63 (3.20)	9.30 (3.35)	9.82 (3.12)	.227
Learner Autonomy	*n* = 257	*n* = 94	*n* = 163	
Mean (SD)	16.03 (2.74)	16.38 (2.80)	15.83 (2.69)	.123
Stat Usefulness	*n* = 255	*n* = 93	*n* = 162	
Mean (SD)	17.58 (3.31)	17.35 (3.12)	17.70 (3.42)	.407
Stat Confidence	*n* = 253	*n* = 94	*n* = 162	
Mean (SD)	20.17 (4.43)	20.48 (4.02)	19.99 (4.66)	.379

[a]Equal variances not assumed. Two-tailed tests.

A comparison of student performance showed that the students in the flipped classes obtained slightly higher average Final Exam and higher Average Individual Quiz scores than those in the control group. A comparison of student satisfaction as measured by Course Rating showed that students in the flipped class sections gave their course a higher rating than those in the control group as shown in Table 3.

Table 3. Summary of student performance and satisfaction by group

Variable	All participants	Flipped class	Control group
Final Exam percent	*n* = 365	*n* = 136	*n* = 229
Mean (SD)	77.73 (14.54)	78.70 (14.76)	76.99 (14.41)
Average Individual Quiz score	*n* = 365	*n* = 136	*n* = 229
Mean (SD)	88.65 (6.77)	89.69 (5.65)	88.04 (7.30)
Overall Course Rating	*n* = 338	*n* = 130	*n* = 208
Mean (SD)	6.23 (1.21)	6.36 (1.14)	6.15 (1.25)

Figures 1 to 3 show boxplots of Final Exam percent, Average Individual Quiz percent, and Course Ratings for the two groups, Flipped and Control, respectively. We observed some outliers which might

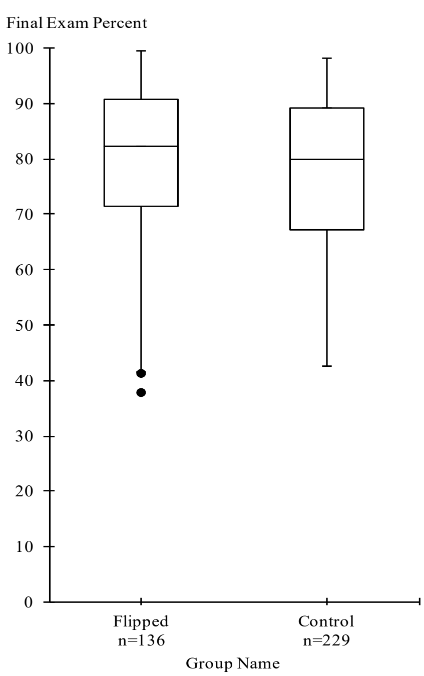

Figure 1. Final Exam percent for control and flipped sections

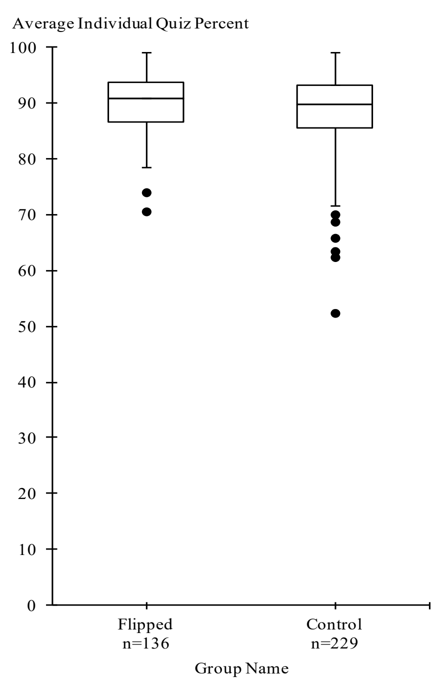

Figure 2. Average Individual Quiz percent for control and flipped sections

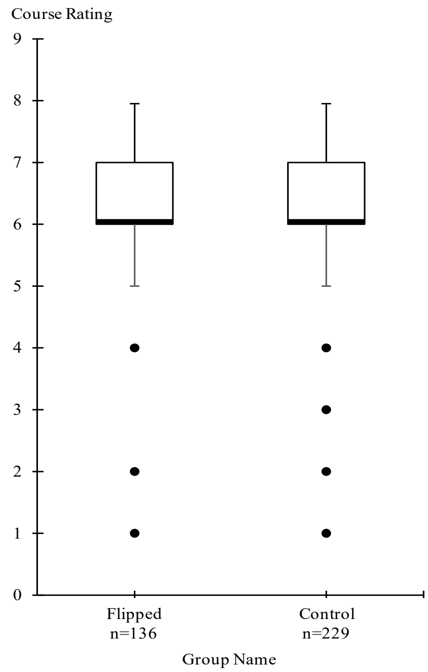

Figure 3. Course Rating for control and flipped sections

complicate the results of the study. Data were analyzed with and without these outliers and resulting findings were similar. The results of the analyses with the outliers are discussed in the following sections.

3.1. MANCOVA RESULTS

To determine whether these differences were statistically significant after the effects of the control variables were taken into account, a General Linear Model (GLM) analysis was used in SPSS 23 and JMP 13. A similar analysis was run using Mplus 7.4 which handles missing data through the full information maximum likelihood method (FIML) and has relaxed assumptions on the covariances of the outcomes. The pattern of results between SPSS and Mplus were very similar, showing that missing data and relaxed assumptions on the covariances of the outcomes did not change the results. Thus, only the SPSS results are shown. The correlation matrix for the three quantitative outcome variables showed that they were highly correlated: the correlation coefficients ranged from .280 for Course Rating and Average Individual Quiz scores to .649 for Average Individual Quiz and Final Exam scores. All of these values were significant at the .01 level. Hence, the decision to use a particular GLM analysis, MANCOVA, was made.

There was also a clustering issue in the design of the study, because students were nested with sections and instructors. Intra-class correlations (ICCs) were calculated to estimate Design Effects (DEFF) which were all less than 2, so the random effects of Section were excluded in the analysis and hierarchal modeling was not used (Table 4).

Table 4. Intra-class correlations (ICC) and Design effects (DEFF) calculations

Outcome variables	ICC	DEFF
Final Exam score	.028	1.397824
Average Individual Quiz score	.037	1.525696
Course Rating	.041	1.582528

The following multivariate model with the three outcome variables combined was run using Group (Flipped and Control), Gender, and Instructor as fixed factors and High School GPA, Math ACT score, Math Anxiety level, Learner Autonomy measure, Confidence in Learning Statistics, and Opinion about the Usefulness of Statistics as covariates:

$$Y_{ij} = \text{Group} + \text{Gender} + \text{Instructor} + \text{HSGPA} + \text{Math ACT} + \text{Math Anxiety} \\ + \text{Learner Autonomy} + \text{Stat Confidence} + \text{Stat Usefulness} + \varepsilon_i$$

Inclusion of all three outcome variables in the analyses would provide the maximum amount of information regarding the effect of the explanatory variable, Flipped class status. A full model with the nine main effects and all two-way interactions involving Group, Gender, Instructor, and the other six covariates was first specified. A best model was selected using the following criteria: high R^2, lower values of AIC and BIC, and as guided by theory. The R^2 values were obtained from SPSS and the AIC and BIC values were obtained from JMP. This best model contained all the main effects and nine 2-way interactions and should have a multivariate R^2 of at least 41% as it is the highest univariate r^2. The univariate values of r^2 are .406, .343, and .247 for Final Exam, Individual Quiz, and Course Rating, respectively. The MANCOVA results for this best model had a significant Box's test of equality of covariance (p = .002) so the Pillai's Trace was used instead of Wilk's lambda. It also had a non-significant Levene's test of equality of error variances: $F(7, 191) = 0.649$, $p = .715$ for Final Exam; $F(7, 191) = 1.861$, $p = .078$ for Average Individual Quiz; and $F(7, 191) = 1.111$, $p = .358$ for Course Rating, so the assumption of homogeneity of variances is met. The assumptions of independence, linearity among the dependent variables, multicollinearity, and equality of regression slopes were looked into and, except for non-normality, no significant violation was found. The Design Effects for the lack of independence were all below the cutoff of 2 indicating that the higher level could be ignored in the analysis (Muthen & Satorra, 1995). In addition, the MANCOVA test procedures are robust against the violation of the normality assumption for large sample sizes. This study's sample size of 365 with nine independent variables of interest is large enough to invoke the Central Limit Theorem. Values of High School GPA equal to 0 for two students were also excluded from the analysis as these values were associated with home-school students. We can therefore continue with the interpretation of the following results.

The MANCOVA tests showed a significant omnibus effect for six of the nine main effects and three of the nine two-way interactions on the multivariate space defined by the three outcome variables. The most significant main effect was Group, followed by Confidence in Learning Statistics, Gender, Math ACT score, level of Math Anxiety and High School GPA. The least significant main effect was Teacher (see Table 5).

These findings show that being assigned to the flipped class sections had a significant difference in students' performance and course satisfaction after controlling for their Gender, Teacher, Math ACT score, high school GPA, their Math Anxiety, degree of learner autonomy, and attitude towards Statistics. Specifically, students in the flipped class sections had higher Final Exam scores, higher Average Individual Quiz scores, and gave the course higher ratings. Table 6 shows the estimated marginal means of these outcome variables for Group after accounting for the six covariates. The significant differences in the Final Exam percentages, Average Individual Quiz scores, and Course Rating between the Control and Flipped class sections indicate that the flipped classroom model has a potential impact on the improvement of student performance. Values of partial eta^2 are included in Table 5 as measures of effect size. They can be defined as the ratio of variance accounted for by an effect and that effect plus its associated error variance in a MANCOVA study after partialling out the effects of other explanatory variables and interactions. Medium effect sizes should have partial eta^2 around .0588 (Cohen, 1969; Richardson, 2011) and the effect size for Group has a greater value than this benchmark. The power for the test involving Group is also greater than .80 which implies that we have sufficient power to detect a significant difference.

Table 5. MANCOVA results for best model combining all three outcome variables

Effect	Pillai's Trace	*p*	*partial eta²*	*Power*
Group	.079	.002	.074	.896
Stat Confidence	.072	.004	.071	.876
Gender	.069	.006	.062	.825
Math ACT	.048	.034	.075	.897
Math Anxiety	.046	.039	.044	.659
High School GPA	.045	.041	.046	.677
Stat Usefulness	.039	.069	.040	.604
Learner Autonomy	.039	.070	.029	.459
Teacher	.032	.124	.026	.419
Group × Math Anxiety	.073	.004	.068	.864
HS GPA × Stat Confidence	.071	.005	.069	.866
Gender × HS GPA	.053	.021	.050	.722
Gender × Stat Usefulness	.034	.103	.037	.572
Group × Stat Confidence	.029	.157	.025	.403
Group × Stat Usefulness	.028	.165	.031	.489
StatCon × Stat Usefulness	.024	.234	.024	.383
Gender × Math ACT	.021	.284	.020	.314
Teacher × Stat Usefulness	.021	.285	.016	.264

Table 6. Estimated marginal means of outcome variables for group after accounting for the covariates

Outcome variable	Group	Mean[a]	SE	95% CI
Final Exam	Control	79.496	1.191	(77.146, 81.845)
	Flipped	80.446	1.432	(77.620, 83.273)
Individual Quiz	Control	88.351	.584	(87.198, 89.504)
	Flipped	89.281	.703	(87.894, 90.668)
Course Rating	Control	6.189	.117	(5.958, 6.419)
	Flipped	6.454	.141	(6.177, 6.732)

[a]Covariates appearing in the model are evaluated at the following values: HSGPA = 3.7536, MATHACT = 26.75, StatCon = 20.23, StatUse = 17.50, MathAnx = 9.43, LrnrAuto = 16.08

Looking at the univariate results for each outcome variable, students' perception of Confidence in Learning Statistics was not associated with Final Exam scores, on average, ($\hat{\beta}$ = -3.081, p = .303); but it had a negative association with Average Individual Quiz scores ($\hat{\beta}$ = -4.071, p = .006) which makes sense as the Average Individual Quiz scores measure effort not necessarily proficiency; and had a marginal association with lower Course Ratings ($\hat{\beta}$ = 0.503, p = .087). Gender also had a significant association with the three outcome variables combined. When looking at the individual outcome variables, male students had higher Final Exam scores, higher Average Individual Quiz scores, and gave higher Course Ratings. Table 7 shows the estimated marginal means of these outcome variables for Gender after accounting for the six covariates.

In addition, students with higher Math ACT scores obtained higher Final Exam scores ($\hat{\beta}$ = 3.09, p = .025) but Math ACT score is not associated with differences in the Individual Quiz scores ($\hat{\beta}$ = .598, p = .374) or Course Ratings ($\hat{\beta}$ = -0.007, p = .957). High School GPA did not have a significant effect ($\hat{\beta}$ = -15.951, p = .326), but was associated with lower Average Individual Quiz scores ($\hat{\beta}$ = -20.705, p = .010), and had a marginal relationship with higher Course Ratings ($\hat{\beta}$ = 2.641, p = .098). Math Anxiety was not associated with Final Exam scores ($\hat{\beta}$ = 0.536, p = .328), or Individual Quiz scores ($\hat{\beta}$ = -0.249, p = .354), but did have a negative association with Course Ratings ($\hat{\beta}$ = -0.136, p = .012).

Table 7. Estimated marginal means of outcome variables for gender after accounting for the covariates

Outcome variable	Gender	Mean[a]	SE	95% CI
Final Exam	Female	79.079	1.399	(76.318, 81.840)
	Male	80.863	1.284	(78.329, 83.397)
Individual Quiz	Female	87.966	.687	(86.611, 89.321)
	Male	89.666	.630	(88.422, 90.909)
Course Rating	Female	6.316	.137	(6.045, 6.586)
	Male	6.328	.126	(6.079, 6.576)

[a]Covariates appearing in the model are evaluated at the following values: HSGPA = 3.7536, MATHACT = 26.75, StatCon = 20.23, StatUse = 17.50, MathAnx = 9.43, LrnrAuto = 16.08.

Three of the ten two-way interactions that were included in the best model showed significant omnibus effect on the three outcome variables. The most significant interaction was Group by level of Math Anxiety (p = .004), followed by High School GPA and Confidence in Learning Statistics (p = .005) and lastly by Gender and High School GPA (p = .021). We can see from Figure 4 that there is a negative relationship between Math Anxiety and Final Exam scores for students in both Flipped and Control groups. However, the slope is much steeper for students in the Control group. We can take this to mean that for these students, their level of Math Anxiety has a greater negative influence on their Final Exam scores, compared to those in the Flipped sections who seem to have a better handle on their anxieties. The same relationship was observed with the other two outcome variables, Individual Quiz score and Course Rating, but the results were not as dramatic for the Final Exam scores (see Figures 5 to 8).

There was also a significant interaction between Gender and High School GPA: Male students who had high GPAs tended to have higher Final Exam scores, higher Average Individual Quiz scores, and tended to give higher Course Rating than female students who had similar GPAs. Figure 9 shows the Gender and High School GPA interaction on Final Exam scores. The interaction plots for Average Individual Quiz scores and Course Rating show the same pattern.

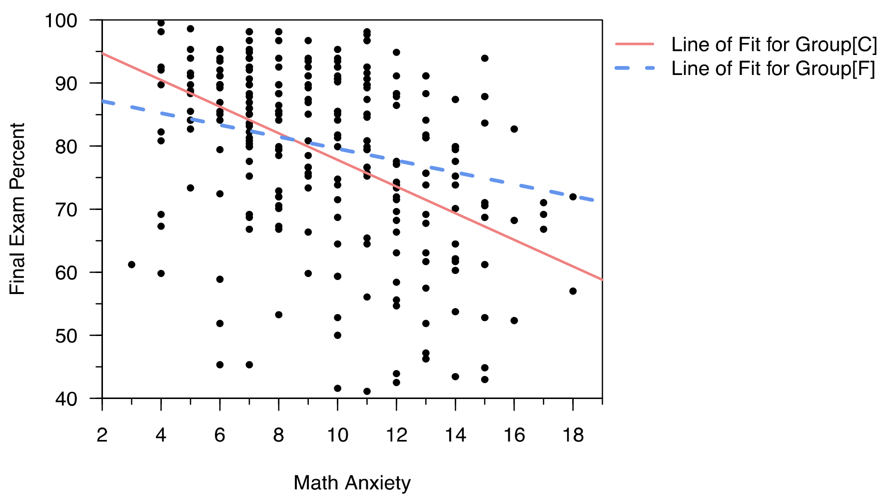

Figure 4. Plot of interaction between group and math anxiety for final exam scores

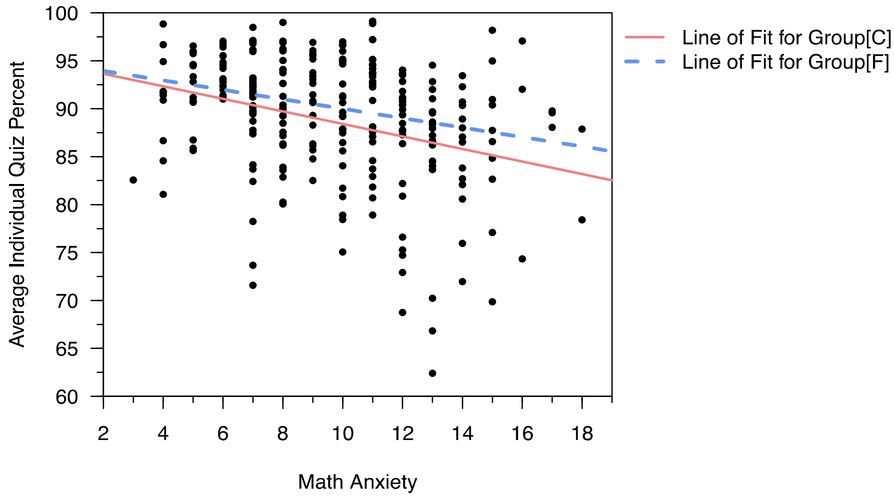

Figure 5. *Plot of interaction between group and math anxiety for average individual quiz scores*

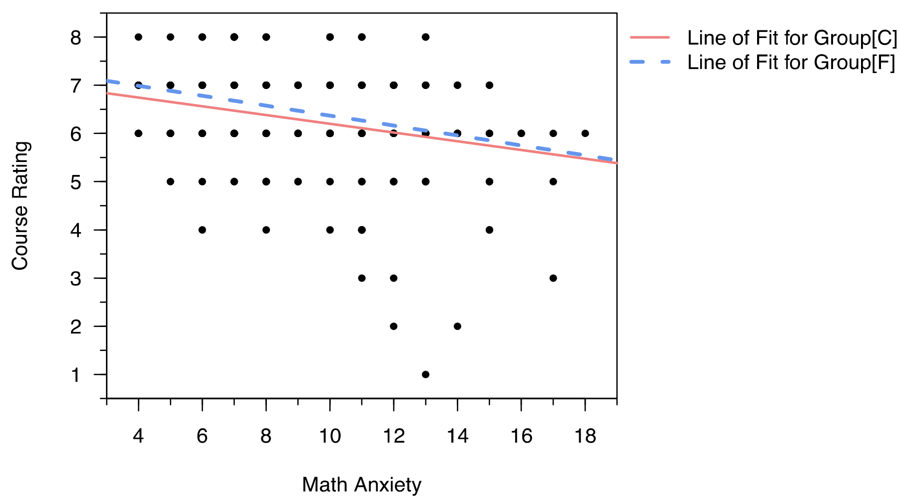

Figure 6. *Plot of interaction between group and math anxiety for course rating*

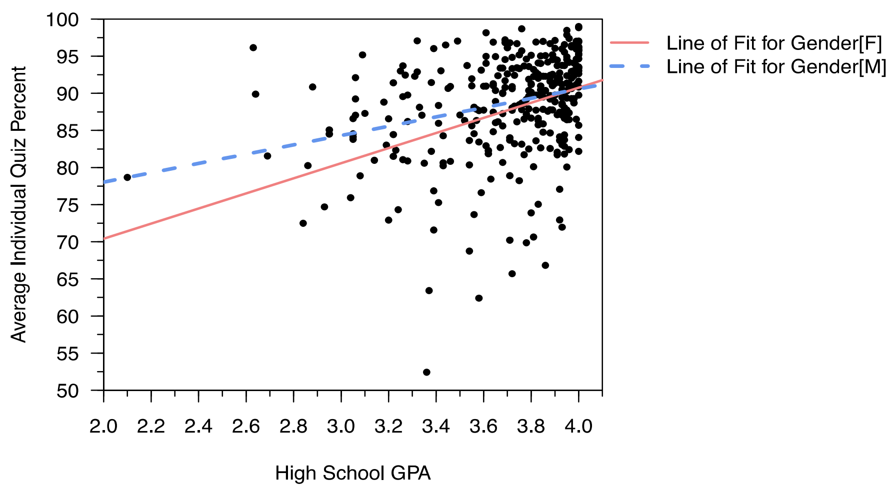

Figure 7. *Interaction between gender and HS GPA for average individual quiz scores*

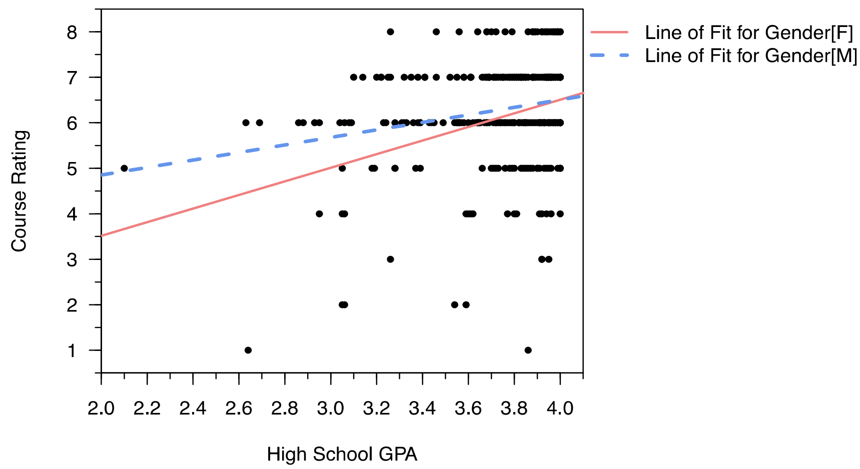

Figure 8. Interaction between gender and high school gpa for course rating

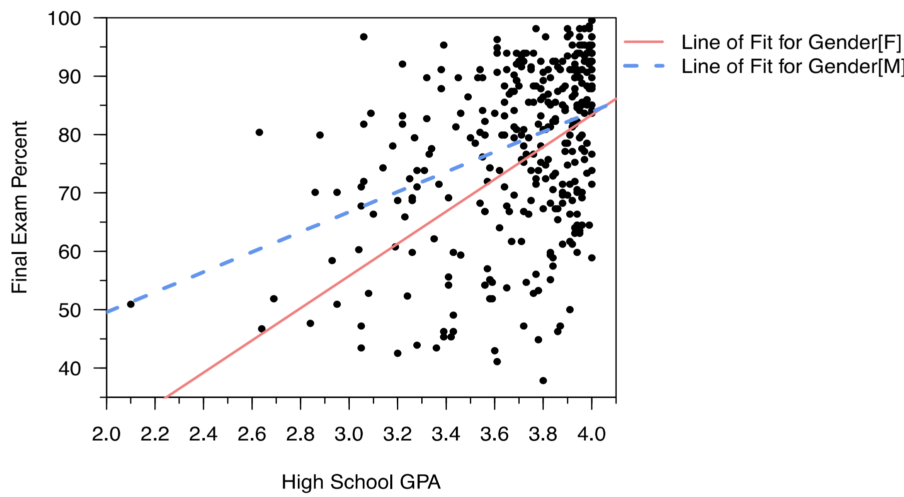

Figure 9. Interaction between gender and high school gpa for final exam scores

To summarize, students in the flipped classes performed better and had better attitudes towards the course after controlling for the effects of gender, teacher, and the six covariates of interest. In addition to Group, the other significant predictors of student performance and course satisfaction were Confidence in Learning Statistics, Gender, Math ACT score, level of Math Anxiety, and High School GPA.

4. DISCUSSION AND CONCLUSION

The most important finding of this study is that our version of the flipped classroom model of learning has the potential to improve student performance in and attitude towards introductory statistics courses. This result holds after controlling for the effects of gender, teacher, Math ACT score, High School GPA, level of Learner Autonomy, Confidence in Learning Statistics, and Opinion about the Usefulness of Statistics. The findings in this study validated the inclusion of four core features of the flipped classroom model of learning in previous successful studies:

- students are held accountable for learning basic content before class
- teachers are aware of student understanding based on formative assessments

- students engage in active learning and testing during class time
- students collaborate in completing group activities or quizzes in class

The results of our study can only be generalized to flipped classroom models incorporating these four features.

Another important finding is that our flipped classroom model can be used in large lecture classes without inflicting harm on the students, with the help of undergraduate teaching assistants and the use of additional lab space. With adequate training and supervision, TAs can facilitate the use of active learning strategies in class or labs, and can monitor group performance. Informal interviews also indicated that students, teaching assistants, and teachers preferred our flipped classroom model to the traditional lecture structure.

The results also showed that students' perceived Confidence in Learning Statistics, gender, Math aptitude, and level of Math Anxiety affected their success in our course and might possibly be a cause for concern. Multivariate results (MANCOVA) sometimes differed from the univariate (ANCOVA) results when looking at individual outcome variables. However, the high correlations among the outcome variables (all significant at the .001 level) indicate that the multivariate results should trump the univariate results.

Winquist and Carlson (2014) suggested that future research "determine if this flipped class approach interacts with instructor characteristics such that the approach is only effective for some instructors" (p. 9). Our study showed no teacher effect with a *p*-value of .124 (see Table 5). The instructors who participated in the study had different genders and widely different ages. Wilson (2013) also suggested future research to "determine whether there are particular demographic or personality variables that might predict reactions to a flipped class" (p. 198). We found that students with higher level of Math Anxiety performed less poorly in our flipped class setting and gave the class higher Course Ratings. Quoting Hughes (2007), Peterson (2015) argued that "blended courses more generally help at-risk learners, but there has yet to be any exploration of this question, specifically within the flipped class framework" (p. 14). One of our findings showed that students with lower high school GPA and lower Math ACT scores tended to have lower exam scores and lower quiz scores but this pattern was similar for students in both flipped and lecture classes. However, the lowest High School GPA and lowest Math ACT score of the participants in this study were 2.10 and 14, respectively, and they would not be classified as at-risk students.

Our results failed to show that level of learning autonomy had a significant effect on student performance and course satisfaction in our flipped classrooms. It may be that the measure being used did not have good psychometric properties. There is also the issue of 29% missing values for the four covariates that measured attitudes toward Math and Statistics, and Learner Autonomy.

We did not randomly assign students to the two teaching methods but we controlled for a number of individual difference variables like gender, high school GPA, Math ACT scores, and attitude towards Math and Statistics. A comparison of means for these variables showed that students in the lecture and flipped classes had similar characteristics regarding these variables. We can therefore attribute the observed difference in student performance and course satisfaction to the different teaching methods and not on student ability or attitude towards statistics.

Furthermore, we had stronger causal claims because the same instructors taught both the flipped and control sections. We believed this controlled for in-class pedagogy in a way that previous studies had not because the treatment and control sections were taught by the same instructor in different semesters (Peterson, 2016; Wilson, 2013). We were able to isolate the effect of using the flipped class format above and beyond any teacher effects.

One issue that was not addressed was the time and cost of developing a flipped class. This study was unique in its approach because it involved the participation, directly and indirectly, of the whole Statistics department. To conduct this quasi-experiment, three faculty members, including the course manager, were given special assignments and reduced teaching load. The department also provided extra lab space, and extra TA and secretarial support. The University also provided funding and confidential information like high school GPA, Math ACT scores, and end-of-semester course ratings for the students who participated in the study. These conditions may not be easily reproduced in other college settings.

Because of the unique characteristics of students at this private research university, at least 86% are Caucasian and the average applicant ACT score is 27, the findings of this study may not be generalizable to the general population of college students, especially to minorities and at-risk students. Future research should focus on the impact of the flipped classroom to these groups. In addition, researchers should look into the possible influence of learner autonomy and motivation on student success in other blended learning environments in higher education. A measure of learner autonomy with better psychometric properties should also be developed, using the construct in this study as a starting point. Future research should further pursue the impact of gender and confidence in learning statistics on student performance and course satisfaction.

The results from our study are encouraging when it comes to maximizing the potential of the flipped classroom to positively impact student performance and course satisfaction, and to allow teaching load to be spread between full-time faculty and undergraduate teaching assistants.

REFERENCES

Abeysekera, L., & Dawson, P. (2015). Motivation and cognitive load in the flipped classroom: Definition, rationale and a call for research. *Higher Education Research. & Development*, *34*(1), 1–14.

Akyol, Z., & Garrison, D. R. (2011). Understanding cognitive presence in an online and blended community of inquiry: Assessing outcomes and processes for deep approaches to learning. *British Journal of Educational Technology*, *42*(2), 233–250.

Bandura, A. (1986). The explanatory and predictive scope of self-efficacy theory. *Journal of Social and Clinical Psychology*, *4*(3), 359–373.

Bean, J., & Metzner, B. (1985). A conceptual model of nontraditional undergraduatestudent attrition. *Review of Educational Research*, *55*, 485–650.

Bishop, J. L., & Verleger, M. A. (2013, June). *The flipped classroom: A survey of the research*. Paper presented at the 120th American Society for Engineering Education Annual Conference & Exposition, Atlanta, Georgia, USA. (Vol. 30, No. 9, pp.1–18).
[Online: https://www.asee.org/public/conferences/20/papers/6219/download]

Brame, C. J. (2014). *Flipping the classroom*. Center for Teaching, Vanderbilt University.
[Online: http://cft.vanderbilt.edu/guides-sub-pages/flipping-the-classroom/]

Butt, A. (2014). Student views on the use of a flipped classroom approach: Evidence from Australia. *Business Education and Accreditation*, *6*(1), 33–43.

Cohen, J. (1969). *Statistical power analysis for the behavioral sciences*. New York: Academic Press.

Crouch, C. H., & Mazur, E. (2001). Peer instruction: Ten years of experience and results. *American Journal of Physics, 69*, 970–977. doi: http://dx.doi.org/10.1119/1.1374249

Davies, R. S., Dean, D. L., & Ball, N. (2013). Flipping the classroom and instructional technology integration in a college-level information systems spreadsheet course. *Educational Technology Research and Development, 61*(4), 563–580. doi: 10.1007/s11423-013-9305-6

Deslauriers, L., Schelew, E., & Wieman, C. (2011). Improved learning in a large-enrollment physics class. *Science, 332*, 862–864. doi: 10.1126/science.1201783

Dove, A. (2013). Students' perceptions of learning in a flipped statistics class. In R. McBride & M. Searson (Eds.), *Proceedings of Society for Information Technology & Teacher Education International Conference 2013* (pp. 393–398). Chesapeake, VA: Association for the Advancement of Computing in Education (AACE). [Online: https://www.learntechlib.org/p/48133]

Duarte, M. (2013). Blended learning and learner autonomy in higher education: A study with mechanical engineering students. *Proceedings of the 7th International Technology, Education and Development Conference* (pp. 3205–3208). Valencia, Spain: IATED.

Dubinsky, E., & McDonald, M. A. (2001). APOS: A constructivist theory of learning in undergraduate mathematics education research. In D. Holton (Ed.), *The teaching and learning of mathematics at university level* (pp. 275–282). Netherlands: Springer

Finney, S. J., & Schraw, G. (2003). Self-efficacy beliefs in college statistics courses. *Contemporary Educational Psychology*, *28*(2), 161–186.

Garfield, J., & Ben-Zvi, D. (2007). How students learn statistics revisited: A current review of research on teaching and learning statistics. *International Statistical Review*, *75*(3), 372–396.

Graham, C. R. (2006). Blended learning systems: Definition, current trends, and future directions. In C. J. Bonk & C. R. Graham (Eds.), *The handbook of blended learning: Global perspectives, local designs* (pp. 3−21). San Francisco, CA: Pfeiffer.

Ginsburg, H. P., & Opper, S. (1988). *Piaget's theory of intellectual development*. Englewood Cliffs, NJ: Prentice-Hall, Inc.

Hamdan, N., McKnight, P., McKnight, K., & Arfstrom, K. M. (2013). *The flipped learning model: A white paper based on the literature review titled a review of flipped learning*. Flipped Learning Network/Pearson/George Mason University.

Hawkins, A., Graham, C. R., Sudweeks, R. R., & Barbour, M. K. (2013). Academic performance, course completion rates, and student perception of the quality and frequency of interaction in a virtual high school. *Distance Education*, *34*(1), 64–83.

Honeycutt, B., & Garrett, J. (2013). The flipped approach to a learner-centered class. *A Magna Publications White Paper*. doi: 10.1080/10875301.2014.975307

Huang, H. M. (2002). Toward constructivism for adult learners in online learning environments. *British Journal of Educational Technology*, *33*(1), 27–37.

Hughes, G. (2007). Using blended learning to increase learner support and improve retention. *Teaching in Higher Education, 12*, 349–363.

Jensen, J. L., Kummer, T. A., & Godoy, P. D. D. M. (2015). Improvements from a flipped classroom may simply be the fruits of active learning. *CBE-Life Sciences Education, 14*(1), 1–12. doi: 10.1187/cbe.14-08-0129

Kerr, M. S., Rynearson, K., & Kerr, M. C. (2006). Student characteristics for online learning success. *The Internet and Higher Education*, *9*(2), 91–105.

Kintu, J. M., Zhu, C., & Kagambe, E. (2017). Blended learning effectiveness: The relationship between student characteristics, design features and outcomes. *International Journal of Educational Technology in Higher Education, 14*(7), 2–20. doi: 10.1186/s41239-017-0043-4

Kohli, A. S., Peng, C., & Mittal, P. (2011). Predictors of student success in undergraduate business statistics course. *Journal of the Academy of Business & Economics*, *11*(4), 32–42.

Lage, M. J., Platt, G. J., & Treglia, M. (2000). Inverting the classroom: A gateway to creating an inclusive learning environment. *The Journal of Economic Education, 31*, 30–43.

Lim, D. H., & Kim, H. (2003). Motivation and learner characteristics affecting online learning and learning application. *Journal of Educational Technology Systems*, *31*(4), 423–439.

Lim, D. H., & Morris, M. L. (2009). Learner and instructional factors influencing learning outcomes within a blended learning environment. *Journal of Educational Technology & Society*, *12*(4), 282.

Macaskill, A., & Taylor, E. (2010). The development of a brief measure of learner autonomy in university students. *Studies in Higher Education*, *35*(3), 351–359.

Mason, G. S., Shuman, T. R., & Cook, K. E. (2013). Comparing the effectiveness of an inverted classroom to a traditional classroom in an upper-division engineering course. *IEEE Transactions on Education*, *56*(4), 430–435.

McLaughlin, J. E., Griffin, L. M., Esserman, D. A., Davidson, C. A., Glatt, D. M., Roth, M. T., Gharkholonarehe, N., & Mumper, R. J. (2013). The flipped satellite classroom: Student engagement, performance, and perception. *The American Journal of Pharmaceutical Education, 77*(9), 196. doi: 10.5688/ajpe779196

Missildine, K., Fountain, L., & Gosselin, K. (2013). Flipping the classroom to improve student performance and satisfaction. *Journal of Nursing Education, 52*(10), 597–599. doi: 10.3928/01484834-20130919-03

Moravec, M., Williams, A., Aguilar-Roca, N., & O'Dowd, D. K. (2010). Learn before lecture: A strategy that improves learning outcomes in a large introductory biology class. *Life Sciences Education, 9*(4), 473–481. doi: 10.1187/cbe.10-04-0063

Muthen, B. O., & Satorra A. (1995). Complex sample data in structural equation modeling. *Sociological Methodology, 25*, 267–316. doi: 10.2307/271070

New Media Consortium. (2014). *NMC Horizon Report 2014: Higher Education Edition*. Austin, TX: The New Median Consortium. [Online: https://www.nmc.org/publication/nmc-horizon-report-2014-higher-education-edition/]

O'Flaherty J., & Phillips, C. (2015). The use of flipped classrooms in higher education. *Internet and Higher Education, 25*, 85–95. doi: 10.1016/j.iheduc.2015.02.002.

Peterson, D. J. (2016). The flipped classroom improves student achievement and course satisfaction in a statistics course: A quasi-experimental study. *Teaching of Psychology, 43*(1), 10–15. doi: 10.1177/0098628315620063

Pierce, R., & Fox, J. (2012). Vodcasts and active-learning exercises in a "flipped classroom" model of a renal pharmacotherapy module. *American Journal of Pharmaceutical Education, 76*(10), 1–5. doi: 10.5688/ajpe7610196

Prunuske, A. J., Batzli, J., Howell, E., & Miller, S. (2012). Using online lectures to make time for active learning. *Genetics*, *192*(1), 67–72.

Richardson, J. T. (2011). Eta squared and partial eta squared as measures of effect size in educational research. *Educational Research Review*, *6*(2), 135–147.

Shea, P., & Bidjerano, T. (2010). Learning presence: Towards a theory of self-efficacy, self-regulation, and the development of a communities of inquiry in online and blended learning environments. *Computers & Education*, *55*(4), 1721–1731.

Strayer, J. F. (2012). How learning in an inverted classroom influences cooperation, innovation, and task orientation. *Learning Environments Research, 15*(2), 171–193. doi: 10.1007/s10984-012-9108-4

Tune, J. D., Sturek, M., & Basile, D. P. (2013). Flipped classroom model improves graduate student performance in cardiovascular, respiratory, and renal physiology. *Advances in Physiology Education, 37*, 316–320.

Vella, E. J., Turesky, E. F., & Hebert, J. (2016). Predictors of academic success in web-based courses: Age, GPA, and instruction mode. *Quality Assurance in Education*, *24*(4), 586–600.

Vidic, N. S., & Clark, R. M. (2016). Comparison of a partially flipped vs. fully-flipped introductory probability and statistics course for engineers: Lessons learned. In *Proceedings of the 2016 ASEE Annual Conference and Exposition.* New Orleans, LA. American Society for Engineering Education. [Online: https://peer.asee.org/26529]

Wilson, S. G. (2013). The flipped class: A method to address the challenges of an undergraduate statistics course. *Teaching of Psychology, 40*(3), 193–199. doi: 10.1177/0098628313487461

Winquist, J. R. & Carlson, K, A. (2014). Flipped statistics class results: Better performance than lecture over one year later. *Journal of Statistics Education, 22*(3), 1–10. doi: 10.1080/10691898.2016.1158017

Worthington, R. L., & Whittaker, T. A. (2006). Scale development research: A content analysis and recommendations for best practices. *The Counseling Psychologist*, *34*(6), 806–838.

PERPETUA LYNNE NIELSEN
223B TMCB
Brigham Young University
Provo, UT 84602