Components of a Flipped Classroom Influencing Student Success in an Undergraduate Business Statistics Course

Lee Shinaberger
Department of Management and Decision Sciences, Coastal Carolina University, Conway, SC

ABSTRACT
An instructor transformed an undergraduate business statistics course over 10 semesters from a traditional lecture course to a flipped classroom course. The researcher used a linear mixed model to explore the effectiveness of the evolution on student success as measured by exam performance. The results provide guidance to successfully implement a flipped classroom course. The largest improvements were achieved by replacing face-to-face lecture with active learning exercises and using quizzes to verify student engagement with offline materials. Using conditional release of course materials to encourage homework completion also provided a significant benefit to students who missed class often.

1. Introduction
The study began as a need to improve student performance in an undergraduate business statistics course. The instructor was concerned with the number of students who failed to successfully complete the course. Compared to other courses at the university, a large percentage of students in this course withdrew prior to completion or did not earn a passing grade. The researcher used reflection to generate ideas to improve student performance within the course at the end of each semester and modified the course over time. This study explored the effect of three types of changes to the course on student performance: formative assessment, active learning, and flipped classroom.

One method identified to improve the course was increased emphasis on formative assessment. Assessment tasks have been shown to improve exam performance (Van Gaal and De Ridder 2013). Students who did not receive feedback on formative assessment tasks were unable to gauge their ability to apply new knowledge (Boston 2002). Extending Boston’s (2002) logic, unprepared students may overestimate their ability to demonstrate new knowledge on an exam.

The instructor identified active learning as a second possible method for improving student performance. Active learning is a broad term that includes discussion, case study, problem-based learning, and many other teaching methods. For the purposes of this article, active learning is considered any teaching method encouraging students to interact with their peers, the instructor, or the environment. Extensive research has shown various active learning activities can improve student performance (Garfield 1993; Jensen 2000; Karpiak 2011; Svinicki and McKeachie 2011).

A final change to the course was made in the form of the flipped, or inverted, classroom. Scholars have studied many uses of the flipped classroom. Some studies found positive results (Day and Foley 2006; Wilson 2013; Winquist and Carlson 2014; Touchton 2015), whereas others did not find such a clear distinction in performance in the flipped versus traditional classroom (Houghton and Kelly 2015). Winquist and Carlson’s (2014) work was especially interesting because it indicates advantages in long-term retrieval of information learned in a flipped classroom. Other research on the flipped classroom concentrates on student perceptions and/or comparing the flipped classroom to traditional lecture (Lage, Platt, and Treglia 2000; Day and Foley 2006; Wilson 2013; Houghton and Kelly 2015; Touchton 2015).

There is a lack of quantitative analysis of the components of a flipped classroom in relationship to their effect on student performance. This article extends the research by exploring the factors that make a flipped classroom effective. Specifically, the article examines the effectiveness of collecting homework and providing feedback, assigning case studies, using conditional release of course materials (related to homework completion), placing lecture content online while using active learning techniques in class, and adding quizzes to verify students’ use of online video lectures. These are a few potential components to a flipped classroom. They are not a required list of components or an exhaustive list of possible elements. The goal is to provide guidance to instructors who wish to implement a flipped classroom and a foundation for future research on the effectiveness of pedagogical components.

KEYWORDS
Active learning; Adaptive release; Classroom strategies; Effective teaching; Inverted class
2. Method

2.1. Characteristics of the Participants

The study participants were primarily undergraduate business majors at Coastal Carolina University, a public, 4-year, independent university. The Coastal Carolina University Office of Institutional Research, Assessment and Analysis provides demographic information for the university. Over the course of the study, the university grew from 8500 undergraduate students in 2011 to 9600 undergraduates in 2015 (Coastal Carolina University Office of Institutional Research, Assessment and Analysis 2015a, b). Within the college of business, students are 34.6% female and 65.4% male. The college’s most prominent ethnic groups are White/European American/non-Hispanic (72.7%) and Black/African American (15.9%). International students comprise 3.0% of the college. Although specific demographics of the participants are not available, they are expected to closely mirror the college of business. The average SAT score of incoming college of business freshmen was 1010 in 2015. An analysis of university SAT scores indicated no significant change over the period of the study.

The course CBAD 291 Business Statistics satisfies part of the university’s liberal arts core; however, similar courses in other departments attract students from other colleges. Approximately 98% of the students in Business Statistics are from the E. Craig Wall Sr. College of Business Administration (Coastal Carolina University Office of Institutional Research, Assessment and Analysis 2015a). The course is typically taken during students’ sophomore year; however, students of all levels have enrolled in the course. A total of 1103 students participated in the study.

Class size is limited to 40 students in the instructor’s course sections. All of the studied sections met through traditional face-to-face delivery.

Required course content includes measures of central tendency and repeatability, normal probability distributions, the central limit theorem, confidence intervals of means and proportions, single sample hypothesis testing of means, correlation, and linear regression. The instructor also includes single and multiple event probability, expected value, single sample hypothesis testing of proportions, analysis of variance, chi-squared applications, and multiple regression. The textbook used was Lind, Marchal, and Wathen (2013). Prior to the Spring 2013 semester, the seventh edition of the textbook was used.

2.2. Initial Course Structure

The course began as a lecture-intensive course. Lecture works well for directing students to areas of emphasis but is less effective than other teaching methods in learning goals such as problem solving and application of knowledge in new situations (Svinicki and McKeachie 2011). The lecture format was used for approximately 85% of the class meetings. The remaining 15% of class time was spent working on a group project. The instructor assigned homework but did not collect it. Students completed exams outside of the classroom. Using timed exams on the learning management system (LMS) resulted in four additional class periods to cover material. This advantage was offset by a loss of control of the testing environment. This format had been used since the spring semester of 2005 prior to the beginning of the course transformation in Spring 2012.

2.3. Course Changes Overview

The data collected were a convenience sample of past participants of business statistics courses. Because the course transformed from traditional lecture to a flipped classroom with active learning content over time, the researcher was able to analyze the effect of each change on student performance.

Pedagogical changes occurred only between semesters. One change was made per semester with two exceptions. During Spring 2013, the content areas for Exams 3 and 4 were rearranged and conditional release of homework was introduced. During Spring 2015, the group project was removed and video quizzes were added. A summary of changes is included in Table 1.

The actual changes are grouped by category for discussion. The categories include formative assessment, active learning and flipped classroom, and other changes.

2.4. Providing Formative Assessment Opportunities

At the end of the Fall 2011 semester, the instructor reflected on student responses to discussion of their progress in the course. Students were often heard discussing that they felt prepared for exams but were unable to perform on them. The instructor identified the need to provide feedback to students on their progress in learning the material through formative assessment. Instead of assigning but not collecting homework problems from the textbook, homework questions were added to the Blackboard Learn (and later Moodle) LMS. Using an LMS allowed automatic grading of the students’ responses so that there was no delay in providing feedback. Students could take the homework multiple times with no penalty. The LMS homework employed question sets so that students would see similar problems but not the exact same problems. If a student got an answer incorrect, the instructor added feedback to the LMS that would provide guidance on how to approach the problem. Questions from

Table 1. Timeline of course changes.

| Semester     | Schedule                  | Course changes                                      |
|--------------|---------------------------|-----------------------------------------------------|
| Fall 2011    | Tu-Th format, Evening     | Lecture Format: Homework Assigned, not Collected     |
| Spring 2012  | M-W-F and Tu-Th format,   | Exams Outside of Class                              |
|              | Morning and Evening       | Homework Collected With Feedback                    |
| Fall 2012    | M-W-F and Tu-Th format,   | Case Studies Added                                  |
|              | Morning                   |                                                     |
| Spring 2013  | M-W-F and Tu-Th format,   | Conditional Release of Course Materials Exam 3b      |
|              | Morning                   | Replaced Exam 3a Exam 4b                             |
|              |                           | Replaced Exam 4a                                     |
| Fall 2013    | M-W-F and Tu-Th format,   | In-Class Exams                                      |
|              | Morning                   |                                                     |
| Spring 2014  | M-W-F format, Morning     | (No Changes)                                        |
| Fall 2014    | M-W-F format, Morning     | Flipped Classroom Replaced                           |
|              |                           | Lecture                                             |
| Spring 2015  | M-W-F format, Morning     | Preclass Quizzes Added Group Project Removed         |
| Fall 2015,   | M-W-F format, Morning     | (No Changes)                                        |
| Spring 2016  |                           |                                                     |
the same random block had similar difficulty, but the overall composition of the questions varied from easy to challenging. Students who did not complete at least 11 of 13 homework sets with a score of at least 70% received a grade penalty. Many students waited until just prior to an exam to begin the homework. This prompted a subsequent change in homework policy.

When the instructor examined the homework submissions, he noticed a large number of students did not begin the homework until a few days prior to the exam. This was interpreted as academic procrastination, also known as cramming. Although cramming has been shown to be just as effective for exam performance in the short term (Kerdijk et al. 2015), spacing multiple study sessions separated by hours or days has been shown to be a more effective strategy for long-term retrieval (Carpenter et al. 2012).

To address academic procrastination, the grade penalty for not completing the homework was removed. LMS settings prevented students from accessing course material until they scored at least a 70% on the homework assigned from the previous section. If an exam covered chapters 5, 6, and 7, students would have to score 70% on the chapter 5 homework to unlock the chapter 6 materials and homework. To unlock the exam, the student would be required to achieve at least a 70% on the homework from chapters 5, 6, and 7. The LMS settings permitted students to retake the homework as often as they wanted without penalty. The mechanism for accomplishing this is conditional release, referred to as adaptive release in Blackboard Learn and access restriction in Moodle.

A policy was created in regard to students’ failing to unlock the exam. Students taking all four exams were permitted to retake up to two exams during the final exam period to attempt to improve their grade. Students missing an exam for a reason approved by university policy would be scheduled for a makeup exam as quickly as possible and were still permitted to retake exams during the final exam period. Students missing exams without an approved absence forfeited the right to retake exams and were only able to take the missed exam during finals week.

Formative assessment was also added through case studies. Students were given a relatively complex problem to work on in class. Informal groups of students were encouraged to work on the problem together. The overall problem was broken down into 15 questions providing structure to guide students’ discovery. These problems were typically much more difficult than what would be encountered on an exam. They were designed to teach critical thinking, problem-solving skills, and collaboration (Garfield 1993). The case studies helped expose knowledge gaps in students’ preparation. Case studies could be completed outside of class.

By requiring formative assessment assignments, students could accurately gauge their progress in the course prior to exams. Homework scores did not directly count toward students’ final grades but became necessary to release additional learning content and exams. Case studies did contribute directly to students’ final grade.

2.5. Active Learning and Flipping the Classroom

Although the changes to formative assessment helped students become more aware of their capability prior to exams, students continued to struggle when asked to apply their knowledge. Students often stated the problems appeared easy when the instructor demonstrated them but were more difficult when working on problems outside of the classroom environment. A method had to be developed to allow students to practice problems and quickly receive feedback.

Scholarship of Teaching and Learning training and conferences led to the instructor’s desire to implement active learning in the classroom to increase student success. Active learning exercises were demonstrated during Coastal Carolina University’s New Faculty Orientation program (Keiner 2012). Scholarship of Teaching and Learning presentations (Adkin 2013; Eison 2013; Kramer 2013; Christensen 2014) demonstrated the success of a flipped classroom strategy. Christensen (2014) warned that one issue with flipping the classroom was finding enough active learning exercises to fill the time normally dedicated to lecture. These educational opportunities led the instructor to the research of Vygotsky (1978) on scaffolding and the zone of proximal development, King (1993) on active learning, and Lage, Platt, and Treglia (2000) on the flipped, or inverted, classroom. Vygotsky (1978) advocated providing expert guidance to build student learning from what is known to what is unknown. King (1993, p. 30) advocated students taking an active involvement in their learning to “produce knowledge rather than reproduce it.” This encourages higher level thinking skills as defined by the revised Bloom’s taxonomy (Anderson, Krathwohl, and Bloom 2001). The work of Lage, Platt, and Treglia (2000) indicated flipping the classroom could result in improvements in student perception of a course, reinforcing the positive effects seen by the Scholarship of Teaching and Learning presentations.

The research indicated active learning techniques could result in improvements in student performance. However, lecture consumed all of the available class time. A flipped classroom was implemented to provide class time for active learning. Students were given the topics covered in the textbook along with videos of the standard lecture. Short topic videos were created by the researcher using Camtasia and Jing software, and were presented as voice over PowerPoint and voice over Excel. The videos were broken into segments 5–10 min long. The pace of the videos was substantially faster than the pace of the same content presented in lecture. No time was spent verifying students’ understanding. The same material covered in 2 hr and 50 min over the course of a week was frequently covered in approximately 1 hr of videos. Additional videos were generated to more carefully explain difficult concepts and provide additional examples of application. These videos were in addition to the material prepared for a traditional lecture.

After the videos were placed on the LMS, 10–20 min of class time would be spent reviewing difficult material covered and the remainder of the time over multiple days was spent in active learning exercises. The most commonly used technique was a variation of think-pair-share (Lyman 1981). Students were posed a problem. They would receive an amount of time to solve the problem alone. The students were then asked to pair up with their neighbors and reconcile any differences in their solutions. Finally, a student would be selected at random to share the answer and describe how it was found.

Variations of think-pair-share were also used. For simple problems, students might raise their hands to indicate their
response between two to four multiple choice answers. This allowed students to rapidly progress from a review of previous sessions to new material at the end of class. This technique was used to follow Vygotsky’s (1978) scaffolding methods to connect new learning to prior knowledge.

At the end of the first semester in which the flipped classroom was implemented, the researcher discovered approximately one-third of the students did not watch the traditional lecture videos. Instead, these students relied upon the short topic, difficult concept, and application example videos to learn all of the content. When prompted for answers, these students were frequently unable to contribute to the learning environment. Action had to be taken to encourage students to watch the lecture videos outside of the classroom. The researcher decided to give quizzes at the beginning of class whenever a new set of material was introduced. Because the course was taught in a computer lab, the quiz could be easily given via the LMS using quizzes at the beginning of class whenever a new set of material was introduced.

### 2.6. Other Changes to the Course

Not all of the changes to the course related to formative assessment, active learning, or the use of a flipped classroom. Evening classes were populated primarily with traditional-aged students, so the classes were moved to the morning during their generally preferred class times. The instructor moved exams back into the classroom to address academic integrity concerns. Finally, after the instructor found decreased performance in classes with a Tuesday-Thursday format, the department chair moved all classes to a Monday-Wednesday-Friday format.

### 2.7. Measures

Student success was measured using exam grades. Four exams were given per semester. Exams were not cumulative, although knowledge of previous concepts was assumed. Exam scores were recorded as percentage correct from 0 to 100.

The researcher included one covariate in the study: the number of class absences for each student. The number of absences was intended to represent student motivation. Highly motivated students would have fewer absences than less motivated students. Absences were included as linear and quadratic interaction effects. Quadratic interactions were included because the effects of absences may not be strictly linear. The researcher theorized the difference between missing zero and one class was likely to be more pronounced than the difference between missing eight and nine classes.

### 3. Results

#### 3.1. Variables

The exam scores ranged from 0 to 100 with a relatively large variation. The descriptive statistics are given in Table 2. All tables include only valid exams. Some students took exams late; those exams were excluded from the study. The researcher theorized these students may have gained an advantage by speaking with students who had already taken the exam; therefore, the results of those exams were considered invalid.

The number of absences were nonnormally distributed, with skewness of 2.04 ($SD = 0.07$) and kurtosis of 6.09 ($SD = 0.15$). They ranged from 0 to 35 absences with a median of 3 ($M = 4.64, SD = 4.73$). Most students (90%) missed 10 or fewer classes out of 42 scheduled class times.

Because of the large number of factors, they are grouped by category for discussion: formative assessment, class format, classroom environment, and content.

The first group of factors describes the formative assessment in the course. Initially, the instructor assigned homework but did not collect it. Beginning in Spring 2012, homework was assigned and collected, and automatic feedback was given via the LMS. In Spring 2013, conditional release was added. This required students to complete homework with a minimum passing grade to release additional course materials such as class notes, additional homework, and exams. The final formative assessment variable indicates if case study was used in the course. Case study problems were the most complex problems assigned. The instructor encouraged students to solve the case studies collaboratively in class. The formative assessment factors are shown in Table 3 with the corresponding frequencies. The frequencies are the number of valid exams taken for each factor (Homework: Collected With Feedback, Homework: Conditional Release, and Case Studies), not the number of students enrolled.

The second group of factors describes the class format. Initially, the class format was lecture based. The instructor implemented the flipped classroom in the Fall 2014 semester. The

| Exam | Mean | SD | n | Max | Min | Skew (SD) | Kurtosis (SD) |
|------|------|----|---|-----|-----|----------|--------------|
| 1    | 83.35| 13.06| 1,063| 100 | 0   | −1.62 (0.08)| 5.71 (0.15)   |
| 2    | 66.96| 17.32| 1,029| 100 | 0   | −0.43 (0.08)| 0.28 (0.15)   |
| 3a   | 66.83| 16.11| 353 | 100 | 0   | −0.55 (0.13)| 0.85 (0.26)   |
| 3b   | 61.97| 17.84| 656 | 100 | 0   | −0.10 (0.10)| −0.10 (0.19)  |
| 4a   | 71.30| 16.20| 339 | 100 | 0   | −0.72 (0.13)| 1.14 (0.26)   |
| 4b   | 74.34| 17.63| 675 | 100 | 0   | −0.84 (0.09)| 0.59 (0.19)   |
| All  | 71.95| 17.98| 4,115| 100 | 0   | −0.65 (0.04)| 0.30 (0.08)   |

Table 2: Descriptive statistics of exam scores.

| Semester | Collect with feedback | Conditional release | Case studies |
|----------|-----------------------|---------------------|--------------|
| Fall 2011 | False | 170 | True | 170 | False | 170 |
| Spring 2012 | 601 | 601 | 625 | 625 |
| Fall 2012 | 625 | 625 | 498 | 498 |
| Spring 2013 | 498 | 394 | 394 | 394 |
| Fall 2013 | 498 | 248 | 248 | 248 |
| Spring 2014 | 403 | 279 | 279 | 279 |
| Fall 2014 | 456 | 456 | 456 | 456 |
| Spring 2015 | 441 | 441 | 441 | 441 |
| Total | 3945 | 1396 | 2719 | 771 | 3344 |

Table 3: Frequencies of formative assessment factors.
The third group of factors describes the classroom environment. These factors may affect student performance but are not of interest to the study. The primary semester to take the course is in the fall, so only 441 exams were valid. Because 1103 students participated in the study, the best possible number of valid exams would have been 4412. Therefore, only 4115 (93%) exams were valid. Students not taking any exams were not included in the study.

### 3.2. Model Selection and Fit

The factors of interest are the formative assessment variables from Table 3 (Homework: Collected with Feedback, Preclass Quizzes). The environmental factors are shown in Table 5 with the corresponding frequencies of exams taken.

A factor for the group project was not included because the removal of the project coincided with the addition of preclass quizzes. The removal of the group project should not have affected student exam performance because the project was conducted after the exam on the relevant material. For example, students began collecting data after the exam that covered sampling methods and began their analysis after the exam discussing hypothesis testing.

The final group of factors is the factors used to account for differences in exam content. The frequency of the factors used to account for content changes in the dependent variable of exam score is shown in Table 6.

The four exams given throughout the semester measured performance in distinct areas of statistics. Student exam performance in one area of statistics may not be similar to exam performance in another area of statistics. The researcher considered each exam content area an additional factor to be analyzed with dummy variables to account for each factor. If changes to the exams were relatively minor, such as exchanging a question for one of similar difficulty, the researcher did not create an additional factor to model the change in the exam. If the change was major, a new factor was assigned. Major changes occurred only before the Spring 2013 semester, when the content from Exams 3 and 4 were rearranged. Beginning in Spring 2013, linear and multiple regression were moved from exam 3 to exam 4 and single sample hypothesis tests of means and proportions were moved from exam 4 to exam 3. The instructor covered the same content before and after the change. Exams prior to Spring 2013 are referred to as Exam 3a and Exam 4a, and those given beginning in Spring 2013 and later are referred to as Exam 3b and Exam 4b in this article.

The valid exams statistic in Table 6 was calculated expecting each student to take four exams. Because 1103 students participated in the study, the best possible number of valid exams would have been 4412. Therefore, only 4115 (93%) exams were considered valid. Students not taking any exams were not included in the study.

### Table 4. Frequencies of class format factors.

| Semester     | Flipped classroom | Preclass quizzes |
|--------------|------------------|------------------|
| Fall 2011    | 170              | 170              |
| Spring 2012  | 601              | 601              |
| Fall 2012    | 625              | 625              |
| Spring 2013  | 498              | 498              |
| Fall 2013    | 394              | 394              |
| Spring 2014  | 248              | 248              |
| Fall 2014    | 403              | 403              |
| Spring 2015  | 279              | 279              |
| Fall 2015    | 456              | 456              |
| Spring 2016  | 441              | 441              |
| Total        | 2536             | 1579             |

### Table 5. Frequencies of classroom environment factors.

| Semester     | Spring | Morning | M-W-F | In-class exam |
|--------------|--------|---------|-------|---------------|
| Fall 2011    | 170    | 170     | 170   | 170           |
| Spring 2012  | 601    | 601     | 601   | 601           |
| Fall 2012    | 625    | 625     | 625   | 625           |
| Spring 2013  | 498    | 498     | 498   | 498           |
| Fall 2013    | 394    | 394     | 394   | 394           |
| Spring 2014  | 248    | 248     | 248   | 248           |
| Fall 2014    | 403    | 403     | 403   | 403           |
| Spring 2015  | 279    | 279     | 279   | 279           |
| Fall 2015    | 456    | 456     | 456   | 456           |
| Spring 2016  | 441    | 441     | 441   | 441           |
| Total        | 2048   | 2067    | 286   | 3829          |

### Table 6. Frequencies of factors to account for differences in exam content.

| Exam | 1 | 2 | 3a | 3b | 4a | 4b | Total valid exams | Number of students | Valid exam % |
|------|---|---|----|----|----|----|-------------------|-------------------|--------------|
| Fall 2011 | 42 | 44 | 43 | 41 | 170 | 44 | 97%               |
| Spring 2012 | 157 | 152 | 147 | 145 | 601 | 160 | 94%               |
| Fall 2012 | 150 | 159 | 163 | 153 | 625 | 165 | 95%               |
| Spring 2013 | 132 | 124 | 120 | 122 | 498 | 138 | 90%               |
| Fall 2013 | 106 | 97 | 89 | 102 | 394 | 109 | 90%               |
| Spring 2014 | 65 | 63 | 60 | 60 | 248 | 68 | 91%               |
| Fall 2014 | 111 | 93 | 97 | 102 | 403 | 111 | 91%               |
| Spring 2015 | 71 | 70 | 69 | 69 | 279 | 72 | 97%               |
| Fall 2015 | 116 | 117 | 112 | 111 | 456 | 117 | 97%               |
| Spring 2016 | 113 | 110 | 109 | 109 | 441 | 119 | 93%               |
| Total | 1063 | 1029 | 353 | 656 | 339 | 675 | 4115              | 1103              | 93%          |
Homework: Conditional Release, and Case Studies), the class format variables from Table 4 (Flipped Classroom and Preclass Quizzes), and interactions with the number of absences. The researcher defined small, medium, and large effect size in this model as 2.5, 5, and 10 points on the 0–100 scale used for exams. These nonstandardized effect sizes were chosen to align with the typical 10-point intervals between letter grades. They are reported as effect size = estimated marginal mean/10, so an effect size of 1 represents an increase of one letter grade. In the interest of standardization, both letter grade effect size and Cohen’s $d$ are presented. Cohen’s $d$ represents the ratio of the effect size to the pooled standard deviation, with a larger value indicating a larger relative effect size (Cohen 1988).

A linear mixed model was used to fit the data using SPSS MIXED version 21. Repeated measures with the general linear model was discarded because there were missing data for every student as they either took Exams 3a and 4a or Exams 3b and 4b. The missing data would have resulted in all subjects being discarded. Within the linear mixed model, a random intercept with restricted maximum likelihood provided the best fit using Akaike’s (1973) information criterion (AIC) and Schwarz’s (1978) Bayesian criterion (BIC). Adding random slope for each student resulted in a poorer fit, as did removing the random intercept.

The observational nature of the study limited the analysis of interactions. Because the changes were introduced over time but not removed, one combination of effects always had no samples. Consider Interventions A and B. A is introduced first, then B. No cases exist in which B has been introduced but A has not. Therefore, modeling A*B is meaningless.

Only interactions with the number of absences covariate resulted in improved model fit. Both linear and quadratic interactions of the number of absences were modeled.

During model fit, factors were eliminated stepwise by testing the removal of each remaining factor individually. The factor whose removal resulted in the lowest AIC and BIC was removed and the process repeated until elimination of any of the remaining factors failed to lower AIC and BIC. This method was used instead of step down removal of factors, beginning with those with the least significance. The homework factor had poor significance. However, removing homework or its interactions resulted in a higher AIC and BIC. Only removing interaction terms with the covariate number of absences improved fit; all of the primary factors remained in the model.

### 3.3. Estimated Marginal Means

With 12 remaining variables of interest, 5 yielded statistically significant results. The estimated marginal means for the variables of interest is shown in Table 7. The model was forced to report the estimated marginal means at zero absences rather than the average to aid in interpretation. This change in centering had no effect on model performance. Interaction effect sizes and Cohen’s $d$ are based on the median of absences.

### 4. Discussion

One limitation of the study is that it relies on a sample of convenience. The opportunities to collect additional samples did not exist. With a low sample size, particularly where homework being collected was false, risks of Type II errors were relatively high. The researcher deemed this acceptable for an exploratory study. Post hoc calculations confirmed an issue with power. A substantial risk of Type II errors exists for almost all factors, with the exceptions Flipped Classroom “Absent” and Preclass Quiz “Absent”. Absent. A summary of the power calculations is shown in Table 8.

Despite the power issues, 5 of the 12 variables of interest were found to be significant. However, no correction was made to adjust for the number of tests. A Bonferroni correction would result in only the linear and quadratic video quiz interactions being significant but would further increase the likelihood of Type II errors. The probability of making three Type I errors without the Bonferroni correction is lower than the probability of three Type II errors with the correction.

Flipped classroom, or moving the lecture outside of the classroom while concentrating on active learning example problems in the classroom, had significant main and interaction effects. With an estimated marginal mean of 5.77 on a 0 to 100 scale ($d = 0.36$), the average improvement in student exam performance was over half of a letter grade for the main effect. This is tempered by a decline of 0.75 per absence, or $-2.25$ for the median number of absences in the class ($d = 0.14$).

| Table 7. Estimated marginal means for the variables of interest. |
|---------------------------------------------------------------|
| **Factor** | **Estimated marginal mean** | **Standard error** | **df** | **Sig.** | **95% CI** | **Letter grade effect size** |
|-----------|----------------------------|-------------------|-------|---------|-----------|-----------------------------|
| Intercept | 71.85 | 1.499 | 1546 | 0.000 | 74.68 | 80.56 |
| Homework Collected | 2.33 | 4.718 | 1012 | 0.622 | $-6.93$ | 11.59 |
| Conditional Release | $-0.46$ | 2.403 | 1191 | 0.850 | $-5.17$ | 4.26 |
| Case Studies | 2.89 | 2.244 | 1038 | 0.199 | $-1.52$ | 7.29 |
| Flipped | 5.77 | 2.240 | 1065 | 0.010 | $-1.37$ | 10.17 |
| Preclass Quizzes | $-2.68$ | 2.391 | 1042 | 0.262 | $-2.01$ | 7.37 |
| Absent | $-0.60$ | 0.334 | 1296 | 0.074 | $-1.25$ | 0.06 |
| Homework Collected “Absent” | 2.24 | 5.746 | 1035 | 0.696 | $-9.03$ | 13.52 |
| Conditional release “Absent” | 0.70 | 0.357 | 1518 | 0.050 | 0.00 | 1.40 |
| Case Studies “Absent” | 0.20 | 0.644 | 1152 | 0.759 | $-1.07$ | 1.46 |
| Flipped “Absent” | $-0.75$ | 0.322 | 1186 | 0.020 | $-1.38$ | $-0.12$ |
| Preclass Quizzes “Absent” | 0.03 | 0.578 | 1143 | 0.970 | $-0.56$ | 2.83 |
| Homework Coll. “Absent” | 0.66 | 1.584 | 1028 | 0.677 | $-3.77$ | 2.45 |
| Preclass Quizzes “Absent” | $-0.08$ | 0.027 | 1292 | 0.004 | $-0.13$ | $-0.02$ |

Letter grade effect size: Small 0.10, Medium 0.25, Large 0.40.
The combination of the two effects is illustrated in Figure 1. Students with fewer absences benefited more from a flipped classroom and students with eight or more absences performed worse with a flipped classroom than with traditional lecture. Strong evidence indicates the flipped classroom resulted in improved student exam performance \((p = 0.010)\) and fairly strong evidence \((p = 0.020)\) implied the exam performance improvements declined as absences increased.

Using preclass quizzes to verify students accessed the lecture videos had significant linear and quadratic effects with absences but without a main effect \((p = 0.262)\). Strong evidence suggested a significant effect with the number of absences \((p = 0.003)\). The coefficient of 1.70 suggests preclass quizzes helped improve the grades of less motivated students. Strong evidence \((p = 0.004)\) of a quadratic interaction was also present at a small effect size. The quadratic interaction is \(-0.08\); missing one class is likely to have more effect on a student with a small number of absences than on a student with a large number of absences. The combination of the linear and quadratic interaction is shown in Figure 1. The intercept of 0 reflects the lack of a main effect. After 11 absences, the quadratic interaction begins to overcome the linear interaction and the expected benefit begins to decline. This curve suggests highly motivated students with few absences did not benefit from the use of the quizzes because they were already watching the lecture videos. Students with more absences were more likely to watch the lecture video when motivated by the quiz. The estimated improvement for a student with three absences is an increase of 4.38 points \((d = 0.27)\) in exam performance using a 0 to 100 scale, calculated by \(1.70(3) - 0.08(3^2)\).

The final significant estimated marginal mean is conditional release of course materials interacting with the number of absences. Because the researcher found no evidence of a main effect on performance \((p = 0.850)\), this again indicates conditional release only affects less motivated students. Fairly strong evidence of a significant interaction effect \((p = 0.050)\) was found. The effect size was less than that of the preclass quizzes at 0.70 per absence, or 2.10 points out of 100 for a student with the median of three absences \((d = 0.13)\). Although this was significant, the letter grade effect size at the median number of absences is below the threshold for small. Figure 1 illustrates the benefit of conditional release. The conditional release factor addresses academic procrastination similarly to the preclass quizzes. However, preclass quizzes required students to watch lecture videos whereas conditional release required students to complete homework assignments.

There was no evidence of a main effect, linear interaction with absences, or quadratic interaction with absences for collecting homework \((p = 0.622, 0.696, \text{and } 0.677)\). Neither was evidence present of a main effect or linear interaction with absences for the use of case studies \((p = 0.199 \text{ and } 0.759)\). However, the significant effect for conditional release discussed in the preceding paragraph required the collection of homework. Case studies were originally added to the class to teach critical thinking, problem solving, and collaboration. The collaborative nature of the problems may have undermined the critical-thinking and problem-solving aspects of the assignments as deficient students may rely upon prepared students to determine the correct answers. Any benefit from the case studies may be overshadowed by other formative assessments, such as the conditional release of course materials, which required completion of homework.

Figure 2 shows the cumulative results of the significant effects. Student exam performance increases overall with each significant change to the course. Traditional lecture shows the expected performance prior to any course changes and is calculated from the intercept and absent coefficients from Table 7. The first significant course change introduced was conditional release of course materials, which modified the slope of the predicted score. Conditional release appears to have removed the expected penalty of missing class. Figure 2 shows a slight positive slope to the conditional release line, which is unlikely. The slope is intuitively expected to be negative or flat. Because Figure 2 shows the predicted scores based on adding effects together, error is also added. This may explain the positive slope. A negative slope is possible within the 95% confidence intervals of the coefficients.

The second significant change introduced was the conversion to a flipped classroom. Moving from the conditional release curve in Figure 2, the intercept moves up when the flipped classroom is added. This illustrates the improvement for students with no absences whereas the negative slope shows the penalty to students that miss class. The predicted

---

**Figure 1.** Change in exam performance of significant effects.

**Figure 2.** Cumulative effects on predicted exam performance.
performance of students with eight or more absences drops below the conditional release line.

The final curve in Figure 2 is the addition of the preclass quizzes. The lack of a main effect means the intercept remains the same as the flipped classroom. The linear effect of the preclass quizzes yields another positive slope. However, the negative quadratic effect flattens the curve over the area of interest. Because the curves in Figure 2 were found by adding the coefficients of the significant effects, stacking errors may explain the unexpected shapes.

The significant effects of flipping the classroom, using quizzes to verify that lecture videos were watched, and conditional release of course material were designed to increase student interaction with the course material.

The use of quizzes to encourage students to watch the lecture videos had the next largest effect. Although the quizzes did not have a significant effect on students with no absences, there was an increasing positive effect as the number of class absences increased. The quizzes were added in response to a survey indicating about one-third of students did not watch the videos. The benefits were likely to be seen only in those students who were unmotivated to watch the videos. Students found ways to bypass the quizzes’ intended result as the researcher witnessed students studying other students’ notes prior to class and observed one student watching the videos by skipping through them without sound and writing down the words on the screen.

The smallest statistically significant effect was found in the use of conditional release of course materials. The instructor added this change with the assistance of the LMS. The short topic, difficult concept, and application example videos; quizzes; and exams were locked in a manner similar to a video game. Students had to unlock the content by achieving a passing score on the homework. The benefit was significant only as an interaction with the number of absences and was below the small effect size threshold (i.e., less than a quarter of a letter grade). After adding the lecture videos and quizzes, students found ways to complete an assignment without achieving its benefits. An examination of some students’ homework patterns in the LMS revealed that some students would take the homework repeatedly in rapid succession. A 20-question homework set might be taken 10 or more times in an hour, as if the students were attempting to reveal all of the questions in a question bank and memorize the answers.

When implementing conditional release, the instructor was concerned about the possibility of a decline in ratings on course evaluations. The change was only made after securing the support of senior faculty and the department chair. In the end, course evaluations were consistent with results prior to conditional release. Some students admitted liking the requirement because they otherwise would not have completed the homework.

The transformation of the Business Statistics course was initiated because a high percentage of students was not successfully passing the course. The changes to the course did result in improved performance on exams. However, the overall pass rate of the instructor’s class did not increase. There are several possible reasons the successful completion rate did not increase, including lower student performance on other assignments. A detailed discussion is outside the scope of this article. This unexpected result will be the subject of future research.

A discussion of limitations of the study must begin with the power of the study and risk of Type I and Type II errors. No corrections were made for testing multiple factors because the low power meant the risk of Type II errors was already high. Only one instructor was included in the study; individual teaching style may affect the impact of one intervention over another. In general, the results provide guidance to other instructors interested in implementing a flipped classroom.

5. Conclusions

Implementing a flipped classroom style of teaching in a business statistics class resulted in a significant increase in student performance on exams. Students who missed class typically realized larger gains in exam performance. Significant gains were found by moving the lectures to the LMS and implementing active learning exercises in class, the use of quizzes to encourage students to watch lecture videos, and limiting access to course materials until a sufficient grade on homework was achieved. No significant gains were found by collecting and grading homework without conditional release or by assigning case study problems in a group setting.

This research is beneficial to instructors interested in implementing a flipped classroom format in their own course. It adds to the scholarship of teaching and learning by analyzing which pedagogical components improved student exam performance and which components did not result in a significant improvement.

The study raises some questions that merit additional research. Given the power issues discussed previously, additional testing could be conducted to verify the components without significant effects are truly insignificant and not the result of Type II error. This would also permit confirmation that the significant effects found in the study were not due to the increased familywise error rate due to multiple comparisons. A research design that permits analysis of interaction effects of the flipped classroom components would be quite useful. Additional research may also include multiple instructors and courses other than Business Statistics.

Acknowledgments

Dr. Lindsey Bell and Dr. Keshav Jagannathan provided review and critical analysis of this article prior to submission. Dr. Robert Sheehan provided guidance on the limitations of ordinary least-square models that led the researcher to use a maximum likelihood model. We thank colleagues in the Coastal Carolina University Center for Teaching Excellence to Advance Learning’s Writing Circle program.
References

Adkin, T. (2013), “Flipping the Classroom,” Presentation at 2013 Lilly Conference on College and University Teaching, Greensboro, NC.

Akaike, H. (1973), “Information Theory and an Extension of Maximum Likelihood Principle,” in Second International Symposium on Information Theory, pp. 267–281.

Anderson, L. W., Krathwohl, D. R., and Bloom, B. S. (2001), A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom’s Taxonomy of Educational Objectives, New York: Addison Wesley Longman.

Boston, C. (2002), “The Concept of Formative Assessment,” Practical Assessment, Research & Evaluation [online], 8(9), 1–4. Available at http://pareonline.net

Carpenter, S. K., Cepeda, N. J., Rohrer, D., Kang, S. H. K., and Pashler, H. (2014), “Using Spacing to Enhance Diverse Forms of Learning: Review of Recent Research and Implications for Instruction,” Educational Psychology Review, 24(3), 369–378.

Christensen, E. (2013), “What is Effective Teaching?” Presentation at Coastal Carolina University New Faculty Orientation, Conway, SC.

Christensen, E. (2014), “Transforming the Learning Experience Through Flipped Learning,” Presentation at the 10th Annual SACS-COC Institute on Quality Enhancement and Accreditation, New Orleans, LA.

Coastal Carolina University Office of Institutional Research, Assessment and Analysis (2015a), Coastal Carolina University Factbook: E. Craig Wall Sr. College of Business Administration, Fall 2015 [online], Conway, SC: Coastal Carolina University. Available at https://www.ccaschal.edu/media/administration/institutionalresearch/pdf/factbooks/entireuniversity/fb2015/Overall_UG.pdf

Coastal Carolina University Factbook: Fall 2015 [online], Conway, SC: Coastal Carolina University. Available at https://www.coastal.edu/media/administration/institutionalresearch/pdf/factbooks/entireuniversity/fb2015/pdf

Cohen, J. (1988), Statistical Power Analysis for the Behavioral Sciences (2nd ed.), Hillsdale, NJ: Erlbaum.

Day, J. A., and Foley, J. D. (2006), “Evaluating a Web Lecture Intervention in a Human-Computer Interaction Course,” IEEE Transactions on Education, 49(4), 420–431.

Eison, J. (2013), “Flipping: Changing Paradigms and Classroom Practices,” Presentation at 2013 Lilly Conference on College and University Teaching, Greensboro, NC.

Garfield, J. (1993), “Teaching Statistics Using Small-Group Cooperative Learning,” Journal of Statistics Education [online], 1, 1. Available at http://www.amstat.org/publications/jse/v1n1/garfield.html

Haughton, J., and Kelly, A. (2015), “Student Performance in an Introductory Business Statistics Course: Does Delivery Mode Matter?” Journal of Education for Business, 90(1), 31–43.

Jensen, E. (2000), Brain-Based Learning: The New Science of Teaching and Training (Rev. ed.), Thousand Oaks, CA: Corwin Press.

Karpiai, C. P. (2011), “Assessment of Problem-Based Learning in the Undergraduate Statistics Course,” Teaching of Psychology, 38(4), 251–254.

Keiner, L. (2012), “What is Effective Teaching?” Presentation at Coastal Carolina University New Faculty Orientation, Conway, SC.

Kerdijk, W., Cohen-Shotanus, J., Mulder, B. F., Muntinghe, F. L. H., and Tio, R. A. (2015), “Cumulative Versus End-of-Course Assessment: Effects on Self-Study Time and Test Performance,” Medical Education, 49(7), 709–716.

King, A. (1993), “From Sage on the Stage to Guide on the Side,” College Teaching, 41(1), 30–35.

Kramer, S. (2013), “Should I Flip My Class? Why and How?” Presentation at 2013 Lilly Conference on College and University Teaching, Bethesda, MD.

Lage, M., Platt, G., and Treglia, M. (2000), “Inverting the Classroom: A Gateway to Creating an Inclusive Learning Environment,” The Journal of Economic Education, 31(1), 30–43.

Lind, D. A., Marchal, W. G., and Wathen, S. A. (2013), Basic Statistics for Business and Economics (8th ed.), New York: McGraw-Hill Irwin.

Lyman, F. T. (1981), “The Responsive Class Discussion: The Inclusion of All Students,” Mainstreaming Digest [online], 3, 109–113. Available at https://www.sciencedopen.com/document?vid=347f5de1-c4d9-46e9-a869-fc37cf19383b

Schwarz, G. (1978), “Estimating the Dimension of a Model,” The Annals of Statistics, 6, 461–464.

Svinicki, M., and McKeachie, W. J. (2011), McKeachie’s Teaching Tips: Strategies, Research, and Theory for College and University Teachers (13th ed.), Belmont, CA: Wadsworth, Cengage Learning.

Touchton, M. (2015), “Flipping the Classroom and Student Performance in Advanced Statistics: Evidence from a Quasi-Experiment,” Journal of Political Science Education, 11(1), 28–44.

Van Gaal, F., and De Ridder, A. (2013), “The Impact of Assessment Tasks on Subsequent Examination Performance,” Active Learning in Higher Education, 14, 213–225.

Vygotsky, L. S. (1978), Mind in Society: The Development of Higher Psychological Processes, Cambridge, MA: Harvard University Press.

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.

Winquist, J., and Carlson, K. A. (2014), “Flipped Statistics Class Results: Better Performance Than Lecture Over One Year Later,” Journal of Statistics Education [online], 22(3), 1–10. Available at http://www.amstat.org/publications/jse/v22n3/winquist.pdf