STATISTICS STUDENT PERFORMANCE AND ANXIETY: COMPARISONS IN COURSE DELIVERY AND STUDENT CHARACTERISTICS

SARAI HEDGES
University of Cincinnati
Sarai.Hedges@uc.edu

ABSTRACT

The statistics education community continues to explore the differences in performance outcomes and in student attitudes between online and face-to-face delivery methods of statistics courses. In this quasi-experimental study student persistence, exam, quiz, and homework scores were compared between delivery methods, class status, and programs of study for students enrolled in an undergraduate introductory statistics course. Student persistence and homework grades were significantly different for course delivery method. Anxiety levels, measured using the Statistical Anxiety Rating Scale (STARS), were compared between delivery methods, programs of study, and gender. One anxiety subscale—Test & Class Anxiety—was significantly different between delivery methods and genders. Implications and suggestions for further study are offered based on the study results.

Keywords: Statistics education research; Online; Persistence; Statistics anxiety

1. INTRODUCTION

The proliferation of online course offerings has likewise encouraged the proliferation of studies on student outcomes in these courses. Alternative course delivery methods, like online formats, introduce an important variable for understanding student success in college statistics courses, classes that are known to be barriers to graduation for some students (Roberts & Bilderback, 1980; as cited in Onwuegbuzie & Wilson, 2003). Student attitudes and anxiety are other variables that have been found to explain statistics students' course performance (Emmioglu & Capa-Aydin, 2012; Onwuegbuzie & Wilson). The purpose of this study was to determine if and how various student characteristics, some previously unexplored, explained differences in performance and anxiety of introductory undergraduate statistics students.

Student outcomes in online statistics courses have been recently studied using various methods. For example, in a literature review of the publications on teaching statistics online from 1999 to 2009, Mills and Raju (2011) found 20 relevant journal articles comparing online or hybrid statistics courses with the traditional face-to-face delivery. Of these 20 studies, only four were empirical studies that compared statistics students' performance and attitudes between asynchronous online and face-to-face delivery. The continued relevance of comparing the two factors of student performance and student attitudes between delivery methods is evidenced by four more publications of such articles recently (DeVaney, 2010; Gundlach, Richards, Nelson, & Levesque-Brisol, 2015; Kirtman, 2009; Shotwell & Apigian, 2015). Various combinations of student populations

and measurement tools were used in these empirical studies that compared statistics students' performance and attitudes between the asynchronous online and face-to-face delivery methods. Only one study (DeVaney) specifically measured students' statistics anxiety, and only two studies (Gundelach et al.; Kreiner, 2006) used undergraduate students enrolled in a variety of programs of study, but neither made comparisons of student outcomes among the students' programs of study. Of the two studies whose students were enrolled in a variety of programs of study, only Gundelach et al. compared some student characteristics across levels of class status between delivery methods.

In this article, I describe the course and delivery methods used in the study as well as the diverse group of students who participated. The study has two components: (1) comparing performance outcomes between delivery methods and across class status and programs of study, using persistence, as measured by course withdrawal rate, quiz grades, and exam grades and (2) comparing statistics anxiety between delivery methods and gender, and across class status levels and programs of study, using the Statistical Anxiety Rating Scale (STARS). The study results and their implications are discussed in light of existing research.

The study described in this article differs in that it compared outcomes of statistics students' course performance and statistics anxiety not only for delivery method—whether face-to-face or online—but also the yet-to-be explored variables of program of study and class status. Another difference is that it compares statistics students' anxiety levels between delivery methods and genders. It is important to fill this niche in the growing body of literature and to learn whether these student characteristics are associated with these outcomes so that educators can improve student learning.

2. LITERATURE REVIEW

Research comparing student performance in asynchronous online and face-to-face college statistics classes have used various measures of performance that further resulted in differing results, sometimes even within studies themselves. For example, Summers, Waigandt, and Whittaker (2005) studied undergraduate nursing students and did not find significant differences in the sum of the four exam grades between the methods of delivery. In contrast, Kreiner's (as cited by Mills & Raju, 2011) study of both graduate and undergraduate students in various programs of study found a significant difference between delivery methods for the change from pretest to posttest final exam scores. This finding of significant differences was echoed in Kirtman's (2009) study of graduate education students, and Gundlach et al.'s (2015) study of undergraduate liberal arts students also found significant differences in various exam grades between delivery methods, although these two studies did not find significant differences between delivery methods for other graded components such as projects, homework, and papers. Similarly, Dutton and Dutton's (2005) study of undergraduate business students also compared homework grades, controlling for GPA, and did not find significant differences between delivery methods.

In addition to projects, homework, and papers, no significant differences between delivery methods were found in studies that used other measures of performance. McLaren's (2004) study of undergraduate business students compared final course grades. Shotwell and Apigian (2015) also studied undergraduate business students and compared final course grades and two variables related to an assignment. With respect to persistence among undergraduate business statistics students, although McLaren found a significant difference with online students withdrawing at a higher rate, Dutton and Dutton (2005) found no significant difference. Dutton and Dutton did note that their

online students were significantly older, worked longer, and had a higher GPA compared to their face-to-face counterparts. McLaren did not make demographic comparisons between the two delivery methods.

In the existing research that compared statistics students' performance in asynchronous online and face-to-face classes, various measures of statistics students' attitudes were used. Several studies used student evaluations and student comments (Gundlach et al., 2015; Kirtman, 2009; 2015; Shotwell & Apigian, 2015; Summers et al., 2005;). Three (DeVaney, 2010; Gundlach et al., 2015; Suanpang, Petocz, & Kalceff, 2004) used the Survey of Attitudes Toward Statistics (SATS), but only DeVaney used both the SATS and the Statistical Anxiety Rating Scale (STARS). DeVaney's study of students in a graduate-level educational statistics course showed "higher levels of anxiety and less favorable attitudes toward statistics" (p. 9) for the online students and showed significant differences between delivery methods for the STARS Test & Class Anxiety and Interpretation Anxiety subscales. None of these studies addressed student gender as a variable to explain differences in attitudes or anxiety although gender has been shown to be related to attitudes in Science, Technology, Engineering, and Mathematics (STEM) disciplines (Wang, 2013). In the STEM discipline of statistics, studies of statistics students showed gender was a significant factor in the attitudes about statistics instructional software (Alldredge & Brown, 2006) and preferences for online instruction (Johnson, Dasgupta, Zhang, & Evans, 2009).

Thus course performance outcomes, using various measures including persistence, were contradictory and the one study of graduate students' attitudes and anxiety found that online students were more anxious. The contradictory results and lack of consideration of student characteristics in the studies comparing delivery methods for statistics students motivated the author to study outcomes of statistics students' course performance not only for delivery method but also for other student characteristics such as class status and program of study. The motivation to study statistics anxiety between delivery methods and the student characteristics of class status, program of study, and gender are similarly motivated by lack of study. If differences exist in performance and anxiety among these variables, educators have more information available to facilitate improved student performance and lowered student anxiety.

3. STUDY AIMS

The study was conducted in two parts. The first part measured performance outcomes using persistence and homework, quiz, and exam grades between delivery methods and across class status levels and programs of study. The second part of the study measured statistics anxiety between delivery methods and across class status levels and programs of study as well as between genders. Specifically, the research questions for part one were:

1. Is there a difference in persistence, as measured by withdrawal rate, between online and face-to-face delivery, among class status levels, and/or among programs of study?
2. Is there a difference in homework grades between online and face-to-face delivery, among class status levels, and/or among programs of study?
3. Is there a difference in quiz grades between online and face-to-face delivery, among class status levels, and/or among programs of study?
4. Is there a difference in exam grades between online and face-to-face delivery, among class status levels, and/or among programs of study?

The research question for part two was:

Is there a difference in statistics anxiety between online and face-to-face delivery, among class status levels, among programs of study, and/or between genders?

4. THE COURSE

This study comes from three sections of a 200-level undergraduate statistics course: two sections were delivered in a face-to-face setting and one was delivered in an asynchronous online format. All three were taught by the same instructor during the fall semester 2013. A fourth face-to-face section that was taught by another instructor was not included. It was the second time the course was offered online but the first time by this particular instructor. The course is required for students in the following programs of study: Athletic Training, Health Promotion and Education, Hospitality Management, Information Technology, Sport Administration, and Substance Abuse Counseling. Few students outside of these programs register for the course. It has no pre-requisites and only three of the programs require students to take the follow-up course, which covers inferential techniques.

Students self-selected their sections and although there were enrollment caps for the face-to-face sections based on the room size, none of the sections were full. The online section had an enrollment cap of 200 students but had 71 students register, which was the same number of students registered in the two face-to-face sections combined. The face-to-face sections met for 80 minutes twice a week in a computer classroom throughout the 16-week term.

The course uses the Agresti and Franklin (2013) text, *Statistics: The Art and Science of Learning from Data*, bundled with MyStatLab (2013) and with access to StatCrunch (2013). Topics covered in the course include graphical and numerical summaries of univariate and bivariate data, experimental and observational studies, probability, and some probability distributions. Both methods of delivery included the same assignments, activities, and examinations. All homework assignments were available from the beginning of the term with one assignment due per week and were completed using MyStatLab. Quizzes and exams were given using the Blackboard web-based course management system. Face-to-face students took quizzes and exams at designated times during class while online students had several days to access them. Lab activities, which supplement the topics covered and include using java applets, identifying elements of a study from a journal article, and analyzing data, are incorporated into the course almost every week. Face-to-face students completed seven labs during class usually in groups of two or three students, although frequently some students completed the labs alone— sometimes during class and sometimes remotely before class. Online students completed four of these labs, each open for one week, and were directed to complete them individually. The lowest lab grade was dropped for each student. The total number of course points designated for labs was the same for both groups as the online students' labs were each worth twice as many points as that of the face-to-face students. A summary of the grade distribution is shown in Table 1.

The courses appear differently in Blackboard for the two delivery methods. In the academic unit in which this course is offered, online courses are required to use a standardized template that organizes the course into weekly modules. All assignments can be accessed within that week's module. As with other similar studies comparing face-to-face and online courses, information provided to the students differed between the delivery methods. Face-to-face students viewed information in Blackboard organized by type such as course documents and assignments. Online students were provided links to video demonstrations of StatCrunch whereas face-to-face students were given StatCrunch

demonstrations by the instructor during class and also provided links to the videos. Another difference is that face-to-face students participated in in-class exercises led by the instructor. The PowerPoint presentations were different for the two delivery methods as well. PowerPoint presentations for the face-to-face students supplemented the in-class exercises whereas online students were provided with narrated PowerPoint presentations. Although access to StatCrunch was provided as an option for statistical calculations, its use was not required by either group. Students could choose to use graphing calculators or by-hand calculations with formulas provided in the text.

Table 1
Distribution of Points for Course Grade

Grade Component	Face-to-face Points	Asynchronous Online Points
12 Homework Assignments @ 5 points each	60	60
3 Quizzes (lowest dropped)	40	40
Lab Activities (lowest dropped)	6 @ 25 each = 150	3 @ 50 each = 150
2 Exams @ 100 points each	200	200
Comprehensive Final Exam	150	150
TOTAL	600	600

5. COURSE PERFORMANCE

5.1. STUDENT DEMOGRAPHICS

Student demographics were gathered from the class list provided by the registrar, which included class status and program of study. The list reflected registration after the ninth day of the term. As described previously, this course is required for students in a variety of programs of study.

Table 2 illustrates the differences in the student characteristics between the two delivery methods. A chi-square test of independence of class status and delivery method was statistically significant ($p < .001$). The most striking difference was in enrollment between delivery methods for seniors. The online section was comprised of more than three times as many seniors (47.9%) compared to the face-to-face offering (12.7%).

Inferential statistical analyses could not be performed to compare program of study and delivery method due to the low number of students in some of the programs of study but it is interesting to note from Table 2 that the Information Technology and Substance Abuse Counseling programs had more students enrolled in the online section than in the face-to-face section. Also of note is that the Substance Abuse Counseling program could be completed completely online, although students could opt to take this or an equivalent course in a face-to-face setting. Table 2 also illustrates that students taking this course from the Sport Administration program were the most evenly split between face-to-face and online delivery.

Table 2
Summary of Student Demographics

Characteristic		Face-to-face	Asynchronous Online
Class Status[a]	Freshman	8 (11.3%)	3 (4.2%)
	Sophomore	27 (38.0%)	16 (22.5%)
	Junior	27 (38.0%)	18 (25.4%)
	Senior	9 (12.7%)	34 (47.9%)
	TOTAL	71 (100.0%)	71 (100.0%)
Program of Study[b]	Athletic Training	2 (2.8%)	0 (0.0%)
	Exploratory	7 (9.9%)	1 (1.4%)
	Health Promotion & Education	25 (35.2%)	19 (26.8%)
	Hospitality Management	3 (4.2%)	2 (2.8%)
	Information Technology	13 (18.3%)	23 (32.4%)
	Sport Administration	18 (25.4%)	16 (22.5%)
	Substance Abuse Counseling	0 (0.0%)	6 (8.5%)
	Other[c]	3 (4.2%)	4 (5.6%)
	TOTAL	71 (100.0%)	71 (100.0%)

Note. [a] $p < .001$, [b] unable to conduct chi-square test, [c] includes registered students whose programs do not require this particular course such as Early Childhood, Journalism, Non-matriculated, Pre-English, and Nursing.

5.2. PERSISTENCE

The first research question of this study asks if there is a difference in persistence, as measured by withdrawal rate, between online and face-to-face delivery, among class status levels and/or among programs of study. There is one way to complete the course but multiple designations for not completing the course. Students who officially withdraw do so through the registrar's office no later than the tenth week of the term and are assigned a grade of W by this instructor if they participated in the class or a grade of WX if they did not participate. Unofficial withdrawal counts students who do not go through the formalities with the registrar, both having participated and receiving the grade of UW (unofficial withdrawal with participation) or X (unofficial withdrawal without participation). Table 3 shows that only 1.4% of students with face-to-face delivery did not complete the course whereas 15.5% of the online students did not complete it. The difference in the rate of course completion between the two delivery methods is statistically significant using Fisher's Exact Test ($p = .004$).

Table 3
Summary of Student Persistence

Course Completion	Face-to-face	Asynchronous Online
Withdrew by Nov. 1	0 (0.0%)	9 (12.7%)
Unofficially Withdrew	1 (1.4%)	2 (2.8%)
Total Withdrawals	1 (1.4%)	11 (15.5%)

Further analysis of persistence including the face-to-face sections was not completed because there was only one student who withdrew. The low number of withdrawals in the online section also precluded inferential statistical analyses of persistence among class status levels and among programs of study. Table 4 illustrates that among class status levels for the online students, a higher percentage of juniors and seniors withdrew from

the course. However, a Fisher's Exact Test based on combining freshmen and sophomores into one group and juniors and seniors into a second group did not produce a statistically significant association between class status level and withdrawal from the course ($p = .715$).

Table 4
Online Student Persistence by Class Status

Class Status	Withdrew From Course	Completed Course	Class Status Total
Freshman	0 (0.0%)	3 (100.0%)	3 (100.0%)
Sophomore	2 (12.5%)	14 (87.5%)	16 (100.0%)
Junior	3 (16.7%)	15 (83.3%)	18 (100.0%)
Senior	6 (17.6%)	28 (82.4%)	34 (100.0%)

Table 5 illustrates that the withdrawal rate varied greatly among programs of study. Further investigation yielded that the high withdrawal rate for the Hospitality Management program was due to an academic misconduct issue, leaving the Substance Abuse Counseling program with the highest percentage of students who withdrew from the course. The Athletic Training program of study was not represented in the online section.

Table 5
Online student persistence by program of study

Program of Study	Withdrew From Course	Completed Course	Program Total
Exploratory	0 (0.0%)	1 (100%)	1 (100.0%)
Health Promotion & Education	2 (10.5%)	17 (89.5%)	19 (100.0%)
Hospitality Management	2 (100%)	0 (0.0%)	2 (100.0%)
Information Technology	4 (17.4%)	19 (82.6%)	23 (100.0%)
Sport Administration	0 (0.0%)	16 (100.0%)	16 (100.0%)
Substance Abuse Counseling	2 (33.3%)	4 (66.7%)	6 (100.0%)
Other[a]	1 (25.0%)	3 (75.0%)	4 (100.0%)

[a] Includes registered students whose programs do not require this particular course such as Early Childhood, Journalism, Non-matriculated, Pre-English, and Nursing.

Thus, with respect to the first research question, there was a significant difference in persistence, as measured by withdrawal rate, between methods of delivery, but statistical significance could not be tested for persistence among class status levels and programs of study due to the low number of withdrawals. Of note was that a higher percentage of juniors and seniors withdrew from the course and the Substance Abuse Counseling program of study had the highest percentage of students who withdrew from the course. Also of note was that among the levels of class status, more seniors chose to register for the online section compared to the face-to-face sections. Among programs of study, more Information Technology students chose to register for the online section.

5.3. GRADES

All components of students' course grades were worth the same point values between the face-to-face and online delivery methods but the lab component varied between delivery methods in the number of lab assignments and individual lab assignment points. Consequently the remaining components—homework, quiz, and exam grades—were

analyzed to determine if significant differences exist between delivery methods and among class status levels and among programs of study for students who completed the course. As described in the previous section, 12 students did not complete the course and were removed from the analysis of grades. Three programs did not have students who completed the course in both delivery methods: Athletic Training, Hospitality Management, and Substance Abuse Counseling. Thus for the analysis, the nine students from these three programs were removed, leaving a total of 121 students for the analysis of grades.

Homework was completed using the MyStatLab online system. Homework assignments were announced weekly, although all homework assignments were posted and available to students at the beginning of the term. For each assignment, students were given two attempts at each exercise and the highest grade for each attempt was recorded. Each homework assignment was worth five points, although the assignments varied in the number of exercises. There was a maximum of 60 points to be earned for the homework component of the course grade. Initially, factorial ANOVA was used to test for differences in total homework scores between delivery methods, among class status levels, and among programs of study for students who completed the course. The model showed heteroscedasticity of the variance. Thus independent samples tests of medians were performed using the k-sample median test for each of the three factors. The delivery method was found to be statistically significant, $p = .001$, class status was not statistically significant, $p = .058$, and program of study was not statistically significant, $p = .309$. Figure 1 illustrates that there does not appear to be an interaction between class status and delivery method. Online students had a median homework score of 51.85 points whereas face-to-face students had a median homework score of 43.8 points. Thus online students scored about 8 points, or about 13 percentage points, higher on the homework grade component than the face-to-face students.
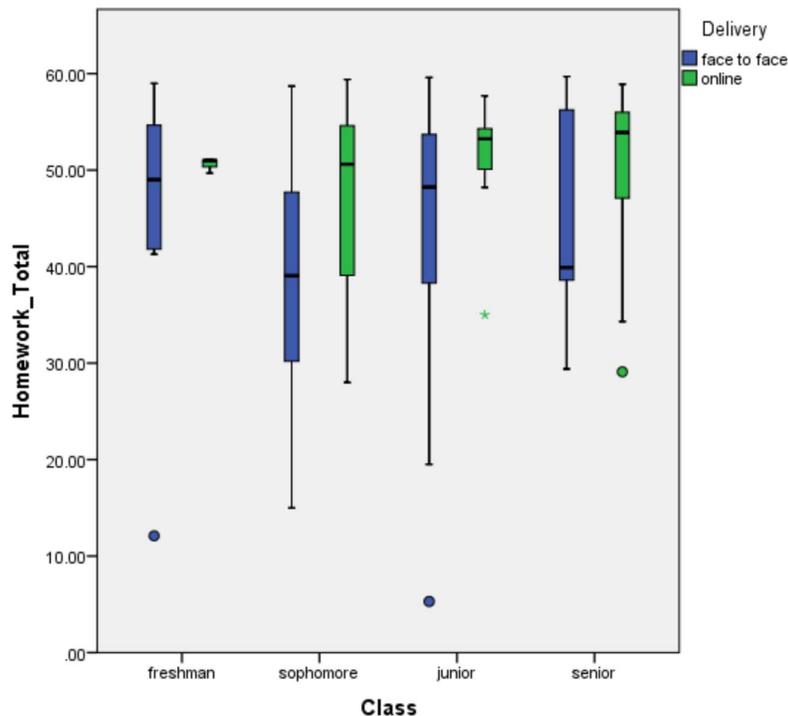

Figure 1. Boxplots of homework scores by class status and delivery method.

Quiz grades were used as another measure of student performance. Identical open book, open note quizzes were given to face-to-face and online sections using the Blackboard course management system. Most questions were multiple choice. All students were given 30 minutes to complete the quizzes, after which time the quiz automatically closed. All students saw one question at a time and they were not able to backtrack, meaning once a question was submitted it could not be viewed again. The only difference in the quizzes between delivery methods was that face-to-face students took the quizzes in class on a designated day with the instructor whereas online students had several days in which the quiz was available for them to take in an unproctored setting. Proctoring for online students was not available for any course offered from this academic unit. Three quizzes were given and the lowest quiz grade was dropped. Factorial ANOVA was used to test for differences in total quiz scores (the sum of the top two quiz grades) between delivery methods, among class status levels, and among programs of study for students who completed the course. The model showed heteroscedasticity of the variance. Thus independent samples tests of medians were performed using the *k*-sample median test for each of the three factors and no statistically significant differences in total quiz scores were found.

Exams were similarly designed and administered as the quizzes. They were open book, open note, and mostly multiple choice questions, the number of which varied. Exams were given to face-to-face and online sections using the Blackboard course management system. All students were given 75 minutes to complete each of the two midterm exams and 120 minutes to complete the final exam, after which times the exams automatically closed. As with the analyses of homework and quiz grades, factorial ANOVA was used to test for differences in total exam scores (the sum of all exam grades) between delivery methods, among class status levels, and among programs of study for students who completed the course. Once again the model showed heteroscedasticity of the variance. Thus again independent samples tests of medians were performed using the *k*-sample median test for each of the three factors and no statistically significant differences in exam scores were found.

Thus—with respect to research questions two, three, and four—among homework grades, quiz grades, and exam grades, only homework grades showed a statistically significant difference and only for course delivery method with online students earning higher homework scores than face-to-face students. There were no significant differences in grades among levels of class status or among programs of study.

6. PART 2: STUDENT ANXIETY

6.1. STUDENT DEMOGRAPHICS

The research question for the second part of the study asks if there is a difference in statistics anxiety between delivery methods, among class status levels, among programs of study, and/or between genders. The student demographics resulting from the statistics anxiety survey do not match those for the analyses of student performance as students responded to the survey anonymously, not all students participated in the survey, and the survey demographics, summarized in Table 6, were self-reported.

To make comparisons of anxiety between delivery methods, students were removed from the analysis that were in programs without students completing the course in both delivery methods. One student from each Athletic Training and Hospitality Management were removed. Upon removing these two students, the freshman class status level only contained one online student, again preventing comparisons between delivery methods

and was thus removed. Further, two students did not complete all survey questions—one not answering 35 of the questions and the other not answering 21 of the questions. Because this would alter the survey scores, these two students were removed from the analysis. Thus for the analysis, five students were removed leaving a total of 103 students for the analysis of statistics anxiety. As summarized in Table 6, there was no significant difference in gender, $\chi^2(1) = 0.081$, $p = .775$, or in program of study, $\chi^2(3) = 2.355$, $p = .502$ between delivery methods for survey participants. Due to small expected counts for one class status level, a statistical test could not be performed. The levels of class status appear similarly distributed between the delivery methods.

Table 6
Student Characteristics for STARS Respondents

Variable		Face-to-face	Asynchronous Online
Class Status[a]	Sophomore	13 (25.5%)	9 (17.3%)
	Junior	21 (41.2%)	22 (42.3%)
	Senior	15 (29.4%)	16 (30.8%)
	Other[b]	1 (2.0%)	5 (9.6%)
	TOTAL	51 (100%)	52 (100%)
Program of Study[c]	Health Promotion & Education	16 (31.4%)	13 (25.0%)
	Information Technology	12 (23.5%)	17 (32.7%)
	Sport Administration	18 (35.3%)	14 (26.9%)
	Other[d]	5 (9.8%)	8 (15.4%)
	TOTAL	51 (100%)	52 (100%)
Gender[c]	Male	28 (54.9%)	30 (57.7%)
	Female	23 (45.1%)	22 (42.3%)
	TOTAL	51 (100.0%)	52 (100.0%)

Note. [a] unable to test, [b] survey choices were freshman, sophomore, junior, senior, and other, [c] no significant difference, [d] survey choices were Athletic Training, Health Promotion, Hospitality Management, Information Technology, Sport Administration, and Other.

6.2. SURVEY INSTRUMENT

The Statistical Anxiety Rating Scale (STARS) developed by Cruise, Cash, and Bolton (1985) was used to measure student anxiety. It contains fifty-one 5-point response items. Although the entire survey was administered, only five of the six subscales were analyzed: Worth of Statistics, Test & Class Anxiety, Computational Self-Concept, Fear of Asking for Help, and Fear of Statistics Teachers. The Interpretation Anxiety subscale was not analyzed because it asks questions related to hypothesis testing, which is not covered in this course. Table 7 provides descriptions of the subscales used in this study with examples of some of the instrument's items.

Table 7

Descriptions of STARS Subscales with Examples of Items

Subscale (number of items)	Score Range	Subscale Description[a]	Examples of Items
Worth of Statistics (16)	16 – 80	...student's perception of the relevance of statistics	I feel statistics is a waste I'm never going to use statistics...
Test & Class Anxiety (8)	8 – 40	...anxiety involved when taking a statistics class or test	Studying for an examination in a statistics course. Doing the homework for a statistics course.
Computational Self-Concept (7)	7 – 35	...anxiety experienced when doing mathematical problems as well as...self-perception of his/her ability to understand and calculate statistics.	I haven't had math for a long time. I know I'll have problems getting through statistics. Since I've never enjoyed math, I don't see how I can enjoy statistics.
Fear of Asking For Help (4)	4 – 20	...anxiety experience when asking for help	Going to ask my statistics teacher for individual help with material I am having difficulty understanding.
Fear of Statistics Teachers (5)	5 – 25	...the student's perception of the statistics teacher	Most statistics teachers are not human. Statistics teachers talk so fast you cannot logically follow them.

[a] Cruise, Cash, & Bolton, 1985, p. 93.

The survey with the addition of several demographic questions was made available to students on Blackboard for one week, ending with the day of the final exam. Participation was voluntary and students earned five bonus points for completing the survey. The instructor was unable to identify individual student responses.

6.3. ANXIETY COMPARISONS

Table 8 shows that typically respondents had a strong-to-neutral computational self-concept and found statistics somewhat worthy. They had a low-to-neutral fear of asking for help and a low fear of statistics teachers. The Test & Class Anxiety score was higher relatively, on average. This subscale includes questions about exams and it is reasonable to assume student anxiety in this area could have been elevated right before the final exam.

Table 8

Summary of STARS Scores

Subscale (Score Range)	*n*	*M* (*SD*)
Worth of Statistics (16 – 80)	100	41.80 (13.84)
Test & Class Anxiety (8 – 40)	100	26.48 (6.94)
Computational Self-Concept (7 – 35)	102	17.63 (6.04)
Fear of Asking For Help (4 – 20)	103	9.76 (3.87)
Fear of Statistics Teachers (5 – 25)	102	10.40 (4.03)

Due to the removal of two students from programs of study not represented in both delivery methods, the variable class status contained no freshman students and only one student who self-designated his/her class status as other. Thus due to incomplete data, class status was not analyzed. Prior to the MANOVA using delivery method, program of study, and gender to explain if differences in anxiety exist, an exploratory inspection of the data was conducted and a recurring outlier was found. One respondent recorded extreme scores for almost every question. As was done with the Carmona, Martinez, and Sanchez (2005) study, the analysis was performed with and without this unusual observation. Results are reported with the unusual observation removed but results with it included yield the same conclusions. Inspection of the Normal Q-Q plots did not show extreme outliers or strong patterns of deviation from normality and thus all subscales remained in the model without transformation. The equality of error variances was verified for each STARS subscale using Levene's Test. Box's Test was not significant and showed that the homogeneity of variance-covariance was fulfilled, $F(105, 3756.807) = 1.1084$, $p = .265$, so the Wilks' Lambda test statistic was used in interpreting the MANOVA results. Results indicated that only delivery method had a significant effect, $F(5,75) = 2.659$, $p = .029$. The program of study, gender and factor interactions were not significant.

Using backward elimination (Neter, Wasserman, & Kutner, 1989), the program of study factor was removed from the model because it provided the smallest contribution in explaining the variability in the subscale scores. For the new model using method of delivery and gender, again the equality of error variances for each factor and homogeneity of variance-covariance were met. Both delivery method, $F(5, 86) = 3.192$, $p = .011$ and gender, $F(5, 86) = 3.323$, $p = .009$, were statistically significant. The interaction of delivery method and gender was not. As recommended in Mertler and Vannatta (2010), univariate ANOVAs were then conducted for delivery method and gender to determine significant differences for each of the five STARS subscales. Homogeneity of variances was verified for the models using Levene's Test. A 0.01 ($0.05/5 = 0.01$) level of significance was used to evaluate statistical significance within each univariate ANOVA. As shown in Table 9, statistically significant effects of both delivery method and gender were found for the Test & Class Anxiety subscale, and the difference between genders approached statistical significance for the Computational Self-Concept subscale.

Table 9

Univariate ANOVA Results for STARS Subscales

Subscale	Delivery Method F (p)	Gender F (p)	R^2
Worth of Statistics	2.768 (.099)	.805 (.372)	.054
Test &Class Anxiety	19.009 (.000)	14.268 (.000)	.256
Computational Self-Concept	4.775 (.031)	6.094 (.015)	.099
Fear of Asking for Help	4.428 (.067)	3.627 (.060)	.068
Fear of Statistics Teachers	1.011 (.317)	.017 (.896)	.017

Note. Per test error rate set at $\alpha = .01$.

Further analysis, summarized in Table 10, showed that on average online students scored about 5 points, or 16 percentage points, higher in the Test & Class Anxiety subscale than the face-to-face students and female students scored about 4 points, or 13 percentage points, higher than males. Of note is the almost 10-point difference in Test &

Class Anxiety subscale score, on average, between face-to-face male and online female students. In other words, this anxiety rating was about 30 percentage points higher for the online females. As shown in Figure 2, when considering medians instead of means, the difference between these two groups is even more extreme.

Table 10
Mean Test & Class Anxiety Scores by Delivery Method and Gender

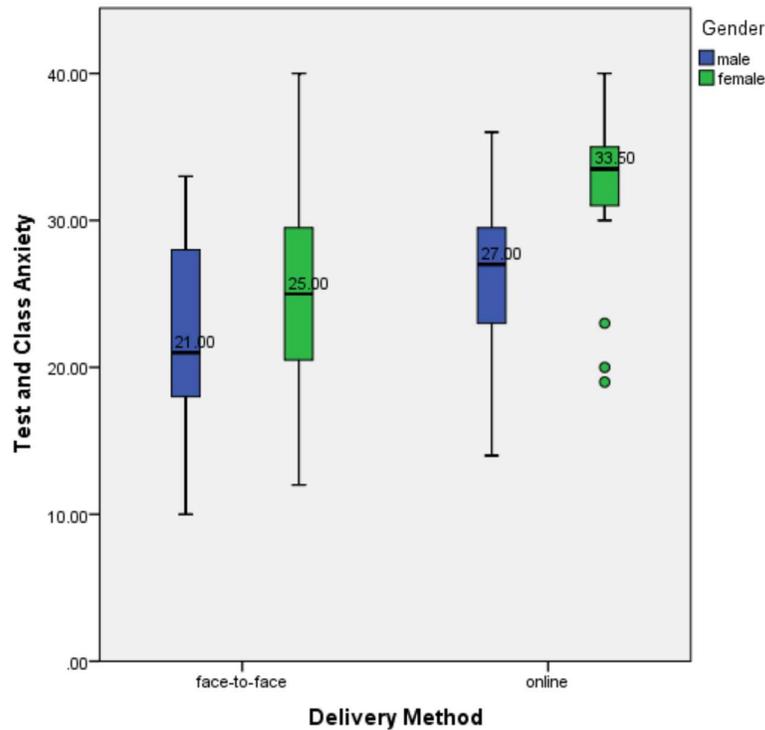|	Overall	Asynchronous Online	Face-to-face
Overall		28.84	23.80
Female	28.76	32.00	25.65
Male	24.33	26.36	22.15

Figure 2. Differences in test and class anxiety by delivery method and gender.

Although Computational Self-Concept was not statistically significant at the 0.01 level, females had more anxiety than males. Females' mean score was 19.23, whereas males' mean score was 16.42. Thus females scored, on average, about 3 points, or 10 percentage points, higher on this subscale.

Thus, with respect to the research question concerning differences in statistics anxiety between methods of delivery, among class status levels, among programs of study, and between genders, significant differences in the Test & Class Anxiety subscale were found for both delivery method and gender. The difference between genders on the Computational Self-Concept approached statistical significance. Program of study was not a significant factor in explaining statistics anxiety and the class status factor was not analyzed due to lack of data.

7. LIMITATIONS

This study had several limitations. The samples of face-to-face and online students were large but were self-selected. As with all observational studies, lurking variables may be present that could explain differences in course performance and anxiety between different types of students.

Although generally the sample sizes were large, the small number of students in some categories of class status and program of study precluded some analyses and reduced the power of the analyses.

All students completed the homework under the same conditions but the administration of the quizzes and exams differed between the delivery methods. All students were directed to take quizzes and exams unassisted. This was verified for the face-to-face students by way of proctored exams but was not verified for online students. Although this may be viewed as a limitation, these are the "natural" conditions under which online students take quizzes and exams in all courses in this academic unit.

8. DISCUSSION

The purpose of this study was to determine if and how various student characteristics, some previously unexplored, explained differences in performance and anxiety of introductory undergraduate statistics students. The importance of understanding these relationships was argued to be a means of improving student learning.

Differences in persistence between online and face-to-face students were statistically and practically significant. Online students withdrew at a rate over 10 times higher than the face-to-face students. Obviously of interest is to know why. For McLaren (2004), who found differences in persistence among undergraduate business students, this question is unanswered. The Dutton and Dutton (2005) study, also of undergraduate business students, did not find differences in persistence but may hint at possible explanatory demographics. Their online student population may have been different than that of McLaren's students: older, working, and with higher GPAs. Age, work status, and GPA were not measured by McLaren or by the current study. Other studies have suggested the increased ability to procrastinate (Johnson et al., 2009) and similarly, the lack of self-motivation and/or time management skills (Kreiner, 2006 as cited by Mills & Raju, 2011) as possible factors explaining less success in statistics offered online. In the current study seniors were over-represented in the online section of the course and comprised over half of the students who withdrew. The vast majority of seniors (79%) taking this sophomore-level course chose to take it online. Almost half (47.9%) of the online students were seniors whereas only 12.7% of face-to-face students were seniors. The withdrawal rate in the online course could not be tested for statistical significance among the levels of class status but the seniors' withdrawal rate was the highest at 17.16%. One possible explanation for the difference in persistence among levels of class status is that seniors often attempt an overly ambitious course load for a timely completion of their degrees. Differences in persistence among programs of study could not be tested but the withdrawal rate among programs of study for online students varied from 0% to 33.3%.

McLaren (2004) inquired if demographic information can predict performance (p. 8). Class status may have been part of the explanation for one aspect of course performance. Total homework points earned were significantly different for delivery method and nearly significant for class status. The interaction between delivery method and class status did

not appear to be significant. On average, online students earned about eight more homework points, or 13 percentage points more, than face-to-face students. These results differ from the Gundlach et al. (2015) study, which did not find significant differences in homework grades among face-to-face, online, and flipped course offerings. Further, unlike both the Gundlach et al. and the Kreiner (2006) studies, the current study did not find significant differences in exam grades between online and face-to-face students. The students were similar in all three studies but Kreiner used a mastery-based approach to testing. No significant differences in exam grades were found in the current study among the other potential explanatory variables of levels of class status and program of study. This also held true for quiz grades.

Although online students who persisted generally did as well as face-to-face students on quizzes and exams and although they did better in the course in terms of homework grades, they were more anxious. Online students rated Test & Class Anxiety about 16 percentage points higher than face-to-face students. This significant difference at the end of the course differs from DeVaney's (2010) results that showed no significant differences in this subscale between delivery methods at the end of the course. The most striking result of this study was that female online student had the most anxiety. The average female online student's Test & Class Anxiety subscale score was over 30 percentage points higher than that of the male face-to-face student. Females also had more anxiety than males in terms of Computational Self-Concept, albeit not significantly more. Previous studies found gender differences in learning statistics but did not compare anxiety levels. For example, Alldredge and Brown (2006) found that gender can be a significant factor in the attitudes about statistics instructional software. Haley, Johnson, and Kuennen (2007) found that introductory students performed significantly worse when the instructor's gender was different from theirs. The Johnson et al. (2009) study of a hybrid versus face-to-face offering for undergraduate mathematics students showed males more likely than females to prefer an internet course.

9. IMPLICATIONS AND CONCLUSIONS

This study strongly suggests a need for research on gender differences in statistic anxiety—and perhaps by extension statistics attitudes—to be expanded into the online realm. Other than the current study and the studies described just previously, little has been done. Further research in this area will improve understanding of gender differences so that interventions can be developed to address them.

The highly significant withdrawal rate of online students suggests the importance of further study of potential explanatory factors and of academic advising interventions. Also, because of the significant differences in homework grades between delivery methods, further study into students' use of homework may provide the most insight into how to improve student learning.

Another area for further research is in statistics students' previous mathematics preparation. Research has shown that previous mathematics course work and performance are associated with statistics students' performance and attitudes (Carmona et al., 2005; Chiesi & Primi, 2010; Green, Stone, Zegeye, & Charles, 2009; Johnson & Kuennen, 2006; Nasser, 2004). Although mathematics requirements varied for students' programs of study, it was unknown if individual students had completed their program's mathematics requirements before taking the statistics course in the current study and thus the relationship between students' previous mathematics course work and statistics course performance and anxiety could not be directly analyzed. This is another important factor that could provide a better understanding of student performance and anxiety.

Understanding factors that influence statistics students' persistence, course performance, and anxiety are important not only to teachers of statistics in terms of improving pedagogy but also to those at an advisory level who help guide students in selecting appropriate course loads. This study's findings furthered that understanding by combining factors only previously explored individually and by providing direction for future exploration based on those findings.

REFERENCES

Alldredge, J. R., & Brown, G. R. (2006). Association of course performance with student beliefs: An analysis by gender and instructional software environment. *Statistics Education Research Journal, 5*(1), 64-77. Retrieved from http://iase-web.org/documents/SERJ/SERJ5(1)_Alldredge_Brown.pdf

Carmona, J., Martinez, R. J., & Sanchez, M. (2005). Mathematical background and attitudes toward statistics in a sample of Spanish college students. *Psychological Reports, 97*(1), 53–62.

Chiesi, F., & Primi, C. (2010). Cognitive and non-cognitive factors related to students' statistics achievement. *Statistics Education Research Journal, 9*(1), 6-26. Retrieved from http://iase-web.org/documents/SERJ/SERJ9(1)_Chiesi_Primi.pdf

Cruise, J. R., Cash, R. W., & Bolton, L. D. (1985). Development and validation of an instrument to measure statistical anxiety. *American Statistical Association Proceedings of the Section on Statistical Education*, pp. 92-98.

DeVaney, T. A. (2010). Anxiety and attitude of graduate students in on-campus vs. online statistics courses. *Journal of Statistics Education, 18*(1). Retrieved from http://www.amstat.org/publications/jse/v18n1/devaney.pdf

Dutton, J., & Dutton M. (2005). Characteristics and performance of students in an online section of business statistics. *Journal of Statistics Education, 13*(3). Retrieved from www.amstat.org/publications/jse/v13n3/dutton.html

Emmioglu, E., & Capa-Aydin, Y. (2012). Attitudes and achievement in statistics: A meta-analysis study. *Statistics Education Research Journal, 11*(2), 95-102. Retrieved from http://iase-web.org/documents/SERJ/SERJ11(2)_Emmioglu.pdf

Green, J. J., Stone, C. C., Zegeye, A., & Charles, T. A. (2009). How much math do students need to succeed in business and economics statistics? An ordered probit analysis. *Journal of Statistics Education, 17*(3). Retrieved from http://www.amstat.org/publications/jse/v17n3/green.html

Gundlach, E., Richards, K. A. R., Nelson, D., & Levesque-Bristol, C. (2015). A comparison of student attitudes, statistical reasoning, performance, and perceptions for web-augmented traditional, folly online, and flipped sections of a statistical literacy class. *Journal of Statistics Education, 23*(1). Retrieved from http://www.amstat.org/publications/jse/v23n1/gundlach.pdf

Haley, M. R., Johnson, M. F., & Kuennen, E. W. (2007). Student and professor gender effects in introductory business statistics. *Journal of Statistics Education, 15*(3). Retrieved from http://www.amstat.org/publications/jse/v15n3/kuennen.html

Johnson, H. D., Dasgupta, N., Zhang, H., & Evans, M. A. (2009). Internet approach versus lecture and lab-based approach for teaching an introductory statistical methods course: Students' opinions. *Teaching Statistics, 31*(1), 21-26.

Johnson, M., & Kuennen, E. (2006). Basic math skills and performance in an introductory statistics course. *Journal of Statistics Education, 14*(2). Retrieved from http://www.amstat.org/publications/jse/v14n2/johnson.html

Kirtman, L. (2009). Online versus in-class courses: An examination of differences in learning outcomes. *Issues in Teacher Education, 18*(2), 103-116.

Kreiner, D. S. (2006). A mastery-based approach to teaching statistics online. *International Journal of Instructional Media, 33*(1), 73-79.

McLaren, C. H. (2004). A comparison of student persistence and performance in online and classroom business statistics experiences. *Decisions Sciences Journal of Innovative Education, 2*(1), 1-10.

Mertler, C. A., & Vannatta, R. A. (2010). Multivariate analysis of variance and covariance. In *Advanced and multivariate statistical methods* (4th ed.) (pp. 117-158). Los Angeles, CA: Pyrczak Publishing.

Mills, J. D., & Raju, D. (2011). Teaching statistics online: A decade's review of the literature about what works. *Journal of Statistics Education, 19*(2). Retrieved from www.amstat.org/publications/jse/v19n2/mills.pdf

Nasser, F. M. (2004). Structural model of the effects of cognitive and affective factors on the achievement of arabic-speaking pre-service teachers in introductory statistics. *Journal of Statistics Education, 12*(1). Retrieved from http://www.amstat.org/publications/jse/v12n1/nasser.html

Neter, J., Wasserman, W., & Kutner, M. H. (1989). Building the regression model. In *Applied linear regression models* (2nd ed.) (pp. 443-483). Homewood, IL: Irwin.

Onwuegbuzie, A. J., & Wilson, V. A. (2003). Statistics anxiety: Nature, etiology, antecedents, effects, and treatments—a comprehensive review of the literature. *Teaching in Higher Education, 8*(2), 195-209.

Shotwell, M. & Apigian, C. H. (2015). Student performance and success factors in learning business statistics in online vs. on-ground classes using a web-based assessment platform. *Journal of Statistics Education, 23*(1). Retrieved from www.amstat.org/publications/jse/v23n1/shotwell.pdf

Suanpang, P., Petocz, P., & Kalceff, W. (2004). Students attitudes to learning business statistics: Comparison of online and traditional methods. *Educational Technology & Society, 7*(3). 9-20.

Summers, J. J., Waigandt, A., & Whittaker, T. A., (2005). A comparison of student achievement and satisfaction in an online versus a traditional face-to-face statistics class. *Innovative Higher Education, 29*(3), 233-250.

Wang, X. (2013). Why students choose STEM majors: Motivation, high school learning, and postsecondary context of support. *American Educational Research Journal*, *50*(5), 1081-1121. doi: 10.3102/0002831213488622

SARAI HEDGES
PO Box 0068
University of Cincinnati
Cincinnati, OH 45219-0068