Learning Outcomes in a Laboratory Environment vs. Classroom for Statistics Instruction: An Alternative Approach Using Statistical Software

Ryan Sterling McCulloch[1]

[1] Assistant Professor Department of Human Physiology, Gonzaga University, USA

Correspondence: Ryan Sterling McCulloch, Assistant Professor Department of Human Physiology, Gonzaga University, USA

Received: September 27, 2017 Accepted: October 12, 2017 Online Published: October 16, 2017

doi:10.5430/ijhe.v6n5p131 URL: https://doi.org/10.5430/ijhe.v6n5p131

Abstract

The role of any statistics course is to increase the understanding and comprehension of statistical concepts and those goals can be achieved via both theoretical instruction and statistical software training. However, many introductory courses either forego advanced software usage, or leave its use to the student as a peripheral activity. The purpose of this study was to determine if there was instructional value in replacing classroom time with laboratory time dedicated to statistical software usage. The first approach used classroom lecture presentations, while the second replaced one classroom period per week with statistical software laboratories. It was hypothesized that replacing classroom time with software based laboratories would increase the level of statistics knowledge as compared to an otherwise identical class with no lab based component. Both pre-course and end-of course surveys were used, as well as identical examination questions. Comparisons within a time point, and longitudinal performance over the course were both evaluated. Survey results indicated that students would recommend lab based instruction significantly more than a primarily lecture based instruction (32% more, p=.020). Additionally, the performance improvement over the course of the semester was significantly higher for those students participating in laboratories (19.2% increase, p=.011). These findings indicate that sacrificing classroom time for a laboratory period improves the educational experience in an introductory statistics course and may help with the understanding and retention of difficult topics.

Keywords: Statistics education, Computers and learning, Classroom research, Statistics software, Pedagogy, Teaching

1. Introduction

Statistics courses are a valuable component to many, if not most, courses of study. The value of statistics knowledge continues well beyond the educational period, as most professions benefit from some level of statistical understanding. Despite the value placed on the courses by the instructors of higher education, few students eagerly anticipate taking a course in statistics (Stork, 2003). In fact, once in the course, students often find the course tedious and difficult and approach them with a level of fear (Stork, 2003) (Ciftci, Karadag, & Akdal, 2014). Furthermore, students express doubt about the applicability of stats to real-world usage (Wells, 2006). It is the aim of most statistics instructors to improve the relevancy and effectiveness of their teaching of the course. In fact, that is one of the principal goals of education research, to discover how students learn most effectively (Larwin & Larwin, 2011).

A key development in the progress of statistics instruction has been the incorporation of statistical software into the learning environment. There are multiple potential benefits to incorporating technology. There is less need for hand calculations that the software can quickly accomplish, thus freeing time for more conceptual instruction (Chance, Ben-Zvi, Garfield, & Medina, 200; Chow, 2015). Also, student's desire to learn may be enhanced when they move from readers to practitioners (Karp, 1995) Furthermore, technology can be used to explore interesting aspects of data, such as the effect of outliers that can't be reasonably accomplished with hand calculations (Mills & Johnson, 2004).

Given the benefit of technology in teaching statistics, there are many studies quantifying its use. One study showed that approximately 60% of Psychology programs use a technology component in teaching their statistics courses (Bartz & Sabolik, 2001). Another meta analysis study showed that students may move from the 50^{th} to the 73^{rd} percentile when the statistics instruction utilizes a computer component to the statistics course (Larwin & Larwin, 2011). The use of technology, along with a more flexible learning schedule, may be even more effective for the lower percentile students (Sherwood & Kwak, 2017) . Many additional studies have shown an improvement in student perceptions, survey results, grades and instructor evaluations (Chance et al. 2007; Karp, 1995; Mills & Johnson, 2004; Mills, 2002; Rodgers & Manrique, 1992; St. George, 1978; Stork, 2003; Wells, 2006; Ciftci et al., 2014). The majority of the existing literature is either qualitative in nature, or if quantitative, it does not employ the use of a control group for comparison. There is a need for more empirical, hypothesis driven research into the effectiveness of statistical software instruction for introductory statistics courses. Furthermore, no studies evaluate the value of replacing classroom time with laboratory periods dedicated to statistical software usage. Given the limited classroom time available during a semester, most instructors are reluctant to sacrifice classroom time for unproven alternative instructional methods.

Our degree program requires an introductory course in statistics. The course occurs in the sophomore year of the degree, and establishes the requisite base level of statistics knowledge needed for the research based courses in the junior and senior year. The course has long focused on theoretical instruction in the classroom, with usage of statistical software (SPSS, IBM Corp. Released 2013. IBM SPSS Statistics for Windows, Version 22.0. Armonk, NY) undertaken by students outside of the classroom to reinforce conceptual understanding. The software usage also introduces the students to a tool that will prove useful for the remainder of their education and post-graduation. More importantly, it is believed that the use of the statistical software training is valuable in reinforcing the understanding of theoretical concepts. However, some studies have suggested that so-called "hands-on" lab time may be the most effective means of learning statistical software (Larwin & Larwin, 2011; Sosa 2011). Some previous work has demonstrated that students strongly prefer a lab component when offered (Adams, Garcia, & Traustadóttir, 2016). A meta-analysis of courses that are laboratory based compared to "theory" based, demonstrated that exam scores are higher in laboratory courses (Zhang et al., 2015). Laboratory, or active instruction often proves to be more valuable in teaching, with success in engineering courses (Ray, Leeper, & Amini, 2014), math courses (Love, Hodge, Grandgenett, & Swift, 2014), and physics courses (Georgiou & Sharma, 2015). With evidence in other fields, there was a need to determine if the challenging concepts in statistics would benefit from more active, laboratory based instruction.

Additionally, anecdotal evidence from the students encouraged the trial of a lab based period for the usage of statistical software instruction. In conversations with students, some motivated learners went beyond the class requirements to spend more time learning the statistical software and felt that their additional time had improved their conceptual understanding. Those experiences partly encouraged the conduct of this investigation.

The aim of the current study was to evaluate statistics teaching methods for two sections of an introductory statistical course and answer the question: 'Does sacrificing classroom instruction time for software-based laboratory periods improve learning outcomes in an introductory statistics course?'. It was hypothesized that a course which replaces a portion of classroom time with weekly laboratory periods would result in a better understanding of statistical concepts than a class that incorporated more classroom time with no laboratory periods.

2. Method

2.1 Participants

Sophomores enrolled in different sections of an introductory level statistics course were recruited for the Institutional Review Board (IRB) approved study. The study was conducted in conjunction with the normal class requirements for the enrolled students. All students were advised of the nature of the study, the confidentiality of their participation, and results, and provided written informed consent to participate. A total of 53 students participated (31 females, 22 males), with a mean age of 18.7 ($\pm$ .7) years.

2.2 Approaches

The two sections were taught with different pedagogical approaches. The overall aim of the course was to introduce students to the fundamental concepts of statistics from both a theoretical and practical approach. Standard introductory statistics concepts were covered, including: distributions, deviation, error, normality, correlations, regression, mean comparisons, and analysis of variance. In both sections of the course, theoretical statistical concepts were taught in a lecture based format and incorporated a textbook for both lecture reference and outside of class assigned readings and homework assignments (Field, 2013). The same text was used for both cohorts, and the text was chosen for its inclusion of software instruction. Therefore, the students had a resource in the text, in addition to the instructor, to seek guidance and solve problems. Any slides used in class were provided for the students to follow along and to use for note taking. Additionally, the slides and lectures provided in each class were identical other than the components pertaining specifically to the statistical laboratories. The same instructor taught both sections to minimize bias related to material presentation, and specifically aimed to provide similar experiences other than the laboratories. Because all students in a section necessarily needed to receive the same teaching approach, the sections were randomly chosen for treatment type. The students were not aware of the study, or the teaching approach of the course for which they were registering at the time of registration. Additionally, because the laboratory sessions were taught during a typical class period, there was no *a priori* reason for students to suspect that their section would have a lab component.

The usage of statistical software in the course (in this case, SPSS) is implemented to both reinforce the understanding of the theoretical knowledge, and to provide students with a tool to use for the remainder of their major. By allowing students to use both real, in-class data sets, and practice data sets, they have the opportunity to perform complex statistical testing in a short time frame to evaluate concepts. Statistical software provides immediate feedback with nearly instantaneous results. Through this means, students are able to test concepts and apply theories.

2.3 Section Treatment

One section (“Class”) included primarily theoretical instruction with some minimal in-class demonstration components of statistical software. As topics were covered from the theoretical perspective (i.e. a student’s t-test) the background information, derivation, assumptions, etc. were taught in a lecture format. Students were encouraged to engage and ask questions, and were observed to do so. After the theoretical component of a statistical concept was presented (sometimes over the course of a few lecture periods) the implementation of a concept or particular statistical test was briefly presented using statistical software in a classroom environment. In the Class section, the instructor demonstrated software concepts and usage in front of the class via an overhead projector, with live usage of the software using provided data sets. However, the time spent on this was minimal, with most of the time available in the class periods dedicated to theoretical teaching. The students were primarily expected to learn the software usage outside of class to reinforce their theoretical understanding.

The other section (“Lab”) employed laboratory-based periods of statistical software usage that allowed the students to work at a computer with the instructor available for questions. The lecture approach was essentially identical other than the software instruction. Theoretical concepts were presented and reviewed with the students in class. However, the fundamental tradeoff in this class, and the crux of the experiment, was that one of three class periods each week was sacrificed to allow for a ‘hands-on’ laboratory. This reduced the time available in-class for theoretical topic presentation as compared to the other section (Class). Therefore, one of the class times each week was conducted in a computer lab. The teaching approach included a brief initial period of didactic software instruction by the instructor in the front of the computer lab. This instruction generally consisted of the instructor demonstrating the most recent statistical concepts via statistical software with a practice data set. This time was kept brief, generally 5-10 minutes. Following the instruction, the lab-based section had in-lab assignments to complete as is typical of college laboratory experiences. The assignment was posted on a class website and usually contained approximately 5 questions that the student was responsible for completing. A data set to use for the assignment was distributed to all of the students, again via the course website. The assignment aimed to step the students through the various theoretical concepts they had learned over the previous week. During the laboratory time, students were strongly encouraged to ask questions both of the instructor and even of other students. To that end, the instructor circulated amongst the computer stations for the duration of the lab period to facilitate guiding the students. The students were allowed the rest of the lab time to work, and the completed assignment was due by the next class period. The lab assignments were graded each week and returned to the students. The lab assignments were designed to reflect the recent lab lecture material and theoretical content since the last lab period. A total of 53 sophomores were enrolled in the study, with 23 in Lab, and 30 in Class.

2.4 Assessment

During the semester, 3 exams were conducted. Exam 1 occurred at approximately one-third of the duration of the semester, Exam 2 at two-thirds, and a comprehensive Final exam was given at the conclusion of the semester. The exams consisted of both multiple-choice questions and free-response/short-answer questions. The multiple choice questions allowed the student to choose from a set of 5 answers to correctly answer the question. For the free-response questions, a question was posed to the student and they were required to answer the question using provided statistical output, their theoretical knowledge, and their ability to interpret findings. Exam 1 was primarily focused on statistical concepts and theoretical understanding as very little statistical software content had been introduced at that point. Both Exam 2 and the Final had significant statistical software content as part of the examination. The intent was not to test students on the software itself, but to provide the students with statistical output from the software and ask theoretical questions that required the use of the output. The exams were intended to be challenging and covered the material in the course up to that point (from the previous exam for Exam 2), with the Final being a cumulative exam reviewing the entire semester.

Both sections were provided with very similar exams to allow for comparisons between sections. Only identical questions between the two sections were used for comparative analysis purposes to evaluate the differences in teaching methodologies. The sections were kept on the same pacing and sequencing of topics, and each section was tested at the same time-point during the semester. The same instructor performed all grading to minimize bias. With all grading being completed by the same individual, the risk of subjectivity was minimized (Adams et al., 2016) .

2.5 Surveys

Furthermore, surveys were used to assess the differences in sections. An initial survey was provided to students to assess their statistical background and experience, comfort with math related topics, anticipations for the course, and anxiety levels. An end of course survey was used to assess their progression from the initial survey, with additional questions to evaluate their development of statistical software proficiency. The surveys were derived from Hasbrouck et al. and included additional questions related to statistical software proficiency (Hasbrouck, Deniz, & Hodges, 2014). Hasbrouck et al. developed a set of pretest and posttest surveys to assess student attitudes related to statistics and the effectiveness of their instruction for business statistics courses. Those surveys were adapted for usage in this study and utilized a Likert-scale (scale of 1 to 5), in addition to a few binomial questions. (See supplemental information for surveys)

2.6 Analysis

Initial vs. final results on both the surveys and performance marks were used to discern differences in the effectiveness of teaching methodologies over the course of a semester’s worth of instruction. All personally identifiable information was removed from the data set to insure confidentiality, and the students were assured that their grades and surveys would not be reviewed until after the final grades were submitted for the semester. SPSS software was used for all analysis (SPSS version 22, IBM Armonk, NY). Categorical questions on the survey were evaluated with non-parametric tests (Wilcoxon Mann-Whitney), and Likert-style questions were evaluated with t-tests (Sullivan 2013). Comparisons between exam scores were compared for sections. An alpha value of .05 was used for all analyses.

3. Results

3.1 Participants

All students in the courses participated: 53 sophomores were enrolled, with 23 in Lab, and 30 in Class. The initial surveys were given during the first week of the course, and the final survey was given on the final day of the course. A total of 106 surveys were collected, with 2 for each student. All survey questions were answered and were used for analysis. Three exams were given for each student, for a total of 159 exams, all graded by the same instructor.

3.2 Initial Survey

The survey results were investigated for any differences between sections. Both Class and Lab had the same median math level achieved (Calculus) and neither had a median level of an introductory statistics course. The two sections were evaluated for any differences in statistics or math level with a Wilcoxon Mann-Whitney test and other comparisons were completed with independent t-tests (Table 1).

Table 1. Results of initial course survey. The survey questions are abbreviated for table presentation, full questions may be found in the supplemental information. The responses are all self-assessments. The first two questions were binomial responses compared with Wilcoxon Mann-Whitney tests and the values in parentheses indicates the percent choosing that answer. All other questions were evaluated with independent t-tests and the values in parentheses indicates the standard deviation.

Initial Course Survey Results			
	Class	*Lab*	*Comparison*
Prior Statistics Course			*p value*
(Yes / No)	No (80 %)	No (78 %)	0.878
Highest Level of Math	Calculus (50 %)	Calculus (65 %)	0.621
Knowledge prior to course			
(1=poor, 2=fair, 3=good, 4=very good, 5=excellent)			
Scientific Method	3.73 (0.87)	3.73 (0.81)	0.980
Central Tendency	2.83 (1.15)	2.04 (1.26)	**0.021**
Standard Deviation	2.33 (1.06)	2.30 (1.18)	0.926
Bias	2.43 (1.07)	2.39 (1.12)	0.890
Correlation Analysis	2.07 (1.08)	2.08 (0.94)	0.943
Linear Regression	2.00 (0.95)	2.61 (0.99)	**0.027**
t-Test	1.47 (0.97)	1.61 (0.94)	0.596
ANOVA	1.10 (0.40)	1.09 (0.29)	0.896
SPSS usage	1.03 (0.18)	1.08 (0.29)	0.412
Anticipations			
(1=strongly disagree, 2=disagree, 3=neither, 4=agree, 5=strongly agree)			
Wish I did not have to take	2.67 (0.84)	2.39 (0.94)	0.268
Nervous about course	3.13 (1.19)	2.65 (0.83)	0.106
Statistics is complicated	3.33 (0.88)	2.86 (0.86)	0.062
Excited about statistics	3.03 (0.89)	3.34 (0.88)	0.207
Statistics will be useful in degree	4.13 (0.97)	4.04 (0.71)	0.710
Statistics will be useful professionally	3.46 (1.07)	3.56 (0.89)	0.724
Most would benefit from statistics	3.70 (0.79)	3.69 (0.76)	0.984
Math is not comfortable	2.53 (1.12)	2.52 (1.16)	0.972
I anticipate doing well	3.53 (0.82)	3.78 (0.52)	0.208

There were no significant differences between sections on the initial course survey other than the questions related to central tendency and linear regression. However, in anecdotal conversations with students, it is believed that those differences were driven by student's lack of familiarity with the terms, with some believing they were standard math concepts.

Questions related to theoretical statistical concepts were grouped (questions concerning familiarity with scientific method, central tendency, standard deviations, bias, correlation analysis, linear regression, t-tests, ANOVA) and evaluated for differences between sections, and no difference was found (p=.949) via an independent t-test. The results of the initial course survey suggested similar samples of students.

3.3 End Survey

An end-of-course survey was given to students on the last day of class to assess their self-assessed knowledge and reflections of the course experience (Table 2).

Table 2. Results of end of course survey. The survey questions are abbreviated for table presentation, full questions may be found in the supplemental information. The responses are all self-assessments. The final question concerning which course students would recommend was compared with a Wilcoxon Mann-Whitney test and the values in parentheses indicates the percent choosing that answer. All other questions were evaluated with independent t-tests and the values in parentheses indicates the standard deviation.

End of Course Survey			
	Class	*Lab*	*Comparison*
Knowledge at end of course			*p value*
(1=poor, 2=fair, 3=good, 4=very good, 5=excellent)			
Scientific Method	4.36 (0.67)	4.00 (0.67)	0.067
Central Tendency	4.28 (0.71)	3.91 (0.90)	0.105
Standard Deviation	4.46 (0.74)	4.17 (0.71)	0.210
Bias	4.21 (0.68)	4.04 (0.71)	0.387
Correlation Analysis	4.29 (0.81)	4.22 (0.74)	0.756
Linear Regression	4.11 (0.78)	3.87 (0.97)	0.338
t-Test	3.89 (0.96)	3.7 (0.97)	0.471
ANOVA	3.71 (1.01)	3.83 (1.03)	0.699
Knowledge of Software usage			
(1=poor, 2=fair, 3=good, 4=very good, 5=excellent)			
Data Entry	3.64 (0.83)	3.91 (0.73)	0.228
Graphing	3.89 (0.83)	3.96 (0.82)	0.786
Exploring Data	3.68 (0.98)	3.87 (0.92)	0.481
Evaluating Normality	3.85 (0.93)	3.87 (0.91)	0.962
Correlation Analyses	3.89 (0.99)	3.91 (0.90)	0.940
Linear Regression	3.89 (0.91)	3.83 (0.98)	0.803
t-Tests	3.43 (1.03)	3.61 (0.83)	0.504
ANOVAs	3.42 (1.03)	3.74 (0.86)	0.257
Reflections			
(1=strongly disagree, 2=disagree, 3=neither, 4=agree, 5=strongly agree)			
Glad I took course	3.93 (0.89)	3.74 (0.96)	0.472
Statistics is complicated	3.36 (1.02)	3.57 (0.95)	0.459
Excited about statistics	3.21 (0.88)	3.17 (1.07)	0.883
Statistics will be useful in degree	4.11 (0.74)	4.13 (0.76)	0.912
Statistics will be useful professionally	3.50 (0.75)	3.13 (1.22)	0.189
Most would benefit from statistics	3.78 (0.63)	3.39 (1.12)	0.119
Math is not comfortable	2.61 (1.22)	2.78 (1.24)	0.616
Believe software will be useful in major	3.46 (1.23)	3.70 (1.26)	0.512
Believe software will be useful professionally	2.93 (1.02)	2.74 (0.92)	0.492
Too much time was spent on software	2.67 (0.90)	2.52 (0.79)	0.518
Too little time spent on software	3.00 (0.86)	2.70 (0.85)	0.206
Software was useful to statistics understanding	3.54 (1.10)	3.78 (0.85)	0.384
Would recommend future students take:	Class (67 %)	Lab (65 %)	**0.020**

The student's perceived theoretical understanding of statistical concepts between sections was no different (p=.252) with grouped theoretical questions. The same was true of their perceived understanding of grouped statistical software questions (p=.578). A final question on the survey was asked to assess which class the students would recommend for future students to take: a lab-based instruction or class based instruction. There was a significant difference between the two sections with Lab recommending the lab based instruction significantly more than Class (32% more, p=.02 via Wilcoxon Mann-Whitney).

3.4 End vs. Initial Survey

With both sections grouped together, the progression of theoretical understanding was evaluated, and the students reported significantly better theoretical understanding of statistical concepts (p<.001), and significantly better understanding of using SPSS (p<.001) from the beginning to the end of the semester. The responses within each section were then evaluated (Table 3), and the difference in theoretical understanding for Lab changed by 1.73 pts (p<.001) and the software understanding changed by 2.75 pts (p<.001). For Class, the theoretical understanding changed by 1.87 pts (p<.001) and the software understanding improved by 2.68 pts (p<.001).

Table 3. End-of-course survey responses compared to pre-course survey responses. The comparisons for Class and Lab are between their respective end/pre surveys.

End vs. Initial Surveys				
Mean diff (end-initial) with SD				
	Class	*Comparison p value*	*Lab*	*Comparison p value*
Knowledge				
(1=poor, 2=fair, 3=good, 4=very good, 5=excellent)				
Scientific Method	0.53 (0.88)	**0.003**	0.26 (0.86)	0.162
Central Tendency	1.39 (1.20)	**<.001**	1.87 (1.45)	**<.001**
Standard Deviation	2.07 (1.21)	**<.001**	1.87 (1.25)	**<.001**
Bias	1.71 (1.08)	**<.001**	1.65 (1.19)	**<.001**
Correlation Analysis	2.17 (1.52)	**<.001**	2.13 (1.14)	**<.001**
Linear Regression	2.03 (1.17)	**<.001**	1.26 (1.28)	**<.001**
t-Test	2.39 (1.34)	**<.001**	2.08 (1.38)	**<.001**
ANOVA	2.61 (1.03)	**<.001**	2.74 (1.01)	**<.001**
Statistics is complicated	0.07 (1.21)	0.758	0.70 (0.93)	**0.002**
(1=strongly disagree, 2=disagree, 3=neither, 4=agree, 5=strongly agree)				

The comparison indicates that students felt their knowledge of the topics were improved over the course of the semester, with both sections reporting primarily significant improvements.

3.5 Grade Comparison

The next point of comparison between sections was that of comparable exam questions. Questions consisted of both multiple choice (MC) and free-response (FR). The two types of questions were evaluated separately to compare sections (Table 4).

Table 4. Grades by section (Lab and Class) as percentage correct for both multiple choice (MC) and free response (FR) questions (± standard error) .

Grades on Exams by section				
		Lab	*Class*	*p value*
Multiple Choice	Exam 1	79.04 (2.53)	77.41 (2.45)	0.652
	Exam 2	82.61 (2.71)	82.44 (2.47)	0.965
	Final	76.55 (2.01)	76.90 (2.72)	0.920
Free Response	Exam 1	59.22 (5.24)	70.01 (4.51)	0.124
	Exam 2	76.19 (3.63)	72.01 (3.54)	0.965
	Final	78.52 (3.44)	75.04 (3.93)	0.520

The change in grades on the MC section were compared between Exam 1 and the final exam. Lab grades dropped by 2.04% (SE 2.57) from Exam 1 to Exam 2 and Class dropped by 0.30% (SE 2.24). The differences in sections were compared with an independent t-test. Based on Levene's test, equal variances could be assumed, and there was no difference between sections (p=.610).

Similarly, the change in FR percentage from the first exam to the Final were compared. Lab grades increased by 19.2% (SE = 4.67) and Class increased by 3.67% (SE = 3.68). This difference was significant with Lab scoring significantly better than Class, (p=.011) (Figure 1).

The differences between Exam 1 and 2 were also compared. There was a significant increase in the FR points earned in section Lab vs Section Class (16.02% (SE = 4.22) vs. 1.95% (SE = 4.10)) with p = .023). However, there was no difference in MC between section Lab (3.99% (SE = 2.81)) vs. section Class (5.04% (SE = 2.62)) with p = .789.

Finally, the differences between Exam 2 and the Final were evaluated. There was no difference in MC between Lab and Class (-6.06% (SE 1.89)) vs. 6.77%(SE2.13) with p=.809). Similarly, there was no difference in FR questions between Lab and Class (2.33% (SE 2.48) vs. 1.00 % (SE 2.30) with p=.698).

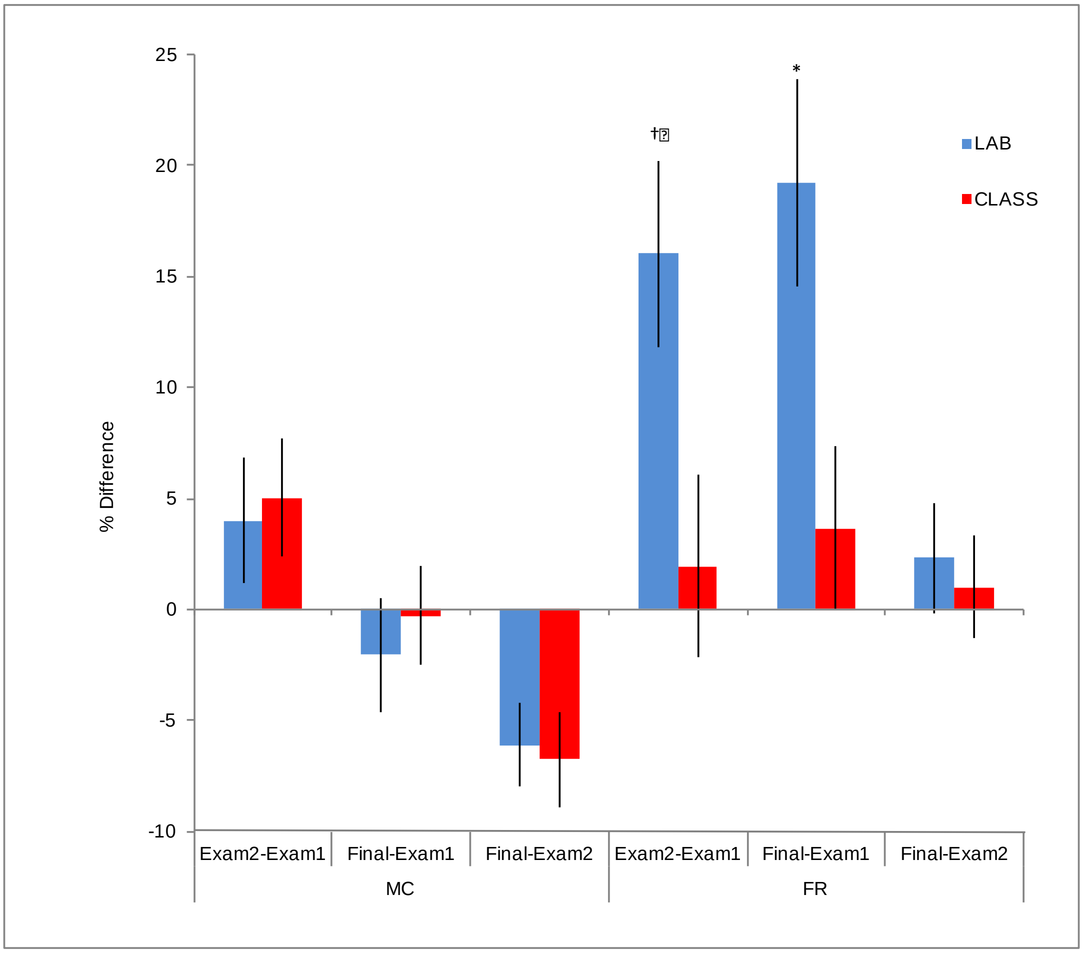

Figure 1. Percent difference in scores between exams by section. Scores are broken into multiple choice (MC) and free response (FR). (*) indicates that Lab improved significantly more between the first exam and the final exam when compared to Class (p= .011). (†) indicates that Lab improved significantly more between Exam 1 and 2 as compared to CLASS (p=.023)

Lastly, the change in student's self-assessed knowledge level from the initial to end survery was computed and compared to the change in the overall percentage score from Exam1 to the Final (Figure 2).

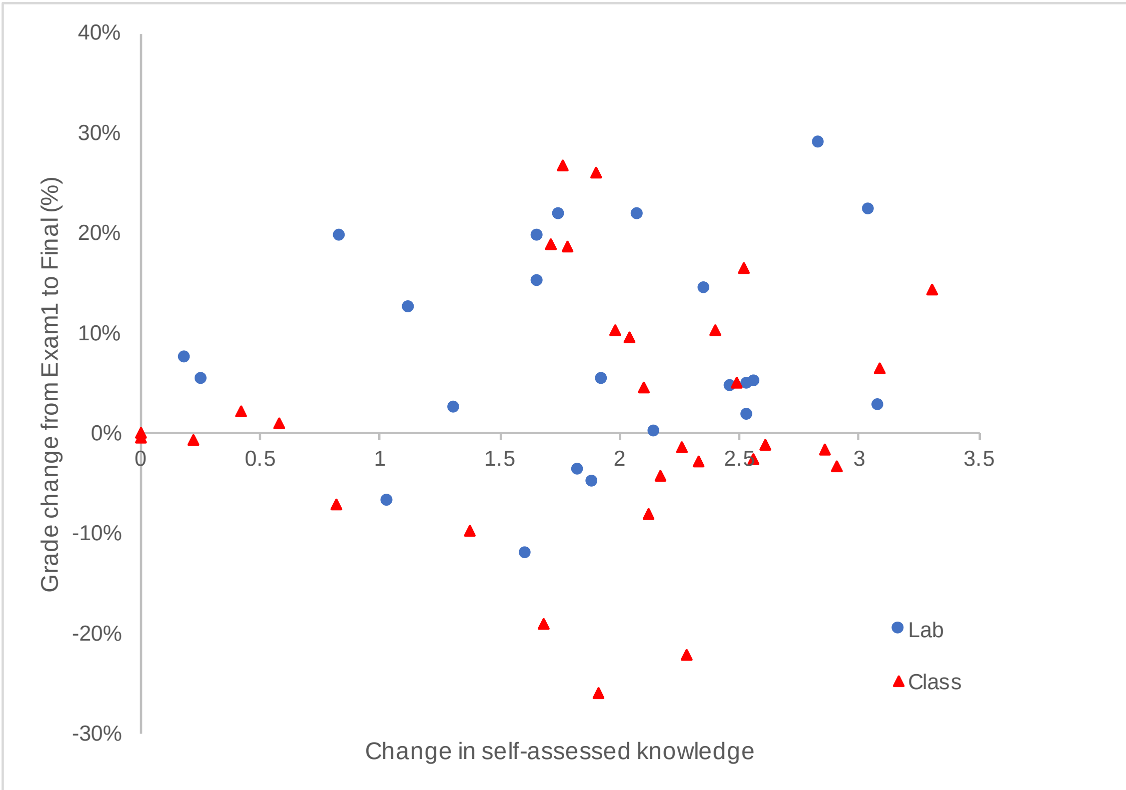

Figure 2. Relationship between self-assessed knowledge changes, and exam performance changes over the semester. The change in the exam scores was computed from both average value of both the MC and FR questions. The self-assessed knowledge scores were the averages of the knowledge sections of the surveys (Table 1 & 2) on a 5 point scale

The figure indicates the larger increase in exam scores for the Lab section as compared to class. Although many of the Lab participants still viewed their knowledge as little to no improvement, their exam improvement indicates an improvement. Conversely, many Class participants indicated large increases in their self-assessed knowledge despite the decrease in their exam scores.

4. Discussion

The surveys were used to assess the student's background at the beginning of the semester and to measure their perception of their own progression at the end of the semester. The two sections (Lab and Class) were very similar in their preparatory courses leading up to the current course, and had similar levels of anxiety and expectations for the course. On the end-of course survey students evaluated their semester. Interestingly, the Lab students were more likely to recommend the Lab section to future students, despite their own perception of their software skills being no better than that of Class students. Both sections reported that their understanding of both theoretical and applied software concepts improved significantly over the course of the semester. Both sections expressed anxiety and concern over the course initially. This finding was in agreement with Stork (Stork, 2003) and Chew (Chew & Dillon, 2014) who found that statistics courses are anxiety provoking. Because learning new statistical software is consistently a challenging component of this course, it is anticipated that continued use of lab time to teach the software will increase the approachability of the class. There is justification for this expectation based on the findings of Ciftci et al. who showed that the use of computer instruction reduced the level of anxiety on the part of students and improved their general opinion of statistics (Ciftci, Karadag, & Akdal, 2014). Even other statisticians have made calls for inclusion of more computational components to statistical instruction (Cobb, 2015). The inclusion of real world questions into the lab component may even encourage enrollment of students from broader backgrounds, as they see the course as more applicable (Cooper & Dierker, 2017) .

The primary aim of this study was to evaluate the value effectiveness of trading class time for a “hands-on” approach, to teach statistical concepts and software. Each of three exams (Exam 1, Exam 2, and Final) had both multiple choice and free response portions, with the majority of questions being identical between sections. For the first exam, no SPSS concepts were introduced or tested. Thus, it was useful for Exam 1 to serve as a baseline. On the first exam, Class and Lab scored similarly on the MC portions, but Lab scored significantly lower on the FR portion.

The two sections continued to score similarly on all of the MC questions for the remaining exams. However, while Class FR grades remained relatively unchanged over the semester, Lab FR grades changed significantly. The FR grades for Lab increased throughout the semester, and the improvements for Lab from their first exam were significantly better than those for Class. Therefore, while Class demonstrated an ability to learn new concepts and succeed in testing on those concepts, Lab was able to make much larger gains in their performance on exams. This is reinforced in Figure 2, where despite each section having similar perceptions of their knowledge level, Lab showed a significantly higher improvement in their scoring on identical exam questions. The use of computers in a lab setting likely effected the observed significant increases in theoretical knowledge as well. The use of computers, and their ability to provide instant feedback while problem solving has been shown to increase outcomes in learning math (Fyfe & Rittle-Johnson, 2016). The significant improvement on the exam performance reflects the benefits produced by conducting a lab period. The students in Lab, spent more hands-on time in a lab setting using statistical software, and this translated to better theoretical understanding of statistical concepts.

The findings presented here are in agreement with those found by Wilson (Wilson, 2013) who used a flipped-classroom approach to aid in the teaching of statistics. Using computers in a lab setting that models a “flipped classroom” may increase motivation on the part of the students (Thai, De Wever, & Valcke, 2017). Similar to this study, having students perform more applied work during the class period resulted in better learning and understanding. Anxiety about math related concepts can be an impediment to learning (Núñez-Peña, Suárez-Pellicioni, Bono, 2013), and it is believed that allowing students to perform statistical work and problems with the instructor present helps to reduce this barrier to learning. This is also in agreement with findings in other areas of science, such as biology, where students show a desire for a lab component to a course when offered (Adams et al. 2016). While the lab based components here were used to reinforce statistical concepts, they were often analogous to typical experimental type labs in that the students had a problem they needed to solve, they had to figure out how to answer the problem, and then had to try their solutions. In an evaluation of the teaching of economics, it has been shown that including experimental methods and technology into the course improves both the learning outcomes and the student experience, much like the current study (Allgood, Walstad, & Siegfried, 2015).

In the implementation of the current study, the Lab section generated more assignments for the students to complete each week. While these proved beneficial to the student’s learning and progress through the course, they did require more grading time. If these components could be linked to an online grading platform, with students perhaps choosing the correct answer after completing their analyses, this could speed the grading process (Chow, 2015).

One of the limitations of this study is the selection of students for each class. The students were not assigned to each group, but were selected based upon their registration for the section of the class. However, the students were given no prior indication of how each section would be taught and therefore did not self-select the type of teaching approach they preferred. Similarly, the poorer performance of Lab students on the FR of Exam 1 indicates that these students were not as well-prepared for the course. However, these students outperformed the Class students on subsequent exams, indicating that the style of teaching (lab-based) was able to overcome differences in proclivities for the subject matter. The final interesting note is that on the end-of-course survey, Lab students did not believe that they were any more well-versed in statistical concepts or software than the Class students. Despite the marked improvement in their grades, and the evident improvement from the standpoint of the instructor, the students did not feel confident in their improved knowledge. The instructor should provide more feedback to the students in regards to their gains and improvements.

The findings here have been shared with other faculty for incorporation in the design of their courses in both statistics and other areas. Any course that has a technical component would likely benefit from using a lab based period of teaching. Often, the typical 2-4 hr time block devoted to a lab instruction period is a limiter for incorporating lab instruction. However, as was shown here, sacrificing one of the available classroom periods each week and using a regular class period as a lab instruction time can prove very beneficial. Anecdotally, it was observed that the students in the Lab course felt much more comfortable in their use of statistical concepts, and that that difference compared to Class students persisted in subsequent semesters. Subsequent offerings of this course have provided a lab-based component and students have continued to thrive in this environment.

5. Conclusion

Improving instruction in an introductory statistics course is an ongoing challenge, as many students approach the course with trepidation. This study aimed to evaluate whether the substitution of classroom time with laboratory based teaching would improve learning outcomes. Despite the fact that the Lab section had less time for theoretical instruction, they still performed significantly better on assessments than the Class section that had more instructional periods dedicated to theoretical instruction. Further research could focus on increasing the conversion of classroom time to laboratory time for statistics instruction. It may be beneficial to do 50%, 75% or more of the class as laboratory based, however, there may be a tipping point where theoretical understanding suffers. Most teachers are hesitant to give-up classroom time, however, the results of this study show that replacing some classroom periods with laboratory periods is beneficial for learning outcomes in an introductory statistics course.

References

Adams, A., Garcia, J., & Traustadóttir, T. (2016). A Quasi Experiment to Determine the Effectiveness of a "Partially Flipped" versus "Fully Flipped" Undergraduate Class in Genetics and Evolution. *CBE life sciences education, 15*(2). https://doi.org/10.1187/cbe.15-07-0157

Allgood, S., Walstad, W.B., & Siegfried J.J. (2015). Research on teaching economics to undergraduates. *Journal of Economic Literature*, *53*(2), 285-325. https://doi.org/10.1257/jel.53.2.285

Bartz, A.E., & Sabolik, M.A. (2001). Computer and software use in teaching the beginning statistics course. *Teaching of Psychology, 28*(2), 147-149. https://doi.org/10.1207/S15328023TOP2802_16

Chance, B., Ben-Zvi, D., Garfield, J., & Medina, E. (2007). The Role of Technology in Improving Student Learning of Statistics. *Technology Innovations in Statistics Education, 1*(1).

Chew, P.K., & Dillon, D.B. (2014). Statistics Anxiety Update: Refining the Construct and Recommendations for a New Research Agenda. *Perspectives on psychological science, 9*(2), 196-208. https://doi.org/10.1177/1745691613518077

Chow, A.F. (2015). Online homework impact in undergraduate mathematics and business statistics courses. *Educational Studies, 41*(3), 244-248. https://doi.org/10.1080/03055698.2014.961902

Ciftci, S.K., Karadag, E., & Akdal, P. (2014). Instruction of Statistics via Computer-Based Tools: Effects on Statistics' Anxiety, Attitude, and Achievement. *Journal of Educational Computing Research, 50*(1), 119-133. https://doi.org/10.2190/EC.50.1.f

Cobb, G. (2015). Mere renovation is too little too late: We need to rethink our undergraduate curriculum from the ground up. *The American Statistician, 69*(4), 266-282. https://doi.org/10.1080/00031305.2015.1093029

Cooper, J., & Dierker, L. (2017). Increasing Exposure to Programming: A Comparison of Demographic Characteristics of Students Enrolled in Introductory Computer Science Programming Courses vs. a Multidisciplinary Data Analysis Course. *International Research in Higher Education, 2*(1), 92-100 https://doi.org/10.5430/irhe.v2n1p92

Field, A. (2013). Discovering statistics using IBM SPSS, 4 edn. *SAGE*, Los Angeles, CA.

Fyfe, E.R., & Rittle-Johnson, B. (2016). The benefits of computer-generated feedback for mathematics problem solving. *Journal of experimental child psychology, 147*, 140-151. https://doi.org/10.1016/j.jecp.2016.03.009

Hasbrouck, R.B., Deniz, B., & Hodges, H. (2014). Examining Potential Latent Constructs in Teaching a Business Statistics Course: Clustering Responses from Attitudinal Survey Data. *Journal of Business and Economics, 5*(5), 708-717.

Karp, D.R. (1995). Using SPSS for Windows to Enhance, Not Overwhelm, Course Content. *Teaching Sociology, 23*(3), 234-240. https://doi.org/10.2307/1319215

Larwin, K., & Larwin, D. (2011). A meta-analysis examining the impact of computer-assisted instruction on postsecondary statistics education: 40 years of research. *Journal of Research on Technology in Education*, *43*(3), 253-278. https://doi.org/10.1080/15391523.2011.10782572

Mills, J.D. (2002). Using computer simulation methods to teach statistics: A review of the literature. *Journal of Statistics Education*, *10*(1), 1-20.

Mills, J.D., & Johnson, E.L. (2004). An Evaluation of ActivStats for SPSS for Teaching and Learning. *The American Statistician, 58*, 254-258. https://doi.org/10.1198/000313004X1530

Núez-Pe-a, M.I., Suárez-Pellicioni, M., & Bono, R. (2013). Effects of math anxiety on student success in higher education. *International Journal of Educational Research, 58,* 36-43. https://doi.org/10.1016/j.ijer.2012.12.004

Rodgers, P.H., & Manrique, C. (1992). The Dilemma of Teaching Political Science Research Methods: How Much Computers? How Much Statistics? How Much Methods?. *PS: Political Science & Politics*, *25*(02), 234-237. https://doi.org/10.1017/S1049096500035472

Sherwood, C., & Kwak, D.W. (2017). New insights into an old problem–enhancing student learning outcomes in an introductory statistics course. *Applied Economics*, *49*(56), 1-11. https://doi.org/10.1080/00036846.2017.1332750

St. George A. (1978). A Computer-Integrated Course in Research Methods. *Teaching Sociology, 5*(4), 423-444. https://doi.org/10.2307/1317281

Stork D. (2003). Teaching statistics with student survey data: A pedagogical innovation in support of student learning. *Journal of Education for Business, 78*(6), 335-339. https://doi.org/10.1080/08832320309598623

Thai, N.T.T., De Wever, B., & Valcke, M. (2017). The impact of a flipped classroom design on learning performance in higher education: Looking for the best "blend" of lectures and guiding questions with feedback. *Computers & Education, 107,*113-126. https://doi.org/10.1016/j.compedu.2017.01.003

Wells, M. (2006). Teaching notes: Making statistics 'real' for social work students. *Journal of Social Work Education, 42*(2), 397-404. https://doi.org/10.5175/JSWE.2006.200400466

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. https://doi.org/10.1177/0098628313487461

Zhang, Y., Zhou, L., Liu, X., Liu, L., Wu, Y., Zhao, Z., ... & Yi, D. (2015). The effectiveness of the problem-based learning teaching model for use in introductory Chinese undergraduate medical courses: a systematic review and meta-analysis. *PloS one, 10*(3), e0120884. https://doi.org/10.1371/journal.pone.0120884