The Relationship Between Learning Mode and Student Performance in an Undergraduate Elementary Statistics Course in the United States

John C. Griffith, Emily K. Faulconer, and Bobby L. McMasters
Department of STEM Education, Embry-Riddle Aeronautical University – Worldwide Campus

Abstract

Faculty have conducted many studies on the relationship between learning mode and student performance but few researchers have evaluated final grades, grade distribution, and pass rates in a sophomore introductory statistics course with a non-traditional student population who self-selected the learning mode from among different course sections. Accordingly, we examined 307 end-of-course grades from four different modes of instruction: (a) online, (b) videosynchronous learning classroom, (c) videosynchronous learning home, and (d) traditional classroom in an introductory statistics course. All data on grades, which included pass rate and grade distribution, were collected from the nine-week January 2019 term. All learning modes used the same text, syllabus, assignments, quizzes, and tests. In this study, learning mode was not significantly related to end-of-course score, final grade distribution, or pass rate. Future researchers should explore the impacts of gender, instructor quality, different term lengths, and the standardized use of textbooks and syllabi on student performance when exploring the impact of learning mode on grades, grade distribution, and pass rates.

Keywords: distance learning, online education, quality in higher education, student performance, grade distribution
Introduction

For several years, the option to complete undergraduate coursework and degree programs online has been increasing in U.S. institutions (Online Learning Consortium, 2016). Educators and administrators who wished to meet the needs of current and future students were the audience for studies evaluating the impact of delivery mode on student performance. However, since the coronavirus pandemic emerged in 2020, the number of academic courses offered online has accelerated, and there are even more administrators and educators grappling with the effective use of distance learning tools. There are several modalities within the realm of online education, including asynchronous and synchronous options. Because the literature has shown mixed results regarding equivalence in student outcomes based on modality (Jahng et al., 2007; Nguyen, 2015; Xu & Jaggars, 2013), it is vital to continue to study the influence of learning mode on student performance. Comparisons between synchronous and asynchronous learning modes can be difficult due to possible confounding variables such as learning management systems, texts, syllabi, and other delivery variables. We believe the mixed findings of studies comparing synchronous and asynchronous learning are possibly due to the settings from which the data were drawn. In other words, it is important to ensure that synchronous and asynchronous courses used for comparison share similar structures such as learning management systems, textbooks, syllabi, grade weightings for assignments, tests, quizzes, homework, and other assignments. Controlling for these variables is done through standardized course delivery so that data generated does not reflect the impact of different course delivery aspects.

Literature Review

The modality (delivery media) of a course does not consistently influence student outcomes. Some meta-analyses have supported no significant difference between student outcomes in online and traditional courses (Jahng et al., 2007; Lundberg et al., 2008), while others have reported significant differences (Means et al., 2009; Sitzmann et al., 2006; Williams, 2006). The outcomes investigated in these studies vary, including short term metrics such as grade distribution, pass rate, and withdrawal, and long-term metrics such as degree completion. For example, a study of community college students reported decreased student outcomes in online courses versus traditional courses; but it also reported that students who had completed some courses online were more likely to complete an associate's degree or transfer to an institution granting 4-year degrees (Johnson & Mejia, 2014). Shea and Bidjerano (2019) found similar results in large scale studies examining data on over 40,000 students from 2013 to 2017, concluding that 40% online to 60% classroom ratio as the best mix to earn an associate's degree.

Some researchers have reported that persistence in online courses has been problematic (Atchley et al., 2013; Jaggars et al., 2013; Murphy & Stewart, 2017). One possible reason is that students report lower instructor presence in online courses (Jaggars, 2014). However, Shea and Bidjerano (2016) conducted an investigation of national (rather than state) trends using data from the 2003-2004 academic year and found that online students had higher rates of associate degree attainment or transfer and lower dropout rates than their classroom-only student counterparts. This trend held true at the 6-year point of a student's academic career as well (Shea & Bidjerano, 2016).
While meta-analyses are imperative for analyzing trends by modality in undergraduate education, it is possible that significant differences may be seen within individual disciplines. There is limited literature exploring modality differences within, for example, undergraduate statistics, with much of the work being somewhat dated. Some studies reported that modality had no significant effect regarding the final course grade (McLaren, 2004; Summers et al., 2005) and successful completion (Dotterweich & Rochelle, 2012; Rochelle & Dotterweich, 2007). Other studies reported significant differences by modality, including higher grades for students completing the course in the traditional modality as compared to online and/or hybrid options (Lawrence & Singhania, 2004; Scherrer, 2011).

Interestingly, only one study reported significant differences by modality based on gender, with female students performing worse online than in the traditional classroom and no significant differences reported for male students in either modality (Flanagan, 2012). Online statistics modality has demonstrated a higher withdrawal rate (McLaren, 2004; Zimmerman & Austin, 2018). Test anxiety can predict course completion for the online modality, while self-concept can predict course completion for the traditional modality (Zimmerman & Austin, 2018).

These studies did reveal potential moderating variables for student outcomes in various modalities. Lower student satisfaction was reported for the online modality, specifically regarding instructor’s explanations, class discussions, quality of problems, and evaluation and grading, even when the instructor was consistent and discussions and problems used in the course were equivalent (Summers et al., 2005). This is contrasted with Scherrer (2011) who reported no differences in satisfaction by modality. Dotterweich and Rochelle (2007) identified grade point average as a predictor for success in undergraduate statistics, regardless of modality, and noted that students who had a history of course failure did better in the traditional modality. Modality has not been linked to self-concept and value of statistics education (Zimmerman & Austin, 2018).

Bourdeau et al. (2018) and Roberts et al. (2019) conducted studies in different disciplines of English and Research Methods respectively. Their argument was that comparisons on student performance factors of end of course grades, pass rates, and grade distribution were made possible due to the standardized method of course delivery, which included learning management system, texts, syllabi, tests, quizzes, homework, and other assignments.

**Modes of Learning**

In this study, student assignments were constant across all modes of learning controlling for another possible source of confounding variables. The university used four primary modes of learning: classroom, online, synchronous video home, and synchronous video classroom (Bourdeau et al., 2018; Roberts et al., 2019).

**Traditional Classroom Instruction (In Person)**

The traditional legacy teaching format, sometimes referred to as *on ground*, is any form of instructional interaction that occurs in person and in real time. Before the advent of audio, video, and Internet technologies that allowed people to interact from different locations and at different times, student-instructor interactions occurred in the same place and at the same time.
Online

Online learning is education that takes place over the Internet. Sometimes described as e-learning, online learning is just one type of distance learning. Distance learning has a long history, and there are several types currently available. For the purposes of this research, courses referred to as online learning were asynchronous.

Videosynchronous Home (Blended)

The term blended learning is generally applied to the practice of using both online and in-person learning experiences when teaching. In a blended learning course, students attend a class taught by a teacher in a traditional classroom setting, while also completing online components outside the classroom. In-class time is supplemented by online learning assignments, and students learn the same topics online as they do in class. Asynchronous learning describes instruction and learning that do not occur in the same place or at the same time. The term applies to various forms of digital and online learning in which students learn, on their own, at their home, with instruction that is delivered via the Internet (Adobe Connect) in real time. For the purposes of this research, blended learning included 3 hours and 20 minutes of synchronous sessions and 90 minutes of online activities each week of the 9-week term.

Videosynchronous Classroom (Blended)

Videosynchronous classroom combines the live, in person traditional experience with additional students located in a “remote” classroom environment. Similar to the Videosynchronous Home mode, blended learning included 3 hours and 20 minutes of synchronous sessions and 90 minutes of online activities each week of the 9-week term.

All four modes of learning used the same textbook, discussion board prompts, homework, quizzes, and tests. A standardized syllabus was used for online (distance) and on ground (classroom) offerings of the course. Because of the uniformity by which the course was delivered, regardless of mode, the possible confounding variables of pedagogical method employed by the instructor were controlled (Bourdeau et al., 2018; Lou et al., 2006; Roberts et al., 2019).

There is a lack of consensus in the literature regarding the influence of modality on student outcomes in undergraduate statistics. This gap calls for studies examining universities that take a standardized approach to their courses, no matter what the delivery mode. Furthermore, online education is rapidly evolving, and a current perspective on this relationship is warranted.

Purpose

The purpose of this research was to explore student performance in multiple modes of instructional delivery. A medium-sized university offered the opportunity to minimize confounding factors by delivering an undergraduate statistics course in the following modes of instruction: online, videosynchronous learning classroom, videosynchronous learning home, and traditional classroom. The research question was: Is student performance in an undergraduate statistics course impacted by learning mode in any meaningful way? To evaluate the problem statement, the following hypotheses were proposed.
The alternative hypotheses are:

Ha1. Student end-of-course scores are not all statistically equivalent among classroom, online, and videosynchronous modes of learning.

Ha2. There is a statistically significant relationship between grade distribution and learning mode.

Ha3. There is a statistically significant relationship between student pass rates and learning mode.

Method
We examined pass rates and grade distribution in an introductory statistics course that had four different modes of instruction: online, videosynchronous learning classroom, videosynchronous learning home, and traditional classroom. The university we studied offered five major terms a year. However, to reduce the probability of confounding variables (e.g., seasonal), only one term was studied.

Study Population Sampling
Data were taken from a mid-sized university that is rated as “more selective” (U.S. News and World Report, 2019). In 2019, approximately 90% of classes at this institution met in non-traditional modalities, including asynchronous online and video synchronous. Students enrolled in the studied sections were non-traditional students, with an average age of 34 years and 80% having a military affiliation (currently enlisted or veteran). Students selected the delivery mode they would use to take the course.

Statistics Course Used in the Study
We examined end-of-course scores, grade distribution, and pass rates in an introductory statistics course. The course covered basic descriptive and inferential statistics using discussion boards and real-world assignments—practical exercises that had been successful in different teaching and learning platforms. Topics included types of data, sampling techniques, measures of central tendency and dispersion, elementary probability, discrete and continuous probability distributions, sampling distributions, hypothesis testing, confidence intervals, and simple linear regression.

The course was augmented by Pearson’s MyLabStat, providing homework, quizzes, and tests. Canvas was the learning management system used in all modes of instruction (Instructure, 2020; Pearson, 2020).

Discussion Boards
Student assignments included participation in eight discussion boards which accounted for a total of 15% of the final grade. Discussions were used to reinforce concepts taught in corresponding modules. Students were required to make one initial post to a discussion prompt, and then to make two additional postings to initial posts made by classmates.
Written Assignments
Two discussion assignments requiring the use of spreadsheets to solve practical problems were precursors to written assignments which accounted for 15% of the overall grade. The first assignment required students to calculate means and standard deviations of airshow scores for eight different types of aircraft. Students then had to write a one-page letter announcing the winner of the air show. The second assignment, involving the effective use of spreadsheets, focused on mean, standard deviation, $z$ scores, and probability when comparing flying squadron costs involving several aircraft and bases. The follow-up assignment was for students to write a one-page letter indicating which bases were the most and least cost-efficient.

Homework Assignments
Nine homework assignments were administered through Pearson’s MyLabStat accounting for 10% of the total grade (Pearson, 2020). Homework assignments typically contained 20 to 30 questions and could be attempted up to three times by students to improve their grades. Homework assignment retakes used randomized problems meaning the students would not see the same homework questions twice. Homework assignments included a direct link to instructors where a student could ask a question about a particular problem. The instructor would receive a screenshot of the problem and could help the student by contacting the student to work through the problem or showing a step-by-step example of how to solve a similar problem. Students also had the option to view a step-by-step example of how to solve a particular problem through the MyLabStat software.

Quizzes and Exams
Eight module quizzes followed corresponding homework and were administered through Pearson’s MyLabStat (2020). Quizzes were worth a total 20% of the overall grade. Quizzes were timed, with a 90-minute limit, and typically comprised of 10 questions. Students did not have the ability to ask for help during quizzes in order to accurately ascertain student retention.

The cumulative midterm examination, administered through MyLabStat, was worth 20% of the final grade. Students had one opportunity to take the midterm with a 120-minute time limit. Students did not have the option of seeing examples of particular problems or asking instructors for help. The midterm exam covered material from the first four modules of the course.

The final examination was also administered through MyLabStat and was worth 20% of the final grade. Students had a time limit of 120 minutes. Like the midterm examination, students did not have the option of seeing examples of how to solve problems, and they could not ask instructors for assistance during the final examination. The final examination all content in the course.

Data Collection
The January 2019 term was randomly selected to examine end-of-course grades. End-of-course scores were pulled in the aggregate (post hoc) from the learning management system after the term had ended. Students were not otherwise surveyed or contacted regarding the study. These scores provided the needed data to evaluate all three hypotheses. This term yielded 307 final grades from the four different learning modes (Gay et al., 2009; Gould & Ryan, 2012; Triola 2018). No student identification characteristics were included.
the data set. The university institutional review board exempted this study from further formal institutional review.

**Analysis**

A total of 307 end-of-course grades were examined from the January 2019 term for the elementary university statistics course. The hypotheses were evaluated using a one-way analysis of variance (for end-of-course numerical scores) and chi-square tests for (letter) grade distribution and pass rates. The traditional alpha level of .05 was changed to a Bonferroni corrected alpha level of .017 because all three hypotheses were tested using the same data set. This conservative change was made to avoid falsely rejecting the null hypothesis (Gay et al., 2006; Gould & Ryan, 2013).

**Results**

The first hypothesis concerned student end-of-course scores in classroom, online, and videosynchronous learning (home and classroom) modes and suggested that scores in the four modes would not be statistically equivalent. Table 1 shows descriptive data broken out by learning mode.

<table>
<thead>
<tr>
<th>Learning mode</th>
<th>N</th>
<th>M</th>
<th>SD</th>
<th>Quartile 1</th>
<th>Quartile 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Classroom</td>
<td>22</td>
<td>88.98</td>
<td>8.78</td>
<td>89.00</td>
<td>94.80</td>
</tr>
<tr>
<td>Videosynchronous classroom</td>
<td>12</td>
<td>88.21</td>
<td>8.87</td>
<td>82.75</td>
<td>94.65</td>
</tr>
<tr>
<td>Videosynchronous home</td>
<td>43</td>
<td>79.43</td>
<td>25.00</td>
<td>79.20</td>
<td>92.30</td>
</tr>
<tr>
<td>Online</td>
<td>230</td>
<td>84.17</td>
<td>19.76</td>
<td>82.50</td>
<td>94.70</td>
</tr>
</tbody>
</table>

Using an alpha level of .05, Levene’s test for homogeneity of variance resulted in a high enough p value to conclude the variances between the learning modes were relatively similar meeting the condition required for the ANOVA test \([F(3, 303) = 2.32, p = .075]\). The ANOVA results were not impacted due to the differences in learning mode enrollment numbers. Differences in student performance \((n = 307\) final course grades) based on learning mode did not appear to be statistically significant \((\alpha = .017)\) as shown in the ANOVA \([F(3, 303) = 1.41, p = .24]\) conducted on end-of-course scores based on learning mode (Gay et al., 2006).

There was not enough evidence to support the alternate hypothesis that student end-of-course scores were not statistically equivalent between classroom, online, and videosynchronous learning (home and classroom) modes. As might be expected, pairwise post hoc tests are unnecessary due to the non-significant finding of the ANOVA test.

The second hypothesis attempted to determine if grade distribution and learning mode were related. The results of a chi-square test of independence \((\alpha = .017)\) are presented in Table 2.
Table 2

Chi-Square Analysis on Association Between Grade Distribution and Learning Mode

<table>
<thead>
<tr>
<th>Grade</th>
<th>Classroom</th>
<th>Videosynchronous classroom</th>
<th>Videosynchronous home</th>
<th>Online</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>14</td>
<td>7</td>
<td>19</td>
<td>126</td>
<td>166</td>
</tr>
<tr>
<td>B</td>
<td>4</td>
<td>4</td>
<td>11</td>
<td>63</td>
<td>82</td>
</tr>
<tr>
<td>C</td>
<td>3</td>
<td>0</td>
<td>7</td>
<td>19</td>
<td>29</td>
</tr>
<tr>
<td>D</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>6</td>
<td>9</td>
</tr>
<tr>
<td>F</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>16</td>
<td>21</td>
</tr>
<tr>
<td>Total</td>
<td>22</td>
<td>12</td>
<td>43</td>
<td>230</td>
<td>307</td>
</tr>
</tbody>
</table>

Chi-square test:

<table>
<thead>
<tr>
<th>Statistic</th>
<th>df</th>
<th>Value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chi-square</td>
<td>12</td>
<td>11.37</td>
<td>.497</td>
</tr>
</tbody>
</table>

Note. N=307, (α=.017). The chi-square test reported a low cell-size warning. Data were collapsed into pass and fail categories in Table 3.

Test results indicated that learning mode and grade distribution were not related. Based on these results, there is no statistically discernable evidence to support the idea that student grades have a significantly different distribution based on type of course delivery.

The final hypothesis attempted to evaluate if there was a relationship between pass rate and learning mode. Data were evaluated using a chi-square test of independence (α=.017; Gay et al., 2006) and are presented in Table 3.

Table 3

Chi Square Analysis on Association Between Pass Rates and Learning Mode

<table>
<thead>
<tr>
<th></th>
<th>Classroom</th>
<th>Videosynchronous classroom</th>
<th>Videosynchronous home</th>
<th>Online</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pass</td>
<td>22</td>
<td>12</td>
<td>38</td>
<td>214</td>
<td>230</td>
</tr>
<tr>
<td>Fail</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>16</td>
<td>21</td>
</tr>
<tr>
<td>Total</td>
<td>22</td>
<td>12</td>
<td>43</td>
<td>230</td>
<td>307</td>
</tr>
</tbody>
</table>

Chi Square Test:

<table>
<thead>
<tr>
<th>Statistic</th>
<th>DF</th>
<th>Value</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chi Square</td>
<td>3</td>
<td>4.05</td>
<td>.26</td>
</tr>
</tbody>
</table>

Note. n=307, (α=.017).

The chi-square test of independence results do not support the notion that pass rate and learning mode were related. Based on these results, we do not have discernable statistical evidence to support the research hypothesis that pass rates were significantly related to type of course delivery.
Discussion

This study contributes to the conflicting body of literature that seeks to determine the influence of course modality on student outcomes. In this study, there were no statistically significant findings regarding end of course grades, grade distribution or pass rates based on learning mode. In this case, “no significant difference” in the data results can be considered good news for researchers attempting to determine if students would be put at a disadvantage by using one learning mode instead of another. Regardless of learning mode, students in this study tended to do equally well. These results align with other studies that report no significant difference for short-term outcomes of final course grade in undergraduate statistics courses (McLaren, 2004; Summers et al., 2005) and successful completion (Dotterweich & Rochelle, 2012). It should be noted that possible confounding variables of the textbook, course design, types of assignments, and testing were controlled for by the use of truly standardized curricula across all four learning modes.

Limitations

Four limitations are identified in this study, all of which could be confounding or moderating variables on the results. Data are from a non-traditional student body whose average age is 34. Most students were working adults affiliated with the U.S. military. It cannot be assumed the same results would occur with a more traditional undergraduate student population (18–22 years old). Data were collected from the January 2019 term. Term length was 9 weeks. It cannot be assumed that similar results would occur in other terms or in more traditional 15-16-week terms without conducting a separate study. Students were not randomly assigned to each type of learning mode. Students enrolled in courses without interference or guidance from researchers. Results could be impacted due to the fact that students self-selected into one of the four different types of courses examined. No controls were built into this research to compensate for students’ prior expertise with statistics.

Recommendations for Future Research

Future work should attempt to explore moderating variables. Only one study reports gender as a moderating variable; a future study that explores this could support or provide counterevidence to the conclusion that females are less successful in online statistics courses (Flanagan, 2012). There are two conflicting studies regarding the influence of the instructor (e.g., explanations and grading) and course design (e.g., class discussions and quality of problems) (Scherrer, 2011; Summers et al., 2005). A future study could explore this variable through student perspectives. A measure of student perspectives could also investigate self-concept in order to explore the assertion by Zimmerman & Austin (2018) that self-concept does not influence course completion in traditional statistics but does play a role in online statistics.

Other variables worth exploring in future studies include the use of common syllabi and texts across learning platforms, faculty experience with course content and technology, and how students select learning modes when taking classes. The variables of timing (what term the course was taken in) and term length (traditional versus compressed term formats) should be included in future research to determine their possible impact on student performance.
Future studies should be designed to ascertain a better understanding of variables associated with why students select specific modes of learning. These variables could range from learning style to demographic and academic preparedness potential as measured through standardized tests. These variables could also be used to determine if groups under examination are different, which could impact analysis and results.

The methodology of this study could be replicated on both traditional (defined here as ages 18 to 22) and non-traditional student populations.
References


