AN EVALUATION OF TEACHING METHODS IN THE INTRODUCTORY PHYSICS CLASSROOM

by

Lauren Michelle Williams Savage

A thesis submitted to the faculty of
The University of North Carolina at Charlotte
in partial fulfillment of the requirements
for the degree of Master of Science in
Applied Physics

Charlotte

2015

Approved by:

____
Dr. Donald Jacobs

____
Dr. Susan Trammell

____
Dr. Faramarz Farahi

ABSTRACT

LAUREN MICHELLE WILLIAMS SAVAGE. An evaluation of teaching methods in the introductory physics classroom. (DR. DONALD JACOBS)

The introductory physics mechanics course at the University of North Carolina at Charlotte has a history of relatively high DFW rates. In 2011, the course was redesigned from the traditional lecture format to the inverted classroom format (flipped). This format inverts the classroom by introducing material in a video assigned as homework while the instructor conducts problem solving activities and guides discussions during the regular meetings. This format focuses on student-centered learning and is more interactive and engaging. To evaluate the effectiveness of the new method, final exam data over the past 10 years was mined and the pass rates examined. A normalization condition was developed to evaluate semesters equally. The two teaching methods were compared using a grade distribution across multiple semesters. Students in the inverted class outperformed those in the traditional class: "A"s increased by 22% and "B"s increased by 38%. The final exam pass rate increased by 12% under the inverted classroom approach. The same analysis was used to compare the written and online final exam formats. Surprisingly, no students scored "A"s on the online final. However, the percent of "B"s increased by 136%. Combining documented best practices from a literature review with personal observations of student performance and attitudes from first hand classroom experience as a teaching assistant in both teaching methods, reasons are given to support the continued use of the inverted classroom approach as well as the online final. Finally, specific recommendations are given to improve the course structure where weaknesses have been identified.

ACKNOWLEDGMENTS

I would like to acknowledge Dr. Pedram Leilabady for his assistance on this project. All of the raw data analyzed here has been collected and preserved by Dr. Leilabady throughout the past ten years. As the instructor of the course examined in this thesis, his dedication to constantly improving this course has created the bedrock on which this evaluation stands.

I would also like to thank my advisor, Dr. Donald Jacobs, for the numerous hours spent discussing this topic and collaborating on this thesis.

DEDICATION

I would like to dedicate this thesis to my wonderful and loving husband, Luke, for all his encouragement and tireless efforts to ensure that I completed this endeavor. Our long discussions, his experience in statistics and teaching, and his useful advice have each been a huge contribution to this work and a great encouragement to me.

TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION 1

1.1. Two Teaching Methods 1

1.1.1. Traditional Classroom Approach 2

1.1.2. Inverted Classroom Approach 2

1.2. Motivation for UNCC Course Redesign 3

1.3. Redesign Pilot Results 4

1.4. ExpertTA Final Exam Platform 5

CHAPTER 2: LITERATURE REVIEW 7

2.1. Introduction to Inverted Classroom Approach 7

2.2. Practices 8

2.3. Criticism 9

CHAPTER 3: METHODOLOGY 12

3.1. Raw Grades 12

3.2. Normalization Condition 14

3.3. Establishing a Baseline 17

3.4. Histograms 18

3.5. Pass Rates 20

CHAPTER 4: RESULTS AND CONCLUSIONS 22

4.1. Traditional vs. Inverted 22

4.2. Written vs. Online Exam 23

4.3. Conclusions 27

CHAPTER 5: FUTURE WORK 30

5.1. Continuing Research 30

5.2. Recommendations for Future Course Design 32

BIBLIOGRAPHY 34

APPENDIX: KOLGOMOROV-SMIRNOV TEST 35

CHAPTER 1: INTRODUCTION

We know that students in Introductory Physics courses struggle to grasp the material[1]. This is evident based on the percentage of students who have either received grades below 70% or have withdrawn from the course at the University of North Carolina at Charlotte (UNCC). To reduce failure and drop rates, the introductory physics course was redesigned in 2010[2]. After 9 semesters, can we call the outcome a success? Are our students showing improved learning under the new system? If so, what factors led to its success? If not, in what ways can we improve the redesign?

First, I will discuss the teaching methods used in this course and in similar courses at other universities and why the task of redesigning the course was undertaken at UNCC. Then, I will compare the past nineteen semesters using histograms of the final exam grade distributions. I will evaluate the different teaching methods using both overall grade distributions for the two methods as well as specific case studies. Finally, using my own teaching experiences along with the data, I will identify areas of weakness and strength to aid future instructors in course design.

1.1. Two Teaching Approaches

Over the past ten years, UNCC has used different teaching methods to deliver content in its introductory physics courses. These methods have included the traditional lecture approach as well as the inverted classroom approach.[2] The traditional approach is referred to as traditional because it is the way courses have been taught for many decades at numerous universities around the world and across many different disciplines. The inverted classroom approach, however, is a newer trend utilizing technology to deliver course content. The term "inverted" (or sometimes "flipped") is used because the approach exchanges the student learning practices of the traditional approach: lecture videos are assigned as homework so that class time can be used to answer student questions and develop problem-solving strategies.

1.1.1. Traditional Classroom Approach

Up until the course redesign in 2010, the calculus-based physics mechanics course, "Physics 2101: For Scientists and Engineers (PHYS 2101)", was taught in the traditional lecture format.[2] Students attended two face-to-face lecture sessions each week. The lectures were primarily instructor-driven with few student interactions and simple demonstrations. The lectures focused on theory combined with example problems, all led by the instructor. The course included online homework sets along with two written two-and-a-half hour midterm exams. The course concluded with a two-and-a-half hour comprehensive final exam.

1.1.2. Inverted Classroom Approach

In 2011, the inverted classroom approach replaced the traditional lecture format.[2] Instead of having two face-to-face lectures a week, a problem-solving session with a graduate teaching assistant (TA) replaced one of the lecture times. Students were expected to come prepared for each lecture having read the assigned text[2], watched an online lecture video, and completed pre-lecture quiz. In class, the instructor focused more on addressing the nuances of the material and engaging students with problem-solving questions rather than giving a minimally interactive lecture. In the problem-solving sessions, the TA taught problem-solving strategies using practice problems and administered a weekly test covering the previous week's material. These weekly tests, around 12 in number and 20 minutes in length, replaced the midterm exams. Students also completed weekly homework sets online. The course was concluded with a comprehensive two-and-a-half hour written final exam. In the fall semester of 2014, the written final exam was replaced with an online final exam using the ExpertTA online service. The following figure shows the basic outline of a typical week in the inverted classroom.

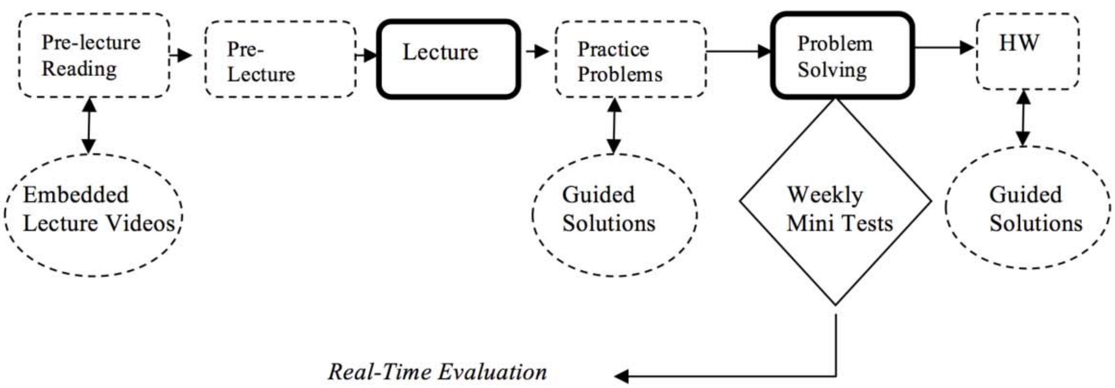

Figure 1: This diagram shows the structure of a typical week in the inverted classroom approach.[2]

1.2. Motivation for UNCC Course Redesign

The motivation for the UNCC Course Redesign was multifaceted. Enrollment had increased by 40% since 2007.[2] The increased number of students put a strain on the availability for large lecture halls to accommodate the growing crowds.[2] By switching to the inverted classroom approach, instructors could teach twice as many students in the same time slot since the classes were now divided between the large lecture hall and the smaller problem-solving classrooms. Additionally, this meant that the cost per student could be reduced, as the number of instructors remained constant while the enrollment rate increased.

Additionally, the course had been experiencing a significant rate of students receiving grades of "D"s, "F"s, and Withdrawal (DFW rate). Between Fall 2006 and Spring 2010, the DFW rate averaged at 46.5%, meaning that nearly half of the enrolled students did not meet the requirement to proceed to Physics 2102, which requires passing this course with a C or better. The hope was that the inverted classroom would foster student learning by engaging them with the material in a smaller learning environment during the problem-solving sessions and giving them 24/7 access to the lecture material via the online resources.[2]

1.3. Redesign Pilot Results

In March 2012, data from the course redesign was compiled and presented to the University.[2] This redesign study included significant details about the strength of the inverted approach. The study concluded that the five preceding semesters in which the inverted classroom approach was used fared better than the semesters in which the traditional classroom was used. This was based on the DFW rates from those semesters. The DFW rates were calculated (by the University) and the average values were compared for each teaching method. The average DFW rates had dropped from 46.5% to 33.3% upon implementation of the inverted classroom approach.

Additionally, it was found that the cost per student was also significantly reduced. Under the traditional approach, the cost per student was $134 per student per section of 110 students. This was based on the hiring costs of one faculty member and one TA. By implementing the inverted classroom method, enrollment increased to 160 students in the same time slot as before, since half of these students would be meeting with the two TAs in groups of 40 each while the other 80 would be in lecture with the instructor. This increase in enrollment meant that the cost per student dropped to $92 each, with the capability to expand the problem-solving sessions by adding another section of 30 students without needing to reserve a larger lecture hall. This represents a 31% reduction in cost per student with the ability to reduce it further to $82 per student, an additional 8% reduction, while also reducing the demand for large lecture halls.[2]

1.4. ExpertTA Final Exam Platform

The ExpertTA online platform mimics the way an instructor would guide a student through a question. The platform has a database of possible student mistakes and hints to aide the student through multiple attempts. The platform first requires students to input algebraic expressions as initial solutions before requiring the students to evaluate the numerical solution. This allows the student to receive partial credit on each question. The platform also gives students multiple attempts at the solution, the number determined by the instructor. The instructor can also decide if multiple attempts will reduce the partial credit awarded.

Migrating the final exam to ExpertTA will eliminate the need for written exams. By reducing the amount of paper used for the exams, a significant portion of natural resources would be conserved. Each exam packet consists of at least 13 sheets of paper per student, totaling a minimum of 6000 sheets of paper per semester including scratch paper, equation sheets, and signature pages. This number could grow as high as 8000 with an enrollment of 500 students and a 10-question (16 page) exam packet. The online exam would reduce this number to 2000 while still providing the students with scratch paper and equation sheets. This would save resources such as paper and toner, and as well as reduce cost.

CHAPTER 2: LITERATURE REVIEW

2.1. Introduction to Inverted Classroom Approach

The inverted classroom approach is a teaching methodology that has rapidly gained popularity since its conception. It was birthed out of a desire to improve student learning outcomes by making lectures more engaging for students, thereby capturing the attention of students with varying learning styles. Dr. Eric Mazur was using inverted methodology in his classroom long before the modern technological advances made it prominent[3]. Later, it was further proliferated by KhanAcademy.org founder, Salman Khan under the backing of Microsoft founder Bill Gates.

The founding principle of the inverted (also known as “flipped”) classroom approach is that the content of homework and class time be reversed. Rather than going to class to sit and listen to a lecture from the professor and then go home and work out problems, the inverted classroom approach “flips” these tasks such that the lecture is assigned as homework and class time is used to work out problems. The key enabler for this method is the advancement of modern technology, multimedia[4] in particular. Lectures are pre-recorded videos that allows students to pause, rewind, and replay the lecture until they have grasped the content, therefore allowing them to move at their own pace, unlike in the traditional classroom setting[5]. This frees up class time to engage the interactive and social aspects of learning[6]. Students are free to ask questions, work together in groups, and engage in classroom discussion about the subtleties of the topic at hand.

2.2. Practices

A pioneering technique of the inverted classroom that was engineered to promote student engagement in the classroom is called Peer Instruction. This teaching method gets the students interacting and discussing with one another while working out problems with end goal of a 90% comprehension rate. The instructor presents the concept in a mini-lecture format, about 8-15 minutes long. Then, the instructor poses a multiple-choice question to the students, generally conceptual in nature. The students work out their solution and submit their answers using a response medium (raised hands, note cards, clickers, etc). The instructor is then able to assess the portion of the class who answered correctly and those who did not. Then, students are encouraged to find someone who answered differently and discuss. After a few minutes of discussion/explanation, the students resubmit their answers. This process continues until roughly 90% comprehension is reached. [3]

This method also works well when integrated with other technologies as well. One of the most valuable tools in applying the inverted method to the flipped classroom is the ability to have instant feedback on student comprehension. This is where technologies like personal response devices ("clickers[7]") come in handy. Also, with the rise of the smartphone, websites and apps like *pollanywhere.com* and *Mentimeter* provide the same functionality for free. These response devices allow instructors to structure their class time around the content that the students struggle with the most. If the majority of students do not comprehend the material, the feedback alerts the instructor and they can re-cover any topics that were not fully understood.[3, 6]

Students can also be assigned pre-lecture quizzes to assess their understanding of the material and insure that they complete it before class time. Just-in-Time Teaching (JiTT)[8] is one method of assessing student completion of assigned pre-lecture material. JiTT provides warm-up activities to prepare students for the content of the lecture and serves primarily as a reading quiz. This ensures that students have a first exposure to the material outside of the classroom and assesses their understanding using the quiz. The professor can then review the results of the quizzes and adjust the lecture according. JiTT also provides after-lecture activities to ensure that the concept has sunk in. JiTT was originally developed for the introductory physics classroom but has been applied in other disciplines as well.

2.3. Criticism

One question that arises from the fast adoption of the inverted classroom and its promotion by educational filmmaker Sal Khan is whether or not these videos are actually improving how students learn physics.[9] One of the main problems that results from moving the lecture to the comfort of the student's home where they can watch it on their computer is the fact that it separates the learner from the teacher. It removes the social aspect of learning. Having a teacher present allows the students to ask questions about the content and be engaged in discussion as the material is being presented.[6]

Physics is a unique discipline to study because we interact with physical concepts on a daily basis without fully understanding how it works. This creates misconceptions in the mind of the students. When these misconceptions are not addressed in the video and then the student is quizzed on their understanding of the concept explained, often, the result is that the student not only gets the concept wrong, but is more confident in their wrong answer after watching a video that explained it correctly. After watching a video that, in addition to the correct conceptual information, also presents and addresses commonly held misconceptions about a physics principle, students scored higher in the post quiz than students who watched the video that simply contained correct information.[10] [11]

Why is this the case? Dr. Derek Muller, creator of *veritasium.com*, discovered that watching the video with the misconceptions included required more mental effort[6] on the students' part. The students had to wrestle with their misconceptions to answer correctly, rather than affirming their previously held mistaken belief. The post-test scores were nearly 50% higher than the students who watched the video without the misconceptions presented.[11]

This isn't to say that what Sal Khan is doing is wrong. It merely suggests that these straightforward tutorial type videos might be better used for review, not first exposure. If we are going to promote student learning and expect to see better student learning outcomes, we ought to use multimedia in a way that is actually beneficial to their education. To this end, KhanAcademy.org has implemented tools for its use in the inverted classroom in order to promote student mastery. For example, students cannot proceed to the next unit without first answering 10 consecutive questions correctly.

Another issue with the inverted classroom approach is that it increases the workload on the student. By adding an online lecture, additional online practice problems, and a pre-lecture quiz, the workload on the student in the inverted classroom nearly doubles. Some would argue that practice makes perfect and yes, that is often the case. In my personal experience, however, I have seen students working for many hours outside of class on the additional material. Often, they get frustrated with the work because no one has explained it in a way that they can understand it. The material in class went over their heads and students do not feel comfortable approaching the professor for help. The added workload on the students and the decreased face-to-face interaction time with the professor makes it harder for students to feel like they can succeed in the inverted classroom.

CHAPTER 3: METHODOLOGY

In order to compare different teaching methods to one another, one must find a common metric by which to analyze the data, *ceteris paribus*. For this course, the only consistent thread of commonality lies in the final exams. Aspects of the course such as online homework service, TA teaching experience and skill, student opinion, intermediate assessment methods, and supplemental instruction availability were excluded from this study. In order to restrict the data to exclude any other possible variables, only data from sections taught by Dr. Leilabady were considered. In addition, this course is one of four introductory physics courses taught at UNCC; PHYS 2101 was selected as the target of this study, although the conclusions here could be extrapolated to any of the four courses taught.

Even when looking only at the final exam, the exam length, difficulty, medium, and contribution to the final course grade were not consistent throughout all semesters. This problem is addressed in the following sections.

3.1. Raw Grades

The following box-and-whisker plot shows all nineteen semesters of raw exam grades. The box-and-whisker plot was used to compare the different semester distributions visually. The whiskers show the minimum and maximum grades in each semester. The upper and lower 25% of the class grades are contained within the range of the whiskers and the boxes show the range of the middle 50% of each class. The two boxes are divided by the median grade of each semester.

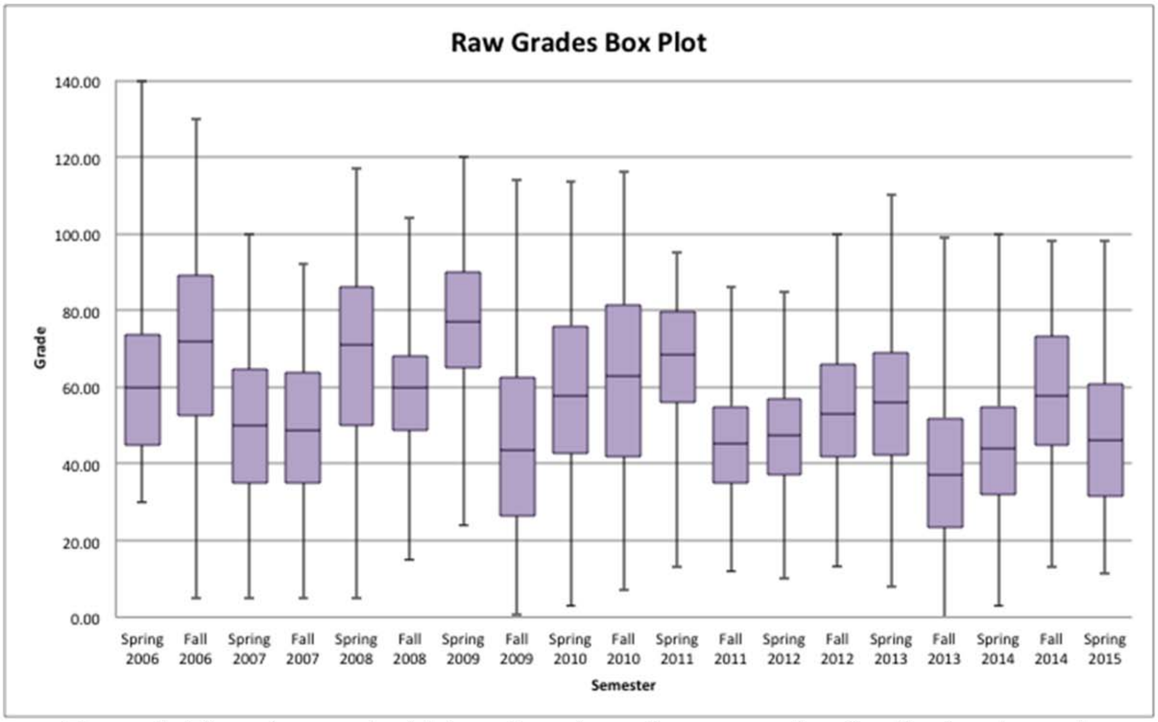

Figure 2: These box-and-whisker plots show the raw grades distribution for each semester.

This figure shows the differences between the distributions of the grades. Nearly all the median values fall between the grades of 40-80 while the majority of minimum grades lie within the 0-20% region. Some of the boxes are tightly clustered around the median grade while other boxes are much more spread out.

The upper 25% reveals a trend about the data. In several cases, the maximum grade scored was well above 100%. This matches up with the fact that on many final exams, extra credit was offered. This seems to indicate that the students needed extra credit on the final exam to boost their overall course grade.

In other cases, the maximum grade on the final exam fell well below the 100% mark. This would suggest that these final exams were tougher than the other exams or that perhaps that the students were not as well prepared as students in other semesters. Thus, it is reasonable to include that the raw exam data as it stands is not sufficient to offer a fair comparison between the different semesters. Thus, some sort of normalization condition is required to make the tests comparable.

3.2. Normalization Condition

It should be noted that different methods of normalizing the data were investigated. First, the raw grades were normalized by the max grade for each semester. This method was difficult to justify, however, because in some cases, the difference between the maximum grade and the semester average was greater than twice the standard deviation of the data set. The Modified Thompson Tau[12] technique was used to identify outliers in the data and, in nearly every semester, the maximum score was an outlier that was not representative of the rest of the distribution. Therefore, data points would have been much lower after normalization than if the outlier had been removed. As a student might say, the outlier(s) "killed the curve".

Another method that was considered was to normalize to the median. This method would serve to equalize all the semesters by mapping the median grade to 50% by dividing all the grades by twice the median value. The problem with this method of normalization was it was no longer possible to set a standard cutoff for the pass rate without dictating that only a certain percentage of students would pass, quite the opposite of how one would want to normalize the data to find the pass rate of each semester.

The second-to-last method that was investigated was to remove outliers in the data using the Modified Thompson Tau technique and then normalize by the new maximum grade. The problem with this method is that it would appreciably change the pass rate of the data, as many of the top-scoring students would be removed from the population. The solution to this problem would be to find a normalization method that would still account for the presence of the outliers without allowing the outliers to cause a misrepresentation of the rest of the data.

To derive the final normalization constant, the median of the top ten percent of students in each semester was calculated and then averaged. Then, this normalization constant was divided by each median value for the top ten percent to create each semester's unique normalization weight. The raw grades for each semester were then multiplied by this normalization weight so that the semesters could be compared on even footing. This is summarized in the following expression, with N=19:

$$W_i = \frac{\sum_{i=1}^{N} \widetilde{x_i}}{N} \widetilde{x_i}^{-1}$$

The following table shows the median value of the top ten percent of the students, the average, and the weight assigned to each semester.

Table 1: This table shows the top 10% median grades and normalization weights for each semester.

Semester	Median	Weights
Spring 2006	73.21	1.14
Fall 2006	88.46	0.94
Spring 2007	87.50	0.95
Fall 2007	82.00	1.02
Spring 2008	87.50	0.95
Fall 2008	80.45	1.03
Spring 2009	87.50	0.95
Fall 2009	73.75	1.13
Spring 2010	82.08	1.01
Fall 2010	91.67	0.91
Spring 2011	87.50	0.95
Fall 2011	77.50	1.07
Spring 2012	69.00	1.21
Fall 2012	89.00	0.94
Spring 2013	87.27	0.95
Fall 2013	89.00	0.94
Spring 2014	67.00	1.24
Fall 2014	92.92	0.90
Spring 2015	88.15	0.94
Average	83.24	

After multiplying the raw grades by the normalization weights, the semesters were plotted in box-and-whisker plots for comparison against the raw grades.

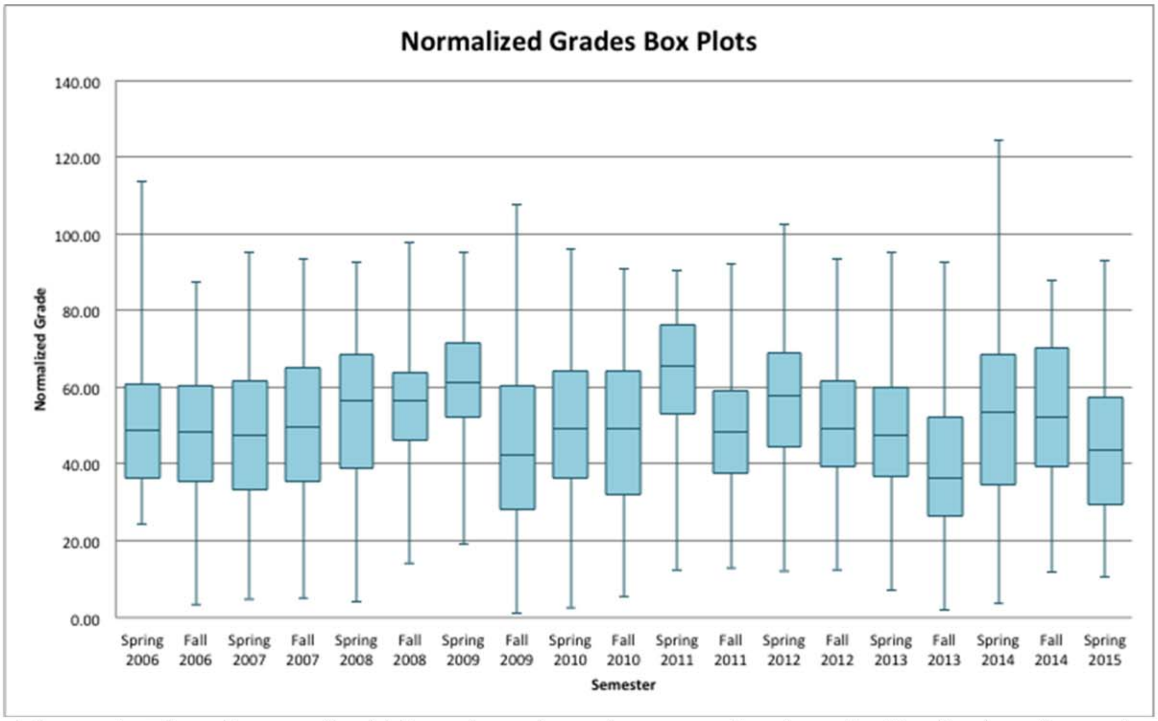

Figure 3: These box-and-whisker plots show the normalized grade distribution for each semester.

After normalization, many of the problems inherent in the raw grades distribution have been mitigated. The normalization constant reduces the effect of outliers on the distribution without removing them completely. It also removes the effect of the extra credit given on different exams as prior to applying the normalization weights, each exam grade was converted to a percentage of the total possible points for that exam by dividing by the maximum possible grade. This way, each exam was scaled to 100 points without affecting the distribution of the data. Finally, this normalization process tamed many of the outliers in the raw data by bringing them into accordance with the rest of the data.

3.3. Establishing a Baseline

Calculating the frequency of each grade and then dividing by the total number of students enrolled in the course over the last ten years defined the baseline. The resulting percentages were then plotted against the grade values to give an overall grade distribution. This procedure was repeated for each semester to calculate the individual pass rates. It was also used to calculate the pass rate for each teaching method. The following graph depicts the frequency of each exam score as a percentage of the total number of students enrolled over the past nineteen semesters:

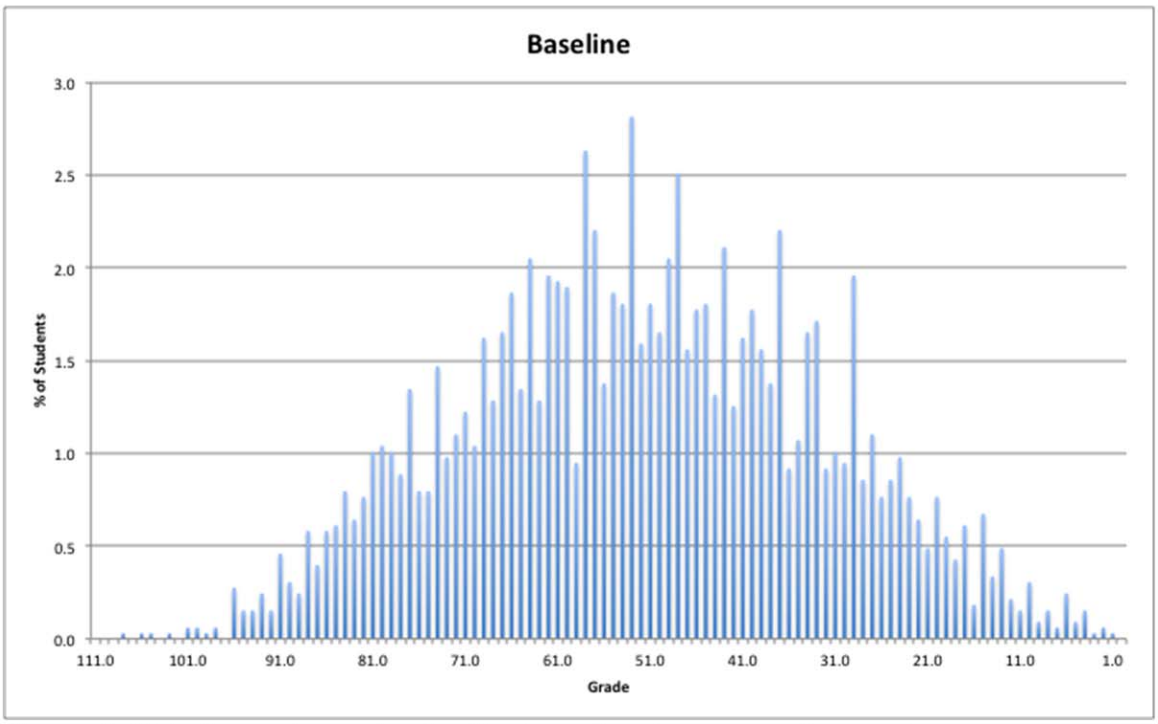

Figure 4: This figure shows the distribution of grades across all semesters as a percentage of the total number of students enrolled in the class.

3.4. HISTOGRAMS

Histograms were created to plot the number of students scoring in each grade category. In some instances, students still scored over 100% due to the normalization weighting. This category is labeled as >100%. Grade categories of A, B, C, D, and F were used in accordance with a 10-point scale, that is:

$$100 \geq A > 89$$

$$89 \geq B > 79$$

$$79 \geq C > 69$$

$$69 \geq D > 59$$

$$59 \geq F > 0$$

The grade category F was further broken down into sub categories of the same size as for those above, such that

$$59 \geq F1 > 49$$

$$49 \geq F2 > 39$$

$$39 \geq F3 > 29$$

$$29 \geq F4 > 19$$

$$19 \geq F5 > 9$$

$$9 \geq F6 > 0$$

The grades for the overall inverted and traditional classroom methods were then plotted against these grade categories. The resulting histogram follows:

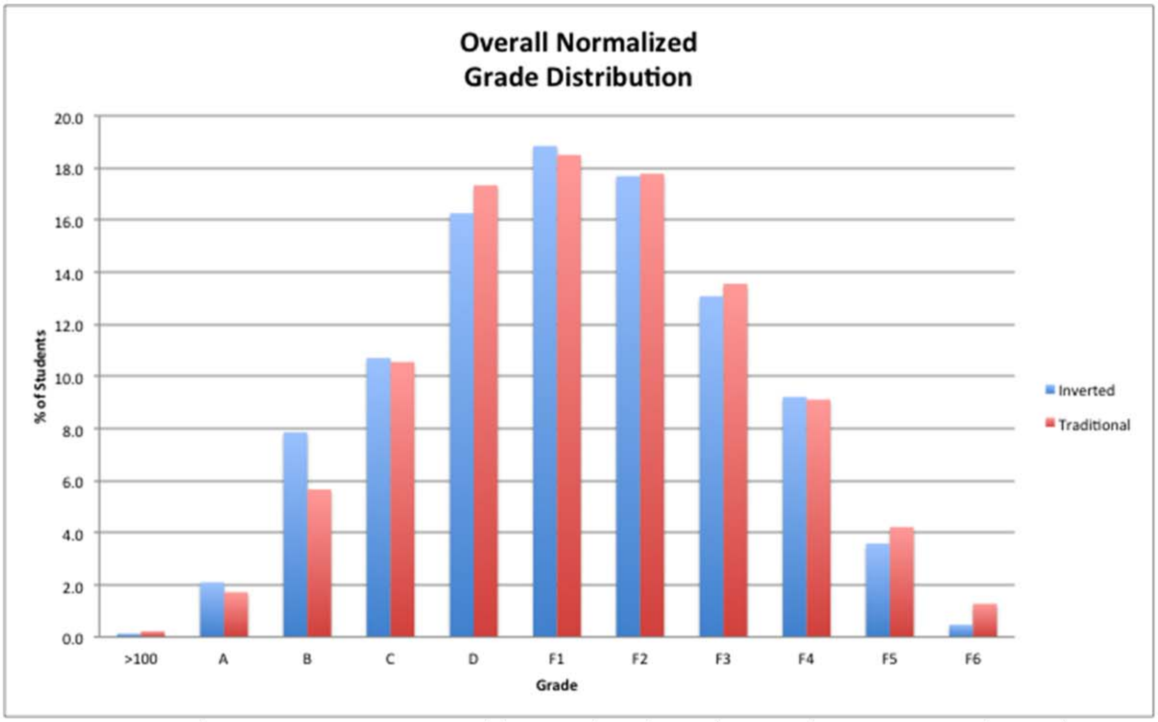

Figure 5: This figure shows overall normalized grade distribution for each method according to the predefined categories.

3.5. Pass Rates

By analyzing the baseline graph (Figure 4 above), a passing cutoff grade of 70% was chosen and applied to all semesters. This is based on the fact the prerequisite condition for many of the subsequent courses taken by the student is to pass PHYS 2101 with a C or better, and it is a common convention to use 70% as the cutoff for a C. The baseline supports this condition since these are the grades after the exam scores have been normalized for difficulty and across all semesters. This standard was then applied to each semester and the pass rate was determined by calculating the percentage of student performing at or above this standard on the final exam. The following graph depicts the final exam passing rate values for each of the nineteen semesters.

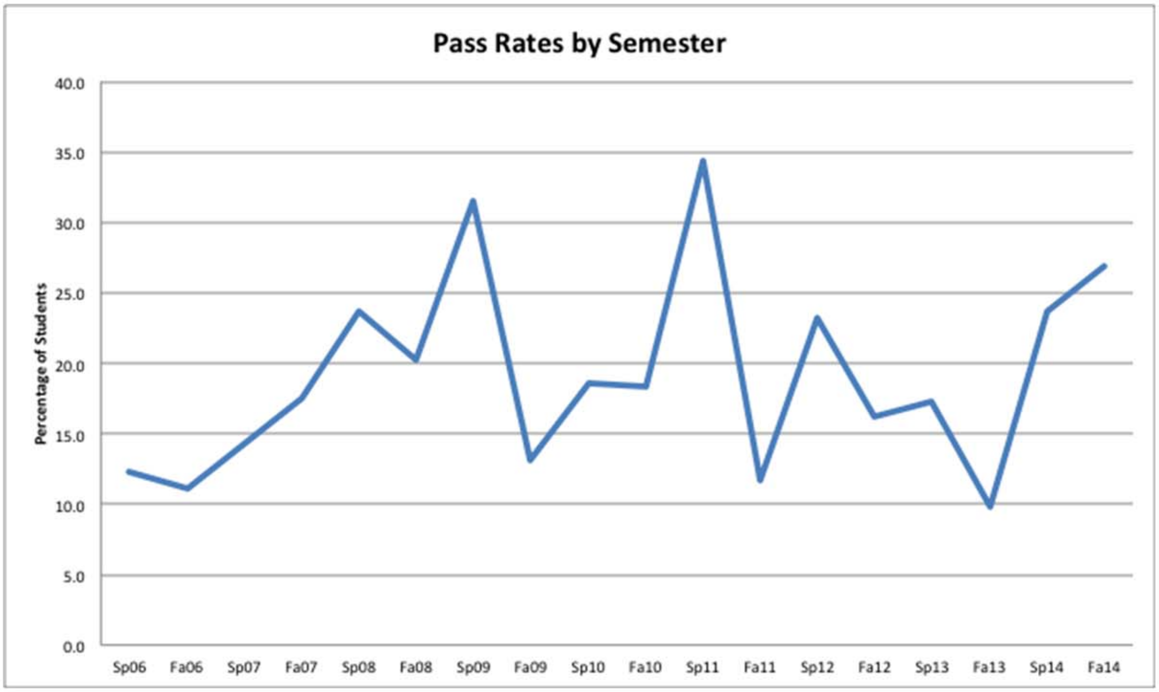

Figure 6: This figure shows the pass rates by semester.

CHAPTER 4: RESULTS AND CONCLUSIONS

4.1. Traditional vs. Inverted

The primary measure available to judge the inverted method's success is to compare its pass rates to those of the traditional method. In order to accomplish this, Figure __ was created to compare the total number of students scoring in each category across all semesters that each method was used. At first glance, it appears as though the inverted method performed slightly better, with a larger percentage of passing students than the traditional method. If we look at the median and average pass rates for both methods, the data supports this.

Table 2: This table shows the average and median pass rates for Traditional and Inverted approaches.

Pass Rates:	Average
Traditional	18.0
Inverted	20.2
Gain	12.2%

These results indicate a 12.2% gain for the inverted approach over the traditional. This difference is significant, but does not reveal the entire picture.

To get a better understanding of the data, we can examine the differences in the overall histograms of the two methods.

This table shows the percentage of students scoring in each grading category for each of the two methods, as well as the gain between the inverted and traditional percentages.

Table 3: **The percentage of students scoring in each of the grading categories for each method and the difference between them as well as the gain.**

Grade	Traditional	Inverted	Gain %
>100	0.2	0.1	-39.0
A	1.7	2.1	22.0
B	5.7	7.9	38.7
C	10.6	10.7	1.4
D	17.3	16.3	-6.2
F1	18.5	18.8	1.8
F2	17.8	17.7	-0.5
F3	13.6	13.1	-3.5
F4	9.1	9.2	1.1
F5	4.2	3.6	-14.9
F6	1.3	0.5	-62.9

This table shows that the percent of students with scores in the A category increased by 22% under the inverted classroom approach. B grades increased by 39% and C grades increased by 1.4%. Also, the percent of students with a grade of less than 10% correct was reduced by 63%. The percent of students with grades of 10-20 was reduced by 15%.

The data shows that the percentage of students scoring a passing grade increased appreciably while the percent of students failing the exam was noticeably decreased. However, a Kolgomorov-Smirnov[13] (KS) two-sample test states that we cannot reject the null hypothesis that the two data sets come from the same normal distribution at the 5% significance level. Based on the KS-test results, there is no significant statistical difference between the two data sets (see Appendix A).

4.2. Written vs. Online Exam

Another aspect of the course that can be evaluated from the data collected is the effect of an online exam as opposed to a written final exam. In order to make this comparison, the Spring 2014 and Fall 2014 histograms were compared to each other. Both of these semesters were taught by the same instructor and problem-solving TA and at the same time of day and location. They were both taught under the inverted classroom approach and used the same resources (ExpertTA and Young's University Physics textbook) and grade breakdown. The only difference between these two semesters is the fact that the Spring 2014 final exam was administered in a written format and the Fall 2014 utilized ExpertTA. Both exams were administered in a proctored classroom setting by TAs and instructors. Two TAs graded the written exam over the course of 3-4 days whereas ExpertTA graded the online exam instantly. ExpertTA also allowed multiple attempts at a solution whereas the written exam was graded based on answers written in the test booklet, for which partial credit can be given. The following histogram shows how each distribution fell in the grade categories.

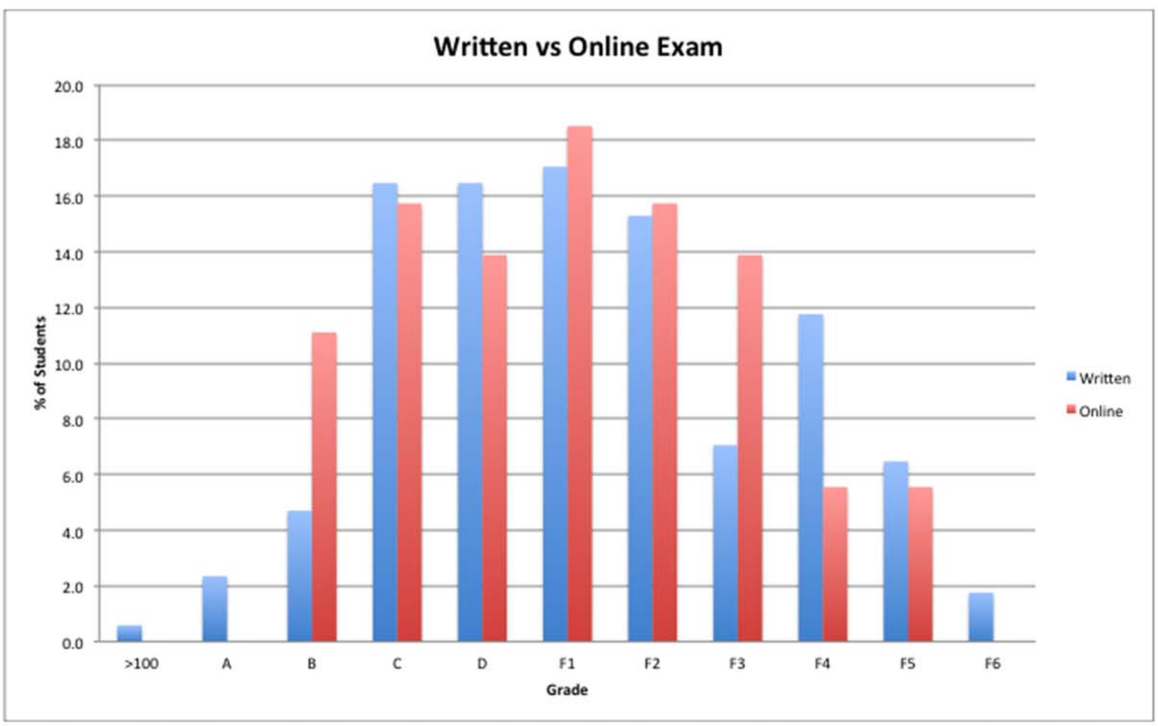

Figure 7: This graph shows the grade distributions for the two assessment methods.

At first glance, this graph seems to indicate that the online exam fared worse than the written exam because no grades greater than 90 were scored. But upon further inspection, one sees that many more students scored a B on the online exam compared to those scoring a B on the written exam. In fact, the pass rate, based on a 70% cutoff grade, was higher for the online exam than the written exam by 2.8%, an 11.6% gain, as the following table shows.

Table 4: This table shows the pass rates for each semester with cutoff grades of 70% and 60%, respectively as well as the gain.

Pass Rates:	70%	60%
Written	24.1	40.6
Online	26.9	40.7
Gain	11.6%	~0

If one moves the cutoff grade to 60%, however, one can see that the passing rates for the two exam methods are nearly identical. The following table shows the percentage of students scoring in each of the grade categories and the difference in percentage between the two methods by category.

Table 5: This table shows the differences in percentage of students scoring in each grade category on the written and online final exams formats as well as the gain.

Grade	Written	Online	Gain %
>100	0.6	0.0	-100.0
A	2.4	0.0	-100.0
B	4.7	11.1	136.1
C	16.5	15.7	-4.4
D	16.5	13.9	-15.7
F1	17.1	18.5	8.6
F2	15.3	15.7	2.9
F3	7.1	13.9	96.8
F4	11.8	5.6	-52.8
F5	6.5	5.6	-14.1
F6	1.8	0.0	-100.0

A Kolgomorov-Smirnov[13] (KS) two-sample test states that we cannot reject the null hypothesis that the two data sets come from the same normal distribution at the 5% significance level. Based on the KS-test results, there is no significant statistical difference between the two data sets (see Appendix A).

While fewer students scored an A or greater on the online exam than the written one, far fewer students scored in the lowest categories as well. The percentage of students with grades between 80-90 increased by 136%. The percentage of students with grades between 30-40 nearly doubled. The percentage of students with grades between 0-30 were drastically reduced. This seems to indicate that while the online test may have been harder than the written test or graded harsher, more students were prepared than the written because a smaller percentage of students scored extremely low. The median and average scores for these distributions were also fairly close, as the following table shows.

Table 6: This table shows the average and median grades for the two assessments methods along with the differences between the two sets of values.

Average and Median Grades for Each Method		
	Average	Median
Written	52.33	53.42
Online	53.38	52.21
Gain:	2%	-2%

A Kolgomorov-Smirnov[13] two-sample test was performed on this set of data and the null hypothesis was not rejected. This means that despite their being some indication that the online test yielded a better outcome, it is not possible to conclude the methods are somehow distinguishable with any statistical significance (See Appendix A).

4.3. Conclusions

Statistically, the final exam scores depict only slight advantages of the inverted classroom method over the traditional method, but there are other factors to consider when choosing a method in which the course should be taught. The principles of the inverted classroom emphasize the benefits of active student engagement with the material during class time, as opposed to the passive learning style of the traditional lecture. This factor, along with more frequent instructor feedback and the ability to review material using effective multimedia at home, support my recommendation to implement the inverted classroom approach at UNC Charlotte.

There are, however, several pitfalls to the inverted method as it has been practiced at UNCC. From my experience interacting with the students, it is apparent that they did not utilize the online material in the manner that the instructor had intended. I even discovered that the online lectures were not available to the students during the semesters that I served as TA. Also, the redesign presupposes that the TAs would be trained graduate students. In my experience working with the other TAs, this was not the case. Many TAs had no teaching experience prior to teaching this course and there was very little communication between TAs and the professors of the course as to what was expected of them. Often, it was left to the students to inform the TAs as to what topics were covered in lecture so that the problem-solving session would match what they had just learned in class.

There was also inconsistency in the grading methods of the TAs and the feedback rate; often students would score rather poorly on weekly tests and get discouraged about the material when TAs offered little partial credit or did not give helpful feedback. The students I taught also struggled with the weekly tests since they were being assessed on the previous material rather than the topic they were currently trying to understand. I found that this caused issues with concept retention as students would study each concept in a vacuum and had difficulty synthesizing the information they had learned.

Despite all these issues, the evidence shows that a larger percentage of students scored higher on the final exam in the inverted classroom format. While this result favors the inverted classroom approach, there is still room for improvement. It is safe to assume that if some of these problems could be rectified, the pass rate under the inverted method should increase.

The data concerning assessments also revealed a similar relationship between the methods. Statistically speaking, the pass rate for each method was very similar. This supports the conclusion that the students will likely perform equally well using either assessment method in the future. The online format, however, possessed several advantages over the written format. ExpertTA graded the exams instantly, providing prompt feedback to the students, which reduces their anticipation and resulting stress. The automated grading does not require human intervention, which reduces the cost. The automation removes two major downfalls of the written exam, which are the subjectivity and inconsistency of the graders. These factors alone support the decision to switch to an online exam format.

There is another benefit for the students as well. The online exam is administered in the ExpertTA format. This format is used for the online homework sets and monthly topical quizzes. The students would be familiar with the program before taking the final exam, as opposed to semesters where all the homework was completed online but the final exam was administered in the written format. Students would be more comfortable with the online exam because they would have grown accustomed to the question style and wording.

CHAPTER 5: FUTURE WORK

5.1. Continuing Research

In order to advance this research further, there are a few modifications to this methodology that could be made.

First, a standardized test could be created in order to remove the normalization condition needed to equalize the exams across multiple semesters. No extra credit would be given on the exams so that they would all be graded out of 100 points. These exams would come from a common question bank and would have equal difficulty, such as exams given on ExpertTA. One bonus of the ExpertTA package is that the creators focus on analytics and provides students with feedback for their incorrect answers. The ExpertTA team is constantly improving their system to help students grasp concepts[14]. The professor can also select the questions that they are confident that their students can answers, helping them to succeed.

Additionally, it would be useful to have student opinion surveys to go along with the exam data. Questions that would be useful for analysis of best practices for this course could include:

- Did you think the course load was too much or too little?
- How much time did you spend studying each week?
- How confident do you feel about your knowledge base in physics after/while taking the course?

- What aspects of the course design helped you the most?
- Which aspects were least effective in helping you understand the material?
- Were the instructors helpful and available when you had questions?
- Was the lecture presentation useful to help you learn the material?
- Did watching the online videos/reading the online lecture notes before class help you grasp the concept better?
- Have you taken physics before?
- How is this class different than your other classes?

This type of questionnaire would provide very insightful qualitative data that is not extractable from quantitative results like final exam grades or pass fail rates. There is a student course evaluation collected at the end of each semester, but the questions presented above are more specific to this research in that they ask directly about the practices used in each methodology rather than general student comments about instructor performance and best/worst aspects of the class.

Rather than doing a historical statistical analysis on the students' grades or simply reacting to high DFW rates for the course, the teaching methodologies implemented in this course could be evaluated using the Force Concept Inventory (FCI)[15] This method uses pre- and post-test data to evaluate whether students actually understood the concept better after the lesson. This is considered to be a "gold standard" evaluation technique that has been used for many years by many different institutions.

One of the assumptions of this research is that the primary contributing factor to the high DFW rates of this course is the teaching method used. It is reasonable to suggest, however, that another factor may have a significant impact on the DFW rate for this course, namely, the mathematical ability of the students. To account for this factor in future research, I recommend administering a math assessment prior to the physics final exam. This exam would reflect the mathematical concepts required for the final exam, but without the physical context. By correlating the results of this assessment with the results of the final exam, the role that mathematical preparedness plays on the success in the introductory physics classroom can be determined.

Another way this research could be extended would be to examine the inverted classroom approach with the online final exam after implementing new practices in the classroom. These practices would serve to further engage the students with the material in the classroom while still falling under the inverted classroom model as mentioned above. The next section will cover a more detailed description of what some of these practices might look like when applied to the PHYS 2101 course at UNCC.

5.2. Recommendations for Future Course Design

Since there are no significant statistical differences in the final exam data for the two teaching methods, I recommend the adoption of the inverted classroom approach for several reasons, with some modifications to its current use at the University of North Carolina at Charlotte.

One of the strengths of the inverted classroom approach is the addition of a recitation/problem-solving session. This meeting time should be better correlated with the material taught in class if an outline of what was discussed in the previous lecture review session was provided to the TAs. That way, TAs would know exactly how the professor introduced each concept and they could then follow that model as they work out the problem-solving sets for the students.

The modern inverted classroom relies heavily on video presentation of the material in the online lecture. This material could be modified to include common misconceptions about physics rather than just the facts. It is also important to make sure that these materials are engaging to students and are not simply recorded lectures or notepad videos. Real life examples would help students connect the concepts they are learning to what they actually see in the real world.

Online quizzes to assess comprehension of the required pre-lecture videos would also be extremely useful in helping instructors prepare for the in class lecture review session. During this time, the instructor could implement the peer instruction methods used to develop comprehension and problem-solving skills. Just in Time Teaching strategies could also be implemented to assess student comprehension and preparedness for lecture.

Finally, it is essential that instructors, TAs and students work together to ensure student success. When one of these key elements is missing from the conversation it is the students' understanding, their grades, and ultimately their careers that suffer. By opening lines of communication between instructors, TA, and students, course concerns can be addressed. There should be regular meetings between faculty and TAs to discuss pedagogy and create a student-centered learning environment. Students should have a way to provide feedback to the instructor on how well the material is being presented and to allow the students to make suggestions that would improve their learning experience.

BIBLIOGRAPHY

1. N. F. Nathaniel Lasry, Eric Mazur, The Physics Teacher **47** (7) (2009).
2. P. Leilabady, internal documents and personal comunication (2014).
3. C. H. Crouch and E. Mazur, American Journal of Physics **69** (9), 970-977 (2001).
4. H. R. Sadaghiani, Physical Review Special Topics - Physics Education Research **8** (1) (2012).
5. "TEDTalks: Salman Khan: Let's Use Video to Reinvent Education" Link: <http://digital.films.com.librarylink.uncc.edu/PortalViewVideo.aspx?xtid=4861>
6. M. K. Kim, S. M. Kim, O. Khera and J. Getman, The Internet and Higher Education **22**, 37-50 (2014).
7. S. W. Draper, Brown, M., Journal of Computer Assisted Learning **20** (2) (2004).
8. G. M. Novak, New Directions for Teaching & Learning **2011** (128), p63-73 (2011).
9. Video: "This Will Revolutionize Education" Link: <https://youtu.be/GEmuEWjHr5c>
10. Video: "Khan Academy and the Effectiveness of Science Videos" Link: <https://youtu.be/eVtCO84MDj8?list=PL772556F1EFC4D01C>
11. D. A. Muller, J. Bewes, M. D. Sharma and P. Reimann, Journal of Computer Assisted Learning **24** (2), 144-155 (2007).
12. Link: <http://mne.psu.edu/me345/Lectures/Outliers.pdf>
13. N. Smirnov, Annals of Mathematical Statistics, 279-281 (1948).
14. "Analytics" Link:<http://site.theexpertta.com/our-product/features/analytics/>
15. D. Hestenes, M. Wells and G. Swackhamer, The Physics Teacher **30** (3), 141 (1992).

APPENDIX: KOLGOMOROV-SMIRNOV TEST

The null hypothesis for the Kolgomorov-Smornov[13] test is that the samples are drawn from the same conclusion. If the null hypothesis is rejected, the test returns a value of 1. If the null hypothesis cannot be rejected, the test returns a value of 0.

The MATLAB code and results used to perform this test are below.

6/19/2015 KSTest_Analysis

```
% Statistical Analysis of Inverted and Traditional Data
clear; clc; close all;

%Import the data
Written = xlsread('D:\Thesis\KS Test Data.xlsx','Spring 2014');
Online = xlsread('D:\Thesis\KS Test Data.xlsx','Fall 2014');
Traditional10pt = xlsread('D:\Thesis\KS Test Data.xlsx','Traditional 10PT');
Inverted10pt = xlsread('D:\Thesis\KS Test Data.xlsx','Inverted 10pt');

%Compare the pairs of data with the kstest2 to see if they are
%stastistically the same.
h_Online = kstest2(Written,Online)
h_10pt = kstest2(Traditional10pt,Inverted10pt)
```

```
h_Online =

     0


h_10pt =

     0
```

Published with MATLAB® R2013a

Figure 8: This figure shows the MATLAB code used to perform the KS test and its results.