VIDEO TUTORIALS FOR INTRODUCTORY STATISTICS

Interactive Qualifying Project Report completed in partial fulfillment

of the Bachelor of Science degree at

Worcester Polytechnic Institute, Worcester, MA

Submitted to:

Professor Joseph D. Petruccelli

Lucas McLaughlin _______________________________

Kerrin Spinney _______________________________

____

ABSTRACT

Online video tutorials have the potential to be a valuable study tool for students, but are a resource largely untapped by college courses. The goal of our project was to investigate whether a series of video tutorials would be a worthwhile addition to traditional classroom courses, specifically to WPI's MA2611 Applied Statistics 1 course. By examining the results of similar studies, we assessed how effective we could expect such a resource to be, and developed a plan to create a series of statistics video tutorials ourselves. After a series of surveys and interviews with past students and TAs of the course, we determined how best to present the material in video format. We then set about scripting and producing tutorials, and testing them to assess not only their effectiveness at conveying the material in an understandable manner, but presenting it in an engaging way that students of the course would accept and use.

ACKNOWLEDGMENTS

We'd like to give our thanks the following people for helping make this project a success: first, to the staff at the Academic Technology Center for being extremely helpful by giving us tips on how to use the recording equipment as well as letting us use it freely; to all the students that participated in the survey we sent out as well as the final experiment; to Elisa Rosales for agreeing to an interview and telling us her opinions on video tutorials in statistics; to the staff of the Gordon C. Library for allowing us to use the tech suites each weekend as recording spaces; to Barbara Jean Spinney for aiding in our final experiment by making a massive number of cookies for all the participants; and finally to our advisor Joseph D. Petruccelli for his continued support, advice, and dedication throughout our project. Without the help of all of these people, this project would not have been possible.

AUTHORSHIP

Although this project was certainly a joint effort, we each have different strengths, and thus contributed to some areas more than others. During the scripting process, Kerrin was very creative in coming up with example problems that illustrated the different concepts of statistics, and her knowledge of the subject was invaluable. As for the recording process, Lucas had a great amount of experience with recording and editing video, and ended doing the editing and recording setup. The paper, however, was not an individual effort of either of us. We used teamwork and planning to split up the work evenly, and we constantly proofread each other's sections making changes as needed.

CONTENTS:

I. INTRODUCTION

Effectively teaching a group of people a new subject is a difficult task. In a large class, even the best lecture can leave some students confused about a topic. Much of a professor's time is used to grade assignments, give feedback, and prepare material for future class times. Meeting with each student that is struggling is a challenge - though professors devote all the time they can, they often find it’s not enough. After all the work that goes into preparing material, it’s extremely discouraging for them to feel their class is failing to grasp the concepts they have taught. This is because they realize the importance of a good education. They see that each course provides a base knowledge for the next course, each building on the previous. Without a good base knowledge, a student is likely to struggle in more advanced courses.

The feeling of failure, of not living up to the standards of our schooling system, is a terrible one that many students face during their time in school. It can lead to a desire to give up, tune out, or simply accept that they aren’t up to the challenge when they very well could be. It’s extremely important for students that feel this way get the help they need when they need it. As technology advances, education should advance with it to find and provide more ways for students to learn or receive help.

Even without technology, the ideal situation would be that there were always enough teachers and resources available for a struggling student regardless of time of day. This, however, is a perfect case and is impossible because of a few interacting factors. Teachers are often faced with large classes to teach. Many students will require attention, perhaps all at once. With the busy schedules teachers keep, there is no way to give each student all the attention they need. On the other hand, students are busy themselves, and may find their own schedules too full to find time to seek help. There are also those students who are not diligent, and put off going for help. They wait until the last second, either the night before an exam or homework is due, and find themselves completely lost.

While education has come a long way in terms of providing resources for help, there is always room for improvement. Are we utilizing the tools we have to their full potential? Perpetually asking this question is important. In the best case scenario, we could discover a resource that could aid the student by providing a new venue for help. This would in turn save time for the professor by cutting down the number of student questions he or she must entertain. We would like to focus on and research one particular tool that could do just that: online video tutorials.

For over twenty years, video technology has been studied as a possible new educational resource. So why then is it not widely used by universities and schools? The answer is that technology previously did not allow it to be useful. An article from 20 years ago by David S. Moore (1993) details the roadblocks faced at the time. He addresses video as a passive medium, stating that “video is unsuitable as the sole medium of instruction”. He explains what he means with this statement: “The time and expense constraints of producing video do not allow the detailed exposition and carefully worked examples that a text can provide.” Of course, after twenty years, technology has come a long way. The “expense constraints” that Moore references are no longer an issue in regards to technology. Today’s computers offer much memory space and processing power, and the internet provides an amazing platform for delivering video content to students. Furthermore, Moore was correct in saying that by itself video is passive. However, the internet has evolved immensely since the time of Moore’s article. Embedding a simple program beside a video could counteract the passivity. Practice problems could supplement videos, allowing students to get instantaneous feedback. This provides the interactivity needed to augment the video's passive nature.

All this is to say that the issue is no longer a lack of technology, but is the way in which it is utilized. A boring, unprofessional video will leave a student snoring, while an entertaining but inaccurate video will leave him or her confused. Furthermore, if some form of interactivity is not provided, then the issue of video's passive nature remains. To make educationally effective videos, much research must be done to figure out what works and what doesn't.

There are many journal articles and reports that explore what does and does not work in an educational video. However, very little research has been done in the way of actually implementing these videos into a class. There are two questions to be answered before online tutorial videos can be declared a viable educational option. First, how much work does it take to create a decent size archive of videos tailored to a specific class? If there is too much involved in the process or too many complications arise, then perhaps video technology isn't the best option. Second, do such videos actually make a difference? After all, if they don't positively affect failure rates, average grades, or understanding, maybe they aren't worth the time.

The goal of our project was to provide and measure the effectiveness of an archive of such videos for a course at Worcester Polytechnic Institute. The course we covered was MA2611, Applied Statistics I. As students who had taken this course, struggled with its ideas, and finished successfully, we were well positioned to provide effective sources of help for students. By making creative, engaging, effective educational tutorial videos, we hoped to give future classes the option to review or re-learn material at the press of a button. To do this, we had to do the following: research, prepare, film, and test. First, we researched and found sources that helped us in our endeavor. There are two components an effective tutorial: great video quality and engaging, educational content. By researching filming and recording techniques, we ensured we had the former. The latter was harder to achieve, but by finding previous studies that were relevant to our project goals, we had a better understanding of how to succeed. After our research, we began to prepare the content of the videos based on what we found. This preparation was done in the form of scripts and storyboards. We did our best to carefully word explanations and give examples for each point we made. This included getting an accurate understanding of each concept we discussed. Once we had finished scripting and storyboarding, we moved to filming the videos. This may seem like it would be a straightforward task, but it required thought and attention to detail. Getting a quality clip was dependent on many factors, such as lighting and sound in each set. After our videos were finished, we needed to test them. Thinking of a realistic way to put our videos to the test was a challenge, and required some creativity. However, with much thought and attention to detail, we were able to design a statistically sound experiment that helped us see whether or not our videos were effective.

Before we put work into writing scripts, storyboarding, and filming our videos, we first had to answer an important question: which concepts would we address in our videos? As previously stated, we both have taken the course, so we know which concepts were confusing to us. However, to address the needs of current and future classes of students, each with diverse levels of ability, we attempted to identify which concepts confuse the majority. We did this by administering a survey to students that had already completed the course. To supplement the findings of the survey, we thought it would be a good idea to also interview professors and teaching assistants that had taught or were currently involved in the class. We found their opinions about which concepts confuse students to be valuable.

After obtaining the answers to the surveys and interviews, we had an idea of the topics with which students struggle most. The challenge was coming up with explanations and examples to make these topics clear. Professors are highly trained to teach these concepts, so their explanations are well thought out and articulate. Finding the correct way to communicate the material to the students was not simple. We ran into this during the scripting and storyboarding process. Our scripts determined the dialog in each video, while our storyboards helped us visualize each scene beforehand.

Filming and editing the videos was the last and most demanding step. Using the results of our research and planning, filming and editing was a matter of dedicating time to the project. Most of our efforts were invested in these steps in hopes that we could cover as many topics from the curriculum as possible. Nearly every student has experienced being stuck trying to grasp a confusing idea in a course. It's the worst feeling to think there's no one available to help. But as long as the students have the internet, these videos can provide another 'someone'. The point of taking a course is not to suffer through it, but instead to learn from it. If these tutorials help even a handful of students, we will have succeeded.

II. LITERATURE REVIEW / BACKGROUND

As preparation for designing and producing our tutorial videos, we first researched the teaching and learning process. We wanted to study different ideas in the field of education in order to learn which teaching strategies have proven most successful. As previously mentioned, video is a passive medium, so one of our concerns was that we would be unable to fulfill the needs of each student.

While researching the prominent ideas in education, one recurring concept was the idea that students have various styles of learning and teachers have various styles of teaching. A learning style is the way a student is best able to learn and retain knowledge. For example, a student may better remember material by reading it rather than having someone convey it verbally. A teaching style is the set of strategies a teacher employs during class sessions.

Numerous reports and studies of learning and teaching styles exist, most of which endeavor to show the ways in which learning styles and teaching styles interact. The issue with summarizing these studies was that each used a different model for the learning styles (most agree on the teaching styles). These models have many slight differences, such as the number of possible learning styles or the wording of each style's definition. Instead of detailing the model used in each report, we found it sufficient to consider the main features of these models. The three common categories of learning styles were visual, auditory, and kinesthetic (Fleming; James, 2009; Roell; IUPUI, 2008). The models in each of the following studies tended to add to the above three categories or split them into more specific styles. This is why the three mentioned styles adequately summarized the models in each of the studies. A visual learner prefers to see things, and learns best from charts, pictures, text, etc. An auditory learner gains most from hearing or saying things, and best retains information presented through voice or sound. Finally, kinesthetic learners enjoy hands-on experiences, and respond well to demonstrations of material or group activities.

While a specific model of learning styles is not widely agreed upon, the model for teaching styles was often very similar across reports. Many of them stated that there are four types of teaching styles: formal authority, demonstrator, facilitator, and delegator (College of South Nevada, 2008; Stein, J., Steeves, L., & Smith-Mitsuhashi, 2001; University of South

Carolina). Teachers who use formal authority are prone to make a lecture center around them, meaning they teach the concepts while students sit, listen, and absorb. They do not include many example problems or require much participation, thus not building any student-teacher relationships. A demonstrator will take a slightly more interactive approach by asking the students questions or solving example problems on the board. While the class is still expected to listen for the majority of the lesson, there is some participation included. A facilitator will often use activities that demonstrate how different concepts work. These activities are either group or single-student oriented, but all strive to give the students hands-on experience. The course content is still present, but it is no longer solely conveyed through lecture. The last type, the delegator, will give the students extreme amounts of responsibility. They will often assign large projects for students or groups of students in which they act merely in a consultative role. In other words, the students are given an end goal and challenged to meet it through their own innovation.

It is important to note that students and teachers often do not adhere to one particular style. A student that learns best through verbal communication will likely be able to learn through text or interaction as well. However, those mediums may not be as effective as a lecture or conversation. Similarly, a professor may favor showing the class example problems of one concept while using group activities to teach another. The fact that students and teachers do not possess a single teaching or learning style made it difficult to study the different styles.

With the above general classifications of the learning and teaching types, we now move on to reviewing the results and findings of the actual studies. The first example of such research was an extensive journal article by Anshu Arora, Reginald Lesean, and Mahesh Raisinghani (2011). Immediately in the introduction, the authors make an overarching statement that “learning is best accomplished when the individual needs of learners are established well in advance” (p. 1). Finding the evidence for the validity of this statement was the challenge.

To face this challenge, the authors chose to first gather a sample. The sample consisted of 161 students, 55 males and 106 females attending business school. The majority of this sample were between 18 and 23 years old, although the ages ranged from 18 to 53. As a side note, the authors say the following in regards to the data collection process and its credibility:

> “Since the data collected for a sample size of 161 observations are from business majors at a historically black college and university during two-semester periods between spring and fall semesters of 2010, the findings should be considered exploratory and preliminary and may not be generalizable to other universities, or other groups of undergraduate and graduate students.”

Following the sample selection, the authors gathered their data.

They first administered a questionnaire called the Felder-Silverman model and Index of Learning Styles (Felder & Silverman, 1988). This allowed them to determine the learning style of each of the students in the sample. From there, they administered the Center for Occupational Research Development (CORD) teaching style inventory to the same students (the CORD is a national nonprofit organization dedicated to leading change in education). In a nutshell, the CORD teaching style inventory helped the authors determine how the students perceived their teacher’s teaching style and teaching effectiveness. The authors explain why they had the students complete the inventory rather than the faculty: “This way, we were able to assess and match the preferred learning and teaching styles from the learners’ perspective.”

The authors then thoroughly analyzed the results of the questionnaire, inventory, and the relationship between the two. They found numerous associations between different teaching styles, different learning styles, and teaching effectiveness. Before the study, they had proposed six different hypotheses:

- Hypothesis 1: There is a strong positive relationship of learning styles on teaching styles.
- Hypothesis 2: There is a strong positive relationship of teaching styles on learning styles.
- Hypothesis 3: There is a strong positive relationship of teaching styles on teaching effectiveness.
- Hypothesis 4: There is a strong positive relationship of learning styles on teaching effectiveness.
- Hypothesis 5: There is a strong positive relationship of learning styles on the student performance.
- Hypothesis 6: There is a strong positive relationship of teaching styles on the student performance.

Of these six, the authors only rejected Hypothesis 4: There is a strong positive relationship of learning styles on teaching effectiveness. They posit that although a professor would like to improve his or her teaching effectiveness by catering to each student's learning style, it is nearly impossible as it requires the teacher to know each of his or her students' learning styles individually. The authors failed to reject every other hypothesis based on their statistical analysis.

Although these results suggest that different teaching styles affect performance based on a student's learning style, the authors themselves admit their findings are unlikely to be conclusive because of the limited variety in their sample. However, there is yet another reason they may be inconclusive: a majority of the data they collected was based on student opinion, such as the teaching style and teaching effectiveness of each professor. Because of this, some of the responses pertaining to a professor's teaching style or effectiveness may have been affected by the student's opinion of the professor himself. Instead of teaching effectiveness, it is more a measure of student opinion of teaching effectiveness.

R. J. Charkins, Dennis O'Toole, and James Wetzel (1985) also studied the possible effects of learning and teaching styles on student learning. Their article, titled "Linking Teacher and Student Learning Styles with Student Achievement and Attitudes", attempts to find a relationship between a student's learning style and a professor's teaching style. By finding such a relationship, the authors believed they could find how teaching styles affect both student performance and attitude in a class. They wanted to use this information to better understand what makes a good teacher.

The authors began by preparing their sample. They gathered around six hundred students and twenty teachers at Purdue University. Their methodology includes statistical and technical vocabulary as well as many details of the study, but a summary provides an adequate understanding of the results. To get a real idea of how the students would perform under the teachers, the authors decided against giving the participating teachers specific course designs to follow. To begin, the authors used "The Grasha-Riechmann Learning Styles Questionnaire" which was created to determine the learning style of a student or the teaching style of a professor based on his or her score. This questionnaire was completed by every student and teacher. The authors then used these scores to create an independent variable meant to measure the "congruence" or "divergence" between different teaching and learning styles. Next, each student was given two special tests before beginning the study: the Test of Understanding College Economics (TUCE, designed to measure achievement levels in economics) and the Karstenson-Vedder Questionnaire (used to each student's attitude towards economics).

Once the courses were completed, the students were asked to once again take the two special tests. Using the difference between the scores of the two tests before and after the courses (and several other variables not mentioned here), the authors were able to use statistical reasoning and come up with the results of their study. The general results according to the authors are as follows:

> "1. The larger the divergence between teaching style and learning style, the lower the student's gain in achievement in economics. 2. The greater the divergence between teaching style and learning style, the less positive the student's attitude toward economics. 3. The results of this study support the findings of earlier studies in this area; thus, we can suggest some interesting educational implications and directions for future research."

Because of the thorough detail in design of this study coupled with its large sample size, these findings are legitimate. They fully support the idea that teaching style can be an important factor in the achievement of a student as well as his or her interest in a subject. With that said, there is still one more aspect of student experience to consider.

While the previous two articles support the idea that catering to a wide variety of learning styles is effective, the next compares the success of a traditional classroom experience to that of a course based around interaction and participation during class. The paper, called "Improved Learning in a Large-Enrollment Physics Class", was written by Louis Deslauriers, Ellen Schelew, and Carl Wieman (2011).

They began by planning the study and collecting a sample. Unlike the authors of the previous article, Deslauriers et. al. decided to much more carefully dictate the teaching styles and strategies employed on their sample of students. They used two different teaching styles: a traditional lecture style (the control group) and an interactive style (the experimental group). These techniques were to be used in an introductory level calculus-based physics course. The control group was taught by a faculty member with good student evaluations and years of teaching experience, while the experimental group was taught by a postdoctoral professor "using instruction based on research on learning". The study was conducted at the University of British Columbia during the second term. The student sample was substantial for both of the two lecture sections, with 267 in the control group and 271 in the experimental group.

For the first 11 weeks of the semester, the sections were taught in the same way. All components of the course (this includes homework, exams, and lab sessions) were the same for both sections save for each having a different professor and slightly different lecture styles. For those 11 weeks, certain aspects of each class section were studied, such as attendance and midterm scores. The researchers also monitored class engagement by having four faculty members observe the class closely.

It was during week 12 that the true experiment began. The control section continued to be taught by the original professor, while the experimental group was taken over by two of the authors of the paper (L.D. and E.S.). They took on a completely new teaching style as follows:

> "The design goal was to have the students spend all their time in class engaged in deliberate practice at 'thinking scientifically' in the form of making and testing predictions and arguments about the relevant topics, solving problems, and critiquing their own reasoning and that of others"

As a result of this change, there were notable differences in the students of the second section. As expected, nothing changed in the control section during the 12th week. However, in the experimental section, student engagement seemed to double, and attendance increased 20%.

Following the week of the experiment, the students of both sections were administered a test (created by the authors and the instructor of the control group) to find out what they had learned. The scores of each section were then compared. To give the reader a better feel for how well the students in the experiment performed, they provided the histogram shown in Figure 1.

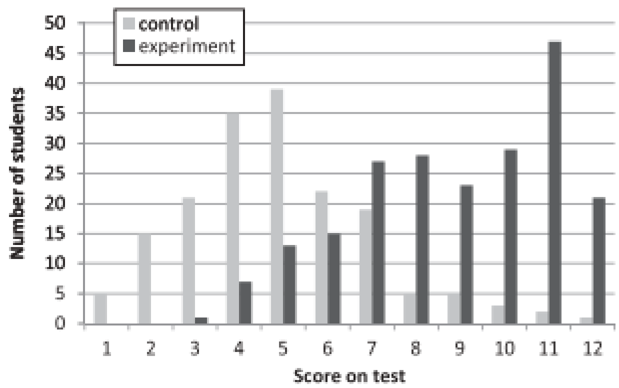


Figure 1: Histogram showing the results of experiment.

The histogram shows a striking difference between the two sections. The experimental section scored significantly higher than the control, supporting the idea that learning can be affected in many ways by different teaching techniques.

For a final measure of the success of the experimental teaching strategy, the authors administered a survey to the students in the interactive class which 150 students completed. 90% agreed with the statement “I really enjoyed the interactive teaching technique during the three lectures on E&M (Electromagnetic Waves)”, while only 1% disagreed. Similarly, 77% agreed with the following statement: “I feel I would have learned more if the whole physics 153 course would have been taught in this highly interactive style.” (Only 7% disagreed).

Although the results of this study seem overwhelmingly in favor of interactivity in the classroom, the study contains a number of glaring flaws. First of all, the design choice in which the authors decided to replace the teacher of the second class with themselves causes confounding. Specifically, they have introduced two confounding variables. A confounding variable is a variable that is related statistically to the independent variable but is not related to the study itself. The independent variable is the teaching style applied to the class. The confounding variables are the professor they replaced and the authors themselves. They introduced not only a new teaching style but themselves as well, removing the original professor. Because of this they cannot be sure to what extent the changes in the other variables (attendance, engagement, etc.) are related to the first professor not teaching well, to themselves teaching better than the first professor, or to the new interactive teaching style.

The second flaw with their study was in the responses to their surveys. The authors issued the survey to the 271 students in the experimental class. However, only 150 of them replied, meaning the opinions of nearly half the class were not recorded. This leaves room for doubt in the validity of the survey results.

In addition to research on teaching and learning styles, we also studied current research pertaining to use of tutorial videos in a course and how effective they can be. However, before creating and implementing any resource into a preexisting environment, it is essential to know how the addition will be received. There is little sense in making a series of tutorial videos if they will go unused by students. H. Brecht , David and Suzanne M. Ogilby's 2008 report, which examines students' responses to and usage of a series of video lectures tailored to and implemented alongside a conventional college course, gave us an idea of whether students would use our tutorials.

In order to test the appeal and effectiveness of video as a study resource, the lectures were made available to a class and the students were asked to complete an anonymous survey before their final exam. Though the videos in question were a series of full-length lectures, unlike the shorter tutorial series we set out to produce, we were able to use the results of their study to inform our decisions in planning our project because the videos were simply made available to the class. Students were free to use or ignore them as they saw fit. Alternative study resources, such as classroom lectures, instructional PowerPoint files, homework assignments and solutions, lab practice problems and solutions, and the textbook, were always available to students. This is similar to how we hope to implement our videos. If students do not believe they are a worthwhile and time-efficient study resource, they will not use them, no matter how effective they may actually be. This study has helped us determine how to create a video resource that students will be receptive to as an addition to their course, rather than one they may deem a waste of time.

Analysis of the survey responses revealed that a majority of students took advantage of the videos and felt as though they helped improve their performance and understanding of the material. Out of the 56% of students (74 out of 132) who responded to the survey, 73% (54 out of 74) reported making use of the videos, suggesting that students will accept and make use of video technology as an addition to a course. The survey then asked more targeted questions to determine usage patterns of the videos. A test of population proportion was conducted on the collected data to determine if the proportion of students who reported using the videos for specific purposes was statistically significant, with the null hypothesis for each question being that an insignificant proportion (defined as a percentage less than 62%, giving $p=50\%$) of students found the video helpful for the specified purpose or category, and the alternative being that a significant proportion (defined as a percentage greater than 67%, giving $p>50\%$) found the videos to be an effective and efficient study resource. A statistically significant number of respondents reported that the videos helped them understand the subject material (68.5%), complete homework assignments (72.2%), prepare for weekly exams (72.2%), and obtain tutoring help (63%), so the null was rejected for these questions and the usage for these purposes deemed significant. Additionally, while the percentage of students who reported using the video lectures to prepare for midterm exams was not statistically significant (61.1%), the percentage of students who felt the videos helped to prepare them for the exams was (68.5%). This could indicate that the students felt the video lectures were more helpful for initial learning than for review – perhaps due in part to the lectures' length and breadth of subject matter. Shorter, targeted tutorial videos could be seen by students as a more viable resource for exam review. Finally, 31.5% of students reported viewing the video lectures prior to attending classroom lectures – a larger percentage than should be expected, according to previous education research that indicates students do most of their studying two days before an exam, regardless of the resources they have available (W. Maki and R. Maki, 2000). This increase could suggest that the availability of an appealing, convenient video tutorial series might influence students to do more than simply cram for exams.

The study also used a Kolmogorov Smirnoff test to compare student performance, as measured by course and final exam grades, in the course given access to the video lectures with a course to which the lectures had not been made available. However, not only are these results not directly applicable to our project, as it is unknown whether full length lectures and short tutorials have equivalent effects on performance, but the results of the comparison are of dubious merit even in their relevance to the study itself. 20.4% of survey respondents reported that the videos helped them pass the course, and a comparison of the two class groups' grade data revealed that, while 24.2% of the class without the videos failed the course, only 6.8% of students with access to the videos failed. However, access to the videos was not the only difference between the two groups. The class with access to the video lectures had more time in their course for interactive labs. No attempt is made in the report to differentiate between the effects of the video series and the effects of the in-class labs - therefore, while the students' high opinion of the videos suggests that they may have had some positive effect, we cannot conclude from this study's findings whether the video tutorials had any real effect.

[Address Comment 2!]

The results of the previous study support our claim that students tend to accept and make use of tutorial videos. We can now examine more closely the effect a video resource will have on student learning. In his recent paper, Nicholas Wyant (2013), a student at Wichita State University, recognized a specific need in his school in regards to using online academic databases to find articles and research papers. He saw that because there is a variety of different databases to use, each with its own format, students could easily get confused when trying to find relevant material. Wyant decided to create a simple tutorial video showing the steps a student should take when looking through a database for a paper.

He based the tutorial on the Sociological Abstracts Database, maintained by a company called ProQuest. To measure the failure or success of his video, he tied the tutorial to an assignment in a sociology course at his college (the assignment being to find three articles about a specific sociology topic using the database). The course consisted of 217 students, all taught by one professor.

Before stating the results of the study, Wyant discusses the methods he used in both writing the script for and making his video. To record his video, he used the screen capture feature of Macintosh's QuickTime Player. Once he had the actual footage of himself navigating through the ProQuest database on the computer, he wrote the script for what he would say behind the video. In recording the audio, he mentions the importance of making sure there is no ambient sound in the track. In the end, the video was only a little more than three minutes long.

Once the tutorial was finished, Wyant sent his video to the teacher of the sociology class. When the assignment to find the articles was given, the teacher showed the students the location of Wyant's video and informed them of its usefulness. The teacher was able to see which students viewed the video and how many times it was viewed. After the assignments were turned in and graded, the final scores of the students were examined. Out of 217 students, 82 did not watch the tutorial. The average score of these 82 students was a 58% out of 100%. The remaining 135 students that viewed the video averaged a 94%, a 36% difference.

This study shows an important advantage of tutorial videos: instead of the teacher taking precious time out of class to go through the steps of searching a database, he simply pointed the students to the video. However, the design of the study itself contains flaws. For example, an argument can be made that says those who watched the video are likely more studious and diligent than those who didn't. Because of this, we cannot be sure whether the better grades of students that watched the video resulted from their studiousness or from their watching the tutorial. Furthermore, there may be students that did not watch video who failed to turn in the assignment. This would cause outliers in the grades of the 82 students that did not watch, and potentially lower the average significantly. With this said, few conclusions can be drawn from the results above.

Steven A Lloyd and Chuck L Robertson's study "Screencast Tutorials Enhance Student Learning of Statistics" tests the viability of a video tutorial as a learning aid for statistics. The video was designed as a screencast – a video recording of work done on a computer screen synced with audio commentary – and demonstrated several steps of statistical analysis in working with IBM's SPSS software. This tutorial was presented to a randomized group of students in a college psychology course, and its effect on their ability to understand and apply statistical knowledge was observed. The style of this tutorial - a short video targeting a specific statistical concept and its application - closely matches that of the videos we produced in our tutorial series. The difference is that the video used in this study was intended as a standalone learning resource, presented to those with minimal prior statistical knowledge and not in conjunction with a classroom environment. That said, its effectiveness as an aid to initial learning of statistics - especially in comparison to other, traditional means of initial learning – informed our ability to design tutorial videos that effectively demonstrate statistics concepts and their applications.

In order to measure the impact of the tutorial video on student performance, the 53 participants were sorted into two groups. One group was shown the video; the other was shown a text tutorial. Great care was taken to ensure there would be no confounding factors in the division of the groups. Students were divided into groups using stratified randomization to ensure an equal distribution of gender, and all those involved in the study had the same level of prior math experience and had taken the same prerequisite course in elementary statistics. Even the students' attitudes toward math and computers were taken into account, using the results of a 10-item Math Anxiety Scale (MAS; Betz, 1978) and a 6-item Computer Anxiety Scale (CAS; Lester, Yang, & James, 2005).

The group shown the video saw a tutorial roughly 12 minutes in length which "demonstrated the following steps of statistical analysis: data entry, conducting an independent samples *t* test analysis, and working with output files in SPSS." The control group was provided with text taken from the SPSS user manual covering the same concepts, including relevant screenshots. Each group was given twelve minutes to study their respective tutorials, and then took part in two experiments where they were required to perform an independent samples t test analysis using SPSS. In the first experiment, the students were required to complete the task in 25 minutes and were not given access to their respective tutorials, in order to test their pure retention of the material. In the second experiment, the time limit was increased to 55 minutes, and students were allowed to freely review their tutorials alongside their task. Because the same student groups participated in both experiments, it is possible their performance in the second experiment was affected by their familiarity with the subject matter gained in the first – however, as students were allowed to view their tutorials while they completed the test in the second experiment, it is not likely that this familiarity affected their performance any more than an increased time limit of 80 minutes would have.

The results of the first experiment seem to indicate that the screencast tutorial had a positive effect on student learning and performance. Not only did the group shown the video take less time to complete the task - an average of 15.20 minutes to the text group's 18.06 ($p = .006$), but they also received higher scores on the completed task - an average of 7.27 points out of 10, much higher than the text group's 4.5 out of 10 ($p < .001$). During the second experiment, where the students were given extra time and the ability to review their respective materials during task completion, the video group still scored higher on the task (7.27 out of 10 compared to the text group's 5.36 out of 10 ($p = .044$)). However, their scores did not improve with the ability to interact with the video and there was less of a gap between the groups' completion times. The text group, on average, finished before the video group, taking 34.27 minutes to the video group's 38 ($p = .41$). Nevertheless, the video group's scores remained significantly higher

($p<.05$) than those of the text group, supporting the idea that video tutorials are an effective resource.

In Dongsong Zhang, Lina Zhou, Robert O Briggs, and Jay F. Nunamaker Jr.'s study "Instructional Video in E-learning: Assessing the Impact of Interactive Video on Learning Effectiveness," the advantages of interactive video over passive video are explored further. The study's authors posit that a learning resource which engages students by frequently challenging them to process and use the material being taught will not only help the students to learn and perform better, but will make the learning process more enjoyable for them.

In this study, interactive video is defined as "[video which] allows proactive and random access to video content based on queries or search targets," essentially allowing users to skip through a long lecture video by using an index to find the parts they wish to review. This eliminates the time-wasting and frustrating process of scrolling through a long video to find the desired part, and reduces the passive linearity of the lecture. We speculate that our well-organized and indexed series of shorter tutorial videos will work on much the same principle.

In order to test their hypothesis, the authors examined student learning and satisfaction in four different environments. "Three were e-learning environments—with interactive video, with non-interactive video, and without video. The fourth was the traditional classroom environment." 138 undergraduate students from the same introductory management information systems course were recruited and randomly assigned to one of four groups, which were then randomly assigned to one of the four environments. Participation was voluntary, and students were offered extra credit in their course based on their individual performance in order to encourage meaningful participation. Additionally,

> "Participants completed a preliminary survey two weeks before the experiment to provide their demographic information such as age, GPA, computer experience, and prior experience of e-learning. A series of analysis (sic) did not find significant difference among four groups on those dimensions. [. . .] Therefore, we could assume homogeneity of pre-experiment skills, e-learning experience, and learner characteristics among groups."

In order to eliminate potential differences in the content of the environment, and to ensure the content was relevant and supplementary to the course, the course's professor was recruited to produce the video and lecture materials provided to the three e-learning groups. The students in the interactive and non-interactive video environments watched the same lecture video, side by side with relevant PowerPoint slides and lecture notes – the only difference being that the interactive video group was able to easily skip around the video using the index as previously defined. The students in the group without video only saw the slides and notes. The course's professor gave the students in the lecture group hard copies of said materials, and delivered a lecture that did not differ in content from the video he had prepared. By controlling the content in this way, the effectiveness of the four presentational styles was isolated for more accurate study.

Before the groups' participants watched the video, read the materials, or sat down to the lecture, they completed a pretest consisting of true/false questions to determine prior knowledge of the subject matter. After the pretests were analyzed, no significant differences were found. Each session lasted 50 minutes. This ensured that the video groups had plenty of time to review and interact with their 29 minute video lecture, and that the traditional-classroom group had time for a question and answer session following their lecture. After each session, participants were administered a posttest, consisting of objective questions more difficult and specific than those of the pretest. They were also given a survey in which they were asked to “rate their satisfaction with learning effectiveness using a 7-point Likert scale, ranging from extremely dissatisfied (1) to extremely satisfied (7).” Potential test scores of both pre and post-test ranged from 0 to 50, and the differences between individual pre and posttest scores were used to determine each environment’s effect on student learning.

Results showed that video interactivity has a positive, statistically significant effect on student performance and learning, and that it also significantly added to the enjoyment of the learning process. The interactive video group had an average 34.1 point improvement from pre to posttest, as well as an average of 6.46 on the satisfaction survey. The non-interactive video group had an average difference of 27.7 and an average satisfaction of 5.94. The group with no video had 26.7 and 5.74, and the traditional classroom group had 23.7 and 5.03. Below, tables II-A through II-D illustrate the results of the study.

Descriptive statistics of learning outcome in different treatments (post-gain)

Groups	**Means**	**Standard deviations**
E-learning group with interactive video (1)	34.1	8.87
E-learning group with linear video (2)	27.7	8.85
E-learning group without any video (3)	26.7	10.02
Traditional classroom group (4)	23.7	8.79

Table II-A

Mean differences (*P*-values) between groups on learning outcome (post-gain)

Groups	2	3	4
1	6.49 (0.005)**	7.41 (0.001)**	10.47 (0.00)**
2		0.92 (0.967)	3.98 (0.184)
3			3.06 (0.417)

Table II-B

** The mean difference is significant at the 0.01 level.

Descriptive statistics of learner satisfaction

Groups	Means	Standard deviations
E-learning group with interactive video (1)	6.46	0.56
E-learning group with non-interactive video (2)	5.94	0.84
E-learning group without instructional video (3)	5.74	0.75
Traditional classroom group (4)	5.03	0.67

Table II-C

Mean differences (*P*-values) between groups on satisfaction

Groups	2	3	4
1	.51 (0.014)*	.72 (0.00)**	1.43 (0.00)**
2		0.21 (0.621)	.91 (0.00)**
3			.71 (0.00)**

Table II-D

* The mean difference is significant at the 0.05 level.
** The mean difference is significant at the 0.01 level.

The small difference in improvement between the linear video and non-video groups' test scores supports the idea that simply integrating a video resource into a pre-existing environment will not necessarily add value to that environment. In order for video to be worthwhile and effective, it must engage the viewer and prompt them to interact with the material. We took this principle to heart, and worked to ensure that our tutorial series would be a valuable addition to a statistics classroom environment.

III. METHODOLOGY

Based on our background research, we are prepared to advocate that video tutorials could prove to be an effective educational resource for students in a college setting. To reiterate our reasoning, we believe that they can provide a resource for independent study or help, which in turn will free up the time of the professor. This sounds like a win-win situation, one that cannot fail. However, video cannot just be thought or spoken into existence. Education tutorial videos take preparation, hard work, and time to create. Behind every video that is worth watching is the time and energy that went into making it. As a society, we have come to expect the videos we watch to live up to high quality standards. During our project, we ended up working very hard to meet these standards.

Additionally, we have mentioned testing the success of our videos. Only a well-designed study can be deemed useful and conclusive. It was critical to ensure the tests of our videos were designed with care and thought. Even the tiniest detail could have rendered the results of our study meaningless. Planning the details beforehand was very important, and in executing the plan, we worked hard to adhere to it.

Before we looked for advice to help us with our filming techniques, it was important for us to decide the style we were going to use in our tutorial videos. First, we agreed that an introduction to each video would be a good way to start off. We thought that by being welcoming and cheerful during a short introduction segment, our viewer would be more likely to continue watching our video. Originally, we were going to include in each introduction a rundown of what would be covered in the video. However, we realized that this was too large of an amount of time to add to every film. Instead, we came to the conclusion that the best option would be to include a rundown in plain text next to or beneath the video. This gives the user the ability to refer to the list at any point while cutting down on the length of the tutorial.

We ended up beginning each video with a simple shot of both of us talking to the camera. We would introduce ourselves and explain that we hoped to help the viewer gain a better understanding of statistics. While the introduction of our video remained virtually unchanged from our original plans, the format of the content in each video changed immensely from what we first planned.

At first, we created videos that consisted of mostly all explanations. We would introduce a simple example and then explain terms that arose throughout. What we discovered was that this style of video did not flow well, and ended up being more of a shortened lecture than an interactive video tutorial. After shooting and completing two or three videos of this style, we decided to try something new. Instead of trying to explain an entire example and all of the terms associated with it, we attempted to create sample multiple choice problems. Generally each problem had four different options, one of which was the correct answer. The example problem would appear on the screen and one of us would read the problem. We then gave the user about five seconds to pause the video and attempt to answer the question themselves. After five seconds, we would go down each multiple-choice option and explain why it was right or wrong. The questions we asked resemble homework problems from past statistics classes.

Aside from the style of each video, we had to ensure that we used captured quality video and audio. We began by researching recommended filming and recording techniques. While we had previous experience with recording and editing video, we by no means considered ourselves experts. In order to avoid making beginner's mistakes, we found several different websites committed to offering advice for beginner filmmakers. In regards to the technical side of video recording, we found three main areas to consider: lighting and setting, sound recording, and video editing.

The first topic, lighting and setting, pertained to the visual aesthetics of our videos. An important factor to consider was the setting of each clip and how it affects the quality of the footage. A particularly helpful webpage created by Paul Grabowicz (2012) aided us in determining how to best prepare for shooting our videos. From the advice on the page, we found some of the tips to be more helpful than others. The first important point was that before shooting, it is important to plan out the details of each scene. The site gave the following advice: "Talk over your shoot with other members of the production team and make sure you're clear on what shots you need to tell the story." This advice coupled with our previous experience with creating videos led us to create storyboards and scripts for our videos, which we will discuss later.

The second important piece of advice given was to be aware of the framing and composition of each shot. The site warns to make sure that there are no distracting items in the background, such as a pole or plant coming out from behind the subject. Even the smallest detail of a video can distract a viewer and dampen the effectiveness. The site also mentions that there should be space for the entire head of the speaker to fit on screen. That is, leaving enough room between the subject and the camera is important.

The third and final important piece of advice we found to be helpful was that changing angles and perspectives occasionally is a good practice. Simple eye level shots are effective, but can be boring if used alone. The tips on this site certainly helped us in creating each video. We consistently planned out each shot, wrote scripts, and talked over each scene before the camera was even turned on.

Moving on from the setting of a video, we searched for sites with advice about attaining good lighting in our shots. What we found was a site called Filmmaker IQ (2011). Specifically, they had a page dedicated to collecting videos of lighting tutorials, many of which we found very helpful. Although there are many different strategies for lighting a shot, one stood out to us as the best option: three-point lighting.

Three-point lighting is most commonly used in interviews or shots with only one person in frame who is talking to the camera. In three-point lighting, there are three main lights illuminating a shot: the key light, the back light, and the fill light. The key light is the main source of light in the shot. Good options for key lights are daylight from a window or a bright lamp. The key light should preferably point towards the subject from behind the camera. As for the back light, its purpose is to separate the subject from the backdrop (whether it be a wall or an open room). To accomplish this, the back light is often placed behind the subject and is weaker than the key light. This creates a soft outline around the subject. The final touch is adding the fill light. A fill light highlights facial features of the subject, such as hair, eye sockets, or jaw line, giving the subject more depth. While the key light does the majority of the lighting, the fill light "fills in" the places the key light cannot reach. Throughout our videos, we stuck to this technique as well as we could. While always ensured we had a fill light, it was the backlight that we did not always have. Because we had only one lamp and no strong spot lights, we often neglected to add the backlight and instead tweaked the lighting during the editing stage to make it appear as though we had one.

With a better understanding of lighting, we shifted our research to the second facet of a good video: sound quality. Recording sound correctly can be a challenge, and the quality often boils down to the equipment used in capturing it. While the equipment we used is discussed later in this section, there were a couple things we considered in preparing our shots with regards to sound.

One website called Longzijun (2013) mentions two important factors of sound recording: eliminating ambient noise and speaking clearly and audibly. The biggest issue with recording sound is that if it is done incorrectly, it can cause a whole shooting session to be wasted. The good news is that with many of the latest editing programs, some mistakes can be fixed by fiddling with volume levels or effects. However, it is preferable to avoid spending time editing by getting the recording right on the first try.

The first issue to consider when recording sound is something called ambient noise. Ambient noise is sound that is made that the filmmaker does not wish to capture. This includes everything from soft, constant noises like the hum of the camera or a heater, to loud, sporadic noises, such as background conversation or wind. To get rid of it, it is important to film in a controlled environment. Instead of recording on the side of a busy road, it is obviously better to record in a soundproof room with the door closed. In our case, we filmed indoors which allowed us to better control our environment. We only had a problem with ambient noise once. We filmed three or four scenes in a room which had a loud heater, and the noise from the heater was evident in the audio of each clip. During the editing stage, however, we were able to remove the noise and make the clips sound much better.

It is also important to speak clearly and audibly into the microphone or camera. If a subject’s voice is too soft, the audience could have trouble hearing them and will not hesitate to leave the video. Similarly, if a subject’s voice is too loud or even approaching a yell, the viewer may find this unsettling or distracting. While recording our videos, we always tried to find a happy medium between yelling and talking too softly.

The final piece to a well-produced video is pouring time into editing what was recorded. For us, this was the most time consuming step, and demanded attention to detail at all times. When recording video, it was impossible to get a flawless shot every time. We are human, and humans make mistakes. From stumbling over words to recording a scene for a bit too long, we needed some way to easily remove mistakes and refine each clip. A simple, personal web page written by Bianca Filoteo (2013) gives several tips about video editing.

Among the most important is to only include the best scenes and takes. We knew we are competing for the viewer’s attention, so we included only the footage necessary to get our point across. Video content must be engaging and contain variety while remaining relevant. Because our videos were meant to be educational, it was an extra challenge to find ways to draw viewers into watching. We tried to be as concise as possible in order to keep our videos short and to the point. We wanted each topic to have its own video instead of cramming a handful of topics into one.

The next tip is similar to the last, but nonetheless important: keep effects in each scene to a minimum. Special effects and animations should be reserved for special occasions, such as an introduction. Effects include transitions (moving from one shot to the next), color manipulations, on screen text, and animations. Extra frills could do more bad than good for a video. In our case, we stayed away from fancy effects and instead stuck to basic text and transitions.

Once we had completed researching the three components of a well-made video, we turned our attention to the content. Learning common recording techniques was an easier task than deciding what should be covered in our films. As we stated in the paper introduction, both of us (the authors) had taken the MA2611 Statistics course and struggled with its concepts. We may have remembered what confused us in the material, but in order to get a better idea of which topics to cover, we consulted students that had recently finished the course by creating and sending them a short survey. Additionally, we contacted professors and teaching assistants (TAs) that were or had been at one time involved in MA2611 and asked them which concepts they believe confused the students in the course.

In creating the survey which we administered to the recent students, we decided to keep it short, as students often don't complete surveys that consume large amounts of their time unless compensation is offered. For the questions, we focused on asking things that would help us learn which topics the students struggle with. We also asked their opinions on any tutorial videos they had seen. The survey was created with a free survey creation tool called SurveyMonkey (surveymonkey.com). See APPENDIX A. Survey Administered to Previous Students of MA2611to view the survey that was sent out.

We then considered the creation of our interviews. Because we wanted to consult both previous professors and the current teaching assistants, we decided to make two interviews with slightly different wording to address either a TA or a professor. Once again, we posited questions that would allow us to gain a better understanding of which part of the material most confuses students. We also included questions about any past experience a teacher or TA has had with tutorial videos. The TA interview questions were as follows:

1. As a TA of the MA2611 course, which topics are particularly confusing to the students?

2. Which concepts in Statistics, if any, are particularly hard to explain?
3. Have you seen or heard of tutorial videos on YouTube that are designed to help students better understand different class topics?
4. Do you think that if provided for a class of MA2611 students, an archive of these videos tailored to the curriculum would help students and improve comprehension and grades?

The interview questions asked of professors were the following:

1. As a professor that has taught MA2611 in the past, what topics were particularly confusing to your students?
2. Which topics in Statistics, if any, were particularly hard to explain?
3. Have you seen or heard of tutorial videos on YouTube that are designed to help students better understand different class topics?
4. Would you consider using such videos in your class?
5. Do you think that if provided for a class of MA2611 students, an archive of these videos tailored to the curriculum would help students and improve comprehension and grades?

Once we had completed the interviews and the survey, we moved to outlining each tutorial. There are two pieces to a video outline: the storyboard and the script. While some filmmakers choose to try and combine them, they are different in their own ways. Both contribute equally to the quality of a video, as they are the blueprints for different aspects of each scene.

It is preferable to begin with scripts, because storyboards can be built directly from them. A script is a guideline for the dialogue and action in any type of video or screenplay. Most readers will certainly be familiar with a typical Hollywood film script, or possibly even scripts used in plays. However, depending on the type of screenplay for which a script is written, the adherence to a script can vary. For example, in a movie, the actors know their lines and are expected to recite them without any change. This contrasts to a game show, in which the host may have a script to follow, but is encouraged to ad lib and react to the contestants.

In our situation, we had to decide whether we were going to follow our scripts to the letter or instead trust ourselves and forget about lines. We believed that a happy medium was the best choice. Our scripts had dialogue and direction for each shot, but we allowed ourselves to stray from them on occasion. This allowed us to be comfortable in our discourse. We knew that if we were focusing on reciting lines we memorized, it would certainly be a challenge to avoid seeming robotic and monotone.

While we always gave ourselves leniency on how closely we stuck to the script, we did not always use the same approach to teaching each topic. To reiterate what was previously explained, in the first couple weeks of the scripting and filming process, we wrote our scripts as if the viewer had absolutely no experience in statistics. We introduced each topic through examples, carefully explaining each term that might be confusing. What we discovered was that videos of this nature came out very similar to a short lecture. It was as if we were re-teaching material that students would learn in class. Here is an excerpt from the original version of the Controlled Experiments Part 1 script (see APPENDIX C. Controlled Experiments Part 1 Script, First Draft for complete script):

> **Introduction**
>
> The three main types of design studies in statistics are observational studies, controlled experiments, and sample surveys. In the this video, we will cover the second of these types:

the controlled experiment. By the end of this video, we will have defined what a controlled experiment is and given examples of how they are used.

Main Discourse

First things first. A controlled experiment is a study in which treatments are assigned to experimental units and then the responses of each unit are observed and recorded. To make this clearer, let's look at an example of a controlled experiment.

Let's say we want to determine whether or not people in America think bacon with brown sugar tastes better than bacon without brown sugar. To do so, we would set up a controlled experiment. In this case, the population from which we are collecting data is all the people in America. Because the population is so large, we cannot hope to conduct an experiment on every person in the population. Because of this, we collect a subset of the population which we call a sample.

To remedy this issue, we completely changed our approach. Instead of assuming that the viewer had no experience in statistics, we instead wrote our scripts assuming the viewer had already been taught the basic concepts in class. As stated earlier, we would come up with three or four example multiple-choice problems that covered key concepts in statistics. The problems we created were akin to problems that might appear on a quiz or test in the class. First, we would introduce and read each problem and its choices. We would then allow the user about five seconds to pause the video and take some time to come up with an answer of their own. After that, we would go through each multiple-choice answer and explain why it was right or wrong. We found this approach to be much easier, and we thought it would be much more helpful for the students. Here an excerpt from the final Controlled Experiments Part 1 script using the techniques articulated above (to view the full script, see APPENDIX E-3. Controlled Experiments Part 1 Script):

Treatments, Responses, Nuisance variables

Oscar Mayer wants to know if bacon with brown sugar on it would sell better than plain bacon. To find out, they conduct the following experiment: they gather a sample and then give everyone in it two pieces of bacon. The first piece has no brown sugar on it while the second piece does. Each person is then instructed to write down which piece they liked better. Identify the following.

1. Treatment **(C)**
2. Response **(A)**
3. Nuisance variable **(B)**

A. Each person's preference

B. The difference in the amount of fat in each piece of bacon

C. The presence or absence of brown sugar

Treatment
In this experiment, there is a single experimental factor: brown sugar. The brown sugar has two levels, level one being the presence of brown sugar and level two being the absence. Because the experimenters either apply or do not apply brown sugar to the bacon before feeding it to the people, they are using combinations of the levels brown sugar. Therefore, the presence or absence of brown sugar is known as the treatment.

Response
The response is the observed or measured reaction of an experimental unit. In this example, the experimenters are finding each person's preference between plain bacon and bacon with brown sugar. Therefore, the response is the preference of each person.

Nuisance variable
In an experiment, a nuisance variable is a factor over which the experimenter has no control. It also has the potential to cause variation in the responses of experimental units. In this example, the difference in the amount of fat in each piece of bacon could be considered a nuisance variable because 1) the experimenters have no control over it 2) many people don't want too much fat in their bacon, and could alter their preference based on the fat content.

After each script was written, we started creating storyboards. While the purpose of a script is to provide instruction on dialogue and action, a storyboard provides a visual of each scene. A storyboard is a collection of sketches (with captions) that depict the setting for each important shot in a video. The sketches are not meant to be detailed, but instead provide a very rough idea of each shot. We intended to build the storyboards according to our scripts so that the two agree with each other. While at first we religiously made storyboards, after our video style changed we found them to be unnecessary, as the content was straight forward and predictable. For an example of one of our earlier storyboards, see APPENDIX D. Storyboard for Original Observational Studies Tutorial.

Once each script was written, the next step was filming and editing our videos. Because we prepared thoroughly for each video, filming was only a matter of dedicating time to recording each scene. During the filming, our biggest concern was using the right equipment to capture the video. The Academic Technology Center (ATC) at WPI had a large collection of sound and video recording equipment of which we made use. After consulting the ATC and discovering the numerous options we had, we decided on the type of camera and microphone we wanted to use.

While the ATC had a wide range of expensive and complex equipment, our videos were all similar in that the shots were stationary and in a controlled environment (little to no ambient noise). This meant that for the majority of our videos we stuck to one specific type of camera and microphone that suited our needs.

The camera we chose was a hard drive camcorder made by Sony. A hard drive camcorder is a special type of camcorder that stores recorded video on a built in hard drive. Each clip is stored as a data file, and thus is very easy to import to a computer for further use. This is a very attractive feature to have in a camcorder. Importing video to a computer can be an issue in some high end video cameras, as they often record to MiniDV tapes (small cassette-like units used for storing video) instead of a memory device.

Yet another attractive feature of the hard drive camcorder is its quality. It produces a sharp, clear image and has numerous settings from which to choose. Two of the most important settings are frame rate (in frames per second), and video definition. Frames per second (FPS) is the count of how many images a camera records each second. The higher the FPS, the smoother a clip will appear to be. Most filmmakers use either 30 or 60 FPS depending on the type of video they are making. Many of today's popular movies are shot at 30 FPS, while a live skit or documentary is often shot at 60. We've chose to shoot at 60 because in our opinion it looked more professional.

Video definition is dependent on the number of pixels in each image of film. A pixel is an illuminated dot on an electronic screen, and is extremely small. All digital images are comprised of numerous pixels. Two common levels of definition are standard definition (SD) and high definition (HD). The dimensions of a standard definition image are 640 x 480 pixels with an aspect ratio of 4:3 (this is the result of dividing 640 by 480). For high definition, the dimensions are either 1280 x 720 pixels or 1920 x 1080 pixels. In both cases, the aspect ratio is 16:9 (again, divide length by width).

Because the dimensions of high definition images are so much greater than those of standard definition, there is a significant visual difference. A high definition image is much sharper than a standard definition, and many more detail can be seen. However, high definition video requires a lot of storage space, and can be difficult to work with.

While high definition would be the obvious choice for a blockbuster film, we decided to shoot in standard definition for two reasons: it was easier to edit and it required less memory space. Some technical complications can arise when editing HD film, so shooting in SD can often save time and frustration. Also, because we expected these videos to be uploaded to a website, we thought it would be a good idea to take up as little memory space as possible.

With our camera selected, we proceeded to select the equipment we would use for sound recording. We found the best option to be a Bluetooth microphone. These microphones are tiny, and hard to see. This makes them easy to hide from view. Bluetooth microphones are also cordless, and can record directly to the camera. Because of this, they are very convenient. In our situation, they were the perfect choice since they allowed us to isolate our voices and eliminate ambient sound.

Once each video was recorded, the next step was editing, which was done entirely on computers. It was important for us to choose our editing software wisely, as each program has its own pros and cons. Some editing software is meant for those who are familiar with computers and video editing. It is harder to learn such programs, but they have many more features than simpler software.

While we originally had chosen to use Adobe After Effects, we later decided to use a program called CyberLink PowerDirector 11. While we had some previous experience with After Effects, we found that PowerDirector was more convenient for our purposes. It has a great text feature and many transitions and tools to use. For the simple nature of our videos, it was the perfect choice.

The final facet of our project to consider was testing our videos. The best option would have been to include our videos in a class and determine a way to measure their effectiveness. However, we didn't end up getting the opportunity to do so.

We instead came up with an alternative method in which we sent out an invitation to previous students in the class to participate in a short, 20-minute experiment. In this experiment, we began by administering each participant a six-question survey to help us determine their background in statistics as well as their opinion of it. To view the survey, see APPENDIX F-1. Survey Given to Experiment Participants**.** Next, we administered a four-problem pretest to test their knowledge of the content. To view the pretest, see APPENDIX F-2. Four Problem Pretest/Posttest**.**

Once the participants had completed the survey, we showed them two of our tutorial videos: “Controlled Experiments Part 1” and “Controlled Experiments Part 2”. Each of these videos covered concepts that appear in the four-problem test mentioned above. Once the participant was finished with the video, we then gave him or her a fresh copy of the four-problem test which they filled out once again. We then compared their answers before and after the videos to see if there were any improvements in number of correctly answered questions.

IV. RESULTS

A. Student Survey Results

Prior to writing scripts for our videos and filming them, we sent out a survey to past students of MA2611 to see which topics confused students the most in the course. As stated in the methodology, we believed this would give us a better idea of which topics to cover and how thoroughly we should explore them. The website we used to aid in the creation and distribution of this survey was called SurveyMonkey. Once we had created the survey, the first step in conducting it was getting permission from our university to use and distribute it. This took about a week, but once we got permission we sent out the survey.

A link to the survey was sent to 89 recipients by e-mail. Each of these recipients was selected because they were students of the most recent MA2611 class. We simply gathered their e-mails from the class mailing list. We waited for five days, and got a total of 12 responses. While this sample size was small, it was to be expected as no reward or incentive was offered to students that completed the survey.

Question 1 asked the respondents: “Which topics or ideas in MA2611 left you confused? Please list them.” Many of the written answers to this were vague. Common answers were “All of them", "pretty much everything towards the second half of the class..", etc. While these types of answers were not of much help, other students gave useful feedback. Three of the twelve answers mentioned probability calculations as a confusing topic. Only one student took the time to elaborate, mentioning p values, hypothesis testing, central limit theorem, and a number of others. These answers were not overwhelmingly helpful, but did show us that the students struggled with a variety of topics throughout the course. While we were unable to cover all the topics mentioned, the fact that so many concerns were brought up supports the idea that there was need for some extra resource for studying the concepts (at least for this particular class of MA2611 students).

Question 2 was a yes or no questions, stated like so: "Did you go to the teacher, a TA, or the tutoring center at any point during the course for help? If so, were your questions answered?" In response, 66.67% (eight people) said they asked questions, while 33.33% (four people) said they did not. Of the eight people that went, five said their questions were answered, one said it was not fully answered, one said it wasn't answered at all, and one did not comment.

While our sample size of twelve people isn't large enough to draw conclusions about the entire population of MA2611 students, these results were somewhat surprising. A majority of students went and asked the teacher questions, which supports the idea that students aren't afraid to seek out and ask for help. Of the eight that did so, two of the students didn't even get their question fully answered. If the videos had been available, these students may have been able to use them for additional help.

In Question 3, we asked, "Have you used video-sharing sites or tutorial sites, (i.e. YouTube, Khan Academy, etc.) to help you get a better understanding of a topic? If so, how did they help?" This was another yes or no question, to which 41.67% (five people) said they had used such sites, and 58.33% (seven people) said they had not. This was a surprise, as I thought these sites were more widely used. Three people commented, each stating that tutorial videos help by providing examples and secondary instruction to supplement a teacher's lecture.

Question 4 was yet another yes or no question: "Do you think tutorial videos would help if provided for a class as extra study material?" All twelve students replied yes to this question. This shows that even students who have not used such videos in the past still believe they could be helpful. It is interesting how these answers suggest there is great demand for such videos, yet schools rarely supply them for the students.

The final question, Question 5, was the following: "If videos existed that were tailored to a course you were in, would you use them?" To this, 91.67% (eleven people) answered that they would use them, while 8.33% (one person) said they would not. This was yet another overwhelming response in favor of the addition of tutorial videos as a course resource. Another intriguing observation is that the one person that answered saying they would not use such videos still believed that videos like this would help if provided as extra study material (possibly because this person believes they could help others if not themselves).

Overall, we were rather pleased with the results of our survey. Moving forward with our project, we now believed that our videos would prove useful to future students. The survey also would have given us a better idea of which topics to cover had we been able to cover more topics than we ended up covering.

B. TA/Teach Interview Results

Having consulted previous students of MA2611, we then decided to reach out to the professors and teaching assistants of the class. We believed the insight of those with years of experience teaching the class would be helpful to our project, and though only one TA agreed to an interview, we were not proven wrong. We came up with four questions to ask that we felt would help our project direction. To view the full MA2611 TA interview, see APPENDIX B. Survey for Past/Current Teaching Assistants of MA2611.

The first question we asked the TA was: “as a TA of the MA2611 course, which topics are particularly confusing to students?” Several topics were identified as especially difficult for students to understand - recognizing differences between observational, controlled, and sample studies as well as randomization use in these practices; the difference between stratification and blocking; basic probability concepts such as independence, mutual exclusivity, and binomial and normal distributions; the central limit theorem; use of the t distribution; and interpretation of confidence intervals and p values. These answers were far more precise than those gleaned from the student survey, and helped us to pinpoint which areas to focus on. Though we initially intended to briefly cover all of these topics, we only included tutorials on the different kinds of studies, strata, and blocking among the videos in our final project, choosing to sacrifice breadth for depth of subject matter.

For question two, we asked: “which concepts in statistics, if any, are particularly hard to explain?” The TA listed the central limit theorem, the difference between sample and population distributions, the p value, and the identification of when to use t or normal distribution as particularly difficult to explain to students. Though we initially intended to cover these concepts in our video series, and even wrote preliminary scripts for some, we came to the realization that this was an unrealistic goal. We simply did not have the time to cover these topics as thoroughly as we felt they needed to be. Rather than risk adding to students’ confusion with a rushed, incomplete explanation, we decided it would be best to narrow our focus and create a video series that tackled a few confusing issues in a thorough manner.

Question three asked the TA: “have you seen or heard of tutorial videos online that are designed to help students better understand class topics?” The TA answered that they had seen Khan Academy’s tutorial videos, and added that they felt the videos were helpful to students looking to get a better understanding of the subject matter. They stated that often, just seeing a problem explained in a different way helped students understand the material presented in the classroom.

The fourth and final question we asked was: “do you think that, if provided for a class of MA2611 students, a video tutorial series tailored to the curriculum would help students and improve their comprehension and grades?” The TA agreed that such a resource would be helpful, but cautioned that it should not be designed to replace traditional classroom learning. This supported and reinforced our intent to create a supplementary resource - something that would build on the knowledge students had gained in the classroom and that they could use as a reference while reviewing or doing homework.

C. Video Scripting Results

With a vision for our finished project and this information in mind, we now had an idea of how to approach the task of creating our tutorial videos. We began planning out each of the videos we would record, writing scripts and drawing storyboards. Each script was written in the style of a brief lecture - introducing and explaining the topic and then presenting it in the context of an example. We would illustrate the example on a whiteboard, using the storyboards to plan this part of the video.

It was during this time that we saw several flaws in our approach. During this stage, we still intended to cover all topics in the course that were difficult for students to understand, but we soon realized that we would not have enough time to record and edit videos for all the scripts we had planned. We briefly considered condensing our explanations, so that we would touch briefly on each of the planned topics, but these scripts barely scratched the surface of the subject matter, and we felt they would not improve student understanding. We felt it would be better to thoroughly cover a few topics, which we could then focus on explaining to the best of our ability.

The second flaw we found was in our manner of writing the scripts and presenting the concepts. From the beginning, we intended for our video tutorials to supplement classroom learning - to build on the knowledge students already had and to reinforce it by presenting it in a different manner. The scripts we were writing accomplished neither of those objectives. They approached the subject matter as if students had never before been introduced to the discussed topic, and presented the material in a manner too similar to the lecture style a professor would use in the classroom. A video series produced using this first run of scripts would have added nothing new to the course and would not have had any significant effect on student learning. Because of this, we decided to try a new approach.

To avoid making our tutorials redundant, we changed the manner in which we presented the topics. We switched to a completely example-centric approach, writing multiple questions for each video similar in content and style to what a student might find on a homework or quiz. Each question presented an example, and challenged the students to think the problem through and deduce the correct answer from among the multiple-choice options. After inviting the viewer to pause and consider which answer might be correct, we would go through each answer in turn and thoroughly explain why it was or was not correct. In this manner, our tutorials encourage the viewer to approach the video not in a passive manner, but from the perspective of solving a problem. Through this new style of script writing, we were able to meet our objective of presenting the material in an engaging way that would give students a fresh perspective.

D. Video Recording Results

Once a script was written and we felt well enough prepared, we would shoot the video. A majority of the filming was done on the weekends, as that offered us plenty of time during the week to write and finalize scripts for the videos we wanted to record. Aside from preparing the scripts, we made sure to reserve recording space and equipment for the times we were open to record. Although at times it was a bit stressful to have time constraints on how long we could use equipment or recording space, it forced us to buckle down and focus on getting the job done.

We recorded our videos in rooms called “tech suites” located on the WPI campus. These rooms are virtually soundproof and have white walls in front of which we recorded. The lighting in these rooms was also a plus, as the lights were not too sharp but still bright enough to adequately light the scene. As mentioned in the methodology section, we decided that the best lighting technique for our purposes was three-point lighting (requires a key light, a fill light, and a backlight). For the key light, we used the lighting in the tech suites. To add a fill light, we brought a desk lamp to each recording session and put it just out of the shot to the left. We then shone it towards the speaker.

As for the back light, we chose not to bring a separate light for this for a few reasons. The first reason was that lighting in the tech suites was bright enough that it seemed almost unnecessary. The second was it was easy to add a backlight in the editing process. The third and final reason was that we did not have a bright or large enough light to use for a backlight. We ended up enhancing the lighting during the editing process to make it appear as though there was a back light.

After each scene was set up, we looked at the script of the video we were filming in order to memorize our lines. We would study a chunk of the script and then record it. This would leave us with segments of content that were later pieced together during the editing process. Recording our videos in this style meant that we did not have to come to each session with all the lines of the script completely memorized. We found that recording the videos in chunks better allowed us to explain each topic as it came up without being rigid or robotic in our speech.

During the entire project, we did not make any modifications to the techniques we used to record. What we did change drastically, as previously mentioned, was our script writing and how we explained and discussed each topic. With this change in style came a shift in how we chose to present the content of each video. At first, we recorded our videos in front of whiteboards so that we could use the board to draw and illustrate what we meant. This made the editing process easy, as there was very little to do in the way of adding text into each video. As there were only two of us doing the recording, we decided to alternate each video, with one of us the presenter of one entire video, and the other the subject of the next. We chose to do this because it made sense to only switch speakers for each topic that came up, and since each video covered only one topic, this strategy seemed most logical.

However, once we decided to modify the way in which we covered each topic, we began alternating as speakers multiple times in a video. We did this because each video now had two to five questions that we were explaining. It seemed natural to alternate speakers for each new question that we discussed. Furthermore, we stopped recording in front of and using a white board. As the length of each question varied, we believed that it would be easier to add text to each scene during the editing phase. This made the questions more legible, and gave us the ability to highlight important words, sentences, or sections.

This new style made the recording sessions much easier and smoother, as one person could study their lines while the other recorded. It also meant we were not frequently turning to a whiteboard to write or draw. The drawback to the new style was the difficulty it added to the editing process. While a great deal of time was dedicated to editing, it turned out being worth it as the quality of each video was much improved.

E. Editing Results

With the completion of each recording session came a large collection of content that was in need of being pieced together and edited. While the recording process was time consuming, the editing process proved to be even more demanding. The editing software we used was called PowerDirector 11, a product made by CyberLink. This software was a powerful tool, and allowed us to easily manipulate each scene, add text, and create high quality videos.

Before editing the first video, we created a quick five second introduction to open each video that we created. The introduction was created by taking the WPI logo and using different layers to make it appear in a subtle yet interesting way. We saved this clip and then used it in each video we created. We then moved to editing the actual content.

Prior the change in the style of our videos, each one took around 2-3 hours to edit. The clips were recorded linearly according to the script, so the editing process was mainly adding each clip to the video in order, cutting off excess and adding background music. While this seems simple in theory, it still required large amounts of time and effort to get everything right.

The time it took to edit each video increased by large amounts after the change in style. Because of the need to add text to each scene, each video began to take around 5-6 hours to edit. Each clip required large amounts of text that had to be typed in and positioned correctly on the screen so as not to cover up the speaker, yet still be legible. Furthermore, the text was often highlighted to show which section was being read, or which answer was correct or incorrect. These complexities made it necessary to be thorough in checking whether or not each text section was written correctly and was displayed at the correct times in the correct positions.

Aside from adding text to each video, there arose some other issues that needed to be dealt with during the editing process. For instance, the fact that we were unable to add a backlight during the filming process made it crucial that we enhance the lighting during editing. Due to the very user friendly interface of PowerDirector 11, enhancing the lighting was often just a matter of clicking a button and letting the software do the work.

Another issue that we encountered only once was ambient noise in a scene. For all but one video, ambient noise was a non-issue for us as we consistently filmed in soundproof rooms. However, there was one recording session in which we used a room that had a rather loud heater running in the background. The noise was somewhat overbearing, and because of this we went about removing the noise in the editing phase. To accomplish this, we used a program called Audacity. This program is a free sound wave editor that has numerous tools for enhancing audio clips. We used a tool called “Noise Removal” which detects constant background noise and removes it without affecting any other sound in the clip. This tool accomplished exactly what we wanted, and we were able to remove the ambient heater noise with minimal effort.

Although the use of editing to enhance our videos was time consuming, it was very successful. We certainly believe that our videos benefitted greatly from the effort we put in to the editing process.

F. Experiment Results

We had very low turnout on the day of the experiment - only two students arrived to take the test. As such, we cannot safely draw concrete conclusions from the results of our testing; however we can use the results to make inferences about what the results may have been had we gotten a larger sample.

The responses to the survey were very homogenized. Of the five questions comprising the survey, the two participants answered differently on only one. The first question asked: “have you ever taken any statistics courses besides MA2611?” Neither reported any classroom experience with statistics beyond the introductory course. The second and third questions asked them to rate their enjoyment of the MA2611 course and of statistics in general on a six-point scale. The students answered that they “slightly enjoyed” both. The fourth question, and the only one on which they differed, asked: “have you made use of any video resources when studying any subject in the past?” One reported that he had, while the other had never done so. The fifth and final question asks, “Would you have used a statistics video tutorial series in studying for MA 2611 had one been made available to you?” Both agreed that they would have made use of such a resource.

These answers, along with the responses from the initial student survey, could indicate that the majority of students feel positively about video tutorials. Even those who have not sought out such resources in the past generally feel open to trying them. We speculate that this is because there are currently no video tutorials available that are tailored specifically to the MA2611 course curriculum at WPI. Students who might otherwise benefit from a video tutorial series might not care to spend the time to seek out a video, especially when the content is not guaranteed to be understandable or consistent with the course material. Their positive opinion of tutorial videos suggests that they would take advantage of such videos if they were provided with a readily available collection of them – one that had consistent content created around a specific course curriculum.

Our participants' answers to the next parts of the experiment - the pre-test and the test - were also very close. On the pre-test, the first participant answered every question correctly, while the second answered only one question incorrectly. Both got perfect scores when taking the test in conjunction with watching the video. Though this might suggest that our videos were helpful in aiding their understanding of the statistical concepts, the improvement, as well as the sample size, is far too small to allow any sort of conclusion.

The final point of reference our experiment gave us was the ability to observe our participants as they watched the videos. Because we gave them complete control over the video, with the freedom to pause, rewind, skip ahead or stop whenever they wanted, we believe our observations in this experiment still give us insight into how the tutorial videos would be used by students, were they to be implemented into the course.

Our first participant watched the videos linearly, filling out the test along with them. Though he did not interact with the video while taking the test, he stopped it halfway through when he had finished the last problem on his test. After stopping the video, he commented that he enjoyed the transition graphics and background music we employed in our videos, stating that they held his attention. The second participant interacted more with the video, skipping through parts once he had answered the question covered in that section. He also worked out and stated the answers to some of the multiple choice questions from the tutorial. Once he had answered all the questions, he too stopped the video when he had finished the test.

It was encouraging seeing that the students were able to easily fill out our questions while watching the videos and gaining insight from them. They used the videos just as we hoped they would be: as a guide in their homework and as a secondary resource for studying.

V. CONCLUSIONS

During our work on this project, we discovered and learned many things that helped us accomplish what we did. Throughout the recording process, we became familiar with writing scripts, recording quality video, and creating polished tutorials with editing tools. We even learned a lot about statistics, and reinforced our own understanding of the topics we covered.

Arguably, the biggest challenge in creating tutorial videos is writing the scripts. It takes creativity coupled with a familiarity with the concept being taught to write a concise and entertaining script. We had to write about four scripts before we truly felt like we knew how to approach the writing process. Coming up with relevant and helpful examples was often tough, but it was very rewarding once we were finished and we felt like we had created something helpful for the students.

A great strategy for creating example problems is to brainstorm ideas as a group. Something to remember when brainstorming example problems is that no idea is stupid or ridiculous. It was so much easier to think of examples once we got used to presenting every idea that came into our heads.

During recording, the first thing we had to learn was how to be comfortable in front of the camera. At first, just being in the shot and speaking into the device was for somereason nerve racking. However, after the first or second recording session, we began to lighten up. While recording oneself may seem a little strange and embarrassing at first, it helps to ignore the camera and to think of what it would be like trying to explain the concept to a friend that was standing in front of you. Another thing to remember is that if a mistake is made it's very easy to start the recording over from scratch. Over time, it becomes easier to speak as if there were no camera at all.

Another crucial step in the recording process is preparation. While this includes scripting, it also includes telling every group member what needs to get done at each session and what each member should bring. It was always an annoyance showing up to the shoot and then remembering that we forgot to bring something.

As for the editing process, everything will be made much easier if the clips are shot according to the script. This way, the editor knows which clip comes next in the video and consequently does not spend any time searching for the clip he needs. Furthermore, if the recording is shot with good lighting and no ambient noise, then the editor doesn't have to put effort into fixing the lighting or the audio. This all goes back to the preparation stage of the recording process. The editing process can be made much more painless if everything is thought out ahead of time.

During our testing phase, we learned the hard way that it is important to offer experiment participants compensation for their time. In our case, we offered each participant as many cookies as they wanted. While originally we thought this would be enough of an incentive, we ended up getting only two participants. We believe that had we offered money or gift cards, we may have gotten more interest. Aside from a better incentive, the only option would have been to advertise the experiment even more than we did.

As for the results of the experiment, we saw them as positive. Although we had a sample size of only two, the two that did participate used the videos with ease and found them useful. They were easily able to follow along, and felt comfortable skipping around the video to find what they needed. Future studies should focus on making tutorial videos to supplement – not replace – an instructor's presentations. Furthermore, they should test their videos on a larger sample size and in a class setting if at all possible.

While we learned many things about recording techniques, scripting techniques, editing techniques, and what makes a tutorial video effective, we also had many things that we would change given the opportunity to go back and do it all again. We created an archive of videos that covers a subset of statistical concepts in a detailed and thorough manner. However, there are many parts of statistics that students have trouble with outside of the sections we covered. Ideally, we would have been able to cover them all, but because we overestimated how many videos we could record in the time we had, we were unable to cover many of the topics.

We believe the reason we were unable to complete more videos is a combination of three factors, the first being time. We had roughly an eight week period to record our videos, and at the beginning of the recording stage, this seemed like a huge amount of time. Because of this, we started with the more introductory topics of statistics, figuring we would have time to cover the technical side of statistics in later videos. This was a severe miscalculation on our part, as we didn't realize just how much effort we would have to put into making the videos. The scripting, recording, and editing process was rigorous, and with other courses and extracurricular activities on top of the project, we found it impossible to create the number of videos we wanted.

The second factor was the amount of time it took us to get familiar with the creation process. At first, we wrote scripts and recorded videos that were more like mini-lectures rather than supplemental tutorial videos as mentioned in the methodology. After two weeks, we decided this format was not going to help the student in the way we were hoping, so we completely revamped the way we approached the topics we were covering. Without using the first two weeks to write scripts we didn't use, we would have been able to create videos covering more topics than we did.

The third factor was the fact that only two of us were working on the project. With more hands, the work would have been split more evenly, and more would have gotten accomplished. Of course, with the addition of more people to a team comes the need to learn to work together. That being said, if a larger group was to work on this project and stay organized, focused, and efficient, they could create a much larger number of videos.

The fourth and final factor was that we began with videos covering the introductory topics of statistics. While we by no means consider these videos a waste, it may have been better to start with the more advanced topics that students struggle with the most. Of course, hindsight is 20/20 and all we can do now is speculate what would have happened.

Following from the issues we had, we hope that others can learn from our project and avoid making some of the mistakes we did. The biggest thing to realize is that videos don't just appear. No matter how much experience one has with creating videos, it always requires creativity and a lot of work to make a quality tutorial.

Allot enough time to complete the number of videos desired, and do not underestimate the difficulty or complexity involved in scripting, filming, and editing the content. Planning out each video is extremely important. If it is done well, then every group member knows where to be, what equipment to bring, and what the video is about. There were a number of times in which we felt very prepared to shoot a video and showed up to the shoot only to realize we had forgotten something. While this is inevitable, a majority of such mistakes can be prevented through thorough preparation and planning.

Working on this project was an amazing experience for both of us, and we hope that our discoveries will aid others in future endeavors of this nature. The use of video as an educational tool is still a relatively new idea. While future research must still be done on the subject, we believe that online tutorial videos can be extremely helpful to struggling students everywhere.

REFERENCES

AP Stats Guy. (n.d.). *AP Stats Guy*. Retrieved September 5, 2013, from http://www.apstatsguy.com

Arora, A., Leseane, R., & Raisinghani, M. (2011). Learning and Teaching Styles for Teaching Effectiveness: An Empirical Analysis. *International Journal of Web-Based Learning and Teaching Technologies*, *6*(1), 1-13.

Biyani, G. (2011, December 2). How To Make A Great Tutorial Video.*Udemy/blog*. Retrieved September 5, 2013, from https://www.udemy.com/blog/how-to-make-a-great-tutorial-video/

Boer, J. d., Kommers, P. A., & Brock, B. d. (2011). Using Learning Styles And Viewing Styles In Streaming Video. *Computers & Education*, *56*(3), 727-735.

Charkins, R. J., O'Toole, D., & Wetzel, M. (1985). Linking Teacher and Student Learning Styles with Student Achievement and Attitudes. The Journal of Economic Education, 16(2), 111-120.

College of South Nevada. (n.d.). Types of Teaching Styles. CSN - Home. Retrieved September 24, 2013, from http://www.csn.edu/pages/2358.asp

Crotty, J. (2012, March 26). 7 Signs That U.S. Education Decline Is Jeopardizing Its National Security.Information for the World's Business Leaders. Retrieved September 16, 2013, from http://www.forbes.com/sites/jamesmarshallcrotty/2012/03/26/7-signs-that-americas-educational-decline-is-jeopardizing-its-national-security/

Deslauriers, Louis, Ellen Schelew, and Carl Wieman. "Improved Learning in a Large-Enrollment Physics Class." Science332.6031 (2011): 862-864.

Felder, R., & Silverman, L. (1988). Learning and Teaching Styles in Engineering Education. Engineering Education, 78(7), 674-681.

Filoteo, B. (n.d.). 6 Things You Should Know About Video Editing. Bianca Filoteo. Retrieved September 26, 2013, from http://biancafiloteo.com/2012/6-things-you-should-know-about-video-editing/

Fleming, G. (n.d.). Learning Styles - Visual, Auditory, and Tactile. Homework Tips - Help With Homework. Retrieved September 24, 2013, from http://homeworktips.about.com/od/homeworkhelp/a/learningstyle.htm

Grabowicz, P. (2012, August 13). Video Techniques. *kdmc Berkeley*. Retrieved September 5, 2013, from http://multimedia.journalism.berkeley.edu/tutorials/shooting_tips/

H. Brecht , David , and Suzanne M. Ogilby. "Enabling a comprehensive teaching strategy: video lectures." *Journal of Information Technology Education* 7 (2008): 71-86.

IUPUI. (n.d.). 3 Learning Styles. Bepko Learning Center | University College, IUPUI | Indianapolis, IN. Retrieved September 24, 2013, from http://blc.uc.iupui.edu/AcademicEnrichment

James, B. (n.d.). Understanding and Identifying Auditory, Visual, and Kinesthetic Learning Styles. Sunburst. Retrieved September 24, 2013, from sunburst.usd.edu/~bwjames/tut/learning-style/styleunder.html

Khan, S. (n.d.). Statistics Course, Khan Academy Mathematics Video Tutorials, Salman Khan . *Free Video Lectures, Online Courses and tutorials from top colleges, universities*. Retrieved September 5, 2013, from http://freevideolectures.com/Course/2564/Statistics#

Lloyd, Steven A., and Chuck L. Robertson. "Screencast Tutorials Enhance Student Learning of Statistics." Teaching of Psychology 39.1 (2012): 67-71. Print.

Maki, W., & Maki, R. (2000). Evaluation of a Web-based introductory psychology course: II. Contingency management to increase use of on-line study aids. Behavior Research Methods, Instruments, & Computers, 32(2), 240-245.

Rebholz, G. (n.d.). Seven Steps to Getting the Most Out of Your Personal Video Camera. *Sony Creative Software - Vegas video - ACID & Sound Forge audio editing*. Retrieved September 5, 2013, from http://www.sonycreativesoftware.com/getting_the_most_out_of_your_personal_video_camera

Roell, K. (n.d.). Three Different Learning Styles: Auditory, Visual and Kinesthetic. Test Prep Tips, Strategies, Practice and Advice. Retrieved September 24, 2013, from http://testprep.about.com/od/tipsfortesting/a/Different_Learning_Styles.htm

Saga, H. (2009). Presentational style of video as an affecting factor in learning.*International Journal of Instructional Media*, *36*(2), 177.

Stein, J., Steeves, L., & Smith-Mitsuhashi, C. (2001, April 13). Teaching Styles Categories. Shaw Webspace. Retrieved September 24, 2013, from http://members.shaw.ca/mdde615/tchstyc

S. Mestre, Lori. "Matching Up Learning Styles with Learning Objects: What's Effective?." *journal of library administration* 50.7-8 (2010): 808-829. *taylor and francis online*. Web. 5 Sept. 2013

S. Moore, David . "The Place of Video in New Styles of Teaching and Learning Statistics." *The American Statistician* 47.3 (1993): 172-176. Print.

Tulbure, C. (2011). Do different learning styles require differentiated teaching strategies?. *Procedia - Social and Behavioral Sciences*, *11*, 155-159.

University of South Carolina. (n.d.). Teaching Guide - Things to Consider About Teaching Styles. University of South Carolina. Retrieved September 24, 2013, from http://www.sc.edu/cte/guide/teachingstyle

Video Recording: Improving Sound Quality – Recording Techniques. (n.d.).*Longzijun | Works, Words & Projects*. Retrieved September 5, 2013, from http://longzijun.wordpress.com/2010/09/26/video-recording-improving-sound-quality-recording-techniques/

Wyant, N. (2013). The Effectiveness of Online Video Tutorials as Supplemental Library Instruction. Kansas Library Association College and University Libraries Section Proceedings, 3, 39-43.

Zhang, Dongsong, Lina Zhou, Robert O. Briggs, and Jay F. Nunamaker. "Instructional Video In E-learning: Assessing The Impact Of Interactive Video On Learning Effectiveness."Information & Management 43.1 (2006): 15-27.

Zapalska, A., & Brozik, D. (2006). Learning Styles And Online Education .*Campus-Wide Information Systems*,*23*(5), 325-335.

20 Lighting Tutorials for Film and Video. (2011, April 25). *Filmmaker IQ*. Retrieved September 5, 2013, from http://filmmakeriq.com/2011/04/20-lighting-tutorials-for-film-and-video/

APPENDIX A. Survey Administered to Previous Students of MA2611

Survey for the students of MA2611

*1. Which topics or ideas in MA2611 left you confused? Please list them.

*2. Did you go to the teacher, a TA, or the tutoring center at any point during the course for help? If so, were your questions answered?

- ○ I went and asked questions
- ○ I did not go and ask questions

If yes, were your questions answered?

*3. Have you used video-sharing sites or tutorial sites, (i.e. YouTube, Khan Academy, etc.) to help you get a better understanding of a topic? If so, how did they help?

- ○ Yes, I've used video sharing sites for school
- ○ No, I have not

If yes, in what ways were they helpful or not helpful?

*4. Do you think tutorial videos would help if provided for a class as extra study material?

- ○ Yes
- ○ No

5. If videos existed that were tailored to a course you were in, would you use them?

- ○ Yes
- ○ No

Done

APPENDIX B. Survey for Past/Current Teaching Assistants of MA2611

Survey for Past or current MA2611 TAs

1. As a TA of the MA2611 course, which topics are particularly confusing to the students?
2. Which concepts in Statistics, if any, are particularly hard to explain?
3. Have you seen or heard of tutorial videos on YouTube that are designed to help students better understand different class topics?
4. Do you think that if provided for a class of MA2611 students, an archive of these videos tailored to the curriculum would help students and improve comprehension and grades?

Survey for Past or Current MA2611 Professors

1. As a professor that has taught MA2611 in the past, what topics were particularly confusing to your students?
2. Which topics in Statistics, if any, were particularly hard to explain?
3. Have you seen or heard of tutorial videos on YouTube that are designed to help students better understand different class topics?
4. Would you consider using such videos in your class?
5. Do you think that if provided for a class of MA2611 students, an archive of these videos tailored to the curriculum would help students and improve comprehension and grades?

APPENDIX C. Controlled Experiments Part 1 Script, First Draft

Controlled Experiments Part 1: What is a Controlled Experiment?

Introduction

The three main types of design studies in statistics are observational studies, controlled experiments, and sample surveys. In the this video, we will cover the second of these types: the controlled experiment. By the end of this video, we will have defined what a controlled experiment is and given examples of how they are used.

Main Discourse

First things first. A controlled experiment is a study in which treatments are assigned to experimental units and then the responses of each unit are observed and recorded. To make this clearer, let's look at an example of a controlled experiment.

Let's say we want to determine whether or not people in America think bacon with brown sugar tastes better than bacon without brown sugar. To do so, we would set up a controlled experiment. In this case, the population from which we are collecting data is all the people in America. Because the population is so large, we cannot hope to conduct an experiment on every person in the population. Because of this, we collect a subset of the population which we call a sample.

Now that we have the sample, we have some design choices to make. Basically, we want to start by feeding each subject, or experimental unit, some bacon with brown sugar and some bacon without brown sugar. Then, we want to measure their opinion of the taste of each. Unfortunately, if we want accurate responses from each subject, we need to carefully design our experiment.

First, we need to design the feeding phase. We need to realize that if one piece of bacon tastes better than another, then a person's opinion might not be based only on the addition of brown sugar. This issue is called confounding. In this example, brown sugar is the treatment because we either add brown sugar to the bacon or leave the bacon alone. So how do we ensure that the only difference between each piece of bacon is our treatment? One simple solution would be to put brown sugar on one half of each piece of bacon. We could then have each person try each half of the same piece of bacon, thus eliminating a difference in the taste of the bacon.

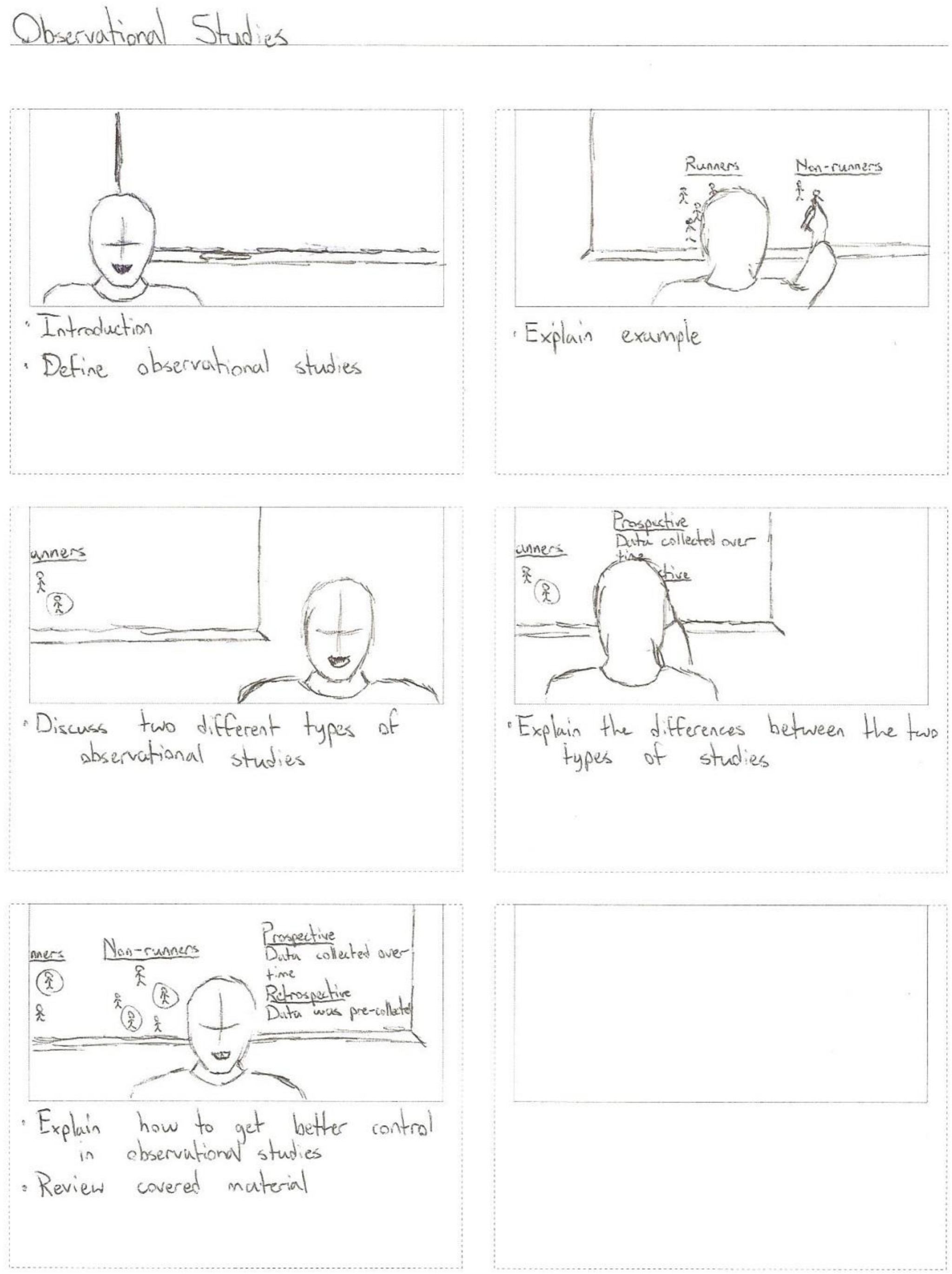
Second, we need to decide how to measure the opinion of each subject. Finding a scale to measure human opinion is not an easy task. While there are many ways to do so, it isn't easy to know which method is the best. To get a thorough response, we might first ask them to rate each taste on a scale from 1 to 10 (1 being worst and 10 being best) and then ask them which taste they prefer and why. By using a number scale as well as asking for an explanation, we can get a both a numerical value and a description of each subjects opinion. The answers given by each individual is known as their response.

Once we have our design, we can conduct the actual experiment. So we've come up with a controlled experiment, executed it, and gotten results. What do we do with the results? The point of a controlled experiment is to try and find some cause-and-effect relationship between a treatment and a response. It is possible for us to do this because we have CONTROL over which experimental units get which treatments and when. In this example, we would try to find a cause-and-effect relationship between bacon and brown sugar. That is, if bacon with brown sugar was more widely preferred, we would say brown sugar causes bacon to taste better. If people preferred bacon without brown sugar, we would say brown sugar causes bacon to taste worse.

That said, it is important to understand that saying something causes another thing to happen is a big statement. In order to make this statement accurately, one must be sure that his or her experiment is designed and conducted with care and detail. For more details on how to properly prepare and conduct an experiment, go to the Part 2 video of controlled experiments.

In review, a controlled experiment is conducted in order to find cause-and-effect relationships between treatments and responses of experimental units. The challenge in finding this is that a controlled experiment must be designed properly to do so. To get the details of different design strategies, go to Part 2 of controlled experiments.

APPENDIX D. Storyboard for Original Observational Studies Tutorial



APPENDIX E. Scripts

APPENDIX E-1. Sampling Strategies Script

Sampling error

We are collecting a sample from all the students in WPI. As we know, the ratio of men to women at WPI is approximately 33% women, 67% men. To collect our sample, we randomly select students at WPI in such a way that each student has an equal chance of being part of the sample. After collecting our sample, we find that it consists of roughly 45% women and 55% men. What has just occurred?

i. Selection bias
ii. Experimental error
iii. Sampling error
iv. Confounding

In this scenario, we collect a sample and end up with one that is not representative of the target population. First of all, this example has to do with sampling strategies, so the answer cannot be either experimental error or confounding, both of which come into play when conducting an experiment or observational study. To decide whether selection bias or sampling error is responsible for our unrepresentative sample, we need to look at the definitions of each.

Our sample collection strategy gives each member of the student body an equal chance of being selected for the sample. We could safely expect that our sample would have percentages of men and women consistent with those of the population, and if we tried collecting a new sample using the same method, we would likely obtain a representative sample. Our selection method was not biased. It was entirely by chance that we collected a sample that was not representative. Therefore, sampling error has occurred. As a side note, do not be deceived by the word "error" in sampling error. As we have said, this was not truly an error as it happened by chance. The term "sampling error" is somewhat misleading in this regard, as there is not actually any error to be found with the way the sampling was conducted.

Cluster sampling

We are trying to survey department managers of Price Choppers supermarkets in Massachusetts to find job satisfaction. We randomly select ten Price Choppers and interview every department manager of each one. What sampling strategy have we used?

i. Randomized complete block
ii. Cluster sampling
iii. Simple random sampling
iv. Case-referent study

Completely randomized blocks
This is incorrect and actually has nothing to do with sampling. Completely randomized blocks is a strategy used in controlled experiments to assign treatments to units.

Cluster sampling

This is the correct answer. In one-stage cluster sampling, we sample the clusters (Price Shopper stores) and then get data from each observational unit (the department heads) in each clust.

Simple random sampling
Simple random sampling is definitely a sampling strategy, but is not the correct answer here. In simple random sampling, we are selecting sampling units at random from the total population. In this example, we could have used simple random sampling by randomly selecting a small number of department managers from all the Price Chopper department managers, regardless of the store they worked at. While both strategies work, cluster sampling allows us to interview people that are geographically near to each other. This could be easier than choosing a bunch of department managers at random only to find they are all over the New England area, forcing us to drive to each one.

Case-referent study
This is another one of those that should be a no right off the bat. Case-referent studies are a type of observational study, and have no relation to sampling strategies whatsoever.

Stratified random sampling

We want to find the public opinion of tax levels in the state of Massachusetts. We begin by splitting everyone in MA into groups based on their income: one group of people that make more than $100,000 a year, one group between $100,000 and $50,000 a year, and one group that makes less that $50,000 a year. We then gather samples of people from each group with numbers proportional to the total population. In this example, what are the three groups called?

i. Blocks
ii. Clusters
iii. Strata
iv. None of the above

Blocks
While the word “blocks” seems like a logical answer for this question, in reality it is a completely silly one. Blocks are not related to sampling at all, as there are used in controlled experiments to help assign treatments to experimental units.

Clusters
Clusters may also seem like a viable answer for this question, but is actually incorrect. Clusters are groups of people or things that are geographically close to one another; hence the word “cluster”. For example, people that make more than $100,000 a year do not necessarily live close to each other. In fact, we have no idea where they are in relation to each other. With clusters, we don’t group units by their traits, but by their location.

Strata
This is the correct answer. We use strata to split population units into groups based on their traits or characteristics. We then take samples inside each strata.

Selection bias

We are collecting a sample from all the students in WPI. As we know, the ratio of men to women at WPI is approximately 33% women, 67% men. However, we neglect this fact and gather our sample from the men in fraternities on campus. What mistake have we made?

i. **Selection bias**
ii. Sampling error
iii. Nonresponse bias
iv. None of the above

Selection bias
Put simply, selection bias is when a portion of the population is omitted or severely underrepresented. In this example we are not only omitting those students not involved in Greek Life, but also women students! Because of this, the WPI population as a whole is not accurately represented by our sample.

Sampling error
As explained in a previous example, sampling error is not a good term as it implies some error has been made. Actually, sampling error happens by chance, and can occur even when a sampling method is valid. However, our sampling method was flawed from the start. In this case, it was not by chance that part of the population was underrepresented, but by our own doing. Although there will inherently be sampling error when taking a sample, it is not the main issue with our sample.

Nonresponse bias
Nonresponse bias actually has nothing to do with sampling. It is an issue that arises during surveys when people chosen to be surveyed refuse to respond. Therefore, it isn't even relevant to this example.

<u>Simple random sampling</u>

We are trying to gather a sample from a course of ninety students. We begin by getting a class list, and then select a sample at random (each possible sample has the same chance of being selected). What strategy have we used?

i. Stratified random sampling
ii. Completely randomized design
iii. **Simple random sampling**
iv. None of the above

Stratified random sampling
In stratified random sampling, we first split our population into groups based on traits and then take samples from each group. In this example, we're gathering our sample from our total population instead of stratifying first, so this is incorrect.

Completely randomized
While completely randomized sounds like a logical choice, it's really one of those answers here to trick you. It has nothing to do with sampling, but actually is used in controlled experiments to assign treatments to experimental units.

Simple random sampling
This is the answer we're looking for. We are randomly sampling from our population, giving each student the same chance of being chosen.

<u>Multistage cluster sampling</u>

We are trying to find job satisfaction of Price Chopper employees. We first randomly select five Price Choppers. We then randomly select two departments inside each of these Price Choppers. We then conduct our study on the employees in the two randomly selected departments in each Price Chopper. What strategy have we used?

i. **Multistage cluster sampling**
ii. Stratified random sampling
iii. Cluster sampling
iv. None of the above

Multistage cluster sampling
This is the correct answer, as we are sampling using clusters and multiple stages. In our first stage, our clusters were Price Chopper stores. In our second stage, our clusters were the departments in the already chosen stores. By doing this add even more convenience by not having entire stores of employees to interview. We keep our sample units geographically close to other sample units.

Stratified random sampling
By now, it should be clear that this is not the answer. We are not stratifying, but instead using clusters, and once more we are doing multiple stages of clusters.

Cluster sampling
While in each of our stages we are using cluster sampling, it is not enough to describe this solely as cluster sampling. We need to capture the fact that we do our sampling in stages. This is why this answer is incorrect.

APPENDIX E-2. Observational Studies Script

Main example

The school board at a local high school would like to see if playing an instrument causes students to get better grades in their classes.

In the following example problems, we are going to discuss different strategies the school could use to accomplish this and ask questions about each one.

Case-referent

The board begins by randomly selecting a sample of twenty students that have attained exceptionally high grades during high school (nothing below a B+). The board then randomly selects another sample of twenty students that have received grades no higher than a B. Next, the board asks every student in each sample whether or not they play an instrument. The numbers of people that play instruments in each group are then compared to see if there is an association between academic performance and playing an instrument. What are we conducting here?

i. Sample survey
ii. Cohort study
iii. **Case-referent study**
iv. None of the above

Sample survey
In a sample survey, the goal is to learn about characteristics of a target population by polling or surveying a sample from that population. However, this study clearly tries to do more than find characteristics of a target population. The board is actually attempting to find an association between playing a musical instrument and performing better in school. They do this observing characteristics of high schoolers. This is the definition of an observational study. The question now is what kind of observational study are we conducting and is that study an option in the answers?

Cohort study/Case-referent study
Although this is a type of observational study, it is not a cohort study. A cohort study is an observational study in which comparison groups are formed based on the presumed cause. In this study, the presumed cause is playing an instrument: the board presumes that playing an instrument causes better academic performance. In this example, the board first gathered two samples of students: one of students that perform well in school and one of students that perform poorly. They then observed whether or not each student plays an instrument. Because the board formed the groups based on the effect - GPA - this is a case-referent study.

Cohort

Say the school board takes a different approach. They begin by collecting a random sample of twenty students that play instruments *they find out which students play instruments by conducting a survey*. They then collect a random sample of students that do not play an instrument at all. Finally, the board compares the overall grade averages of each sample and tries to draw conclusions based on which sample's average GPA is higher. What kind of study is this?

v. Controlled Experiment
vi. Case-referent study
vii. **Cohort study**
viii. None of the above

Controlled Experiment
This study is not a controlled experiment. We know this because the board never assigns treatments to the students in either sample. In other words, the board doesn't choose which students play instruments and which don't. They merely observe and record whether or not each student plays an instrument.

Case-referent study/Cohort study
As we discovered in the last example problem, in case-referent studies comparison groups are formed based on an effect, and in cohort studies they are formed based on a presumed cause. Here, the presumed cause and effect are the same as in the first problem. The presumed cause is playing an instrument and the effect is getting better grades. Because in this study the board gathered a sample of students that play instruments and a sample of students that don't, they formed comparison groups based on the presumed cause. Therefore, this is a cohort study.

Confounding

Let's say that the school board conducts a cohort study in the way we described in the last example and gets the following results:

Average GPA of the group of students that play instruments
3.65
Average GPA of the group of students that do not play an instrument
3.21

Can the school conclude that playing an instrument causes a student to perform better in school?
Yes
No

The answer
Without knowing anything about observational studies, it might sound logical to answer yes. However, when conducting an observational study, we can NEVER conclude that there exists a cause and effect relationship. The best one can do is show an association between two factors.| The reason for this is confounding. Confounding occurs when the effects of two or more factors cannot be separated from each other. For instance, in this example how does the board know that playing an instrument is the cause of the high grades of the students? Perhaps a majority of the students that play instruments are younger and are taking easier courses, making it less of a challenge to do well. Maybe the students that don't play an instrument are just inherently lazy, and are less likely to do homework? The point is that the board cannot know for sure the cause of the good grades of the students who play instruments. In conclusion, understand that observational studies can only show association, not cause and effect. Only controlled experiments are capable of showing cause and effect.

APPENDIX E-3. Controlled Experiments Part 1 Script

Treatments, Responses, Nuisance variables

Oscar Mayer wants to know if bacon with brown sugar on it would sell better than plain bacon. To find out, they conduct the following experiment: they gather a sample and then give everyone in it two pieces of bacon. The first piece has no brown sugar on it while the second piece does. Each person is then instructed to write down which piece they liked better. Identify the following.

1. Treatment **(C)**
2. Response **(A)**
3. Nuisance variable **(B)**

A. Each person's preference
B. The difference in the amount of fat in each piece of bacon
C. The presence or absence of brown sugar

Treatment
In this experiment, there is a single experimental factor: brown sugar. The brown sugar has two levels, level one being the presence of brown sugar and level two being the absence. Because the experimenters either apply or do not apply brown sugar to the bacon before feeding it to the people, they are using combinations of the levels brown sugar. Therefore, the presence or absence of brown sugar is known as the treatment.

Response
The response is the observed or measured reaction of an experimental unit. In this example, the experimenters are finding each person's preference between plain bacon and bacon with brown sugar. Therefore, the response is the preference of each person.

Nuisance variable
In an experiment, a nuisance variable is a factor over which the experimenter has no control. It also has the potential to cause variation in the responses of experimental units. In this example, the difference in the amount of fat in each piece of bacon could be considered a nuisance variable because 1) the experimenters have no control over it 2) many people don't want too much fat in their bacon, and could alter their preference based on the fat content.

Confounding

Let's say we want to find out if Red Bull causes twenty-year-olds to run faster. We gather two samples: a sample of female twenty-year-olds, and a sample of male twenty-year-olds. We instruct the females to run a mile, record each of their times, and find the average time of the females. We then give each male in the other sample a 20oz can of Red Bull. After half an hour, we tell each of them to run a mile, record their times, and find the average time of the males. Finally, we compare the two averages to see which sample had a faster average time. Why is this experiment flawed?

i. Nonresponse bias
ii. **Confounding**
iii. Selection bias
iv. None of the above

Nonresponse bias/Selection bias
Nonresponse bias occurs in sample surveys when people in a sample either refuse or are unable to respond. This has nothing to do with controlled experiments, and therefore is not the answer. The same goes for selection bias. Selection bias occurs when a part of the population is underrepresented during sampling. Again, this has nothing to do with controlled experiments. This leaves us with either Confounding or None of the Above.

Confounding
Let's think about this for a second. We know that men and women are inherently built in different ways, and that they perform differently in many different activities. Running is no exception. In this example, we have given Red Bull to all men while leaving the sample of all females to run without consuming any of it. Therefore, we do not know if the difference in the average mile times will be due to gender or due to the consumption of Red Bull! This uncertainty of which factor is the cause of a response is called confounding. In this case, we should have gotten two samples of mixed genders in order to get rid of confounding.

APPENDIX E-4. Controlled Experiments Part 2 Script

Main Example

A cosmetics company has been working on a new formula for temporary blue hair dye, and wants to know if it washes out more completely than their current formula. They have gathered a sample of 500 people, and plan to randomly assign them to one of two treatment groups – they will apply either the original or the new formula, wash their hair twelve hours later, and report whether any dye failed to wash out. However, there is some question as to whether hair length could affect the results. Three plans are suggested for assigning treatments:

1. Randomly assign the 500 people to receive either the new formula or the old formula such that 250 are treated with the old and 250 are treated with the new
2. Randomly assign 250 short haired people to be treated with the new formula and randomly assign 250 long haired people to be treated with the old
3. The 500 people are split into two groups: one of people with short hair (above their shoulders) and one of people with long hair (below their shoulders). Within each group 50% of the people are randomly assigned to be treated with the new formula and 50% are randomly assigned to be treated with the old formula.

Randomized complete block design

Which of these is a randomized complete block design?

i. Plan 1
ii. Plan 2
iii. Plan 3
iv. None of the above

Plan 1
So what is blocking? Blocking is the arranging of experimental units into groups of similar units. We call these groups blocks. While Plan 1 is valid in its own right, it makes no use of blocking, and therefore is not the answer.

Plan 2
This plan does make use of blocking, but in the least effective way possible. It divides people into blocks based on hair length – one block for people with short hair, a second for those with long hair – but goes on to assign everyone in a single block to a single treatment. We couldn't conclude that either hair dye washes out more completely than the other from this experiment design because of confounding. For instance, say a majority of the group of people with short hair using the new formula reported that all the dye washed out with ease, and that the other group found much of the dye didn't wash out. We wouldn't be able to tell if these results were because the new formula is better than the old, or because long hair traps the dye better.

Plan 3
This plan follows the randomized complete block design, and is the only plan that makes effective use of blocking. The people are assigned to blocks based on the variable we want to control for – hair length. Then the people within

the blocks are randomly assigned to one of the treatment groups – this way we know an equal number of short and long haired people were assigned to be treated with both the new and old formula. Blocking allows us to assign treatments in a targeted way to homogeneous units. The reason why blocking is so effective is because it allows us to assign treatments in a targeted way to units that are the same. We'll know for a fact that an equal number of short and long haired people were treated with the new and old formula because we designed the experiment that way – and we will be able to more accurately compare the treatments.

Complete randomization

Which of the plans uses a completely randomized design?

- **i. Plan 1**
- ii. Plan 2
- iii. Plan 3
- iv. None of the above

Plan 1
The first plan is actually an example of a completely randomized design. It does not make use of blocking to eliminate confounding from variation between hair lengths, but instead relies on randomization and a large sample size to evenly distribute people with different hair lengths among the two treatment groups. The reason this plan is a viable option is because the randomization helps to ensure that both groups have an even mix of hair lengths. This will help eliminate bias from unknown sources, like nuisance variables. Although this option is viable, it isn't optimal. In designing a controlled experiments, if we know about a potential cause of variation or bias, we should block for it. Basically, block for what you know (as long as it's possible), and then randomize inside each block for what you do not know.

None of the Above

Which of these uses strata?

- i. Plan 1
- ii. Plan 2
- iii. Plan 3
- **iv. None of the above**

None of the above
So from previous answers we've used up Plan 1 and Plan 3. From this it may seem logical that the next answer would be Plan 2. However, we must think critically. As we said in Problem 1, Plan 2 is actually flawed and contains confounding. Furthermore, strata is not even a term that has to do with experimental design. It is actually is a term that has to do with stratified sampling, a sampling strategy. This is why the answer is actually None of the above.

APPENDIX E-5. Controlled Experiments Part 3 Script

Control group, Treatment group, Placebo

We're testing the effect of a new drug meant to prevent migraines. To do so, we gather a sample of people and treat them with the drug, then follow them for a time and see if they experience migraines. If we conduct this experiment and gather our results, would we be able to conclude that the medicine prevents migraines?

Yes
No

No
People often respond positively to receiving treatment, even when the treatment itself has physical no effect. This is due to psychological belief that something may happen. In order to control for this and eliminate confounding from the placebo effect, we need to have two groups: a control group and a treatment group. In controlled experiments, a control group is a group that receives either no treatment or a standard treatment. The responses of the control group are then compared to that of the treatment group. Here, the people in the treatment group will receive the real drug, while the people in the control group will receive a placebo. A placebo is a pill, medicine, or procedure that does nothing. For instance, in this experiment we might give people in the control group pills filled with water. This will allow us to see how much of the effects we observe are due to the drug and how much are due to people's natural positive response to being treated.

Blind and Double-Blind

We change our experiment design so that now we have a treatment group and a control group. The drug administrators are instructed to give the people in the treatment group the real drug and the people in the control group a placebo and then analyze the results. Is there anything we could do to further improve our new plan? Check all that apply.

i. Don't tell the recipients whether or not they are receiving a placebo
ii. Tell the recipients whether or not they are receiving a placebo
iii. Don't tell the administrators which recipients are getting a placebo
iv. Tell the administrators which recipients are getting a placebo

First two options
Let's look at the first two options together. Would it be better for us to inform the members of the control group that they are receiving a placebo, or should we let them think they are receiving the experimental drug? Think about it logically – we introduced a control group to control for the placebo effect, where people respond positively to any treatment, even if it's just a sugar pill. If we tell them they are receiving a placebo, they will no longer believe that the treatment will help them, and the placebo effect will no longer occur. We should instead ensure that the participants in our experiment do not know whether they are receiving a placebo. This practice is called blinding, and is essential when working with human subjects in a control group – especially in a medical setting.

Second two options

So we've made sure our experiment is blind – is there anything else we can do to improve our design? Should we tell the people who are administering the treatments and analyzing the results which subjects received a placebo, or should we keep them in the dark too? It might seem strange, but it is actually best that we do not tell these people which subjects are members of the control group. If we tell our administrators, there is a chance that they might slip up and accidentally reveal that a subject is receiving a placebo. If we tell our analyzers, they might be biased in their analysis of the results. It would be best if these people did not know – and the practice of designing our experiment to ensure that they do not is called double blinding. Double blinding allows us to be as confident as possible in the results of our experiment.

APPENDIX E-6. Sample Surveys Script

Main example

A university is attempting to get an idea of the workloads of their students. To do so, they select a sample of 500 students and send a 10 minute survey out to them. Assume that the sample is representative of the entire population, and was collected correctly.

Nonresponse bias

First, assume that every student receives and sees the survey invitation. Next, imagine that around 50 students have such a high workload that they are too busy to even complete it, and that around 50 more students with light workloads are nervous the school will disapprove and choose not to fill out the survey. What issue has arisen?

i. Response bias
ii. Selection bias
iii. **Nonresponse bias**
iv. None of the above

Selection bias
Selection bias occurs during sampling, and is when a portion of the population is omitted or severely underrepresented by our sample selection strategy. In this example, we assumed that the sample was representative of the entire population. Therefore, selection bias is not the issue here.

Response bias
Simply put, response bias occurs when the response of a sample unit is not representative of his/her/its characteristics or opinions. Assuming all of the students that take the survey answer truthfully, response bias won't be the issue here.

Nonresponse bias
Nonresponse bias occurs when sample units are unable or simply refuse to respond to a survey. As around 100 students have chosen or were not able to answer, nonresponse bias will definitely be an issue. So why is this a big deal? By sending the survey, the school hopes to get an idea of whether or not the workload of its students is too high or too low. However, because students with high and low workloads have refused to respond, the school won't even know about their woes. In the worst case, the school would go on with assigning the same amount of work to its students, even though some might be completely swamped while others aren't challenged at all.

Response bias

Now let's look at a different scenario. Imagine the entire sample of students take the survey, but that approximately 25% of the students do not answer the question truthfully. For example, students with light workloads answer the survey saying they have heavy workloads because they think the school will lighten the workload if a lot of people say it is too high. What is the issue that the school will run into this time?

i. **Response bias**
ii. Selection bias
iii. Nonresponse bias
iv. None of the above

Selection bias
As we saw in the last question, selection bias has to do with sampling strategies and is when a portion of the population is either omitted or underrepresented by the sample. For the same reason as in the first problem, the sample chosen by the school isn't the problem, nor is selection bias the answer.

Nonresponse bias
Again, nonresponse bias occurs when sampling units either are unable or refuse to answer the survey. In this question, all the students in the sample answer the survey, so nonresponse is not an issue.

Response bias
In this question, response bias has occurred as there are students that have not answered truthfully. This is the definition of response bias. Just as with nonresponse bias, this is a major issue because the school will not get the true opinions of its students, and is getting false information. Imagine that all the 25% of the students lied and said they have high workloads even though they have low ones. The school might make a push to lower the workloads of its students even though many of the students aren't even facing a lot of work.

APPENDIX F. Experiment Documents

APPENDIX F-1. Survey Given to Experiment Participants

Thank you for taking the time to come in today. Your responses to this survey and sample quiz questions are invaluable to our project.

Have you taken any statistics courses besides MA 2611 Applied Statistics I?

- MA 2612 Applied Statistics II
- MA 3627 Applied Statistics III
- MA 3631 Mathematical Statistics
- MA 4631 Probability and Mathematical Statistics I
- MA 4632 Probability and Mathematical Statistics II
- Other class (please list)
- Other experience (please describe)

On a scale of 1 to 6, where a 1 indicates hatred and a 6 indicates great enjoyment, how much do you like statistics?

On a scale of 1 to 6, where a 1 indicates hatred and a 6 indicates great enjoyment, how much did you like MA 2611?

Have you made use of any video resources when studying any subject in the past?

Y/N

Would you have used a statistics video tutorial series in studying for MA 2611, had one been made available to you?

APPENDIX F-2. Four Problem Pretest/Posttest

In a pharmacological experiment on eating behavior in rats, 18 rats are to be randomly allocated to three treatment groups: T1, T2, and T3. While under observation, the animals will be kept in individual cages in a rack. The rack has three tiers with six cages per tier. In spite of efforts to keep the lighting uniform, the lighting conditions vary somewhat from one tier to another (the bottom tier is darkest), and the experimenter is concerned about this because lighting is thought to influence eating behavior in rats. The following three plans are proposed for allocating the rats to positions in the rack:

Plan I: Randomly assign the 18 rats to the 18 positions in the rack
Plan II: Put all T1 rats on the first tier, all T2 rats on the second, and all T3 rats on the third tier
Plan III: On each tier, put two T1 rats, two T2 rats, and two T3 rats

1. **Match the following:**
    1. Target population
    2. Sample
    3. Response variable
    4. Nuisance variable

    A. Lighting
    B. Rats
    C. The 18 rats used in the experiment
    D. Eating behavior

2. **Which of the three plans uses a completely randomized design?**
    A. Plan I
    B. Plan II
    C. Plan III
    D. None of the above

3. **Which of the plans has a problem with confounding?**
    A. Plan III
    B. Plan I
    C. Plan II
    D. None of the above

4. Which design makes appropriate use of blocking?

A. Plan III
B. Plan I
C. Plan II
D. None of the above