February 2012

WeBWorK Tutorials For Statistics II

Christopher J. Donnelly
Worcester Polytechnic Institute

Erika S. DiLorenzo
Worcester Polytechnic Institute

Noelle M. Richard
Worcester Polytechnic Institute

WeBWorK Tutorials For Statistics II

Interactive Qualifying Project Report completed in partial fulfillment
of the Bachelor of Science degree at
Worcester Polytechnic Institute, Worcester, MA

Submitted by:

Erika DiLorenzo
Christopher Donnelly
Noelle M. Richard

Submitted to:

Project Advisor:

Professor Joseph D. Petruccelli

February 28, 2012

Abstract

The goal of this project was to create enhanced WeBWorK problems for the topics of One-Way and Distribution Free models, which are covered in Applied Statistics II at WPI. Current literature was reviewed and interviews were conducted in order to determine a tutoring method that resembles how professors tutor students. A new Perl package was created to facilitate the creation of these problems. The new problems were tested against existing WeBWorK problems and recommendations were made based on the results.

Authorship

Our IQP team collaborated on most tasks in order to maximize the input and ideas from each team member. For the research and planning of the experiment, everyone contributed equally. We also evenly split the analysis of the experiment results and the conclusions drawn from these results.

However, some specialization was helpful because certain group members had prior experience with certain tasks. Christopher Donnelly had experience with programming and was able to teach himself the Perl code. Chris was responsible for adding the necessary functionality to the Perl package, with ideas and support from his team members. Chris also created the homework problem outline that explains how to set up a problem using our format. Finally, Chris created a document that discusses issues he came across while coding problems and editing the package.

Erika DiLorenzo and Noelle Richard focused on writing the different portions of the experiment. They wrote the pre- and post- tests, along with the experimental and control homework problem sets. Erika and Noelle also checked the WeBWorK problems for errors and attempted to find any loopholes students could have found during the testing. While Chris worked on the problem outline and documentation, Erika and Noelle performed the statistical tests and wrote about the results, in addition to conclusions and recommendations, in the final report.

Table of Contents

Table of Figures

Executive Summary

A good foundation in statistics allows students to collect, produce, and analyze data, valuable skills for engaged citizenship and future jobs. Colleges and universities, including Worcester Polytechnic Institute, understand the importance of statistical education, and require introductory courses in statistics for many majors. The goal of these classes is to introduce students to the main topics of statistics and get them thinking about the world statistically. However, many students have difficulty learning statistics because it requires them to change the way they think about problems.

One potential way to improve student understanding is by requiring the students to practice these skills through homework. Homework allows instructors to provide feedback on student's work and help them understand their mistakes. However, the frustration many students feel when doing problems without assistance can be counterproductive. Ideally, students should have a tutor available to guide them through each problem. Online homework allows the flexibility to give every student this guidance.

Both instant feedback and multiple trials in online homework sets have been shown to have a positive impact on student learning, but this project team wanted to determine if adding other tutoring methods could improve online assignments. It was determined through interviews with statistics professors at WPI that scaffolding, a tutoring method that breaks a difficult problem into smaller questions and provides hints, would be a good analog for a tutor.

In order to test this hypothesis, a test was developed which would measure the improvement of a student's score on a quiz after completing a homework assignment. Using WeBWorK, an online math homework system in use at WPI, code was written which turned regular statistics homework problems into scaffolding problems. The scaffolding problems, unlike the regular problems, contained hints and extra help questions to aid the struggling students. The homework assignments were created for two topics in Statistics II, One Way Models and Distribution Free Models.

The experiment was completed during the two lab sections of a Statistics II class. Students were given a pre-test, completed the assignment, and then took a post-test, but in one section the assignment included the extra help. In order to obtain their thoughts on the extra help, we asked students to complete a survey after the testing ended.

With the scores for all of the students collected, statistical tests were run on the data in order to determine if the scaffolding made a difference in student learning. It should be noted that the sample size was too small for any of the tests to have much power, so the likelihood of detecting a difference was very low.

A difference was detected in the mean score of the pre-tests between the students in the section that met at 8 AM and at 9 AM for both weeks of testing. A two-sample t-test showed that the students who took the pre-test at 8 AM didn't score as high as those in the 9 AM section.

A two-sample t-test did not find a difference in the mean score improvement between those students who had scaffolding and those who didn't. Additionally, a student's use of scaffolding (how many problems they used the extra help for) showed no significant correlation with their mean score improvement.

While the data for this test failed to show that scaffolding improved student learning, positive student and teacher response to the idea of scaffolding would make future trials of this experiment worthwhile. Including more students in the testing, providing more time for a longer assignment, and testing other statistics topics may add value to this testing.

Chapter 1: Introduction

"The study of statistics provides students with tools and ideas to use in order to react intelligently to quantitative information in the world around them." (Ben-Zvi & Garfield, 2008). Whether it's simply finding the mean of a set of data, or more complexly, attempting to model the relationship between two variables, statistics aids in analyzing different situations and occurrences in everyday life. By learning statistics, students become functioning and engaged citizens, assessing data presented to them and making decisions based on that data. They also prepare themselves for future jobs, which nowadays increasingly require the collection, production, and analysis of data.

Because of its large presence in the various branches of the scientific world, statistics is an important subject to learn and master, especially for students of the sciences. Introductory courses in statistics are offered at colleges and universities, including Worcester Polytechnic Institute. The goals of these classes are not only to teach the main concepts of statistics, but also to teach students how to think statistically and to "better understand and evaluate information in the world." (Garfield, 1995). However, according to Ben-Zvi and Garfield (2004), even the best students have difficulty learning and understanding statistics.

Certain topics in statistics seem complex and involved to some students, and learning the correct statistical reasoning can be difficult. Students also sometimes have preconceived notions and misconceptions about statistics, and changing the students' thought processes can be challenging (Garfield, 1995). A student may think incorrectly about a certain situation, which results in the use of incorrect statistical methods.

Garfield (1995) points out that "students learn to do well only what they practice doing." One feasible way to supply this essential practice is by providing homework. Homework allows students to practice statistics outside of the classroom and to explore the subject in more depth. Instructors can provide feedback on student work, which allows students to recognize their mistakes and correct any of their misconceptions. Homework also presents instructors with an assessment of student performance.

While homework problems can help students acquire knowledge of statistical concepts and special skills necessary to solve statistics problems, homework can be frustrating for many students. Since the steps and skills needed to successfully complete a problem are not intuitive to many students, some researchers, such as Gal and Garfield, recommend that students receive rapid and individualized feedback to help them recognize and overcome stumbling blocks (1999). Ideally, each student would be carefully guided by a tutor through practice problems to ensure a true understanding of the material. However, this is nearly impossible, even in the smallest classes. Luckily, research shows that, in addition to tutoring, (Chi, 1996) feedback can be provided successfully through group work (Goos, 2004) or online homework (Palocsay & Stevens, 2008).

Online homework seems to be the most practical option for many instructors, since it can be used to provide the recommended feedback to every student instantly, while at the same time assisting in homework management and grading. Online homework allows instructors to assign more problems, without the worry of tedious grading. Online homework systems can also provide immediate feedback to students, and usually give multiple tries for each problem. Both instant feedback and multiple trials in online homework sets have been shown to have a positive impact on student learning (Brewer, 2009).

However, online homework systems are often not used to their full potential. Setting up good statistics problems in a web-based system can be time consuming and frustrating for professors. Professors not only have to learn how to code problems, but they have to take the time to set up each individual problem. Because of this, homework problems are typically set up simply. The problems lack tutoring methods and hints, and give little feedback when a wrong answer is submitted. As a result, students do not receive the full benefit of online homework.

Recently, IQP teams at WPI have attempted to confront the issues regarding the use of online homework in statistics classes. Last year, an IQP team focused their efforts on improving online homework for Statistics I at WPI. The team created a new Perl package that expanded the functionality of the online homework system, WeBWorK, which is used in statistics courses (Li & Xia, 2011). The new package allowed for improved homework problems that were intended to help students learn the subject of statistics more efficiently. The package provided a template for professors to create enhanced multi-part problems and to include hints. This group also created and tested a new learning model, called the "Forward Fading Model." "The Forward Fading Model," which combines two already existing learning models, gradually gives students more help as a problem progresses (Li & Xia, 2011). Another IQP team at WPI wrote new

WeBWorK problems specifically for Statistics II, giving instructors a bank of questions to easily implement into assignments.

Our IQP team continued the research and effort of improving online homework problems for statistics courses at WPI, with the goal of allowing more students to effectively master the subject. We focused specifically on improving online homework for the Statistics II course. We created homework problems that involved tutoring methods, which we hoped would help students learn statistics more efficiently. In order to create these problems, we needed to improve upon the previous IQP's Perl package, which allowed for hints and multiple part problems. We constructed an outline for professors to use to make the set-up of good homework problems using our updated Perl package easier. In order to accomplish these tasks, our group focused on six major goals:

- Further test and fine-tune the functionality of Li and Xia's package for WeBWorK, which allows for the use of different learning models
- Research effective homework problem types and tutoring methods, specifically tutoring methods used by professors at WPI.
- Determine a plausible and valuable question type/tutoring method to test
- Create a set of WeBWorK homework problems using the researched question type/tutoring method and test them against the current WeBWorK problems
- Determine if the new homework questions are more effective than the current homework problems
- Create a homework problem outline using our improved Perl package

With the completion of these goals and the help of our homework problems, we attempted to improve student learning of statistics at WPI. We have created an improved Perl package that adds functionality to WeBWorK, and provided an outline for professors to easily implement scaffolding into problems. Our research represents a step towards more efficient statistics homework.

Chapter 2: Background

2.1 Statistics

2.1.1 What is Statistics?

Statistics is "the science that deals with the collection, classification, analysis, and interpretation of numerical facts or data." (Dictionary.com, 2011). Statistics helps people make sense of information gathered from the world around them. Properly designing and conducting scientific studies requires the understanding of statistics in order to prevent bias and to correctly interpret data. Statistics also allows people to provide evidence that makes arguments and research credible (Ben-Zvi & Garfield, 2008). Patterns and relationships in everyday life can be analyzed through statistical methods. Although there is a level of uncertainty when analyzing data, the correct use of statistics can provide a rational basis for decision-making.

2.1.2 Importance and Applications of Statistics

"The applications [of statistics] are as diverse as improving Internet search and online advertising, culling gene sequencing information for cancer research and analyzing sensor and location data to optimize the handling of food shipments." (Lohr, 2009). Statistics appears in various aspects of everyday life, and allows people to explore past events, predict future events, and make decisions to better the world. For example, businesses use statistics to figure out which products customers demand the most and to decide how much of a certain product to make (Emathzone, 2008). Businesses also collect and analyze data on customer satisfaction in order to make improvements and to maximize profit. Pharmaceutical and medicinal researchers (Smith, 2001) use statistics to examine whether a certain drug is effective for curing a disease.

Statistics can be applied in other areas as well. Astronomers use statistics to estimate "distances, sizes, masses, and densities" of planets, stars and other matter in space (Emathzone, 2008). Environmental and ecological scientists (Smith, 2001) utilize statistics to analyze not only local, but global changes on Earth. Statistics supports and gives evidence towards ideas such as global warming and climate change. People frequently come across newspaper articles that mention statistical studies and use graphs based on data. It seems only logical, that in order for a person to "react intelligently to quantitative information in the world around them," he or she must understand statistics to some degree (Ben-Zvi & Garfield, 2008).

2.1.3 Statistics at College and Pre-College Levels

Because of the increasing importance of statistics in everyday life, colleges and universities provide statistics courses to their undergraduate students (Ben-Zvi & Garfield, 2008). Several different departments at universities, not just mathematics departments, offer the subject because "elements of statistical reasoning have become requisite for a wide range of fields of study." (Ahlgren & Garfield, 1988). Prior to the late 1980's, according to Ahlgren and Garfield (1988), few students learned statistics before they reached college. Nowadays, however, more students see statistics in middle school and high school (Ben-Zvi & Garfield, 2008).

2.1.4 Statistics at Worcester Polytechnic Institute

Worcester Polytechnic Institute (WPI) offers introductory statistics courses to its undergraduate students. MA 2611, or Applied Statistics I, "introduces the student to data analytic and applied statistical methods commonly used in industrial and scientific applications as well as in course and project work at WPI." (WPI Undergraduate Catalog 2011-2012, 2011). Statistics I teaches students how to graphically represent data, how to properly design and perform observational and experimental studies, and how to correctly analyze data through appropriate summary measures. The course also acquaints students with, among many other things, the central limit theorem, and the basics of confidence intervals and hypothesis tests, for one and two populations. Instructors often show students in class and labs how to use computer programs, such as SAS, to graph and analyze data.

Applied Statistics II, or MA 2612, is a continuation of Applied Statistics I (WPI Undergraduate Catalog 2011-2012, 2011). According to the course catalog description (2011), Statistics II covers "simple and multiple regression, one and two-way tables for categorical data, design and analysis of one factor experiments and distribution-free methods." This class assumes students understand the topics presented in Statistics I. WPI professors often utilize computer programs in this class as well.

Some undergraduate majors at WPI, including, Biomedical Engineering, Business, Computer Science, and Mechanical Engineering, require students to take the introductory statistics course(s). Other majors, such as Actuarial Mathematics, suggest students enroll in statistics courses in order to prepare for future endeavors at WPI and to prepare for the workforce.

2.2 Difficulties Learning and Understanding Statistics

2.2.1 Statistical Thinking

Ben-Zvi and Garfield (2008) consider statistics as "a type of bridge that connects mathematics and science, in that it provides the mathematical foundations for analyzing data gathered in the real world (science)." A lot of mathematics involves deductive reasoning- one explains a general idea and then moves on to explain more specific ideas. Science, on the other hand, tends to use inductive reasoning- one collects information about specific instances to try and draw general conclusions. For example, Gregor Mendel cross-bred around 29,000 pea plants and then used the results from the experiment (specific) to develop his laws of inheritance (general). Statistics combines aspects of both deductive and inductive reasoning. It allows for the application of mathematics to solve real life problems.

A different way of thinking arises when dealing with statistics. This reasoning is known as statistical thinking. Statistical thinking requires people to analyze data and make general conclusions, similar to how scientists collect information and propose theories. At the same time, because statistics uses mathematics to analyze information, general topics can be broken down into more specific ideas. For example, the use of confidence intervals is a general topic in statistics. However, different aspects, such as the distribution of a population, or the size of the population, can affect how the confidence interval is calculated. Thus, the main idea of confidence intervals must be broken down into smaller, more specific topics to deal with the differences.

Statistical thinking also involves the ideas of randomness and variation, as well as likelihood and probability. One must learn the proper methods for drawing conclusions on events that have variation. For example, a student might observe an inverse relationship between gas prices and crude oil supply; the less oil available, the higher the gas prices. However, the student must realize the relationship between oil amount and gas price is not deterministic, since there are many uncontrollable variables that influence the relationship. One must also learn to deal with randomness through the use of probability and understand that certain events may be more likely to happen than others.

Many students have little to no experience drawing conclusions from data, handling randomness, etc., prior to taking a statistics course. They must learn a whole new way of thinking, which can become problematic. Students see similarities between topics they learned in previous mathematics classes and topics presented in statistics classes. Because of this, and because both mathematics and statistics deal with "numbers," students believe they can apply the same thought processes to both types of problems (Ben-Zvi & Garfield, 2008). However, using the same thought process can result in the incorrect collection of data, the misuse of statistical methods, and the misinterpretation of results. Instructors must point out the key differences between the two subjects, and teach students to use statistical reasoning when exploring data (Ben-Zvi & Garfield, 2008). Nevertheless, even if instructors point out differences, changing a student's thinking process can be difficult (Garfield 1995).

2.2.2 Common Misunderstandings and Misconceptions

In addition to lack of statistical reasoning and thinking, students often misunderstand certain statistical ideas (Garfield 1995). Garfield (1995) summarizes some of the most prominent mistakes students make, some of which are:

- "Representativeness: People estimate the likelihood of a sample based on how closely it resembles the population." (Garfield 1995). Under this misconception, students believe an outcome is just as likely for a small sample as it is for a large sample. For example, Garfield (1995) mentions how students think the chance of getting 70% heads in 10 coin tosses is the same for 1000 coin tosses.
- Gambler's fallacy: Students believe that "chance is a self-correcting process." (Garfield 1995). Garfield offers a simple example of this misconception. If a person tosses a coin and observes heads for several tosses, the person thinks that tails should appear soon. The person believes this because it seems to better represent the 50% chance of observing tails. However, thinking the process will correct itself is incorrect.

Other common mistakes deal with biases, confusing correlation with causality, and making "yes or no decisions based on single events rather than a series of events." (Garfield 1995). Students also sometimes have difficulty obtaining data correctly (Knowledge@Wharton, 2008).

WPI Professor Petruccelli noted more specific issues students have when learning statistics at WPI (2011). In Applied Statistics I, students have trouble identifying different types of studies and have difficulty making interpretations from the studies. Students also have trouble with the meaning of "confidence" in a confidence interval, what a p-value is, and how to interpret results of confidence intervals and hypothesis tests. In Applied Statistics II, students still sometimes have difficulties with p-values and interpreting confidence intervals and hypothesis tests. Students also have problems with learning the more involved statistical models introduced in the class.

2.2.3 Dealing with Misconceptions

Many researchers, including Garfield (1995), Ben-Zvi (2008), and Ahlgren (1988) have proposed ideas to help students overcome their misconceptions. For example, Garfield (1995) suggests that instructors actively help students confront their mistakes and use computers to help students visualize problems. She also advises that students should, in addition to performing mathematical computations, practice analyzing statistical information to gain a "deeper understanding." Receiving feedback constantly and in a timely fashion also allows students "to reflect on the feedback they receive, make adjustments, and try again." (Garfield 1995).

2.3 Homework in Education

2.3.1 History of Homework

Until the mid-twentieth century, homework was generally frowned upon by parents and educators. Since most children finished their education before high school, homework was seen as unnecessary, especially for young children. It was not until the launch of Sputnik that homework was seen as an effective tool to increase student learning in a competitive age.

Today, while the amount of homework assigned to school-aged children is still under debate, it has been generally shown that the effectiveness of homework increases with age. Also, older students can handle and get more out of greater amounts of homework (Cooper, 2008).

College students are expected to complete even more homework then they did in high school. For example, Worcester Polytechnic Institute expects approximately 36 hours of work (including homework, study time, reading and preparation) to be completed per week outside of class for a normal course load (WPI Undergraduate Catalog 2011-2012, 2011). This is more than twice the amount expected of high school students, 15 hours a week, assuming 30 minutes of work a night for 6 courses. (Cooper, 2008 “A Brief History of Homework”) Since college students spend so much time on homework, it is important to consider the advantages and disadvantages of these assignments.

2.3.2 Benefits and Challenges of Homework

While research on homework in the university setting is limited, much has been written about what homework can do to help younger students learn. The American Federation of Teachers cites the following benefits of assigning homework: “gains in skills, abilities, and test scores..., a positive attitude toward schoolwork, more cohesion between school and home environments and an improved self-concept in terms of the student’s ability as a learner.” (American Federation of Teachers, 2010). Supporters of homework believe that it helps students realize they can learn outside of the classroom and prepares them for the workload of higher education.

Concerns about homework are raised when some recent research is considered- according to a 2004 national survey, the amount of time spent on homework by school-aged children is up 51% since 1981. The data also suggests that more than two hours a night of homework for high school students can in fact have a negative effect on test scores. Those who would abolish homework argue that it can stress parent-child relationships, take away from family time and make students resent school (Wallis 2006).

2.3.3 Assigning Effective Homework

While there may be valid arguments in favor of and against the use of homework, it is more relevant to this research to understand what makes a "good" or "bad" homework assignment. Research shows that effective homework assignments have the following qualities:

- Are short and assigned frequently
- Include material from prior and future classes, not just one day's worth
- Include both hard and easy problems in an assignment
- Take into account different learning styles of students
- Provide feedback, not just a grade
- Have a reward after completion (Cooper 2008: Effective Homework)

However, it must be taken into consideration that most of these characteristics have been tested individually and on high school, not college students.

Combining all of these factors into traditional homework assignments can be a challenge for teachers. Teachers would need to first create many assignments that include diverse material, consider all learning styles, and include different levels of difficulty. Then they would individually grade and provide feedback and a reward for every student, even in classes with hundreds of students, such as in a university.

2.4 Online Homework

2.4.1 Advantages of Online Homework

Online homework has many benefits, for both instructors and students. Online homework allows instructors to easily create and grade assignments. Instructors and teaching assistants (TAs) do not always correct paper based homework, mostly because it can be very time consuming. As a result, many students do not complete assigned paper homework because they know their work will not be graded. Students might be more inclined to complete online homework because it will be graded.

Online homework also provides students with instant feedback, which researchers show is helpful when learning something new (Palocsay and Stevens, 2008). A time lag occurs when instructors collect paper-based homework and when they return it graded. The student cannot address any mistakes or errors he or she made immediately, which inhibits learning. Online homework allows the student to address problems instantly. The immediate feedback also helps prevent students from miscalculating their abilities and understanding of the material, which can lead to poor learning strategies (Brewer, 2009, "Effects of Online Homework").

Online homework has other benefits as well. Online homework is non-judgmental, which students find more comfortable. Students can attempt and fail problems multiple times, without having to feel embarrassed because they did not answer correctly the first time. In addition, the use of online homework can help prevent cheating. Random number generators allow students to receive the same type of problem with just different numbers to work with. Instructors can also give each student a set of homework problems randomly selected from a larger set, thus making each student's homework assignment unique.

2.4.2 WeBWorK

WeBWorK is one of the many online homework systems created for mathematics. It is one of the most extensively studied online systems and is used at more than 150 colleges and universities. WeBWorK is able to display and grade a wide range of problems from multiple-choice to open response math equations. Essentially, WeBWorK can grade any type of problem, as long as one can write the programming code to determine if an answer is correct.

WeBWorK provides a different way of helping students with an "Email the Instructor" button, which allows students and instructors to stay in contact. WeBWorK permits an instructor to see where a student got stuck on a problem and the instructor can see previous answers the student submitted. From this, the instructor has the ability to give as much help as he or she desires, as well as target help to address a specific issue. In addition, instructors can monitor class progress at all times, pinpoint students struggling in the class, and see how well the class as a whole understands a given topic. WeBWorK can supply students with solutions to problems after the assignment has closed.

Other advantages of WeBWorK include its simple setup, the large library of pre-existing problems created for a variety of courses, and the ability to easily create new problems using its Perl-based language. Columbia University, Morgan State University, Alfred University, The College of New Jersey, University of Portland, and Eastern Michigan University have seen improvements in student performance because of the use of WeBWorK (Bressoud, 2009).

Currently, a WeBWorK Wiki and a WeBWorK Forum exist. These websites provide instructions on how to use WeBWorK, answer common questions, and supply other information for WeBWorK users.

2.5 Learning Models and Tutoring Methods

"Learning style is an individual's unique approach to learning based on strengths, weaknesses, and preferences." (Dictionary.com, 2011). A learning model is a theory of how to go about teaching students in the most effective way based on their learning style. Learning models recommend tutoring methods, which cater to an individual's needs. There are many different learning models, and we researched three that could be easily implemented into WeBWorK problems for this project.

2.5.1 Catrambone's Subgoal Learning Model

Catrambone's Subgoal Learning Model hypothesizes that student learning can be improved by solving problems in a step by step fashion. Each problem has one main goal, but the student can break the problem down into a series of steps, called "subgoals." By working on one subgoal at a time, the complexity of a problem decreases and the student gains a better understanding of how to go about solving a more difficult problem (Catrambone 1998). According to Catambone, "subgoals guide problem solving by helping learners focus on the steps to modify in novel problems that involve the same subgoals but require new steps to achieve them." (1998). So instead of feeling overwhelmed by what feels like a new problem, a student recognizes similarities between problems he or she already solved and the new problem. The student can then make small adjustments in order to solve the new problem.

Catrambone determined through a set of four experiments that the grouping of certain steps together did not directly improve student performance on problems different from the original. Rather than memorizing one specific way to solve a problem, the process of a student self-explaining the purpose of a set of steps improves his or her understanding (Catrambone, 1998). Further experimentation showed that providing a conceptual, not computational subgoal learning example to students was preferable (Merrill, 2003).

2.5.2 Renkl, Atkinson and Maier's Backward Fading Model

This learning model relies on the theory that successively removing more and more worked example steps as a student completes problems can improve student understanding, without increasing learning time. The experiments carried out introduced "prompts designed to encourage learners to identify the underlying principle illustrated in each worked-out solution step." (Atkinson, 2003). This method was shown to create mid to large size gains in student learning. This method is seen as desirable because it is relatively simple for an educator to implement, does not increase learning time, and is effective.

2.5.3 Scaffolding

Scaffolding is a tutoring process that is used to help a learner increase his or her understanding by identifying his or her "zone of proximal development." A student's zone of proximal development is the new knowledge that can be obtained with support from a teacher, peer, tutor or computer, measured from the current knowledge level (Bull, 1999). This method of tutoring is effective because the level of support for each topic is matched exactly to each student, and this matching can even be reproduced within a computer-based assignment (McLoughlin, 1999).

The first step of scaffolding is to determine the knowledge level of the student. This is accomplished with an example problem, where the work a student is able to complete unassisted for that problem would be his or her current knowledge level. If a student is able to complete an entire example, no scaffolding for that topic is needed. However a student who gets this initial problem incorrect may need more assistance.

The next step of scaffolding is to provide the student with new information to help the student work through the given problem. This support can be provided through hints, suggestions of next steps, asking questions of the learner, references to other materials, etc. The idea is to support the process of understanding where a student went wrong in the problem initially and to correct any misconceptions or gaps in knowledge.

The final step is to return the student to the original place of difficulty and allow him or her to answer the question unassisted. If he or she is able to, there is no more scaffolding needed and the student can move on to the next topic or problem. If there is still difficulty or confusion, further support may be needed (Bull, 1999) (McLoughlin, 1999).

2.6 Previous IQP Work

2.6.1 WeBWorK Problems for Statistics I

From A-term of 2010 to C-term of 2011, the IQP group of Zehao Li and Yizhou Xia worked on improving WeBWorK homework problems for Statistics I classes at WPI. The group researched two learning models, which help students solve "novel problems:" Catrambone's Subgoal Learning Model and Renkl, Atkinson and Maier's Backward Fading Model (Li & Xia, 2011).

Li and Xia then modified the Backward Fading Model to create their own model, which they named the Forward Fading Model (Li & Xia, 2011). Instead of reducing help gradually, the Forward Fading Model provides more help gradually. The problem is broken down into "steps," as in the Subgoal Learning Model. If a student answers a problem part correctly the first time, the student receives no help and can move onto the next question part. However, if the student answers incorrectly, hints are made available to the student. The first hint helps the student approach the first step of the problem; the second hint helps the student approach the second step, and so on. A student can continue asking for hints, until the student solves the problem part, or until he or she views all available hints (Li & Xia, 2011).

Combining the Subgoal Learning Model and the Forward Fading Model together, Li and Xia developed new homework problems for Statistics I classes. The problems involved multiple parts, which act as subgoals to one main goal. The new homework problems also included hints to help the student through the problem. In addition, Li and Xia state that "when students went to a new part of the problem, all solutions of previous subgoals were shown in the problem statement part as help." (2011).

In order to test their new problems, Li and Xia had to modify an existing Perl package to implement into WeBWorK. WeBWorK problems are written in Problem Generation (PG) language, which uses Perl macros and LaTeX code. Perl macros perform specific tasks, such as "inputting data and mathematical formulas, and computing and outputting the results." (Li & Xia, 2011). Collections of Perl macros are known as packages, which load when a WeBWorK session begins. However, no Perl package in the WeBWorK library performed all the tasks required by the learning models (i.e. multiple part questions with "next" buttons, hints, etc.). Li and Xia's new package allowed for the implementation of the Subgoal and Forward Fading models (2011).

Li and Xia tested their new problems in the Statistics I lab periods in B-term of 2010 (2011). They tested problems on estimation and confidence intervals one week, and problems on hypothesis tests a second week. Li and Xia randomly chose half of the students in each lab section to solve traditional WeBWorK problems; these students were the control group. The other half of the students, the treatment group, solved the new problems. Each student completed a pre-test and post-test to help the IQP group evaluate whether or not the new problems were effective (Li & Xia, 2011).

Li and Xia (2011) "found that there were no statistically significant differences in score improvement between students in the control group and students in treatment group." In other words, their new problems were not more effective than the old problems. However, the group did find the following results for students who scored 50% or less on the experiment's pre-test:

> *"1. A significant positive relationship between the score on the instructional material (either treatment or control) and the increase in the post-pre score (as a proportion of instructional material score.*
>
> *2. The increase for those in the treatment group was greater than for those in the control group."* (Li & Xia, 2011).

This result held for both all the material the group tested.

The IQP group suggested future improvements for their Perl package and homework problems. For example, they suggested modifying the Perl package for multiple choice problems to "allow targeted hints or help based on the wrong answer chosen." They also suggested ways to improve testing methods and to decrease noise in the resulting data.

2.6.2 Designing WeBWorK Problems for Applied Statistics II

From B-term 2010 to D-term 2011, the IQP group of Minh Le and Vorayos Roungrojkarnranan worked on improving WeBWorK homework problems for Statistics II courses at WPI. The group researched the advantages and disadvantages of different types of homework problems, and created a new WeBWorK homework set which was tested against previous WeBWorK problems.

Le and Roungrojkarnranan compared the benefits and challenges of using different types of homework problems in WeBWorK, including multiple choice, short answer and essay questions. They also found that an effective homework set contains computation, conceptual and interpretation questions. Through research, the group determined that multiple choice problems should be used to test conceptual and interpretation skills, while short answer problems should be used to test computational abilities (Le & Roungrojkarnranan, 2011). They created problem sets which met these qualifications for simple and multiple linear regression.

In order to test the effectiveness of their new problem sets, Le and Roungrojkarnranan tested them against problem sets on the same material from the WeBWorK National Problem Library. After analyzing the data, the group concluded that "it was seen that our simple linear regression problem set was significantly more effective than the control set in improving students' scores while the multiple linear regression set was not significantly more effective than the control set." (Le & Roungrojkarnranan, 2011). They identified potential sources of bias, including differences in length and difficulty between the problem sets and the length of the pre- and post-tests.

2.7 Conclusion

With the knowledge gained from our background research, our IQP team moved forward with our own study. We used the insight from prior experiments and research to develop new WeBWorK problems in order to test the effect of scaffolding on student learning.

Chapter 3: Methodology

3.1 Introduction

In order to achieve our goal of improving WeBWorK through the addition of tutorials, we carried out a number of steps leading up to and following a testing phase.

3.2 Detail of Steps

We will now outline and explain each of the steps in more detail, and identify background research, decisions we made, and steps for each of these components of our project.

3.2.1 Research Effective Homework Problem Types, Tutoring Methods, Relevant Literature, and Past IQP Reports

In order to identity a new problem type to create and test, we researched many different theories of how students learn and the experiments that had tested these theories. We identified a few of these theories to focus on through conversations with professors, previous IQP projects, and additional background research.

We found it relevant to review previous IQP projects that dealt with improving online homework for statistics classes at WPI. By doing so, we learned about the attempts other groups made to improve problems. We noted specific tutoring methods previously tested and whether or not these tutoring methods had an impact on statistics students at WPI. Our group also gained ideas for our project, as well as insight on how to avoid issues with testing from suggestions and comments made by the previous IQP groups.

In addition, we researched the main issues students have with learning statistics and WeBWorK. By understanding the most common mistakes students make when learning statistics, our IQP group attempted to focus on fixing these issues through our homework problems.

Finally, our group researched information about WeBWorK. Knowing how WeBWorK operates allowed us to try and utilize all of its functions to our advantage, and minimize any technical issues. By learning about WeBWorK, we were able to design our own package that incorporates scaffolding into homework problems. We also were able to design our own problems as well as an outline to help make the set-up of complex problems easier for professors. We have provided the code needed to write a problem using our package, and a simple example problem that should allow professors to easily create their own problems. We have also included a write up of challenges that we faced while editing the code and the changes we made, in order to make future work on this package more efficient.

3.2.2 Determine a Plausible and Valuable Question Type to Test

We decided to interview professors who had previously taught Statistics at WPI because we felt that they could provide the best insight into the tutoring and teaching methods that help WPI students learn statistics. We specifically spoke with Professors Zheyang Wu, Joseph Petruccelli, Jacob Gagnon and Gloria Tashjian, all of whom had taught Statistics I or II in the past few years. We asked the professors how they help students with specific homework problems, as well as homework topics in general. Our group also asked the professors what they think makes a good homework problem, and questioned them about their use (or non-use) of WeBWorK. In addition, we interviewed Professor Heffernan, the co-creator of the online homework platform ASSISTments. We asked Professor Heffernan about the tutoring methods implemented into ASSISTments and how online homework can be beneficial to students. The interview questions can be found in Appendix A.

In general, professors told us that they help a student with a homework problem by diagnosing the source of the misunderstanding causing the student to get the problem wrong. For a more competent student, the professor will point out the mistake and encourage the student to work the problem again. If the student needs more assistance, the professor will help the student more by either working through most of the problem with him or her (never actually giving the answer) or providing a similar worked example.

We identified scaffolding as a good analog for this process with the help of Professor Heffernan, who uses scaffolding in ASSISTments. Scaffolding identifies the ability level of a student by presenting the entire problem to complete. More competent students, who can correctly answer the initial problem in one or two tries, are not interrupted by scaffolding in ASSISTments. For the less competent students, who cannot answer the problem correctly in a few tries, scaffolding breaks the main problem down into steps and helps the students identify their weak spots. These features were implemented in our experimental problem set.

3.2.3 Further Test and Fine-Tune the Functionality of the New Package for WeBWorK

The WeBWorK package created by the IQP team of Li and Xia allows for the use of different learning models, such as the "Forward Fading Model" that the team created. The team used an existing WeBWorK package called "Compound Problem" which created a multi-part problem, and added a maximum number of trials per part and a hint feature.

Since we used scaffolding, we required further functionality from this package. We added the following functions to our new package, called "Scaffolding Problem":

- Fixing the error that automatically carries over and checks answers from one part to the next
- Allowing for a problem to have multiple graded parts, but only have the first part (the general question) count towards an overall score
- Permitting students who get the general question correct to move on to another problem without forcing them to use help
- Allowing a student who gets the general question incorrect and has completed scaffolding, to re-try the general question
- Enabling an instructor to give partial credit to a student who completes scaffolding but still cannot complete the general problem
- Allowing a different number of attempts for each part of scaffolding
- Allowing a student who has completed scaffolding, to re-try the original problem, but denying the student the option of completing scaffolding a second time

We researched the Perl code needed to add each of these functions to our experimental homework set, using the WeBWorK Wiki, WeBWorK Forum, the help of professors who have used WeBWorK in the past, and the Li and Xia IQP report. Essentially, we used a "guess and check" method to modify the code. In other words, we made a change to the code, investigated what the change did, and then made more changes until the problem functioned as desired. We used this "guess and check" method because none of us were familiar with Perl and there was no documentation for Li and Xia's package.

In addition to the package modifications, we also modified features of the individual problems. We adjusted and organized the appearance of each problem in order to minimize confusion. For example, we increased the size of the "Submit" and "Click for Help" buttons so students could easily find buttons for submitting answers or getting help. We also added color to the scaffolding help so students could differentiate between the main question and the help.

3.2.4 Create a Set of WeBWorK Homework Problems

In order to test the effectiveness of the experimental problems that use scaffolding, we created two homework sets: control and experimental. We modified existing WeBWorK problems to create the control set, then added scaffolding to those problems to create the experimental set. By having both problem sets cover the same exact material, we were able to control for changes being tested (the addition of scaffolding). Appendix B includes figures which outline how a current WeBWorK problem was converted into a scaffolding WeBWorK problem.

3.2.5 Test New Problems Against the Current WeBWorK Problems

In order to identify a process to test the experimental problems against the control, we looked to similar experiments that had been completed, specifically the previous IQP reports. Both previous IQP teams used the same general procedure. We followed this model, but implemented some changes that were recommended by the previous IQP teams and some suggested by ourselves and Professor Petruccelli.

The experiment was a randomized controlled test. The previous project groups began by having the entire lab section of students take a pre-test. Li and Xia incorporated the pre-test in WeBWorK, while Le and Roungrojkarnranan gave a paper pre-test. Students were randomly assigned to the experimental or control problems, and were allowed to use whatever resources they would typically use on a homework set to complete them (books, internet, SAS, etc.). The teams then gave them a post-test, and the increase in pre- to post-test score was used as an indicator of increased understanding. The experimental and control problems were also graded and used as part of the analyses.

We had 50 minutes for the students to complete the pre-test, problem set, and post-test. Because of the limited amount of time, the pre- and post- tests consisted of only two questions each with two parts. The WeBWorK portion consisted of 3-4 questions. Initially, we wanted to have the pre- and post-test be separate WeBWorK assignments that would only be open for a short amount of time. However, if a student arrived late to class, we would have had to find out his or her name and then go into WeBWorK to extend the time on the pre-test. This process would consume more time than handing a paper pre-test to the student and taking it from him or her 10 minutes later. Because of this, and the fact we wanted the ability to give partial credit, we opted to use paper tests instead. In each 50 minute lab period, we used the first 5 minutes to prepare students for the activity, followed by 10 minutes for the pre-test, 25 minutes for the WeBWorK problems and 10 minutes for the post-test. The schedule was slightly adjusted for the students who arrived late.

We chose to test the topics of One-Way Models and Distribution-Free Models for two reasons. First, these are the last two topics on the syllabus for Statistics II, which allowed us to test at the end of B-term and gave us more time to perfect the problem sets. Second, the previous IQP group for Statistics II did not create nor test problems for these topics. By focusing on One-Way Models and Distribution-Free models, we expanded the number of new homework problems available to professors related to these topics.

We tested over a two-week period, with two lab sections that each met once a week, one at 8a.m. and one at 9a.m. The first week, all students were tested on material related to One Way Models, but the 8a.m. section had the experimental problems and the 9a.m. section had the control problems. The second week, the problems were on topics dealing with Distribution Free Models. This time, the 8a.m. section had the control problems and the 9a.m. section had the experimental problems. We chose to assign the experimental and control problems by section in order to prevent students from looking at another screen which had a different problem type. However, if students came to the wrong lab section, we had them sit in the back row so they could not look on with another student.

For both lab sections, we handed out a set of instructions to each student at the very beginning of class. The instructions informed the students that they would complete a pre-test, followed by a WeBWorK problem set, and a post-test. We provided an explanation of how to submit answers in WeBWorK and gave even more detailed instructions on how to navigate through the "Help Mode" for those students with scaffolding. The instructions, which can be found in Appendix C, were intended to reduce confusion about the testing and how to proceed. By receiving the instructions ahead of time, the student could ideally start the pre-test exactly at

8:05am (or 9:05am) and then move directly to the WeBWorK problems at 8:15am (or 9:15am) without delay.

For each week of testing, our group created two paper tests that would serve as the pre-test and post-test. In each lab section, about half of the students chose to sit on the left side of the room and the other half on the right side of the room. We gave the left side of the room test one as the pre-test and the right side of the room test two as the pre-test. For the post-test, the students received the other test; so those students who received test one for the pre-test received test two as the post-test. We chose to do this in case we made one test easier than the other. If one test was easier than the other, it would have been difficult to determine if a student's improvement was because of scaffolding or not.

This was a single-blind controlled experiment, since we were aware of which set the student had been assigned, but the student was not.

3.2.6 Determine if the New Homework Questions Are More Effective

Like the previous IQP teams, we needed to conduct statistical tests to determine if the experimental homework set was indeed more effective than the control set. At the end of testing, we compiled data, including not only the pre- and post- test scores, but also the students' WeBWorK scores on our problem sets. We ran tests to determine if there was:

- a difference in the mean score between Pre-test 1 and Pre-test 2
- a difference in the improvement on the test scores between the experimental and control group
- a correlation between the number of problems that students used scaffolding for and their test score improvement

Our group also decided to send out a survey to the students who participated in both weeks of the testing. Through the survey, we were able to receive more specific feedback from the students, such as their opinions of the new questions.

3.2.7 Create a Homework Problem Outline

We want it to be as simple as possible for professors to implement this improved problem type or for other research groups to continue the research. We also want to make professors' jobs easier, since many expressed frustration with WeBWorK coding in our interviews.

We have provided the problem outline we used to write our problems, along with a clear tutorial on how to take advantage of the added functionality of our package, so that professors can write scaffolding problems for other classes or topics. We also provided a write-up of the changes that we made to the package and problem outline so that professors and future IQP groups will have to spend less time learning how to use the package.

Chapter 4: Results and Discussion

4.1 Interviews with Professors

4.1.1 Interview with Professor Wu

The interview with Professor Wu took place on September 13, 2011 in his office. Professor Wu taught Statistics I during both the 2009-2010 and 2010-2011 academic years.

Professor Wu stated that the type and amount of help he gives on a specific homework question varies among students. He said the type and amount of help depends on the level of student. Professor Wu tends to give the better students, or those who understand most of the material, more indirect help and hints. He works out an example similar to that of the homework problem for the students who are struggling more. However, no matter what the level of the student, Professor Wu will only help the student through about 80% of the problem. He expects that the student tackle the other 20% on their own.

When a student doesn't understand a general topic, Professor Wu will try to give the "bigger picture" to help the student understand the main points and ideas. He will discuss previously covered topics that lead up to the topic that the student doesn't understand. He does this in order to make sure the student understands the necessary building blocks. In addition, Professor Wu will point out key topics and ask that the students review their notes carefully.

Professor Wu used WeBWorK for one of his statistics classes and traditional paper-based homework for another. Professor Wu stated that the set-up of homework problems in WeBWorK was time consuming, mostly because the homework problem template was vague and difficult to use. He also noted some technical problems when using WeBWorK, such as problems with rounding and small decimal place differences. However, Professor Wu valued the immediate feedback students received when submitting an answer. He also stressed the fact that homework is only practice, so providing students with multiple tries was beneficial. In addition, Professor Wu pointed out how online homework is non-judgmental, which he sees as another benefit.

Professor Wu balances between conceptual and calculation problems when assigning homework. He chooses problems that are relevant and closely related to topics he discussed in class and believes good homework problems cover major topics, rather than minute details.

4.1.2 Interview with Professor Petruccelli

The interview with Professor Petruccelli took place on September 15, 2011 in his office. Professor Petruccelli taught both Statistics I and Statistics II during 2009-2010 and 2010-2011 academic years. He also taught Statistics I during the 2006-2007, 2007-2008, and 2008-2009 academic years.

When helping students with a specific homework question, Professor Petruccelli tries to diagnose the source of the problem. He asks for the student's work and tries to pinpoint where the student went wrong. He avoids going through the entire homework problem, but will work out a parallel problem to the homework question. Professor Petruccelli makes sure the student understands the concepts associated with the problem as well. When a student does not understand a concept, he will word the main idea differently than how it was explained in the class lecture.

Professor Petruccelli used WeBWorK in both statistics classes. He liked how the online homework platform was nonjudgmental, allowed multiple tries, provided instant feedback, and automatically graded homework. WeBWorK also collected data on student performance in a class. It showed how many times a student attempted a problem, what answers the student gave, and how long they worked on a problem. Professor Petruccelli also noted downsides to using online homework. WeBWorK limited the types of answers a student could provide and gave limited feedback to the students. Although WeBWorK automatically grades, it cannot view a student's work and give partial credit like a professor could. WeBWorK can show students worked out solutions to problems; however the task of setting this up can be time consuming.

Like Professor Wu, Professor Petruccelli also gives a mix of conceptual and calculation problems. Each homework assignment covers a major, relevant topic in the class. Professor Petruccelli feels good problems reinforce the ideas covered in class and require the student to interpret findings.

4.1.3 Interview with Professor Gagnon

Our IQP group interviewed Professor Gagnon on September 23, 2011 in his office. Professor Gagnon taught Statistics I and II in the 2010-2011 academic year.

Professor Gagnon uses a similar approach to helping students as Professors Wu and Petruccelli. If a student has a specific homework question, Professor Gagnon will ask for the student's work and tries to understand what the student is thinking. For the better students, he will point out mistakes or errors in the work. Professor Gagnon will work through the problem step by step with the students who are having a lot more difficulty. He also has the TAs work out problems similar to those in the homework during conferences/labs. When a student doesn't understand a general concept, Professor Gagnon will try to explain the concept using different words than in the notes, and will go through a couple of examples with the student.

Professor Gagnon did not use WeBWorK in his statistics classes. He did, however, use online homework to an extent. He assigned both daily and weekly homework assignments. Students answered a quick daily problem through Blackboard, another online homework platform. The weekly homework assignments were longer in length and students completed these on paper.

Although he does not use WeBWorK, Professor Gagnon was able to offer insight on online homework in general. He mentioned how students enjoyed multiple tries and instant feedback. In addition, Professor Gagnon found it useful that Blackboard automatically graded homework. Like WeBWorK, Blackboard also records some information for instructors, such as how many times a student attempted a problem. Professor Gagnon pointed out that online homework allows professors to easy implement visualization, such as graphs, into a problem. Worked out sample problems can be easily and readily available to students through website links. Professor Gagnon did note how online homework systems cannot give partial credit because the system cannot view a student's work.

Like the other professors, Professor Gagnon incorporates both conceptual and calculation problems into his homework assignments. Most of the homework questions he asks students, the students must solve by hand. He also creates problems for students to specifically solve in the computer program MATLAB.

4.1.4 Interview with Professor Tashjian

On September 27, 2011, our IQP group interviewed Professor Tashjian in her office. Professor Tashjian taught Statistics I in the 2006-2007 and the 2009-2010 academic years.

When students have trouble with a specific homework problem, Professor Tashjian, like the other mathematics professors, tries to diagnose the source of the student's difficulty. She will go through the problem step by step with the student or work out an example similar to that of the homework question. When a student doesn't understand a general topic, Professor Tashjian will go over the background leading up to the topic. She will discuss relevant concepts and repeats the most important concepts several times.

Professor Tashjian does not use WeBWorK or any online homework system in her classes. She stated she would have trouble picking out good problems from a large bank of questions. She creates her own homework problems while writing her lecture notes, so she knows they are closely related to the topics presented in class. Professor Tashjian feels not being able to view a student's work is also a downside to using online homework. She usually looks for certain steps when correcting problems and not just for a final answer. If online problems implemented a step-by-step set-up, Professor Tashjian said she might be more inclined to use online homework.

When creating homework problems for her class, Professor Tashjian makes sure they tie in with the topics she presented in lecture. She told us that good problems reinforce the main concepts and give students a firm idea of the definition of a topic. She tries to incorporate a combination of both calculation and conceptual questions if possible.

4.1.5 ASSISTments and Interview with Professor Heffernan

Our group interviewed Professor Heffernan in his office on September 15, 2011. Professor Heffernan helped created the web-based homework platform ASSISTments. Individual ASSISTments are questions composed of several parts, hints, and various tutoring methods (Heffernan, C., 2011).

One of the most common tutoring methods used within the ASSISTments questions is scaffolding. Professor Heffernan stated in our interview that scaffolding helps most with the students who really struggle with a certain topic. Scaffolding breaks down a question into several parts and provides hints to help students through the problem. Through scaffolding, a student can learn the process of how to go about solving a certain problem, rather than just memorizing the steps of a specific example done in lecture. Professor Heffernan also stated, however, that scaffolding is less effective for students more competent in a certain topic. The competent students tend to learn better if they are told their answer is incorrect and are given the correct answer. These students have the ability to go through their work and figure out exactly where they went wrong.

4.1.6 Interview Results and Conclusions

These interviews gave us insight into how professors at WPI tutor students in statistics. Generally, the professor will try to determine the student's point of misunderstanding by working through the problem with him or her. Then, the professor can provide hints or work through that step of the problem with the student to help him or her complete the problem correctly. Through our research, we identified scaffolding as a good representation of this process because it lets students complete a problem without interruption unless they ask for help.

4.2 Improved WeBWorK Problems

In order for WeBWorK to mimic the tutoring that a teacher can provide, we needed to add several functions to our WeBWorK problems. Using the Perl package created by the IQP team of Li and Xia as a guide, we created a new package which allowed a student who got a problem wrong to enter a "help mode." When a student enters this mode, he or she is walked through the steps of the problem in order to determine areas of weakness, and hints are provided to help the student solve the problem.

Since the research on improving WeBWorK problems is ongoing, we wanted to give clear documentation of our package to future IQP groups. That way, they will be able to easily understand how our package works, and modify it for their own projects. Also, since professors expressed interest in adding hints and help questions to their own WeBWorK assignments, we have provided an outline of how to write a problem using our package that implements scaffolding.

Both the package code and problem outline are available as attachments to this report on the WPI Gordon Library website.

4.3 Data Analysis

Following the completion of the testing phase, our group collected as much data as possible. We collected each student's pre- and post-test scores and made sure to note which test the student received as the pre-test and which test as the post-test. We also calculated the difference between the pre- and post-test scores as a way to measure the student's improvement. In addition, our group collected the overall WeBWorK scores and recorded how many times the students with scaffolding clicked the "Click for Help" button, serves as a measure of how much help the student actually used. We also had available all data that WeBWorK recorded, such as the number of attempts at a problem or the answers submitted.

After collecting the data, we had to remove certain students from the data before beginning our analysis. Some students did not complete all three steps of the testing, while others had computer program errors preventing them from completing all the testing. By removing these students, we hoped to remove some of the bias in our data. In some cases, we also chose to remove students who answered the WeBWorK problems correctly on the first try. These students already had a good understanding of the material and did not need nor use scaffolding; therefore, their scores wouldn't be appropriate when analyzing the effectiveness of scaffolding.

Since we knew that our sample size was small, there was the possibility that our statistical tests would have low power. We calculated the power for several of our statistical tests to determine the likelihood that these tests would detect a mean score improvement of specified size. We have included this information where relevant.

For each week of testing, our IQP group created two paper tests that would serve as the pre- and post-tests. As said before, half of the class received test 1 as the pre-test, while the other half received test 2 as the pre-test. One of the first things we did was compare test 1 with test 2. If one of the paper tests had a higher mean score, then it would have introduced unnecessary bias into our results. We would not know if scaffolding truly helped the student, or if the student's improvement was a direct result of which pre-test the student received. Our group ran t-tests, assuming unequal variances, to determine if the mean pre-test scores differed significantly between the students with test 1 and the students with test 2. As shown in the figures below, there was no significant difference between the two paper tests for both One-Way Model questions ($p=0.79$) and Distribution Free Model questions ($p=0.84$). However, if there was a 10% difference between the two means for the One-Way testing, our statistical test only would have a power of 0.1415. This means there is about a 14% chance that our statistical test would detect a 0.1 difference in means. Similarly, if there was a 10% difference between the means for the Distribution Free testing, our test would have a power of .1626.

4.3.1 Figure: Comparison of Test 1 vs. Test 2 for One-Way Models

t-Test: Two-Sample Assuming Unequal Variances

	OW Test 1	*OW Test 2*
Mean	0.390625	0.422222222
Variance	0.106622869	0.075343915
Observations	12	15
Hypothesized Mean Difference	0	
df	22	
t Stat	-0.267925444	
P(T<=t) one-tail	0.395626362	
t Critical one-tail	1.717144335	
P(T<=t) two-tail	0.791252724	
t Critical two-tail	2.073873058	

These t-tests fail to show that there was a significant difference in the mean score between Test 1 and Test 2 for One Way Models.

4.3.2 Figure: Comparison of Test 1 vs. Test 2 for Distribution-Free Models

t-Test: Two-Sample Assuming Unequal Variances

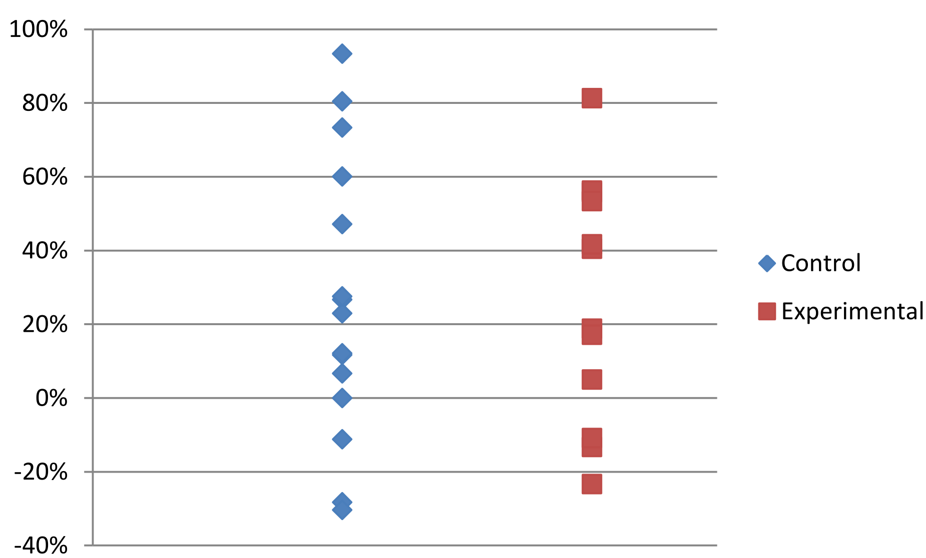
	DF Test 1	*DF Test 2*
Mean	0.690476	0.714285714
Variance	0.114409	0.040816327
Observations	12	11
Hypothesized Mean Difference	0	
df	18	
t Stat	-0.20689	
P(T<=t) one-tail	0.419211	
t Critical one-tail	1.734064	
P(T<=t) two-tail	0.838422	
t Critical two-tail	2.100922	

These t-tests fail to show that there was a significant difference in the mean score between Test 1 and Test 2 for Distribution Free Models.

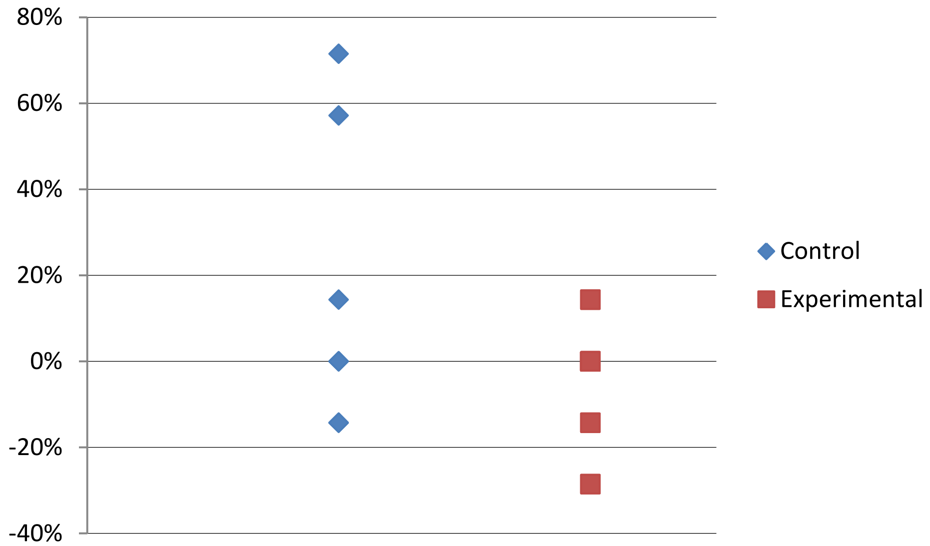
We had heard that some students in the 8 AM lab section struggled to complete the testing. Because of this, we wanted to determine if there was a difference between the mean pre-test score of the 8 AM section compared to those who tested at 9 AM. The pre-test scores serve as each student's knowledge level going into the WeBWorK assignment. We compared the mean pre-test scores of the 8 AM section and the 9 AM section using a two-sided t-test. We included all students in each section; because separating them by which pre-test they had would have made the sample size would have been too small. We found that the 9 AM performed significantly better than the 8 AM group. For the first week of testing, our p-value for the t-test was 0.076 (marginally significant) and the second week was 0.031 (significant at the .05 level). We assumed unequal variances once again. Although the true source of difference in means is unknown, we know it was possible that the 9 AM section was better prepared going into the WeBWorK problems.

One of the main goals of our project was to see if scaffolding helped students learn statistics more effectively. We decided to examine the improvement of the students who used scaffolding compared with those who did not use scaffolding. We used the post-test score minus the pre-test score for each student as a measure of improved understanding. Figures 4.3.3 and 4.3.4 show student score improvement for each week of testing. T-tests comparing the mean improvement of the scaffolding group and the control group show that there is no significant difference between the two. This applies for both weeks of testing.

4.3.3 Figure: Comparison of Improvement- One Way

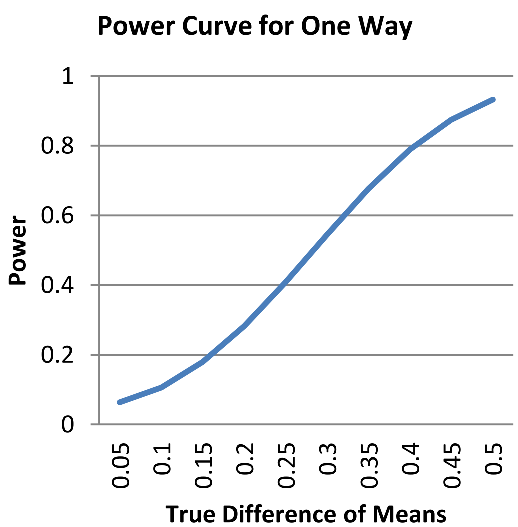


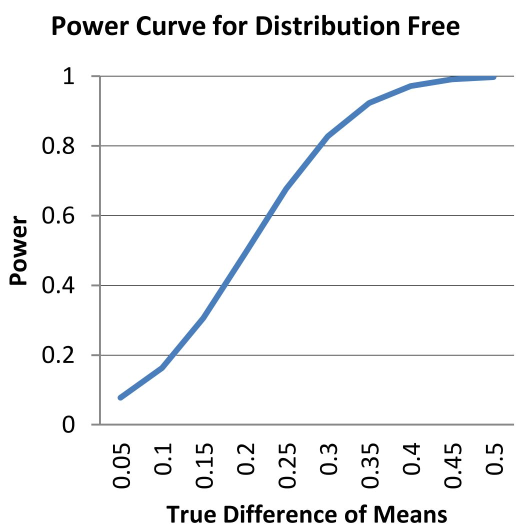
4.3.4 Figure: Comparison of Improvement- Distribution Free



Once again, the powers of the t-tests are low due to the small sample sizes. If there was a 10% difference between the control group and experimental group for the One-Way Models, our t-test only would have a power of 0.1063. The power curves in Figure 4.3.5 show that the chances of detecting a small difference between control and experimental means is low.

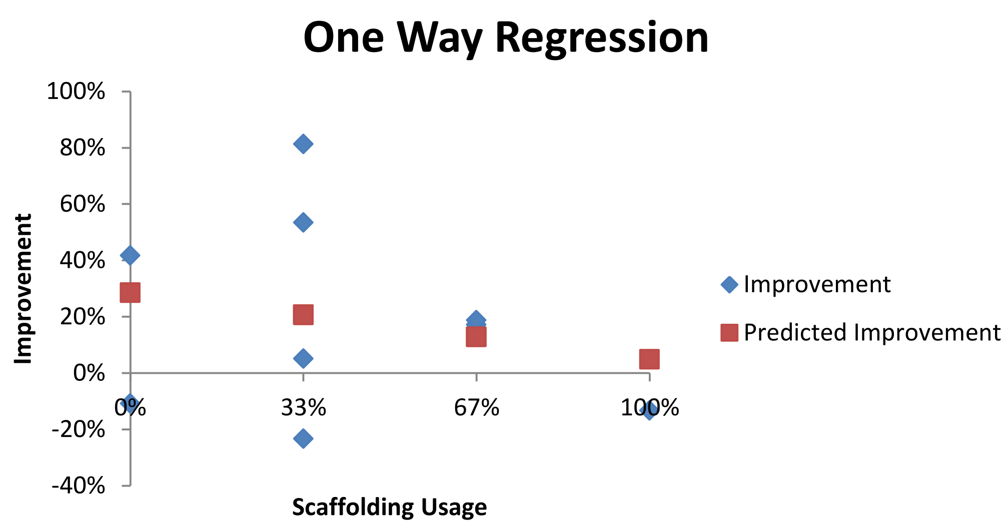
4.3.5 Figure: Power Curves for Improvement t-tests





To further examine if scaffolding had a positive impact on students, we looked at the amount of scaffolding the student used versus their improvement from pre-test to post-test. We used a regression line, with the amount of scaffolding the student used as the predictor variable and the student improvement as the response. The amount of scaffolding the student used was determined by how many problems they used the extra help on. In other words, if a student used help on 1 out of 4 problems, they used scaffolding 25%. For both weeks of testing, the regression showed there was no significant correlation between the amount of scaffolding used and improvement from pre-test to post-test. Figures 4.3.6 and 4.3.7 display the regressions.

4.3.6 Figure: Regression of % Used Scaffolding vs. Difference- One Way Models

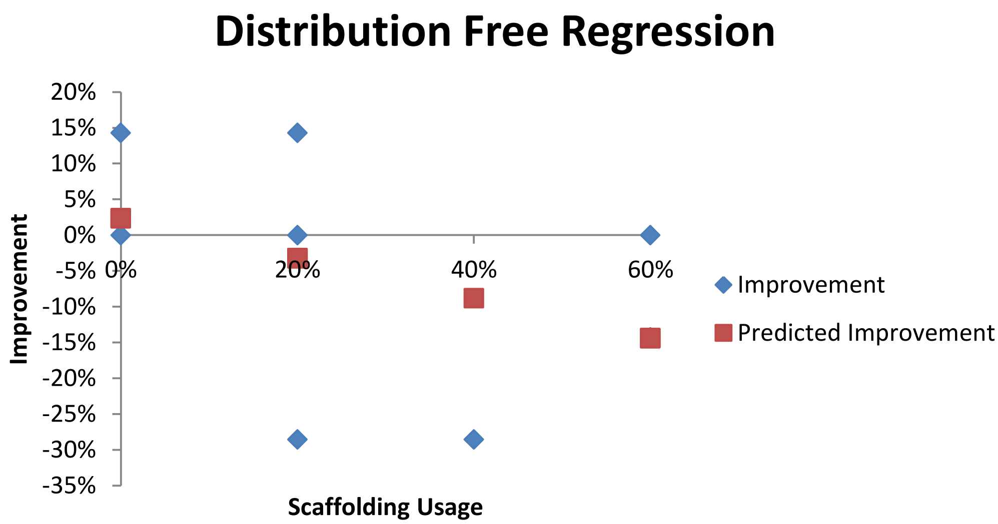


Regression Statistics	
Multiple R	0.221595082
Observations	9

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	0.046623207	0.046623207	0.36148096	0.566637065
Residual	7	0.902848243	0.12897832		
Total	8	0.949471451			

4.3.7 Figure: Regression of % Used Scaffolding vs. Difference- Distribution Free



Regression Statistics	
Multiple R	0.42744299

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	0.038115852	0.038115852	1.564865192	0.251142814
Residual	7	0.170500928	0.024357275		
Total	8	0.20861678			

Following the testing, our IQP group wanted to get feedback from the students about their opinions of scaffolding, among other things. We sent out a survey to the students who had completed both weeks of testing because these students would have seen both the control and experimental problems.

We asked the students a number of questions, but some responses were of particular interest. One question we asked the students was “Did you notice that there were hints and help questions available during one of the WeBWorK assignments?” Ten students out of the thirteen who responded said they did notice the questions. Then, when asked, “Did you use the hints?”, only four out of the thirteen used them. Six out of thirteen used the help questions. We thought it was interesting that even though the help was available, several students did not use it. We also found it interesting than some students didn't even know help was available, considering we increased the size of the "Click for Help" Button.

We also asked the students to rank several statements from 1 (for Disagree Strongly) to 5 (for Agree Strongly). Only 1 student agreed with the statement, "I got frustrated working through the problems," and ten out of thirteen agreed or strongly agreed with the statement, "It was easy to navigate the problem." In addition, half of the students agreed or strongly agreed with the statement, "I would like it if more classes used hints and help questions within WeBWorK problems."

4.4 Testing Issues

During the course of our testing, we encountered unexpected issues that might have affected our data. Less than half of the students registered for each section actually came to lab on the day of testing. This gave us about 25 data points for each week of testing. We would have liked to see more students participate in the testing, because it would have allowed our statistical tests to better represent the population, and because with so few data points, it was difficult to see if scaffolding was effective.

Another challenge was the scheduling of the lab period. The sections ran from 8 AM to 9 AM and 9 AM to 10 AM. We received feedback, both during the lab period and in the survey we sent out after testing, that students did not participate as actively in the testing as they might have in a lab later in the day. In our survey, one student commented, "If the lab were an hour later, I probably would have noticed the help section."

In some cases, students arrived late or came to a different section than the one they were registered for. While we did plan for students to arrive a little late, we didn't expect students to attend a different lab section than the one they were registered for. Because of this, we were pleased with our decision to use paper pre and post-tests, as well as a printout of instructions. This allowed us to adjust the schedule of the students who came late without having to change anything in WeBWorK, and give those who came to the wrong section the proper set of instructions. While we were able to modify our methodology to meet the needs of the students who arrived late, in many cases we did not use their scores in our data because they did not have time to complete all of the testing.

While we were pleased with the instructions we provided for each section, we found that some students may not have read them. In our survey, students told us that they didn't notice that there was help available. We sent the survey only to students who attended both weeks of testing, so we could be sure they had completed problems that implemented scaffolding. Also, while the instructions clearly stated that the work completed during testing did not count toward the student's course grade, some tried to keep the tests as long as possible because they thought it would be graded and needed more time.

Finally, during the WeBWorK portion of the testing, one student was unable to get the SAS software to work correctly and she could not complete the assignment. While she was able to get help from the professor, she did not get a fair chance to attempt all of the problems. We did not plan for this, and we were lucky that the professor was available to help her.

Chapter 5: Recommendations and Conclusions

5.1 Recommendations

Throughout our entire experiment, we encountered both avoidable and unavoidable obstacles that limited the results of our testing. If we were afforded the opportunity to perform this experiment again, we would make some changes to our procedure in order to avoid entirely or mitigate the effect that these obstacles had. We also have come up with suggestions for future projects and research on finding a valuable homework type.

The factors that limited our results the most were the number of students we were able to test (and therefore, the amount of data collected) and the time constraints placed on our testing. We tested on Statistics II, a class that is not as popular as Statistics I. There were about 50 registered students, and only half came to lab each week. The lab section was only 50 minutes long and attendance was not required in order to pass the class. We think that the students who skipped lab may have been the ones who could have benefitted from scaffolding the most, because they would also have been likely to skip lectures or homework. Finally, the testing period was at 8 AM and 9 AM on Wednesday, a day that students have fewer required classes.

The changes that we would make in an ideal situation include:

- Test during a term where more students are registered for Statistics II
- Find a way to have the testing run longer than one hour, or even give students unlimited time to complete the WeBWorK
- If the lab section is at an inconvenient time for students, schedule a different time for the testing
- Change how students are introduced to the new problem type so they will better understand how to use them (include a tutorial, show an example on a projector, etc.)

Future projects might look into giving more detailed feedback to students when they answer a question incorrectly, rather than just a correct or incorrect response. This feedback could reflect the answer the student has submitted for a question as well. For example, if a common mistake on a problem is typing in a +5, but the answer is -5, the feedback would suggest looking at the sign of the answer to make sure it is correct.

5.2 Conclusions

Our data did not show a statistically significant difference in mean score improvement between the students who used scaffolding and those who didn't. There was no significant correlation between the amount of scaffolding used and the student's improvement as well. While our experiment did not provide the results we were hoping for, we feel that other positive things have come out of our efforts.

We received encouraging feedback from students regarding their use of scaffolding. Nine out of twelve students agreed with the statement "I would like it if more classes used hints and help questions within WeBWorK problems." Even though our data doesn't show a difference in student learning through scaffolding, it could still improve a student's homework experience by reducing frustration.

The professors we interviewed told us that if it was simple to implement, they would consider adding hints, help questions and partial credit to WeBWorK problems. We have created an outline and other documentation in order to assist those professors who might want to use the scaffolding package. In the future, this might increase the number of professors who will add these features to WeBWorK.

We chose scaffolding because we felt that it most accurately mimicked the tutoring a professor provides his or her students. There may be modifications to scaffolding or an entirely different problem type that would improve student learning. It is also worth noting that statistics problems are very different from other math problems because they require detailed interpretation. Therefore, a tutoring method that has been shown to work for algebra, such as scaffolding, might be less effective in this context. We are encouraged by the fact that future groups can use our results to further the research of finding an ideal homework problem type for Statistics students.

While our experiment was not a successful in proving our hypothesis, we are proud of the many other things we accomplished in the course of this project. We created an original idea and ran a well-organized experiment to test that idea. We are also providing a useable Perl package and outline that is well documented for the use of professors and other project groups. This experience has taught us that proving a hypothesis is not necessarily as valuable as making a contribution to larger research.

Additionally, we learned that it's almost impossible to plan for everything, and that one must be able to adjust to any complications or adversities that arise. For example, we did not expect students to arrive to a lab section they were not registered for. However, our group was able to quickly adjust and made sure these students still completed the testing.

Large samples of data are important because mistakes and errors can occur. They are also important because the larger samples can detect differences that smaller samples can't. Our statistical tests had very low power. If there was a small difference between mean score improvement between the experimental and control groups, it would be very unlikely that our t-test would pick up on it.

Our IQP project has helped us learn how to understand a problem in the world and how to go about solving this problem. Among many things, we learned how to properly research information and how to define and carry out achievable goals. We sharpened our oral communication skills as well as our writing skills. We learned about the importance of collaboration with each other as well as our advisor.

With the completion of our project, we hope we have positively contributed towards the goal of finding a homework problem type that helps WPI students learn statistics efficiently.

Appendix

A. Interview Questions

A.1 Questions for Statistics Professors

1. If I was a student coming to you with a homework question, what tactics would you use to help me?
   Examples: Would you have me work through the problem as far as I could go, and then help me when I either got stuck or did something incorrectly? Or, would you work out the problem step by step with me? Or, would you work out an example similar to the homework question, and then have me tackle the problem?
2. If I was a student coming to you because I did not understand a topic in class (NOT a specific homework problem), how would you help me?
   Examples: Would you have me go through the notes again? Would you have me review previous topics I need in order to understand the new topic?
3. What makes a good homework problem? Why do you choose certain problems to assign over others?
   Examples: Examples of problems we are considering
4. Have you ever used WeBWorK or another online homework system for assignments? If so, why did you use them? In your opinion, what are some pros and cons of using online homework systems for assignments?
   If you have NOT used an online homework system, have you ever considering starting to use one? If not, why?

A.2 Questions for Professor Heffernan

1. Can you give a description of what ASSISTments is and how it's beneficial? How is it different from WeBWorK?
2. What types of homework problems/tutoring methods do you think are most effective? How did you determine the learning methods to be implemented in ASSISTments? (ex. questions with hints, step by step problems, etc.
3. We're trying to improve homework questions on WeBWorK for Statistics courses. Do you have any suggestions of how to make the problems effective?

B. Conversion from Current to Scaffolding Format

B.1 Current Format

(1 pt) **local/setiqpstats/one3.pg**
Non-denominational evergreen tree farmers are interested in the growth of different types of trees on certain tree farms. An experiment is run in which four different tree types are planted, Douglas Fir, Blue Spruce, White Pine and Balsam. One of each of these trees is planted on three different tree farms. Each tree is given the same care, and the increase in tree height (in cm) after one year is recorded in the table below.

	Douglas Fir	Blue Spruce	White Pine	Balsam
Farm 1	56	56	40	47
Farm 2	59	50	35	40
Farm 3	85	73	58	56

Estimate the model effects:

$\hat{\mu}$	$\hat{\tau}_1$	$\hat{\tau}_2$	$\hat{\tau}_3$	$\hat{\tau}_4$
Next				

$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$

Compute the ANOVA table (give the p-values to four decimal places):

Source	df	Sum of Squares	Mean Square	F	p-value
Trees					
Farms					
Error				× × × × ×	× × × × ×
Total			× × × × ×	× × × × ×	× × × × ×

Does the test reject the null hypothesis of equal tree growth at the 0.05 level?

- **A.** No
- **B.** Yes

Note: *You can earn partial credit on this problem.*

Preview Answers | Check Answers | Submit Answers

The student is presented with a problem description and data to work with. The student is shown steps needed to complete the problem all at once. The problem presents these steps in a suggested order, but the student does not need to complete the steps in the presented order. The student may submit answers for all or some parts of the problem. The student can submit answers for however many tries the professor allows. Once the allotted trials are exhausted, or the student answers all parts correctly, he or she will move on to the next problem.

B.2 Experimental Format (Scaffolding) - First Question

(1 pt) **local/setiqpstats/one3.pg**
Non-denominational evergreen tree farmers are interested in the growth of different types of trees on certain tree farms. An experiment is run in which four different tree types are planted, Douglas Fir, Blue Spruce, White Pine and Balsam. One of each of these trees is planted on three different tree farms. Each tree is given the same care, and the increase in tree height (in cm) after one year is recorded in the table below.

	Douglas Fir	Blue Spruce	White Pine	Balsam
Farm 1	56	56	40	47
Farm 2	59	50	35	40
Farm 3	85	73	58	56

Estimate the model effects:

$\hat{\mu}$	$\hat{\tau}_1$	$\hat{\tau}_2$	$\hat{\tau}_3$	$\hat{\tau}_4$
Next				

$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$

Compute the ANOVA table (give the p-values to four decimal places):

Source	df	Sum of Squares	Mean Square	F	p-value
Tree					
Farm					
Error				× × × × ×	× × × × ×
Total			× × × × ×	× × × × ×	× × × × ×

Does the test reject the null hypothesis of equal tree growth at the 0.1 level?

- **A.** Yes
- **B.** No

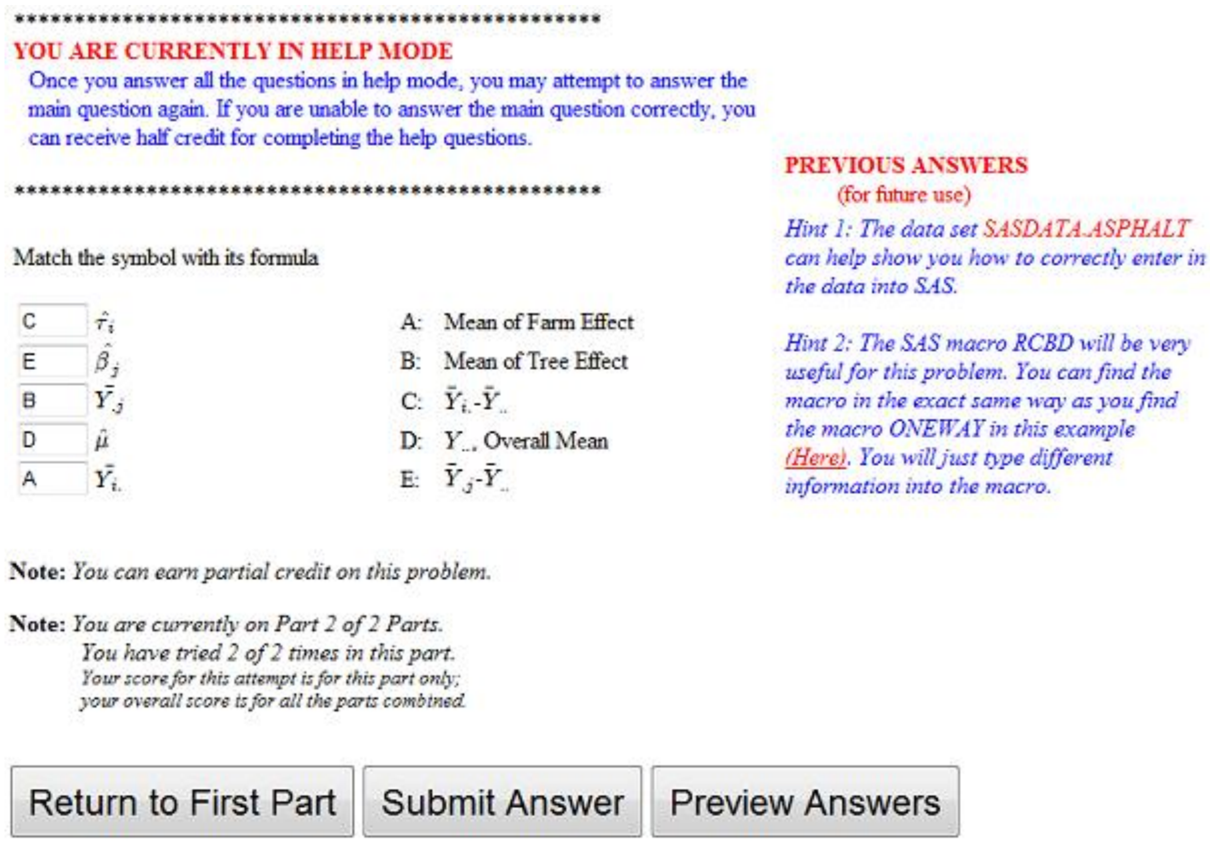
Note: *You can earn partial credit on this problem.*

Note: *You are currently on Part 1 of 2 Parts.*
You have tried 1 of 2 times in this part.
Your score for this attempt is for this part only;
your overall score is for all the parts combined.

[Click for Help] [Submit Answer] [Preview Answers]

The student begins with a problem description that from "Note: You can earn partial credit on this problem." up is identical to a current WeBWorK problem. However, this problem contains larger "Submit Answer" and "Preview Answers" buttons, as well as a note that tells the student which part of the problem he or she is on. The different parts of the problem are the main problem (shown above) and the different questions asked in the scaffolding. This problem also shows how many tries the student has remaining. Like a normal WeBWorK problem, the student *can answer the different sections to the main question in any order and can submit answers for all or some sections. If the student does not answer the entire problem correctly on the first try, a "Click for Help" button appears. If the student does answer the problem correctly on the first try, the "Click for Help" button does not appear, and the student can move onto the next problem.*

B.3 Experimental Format (Scaffolding) - Help Mode



When the student clicks the "Click for Help" button, he or she enters a "Help Mode." The "Help Mode" contains hints, help questions, references to outside materials, etc. This "Help Mode" aims to mimic how a tutor would help a student solve the main problem. Whenever a student answers a help question, the question and answer will appear in the WeBWorK window under "Previous Answers." Hints and other materials also appear under "Previous Answers." The "Previous Answers" column retains all the help provided to the student and allows the student to use this help when attempting the main problem again. If the student was unable to answer the scaffolding questions correctly, the help will still appear until "Previous Answers," however the student will not be able to receive partial credit for answering the scaffolding (if the professor chooses to give partial credit for answering scaffolding.) The "Return to First Part" button returns the student to the original question.

B.4 Figure: Experimental Format (Scaffolding) - Return to Original Question

(1 pt) **local/setiqpstats/one3.pg**

Non-denominational evergreen tree farmers are interested in the growth of different types of trees on certain tree farms. An experiment is run in which four different tree types are planted, Douglas Fir, Blue Spruce, White Pine and Balsam. One of each of these trees is planted on three different tree farms. Each tree is given the same care, and the increase in tree height (in cm) after one year is recorded in the table below.

	Douglas Fir	Blue Spruce	White Pine	Balsam
Farm 1	56	56	40	47
Farm 2	59	50	35	40
Farm 3	85	73	58	56

PREVIOUS ANSWERS

(for future use)

Hint 1: The data set SASDATA.ASPHALT can help show you how to correctly enter in the data into SAS.

Hint 2: The SAS macro RCBD will be very useful for this problem. You can find the macro in the exact same way as you find the macro ONEWAY in this example (Here). You will just type different information into the macro.

Match the symbol with its formula

$\hat{\mu}$: $\bar{Y}_{..}$, **Overall Mean**

$\hat{\tau}_i$: $\bar{Y}_{i.}-\bar{Y}_{..}$

$\hat{\beta}_j$: $\bar{Y}_{.j}-\bar{Y}_{..}$

$\bar{Y}_{i.}$: **Mean of Farm Effect**

$\bar{Y}_{.j}$: **Mean of Tree Effect**

Hint 3: Click me

Estimate the model effects:

$\hat{\mu}$	$\hat{\tau}_1$	$\hat{\tau}_2$	$\hat{\tau}_3$	$\hat{\tau}_4$
Next				

$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$

Compute the ANOVA table (give the p-values to four decimal places):

Source	df	Sum of Squares	Mean Square	F	p-value
Tree					
Farm					
Error				× × × × ×	× × × × ×
Total			× × × × ×	× × × × ×	× × × × ×

Does the test reject the null hypothesis of equal tree growth at the 0.1 level?

- **A.** Yes
- **B.** No

Note: *You can earn partial credit on this problem.*

Note: *You are currently on Part 1 of 2 Parts.*
You have tried 0 of 2 times in this part.
Your score for this attempt is for this part only;
your overall score is for all the parts combined.

Submit Answer | Preview Answers

After clicking "Return to First Part," the student can attempt to answer the original question again. All of the help provided in the "Help Mode" is available under "Previous Answers" for the student to use.

C. Testing Instructions

C.1 Regular (Control)

Instructions

We are testing the effectiveness of different WeBWorK problems on improving student scores. You will complete a paper pre-test, do some WeBWorK problems and finish with a post-test. Thank you for helping us!

Step 1

Complete the pre-test and raise your hand so we can collect it from you. Please don't forget your name!

Step 2

A.

Go to webwork.wpi.edu and click into the class called MA2612IQP. Sign in and you will have an assignment called One Way 1 that will open at 10 minutes past the hour.

B.

The problems work exactly like the ones you're used to. Just try your best- use your notes, textbook, SAS, anything you would usually use for homework.

C.

If you finish the WeBWorK problems before we hand out the post-test, raise your hand so we can give you the test early.

Step 3

Complete the post test and raise your hand so we can collect it from you. Please don't forget your name!

THANK YOU!

C.2 Scaffolding (Experimental)

Instructions

We are testing the effectiveness of different WeBWorK problems on improving student scores. You will complete a paper pre-test, do some WeBWorK problems and finish with a post-test. Thank you for helping us!

Step 1

You will have 10 minutes to complete the pre-test. Please don't forget your name!

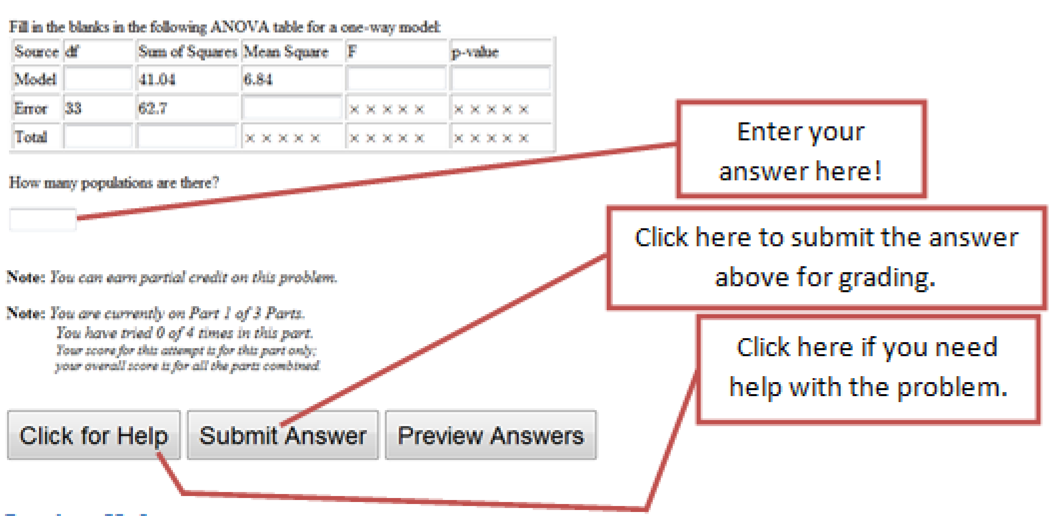
Step 2

Go to webwork.wpi.edu and click into the class called MA2612IQP. Sign in and you will have an assignment called One Way 2 that will open at 10 minutes past the hour and will be open for 30 minutes.

These problems work differently than the ones you're used to. Please read the steps below and feel free to ask any questions.

Starting a Problem

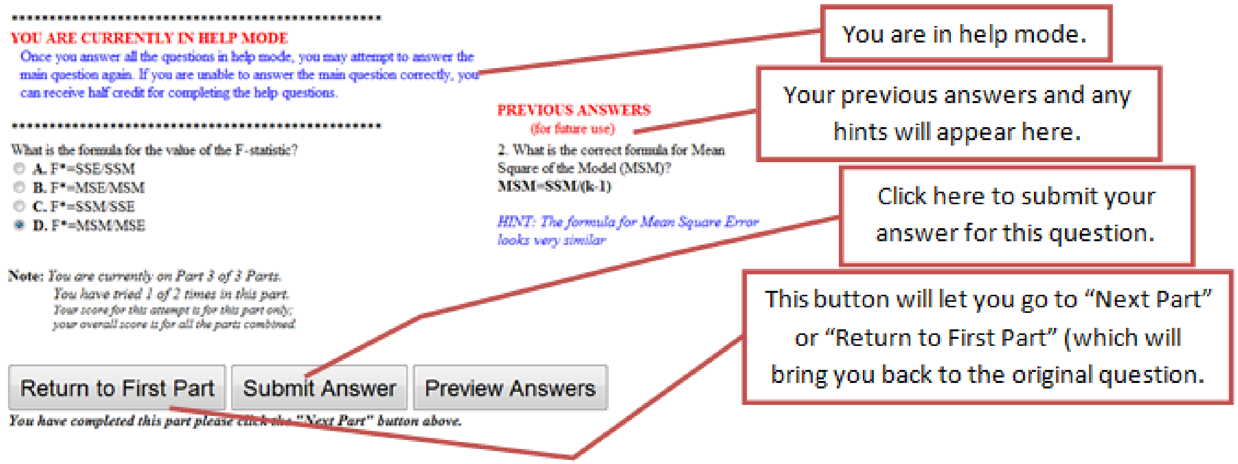
You will have two chances to answer the problem on your own before you will receive a score of 0 for the problem. At any time before you use both attempts, you can get help solving the problem



Getting Help

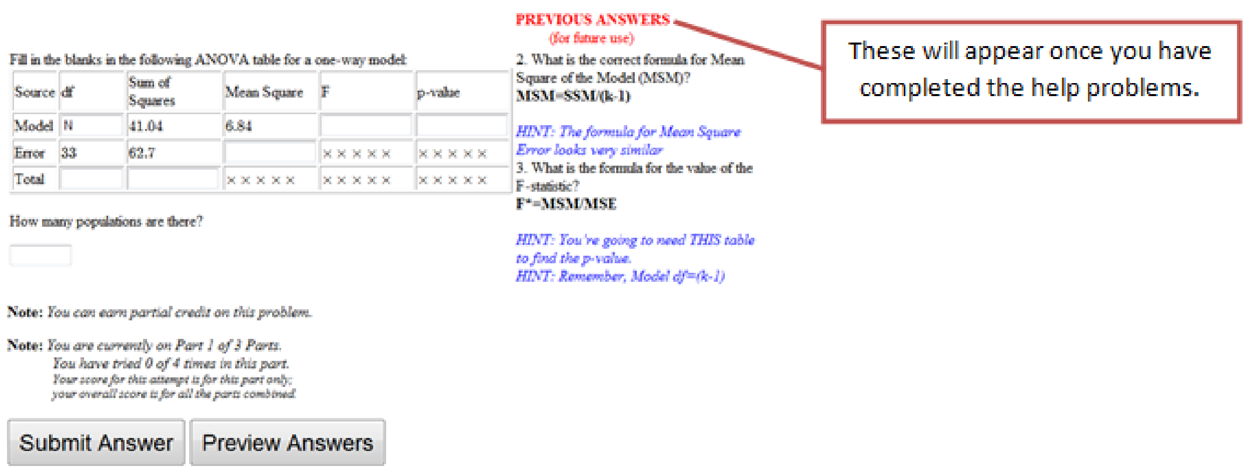
When you "Click for Help," you will go through a set of problems and hints meant to help you understand and solve the original problem. If you get them right you will earn partial credit for the problem (up to 50%). The correct answers and any hints will appear under "Previous

Answers" and will remain there for the rest of the problem. You will need to click "Next Part" after each question, and the "Return to First Part" to exit help.



Finishing the Problem

Once you have attempted all of the help questions and have earned partial credit (check your grade at the bottom of the page) you will get two additional attempts at the original problem.



The assignment will close automatically 40 minutes into the lab period.

Step 3

Complete the post test and raise your hand so we can collect it from you. Please don't forget your name!

THANK YOU!

D. Pre- and Post- Tests

D.1 One-Way Models

Name: ____ Circle One: Pre-Test Post-Test

A company would like to determine whether there is a difference among the gas mileage of the same car produced at three different factories. The factories are located in Albuquerque, Boston and Cleveland. Six cars were tested from each factory, and resulting values are given in the table below.

Albuquerque	Boston	Cleveland
33.3	34.5	37.4
33.4	34.8	36.8
32.9	33.8	37.6
32.6	33.4	36.6
32.5	33.7	37.0
33.0	33.9	36.7

1. Estimate the model effects: (4 pts)

$\hat{\mu}$	$\hat{\tau}_a$	$\hat{\tau}_b$	$\hat{\tau}_c$

2. Compute the ANOVA table (give the p-value to four decimal places): (5 pts)

Source	DF	Sum of Squares	Mean Square	F	P-Value
Model					
Error				xxxxx	xxxxx
Total			xxxxx	xxxxx	xxxxx

3. Does the test reject the null hypothesis of equal gas mileage for all factories at the 0.05 level? (1 pts)

Name: ____ Circle One: Pre-Test Post-Test

An experiment is conducted to determine which of four bacteria nutrient brands is the best. Five petri dishes are prepared with each type of nutrient, and the same amount of bacteria is added to each. The dishes are kept in the same environment for 24 hours, and then the colonies on each are counted. The resulting values are given in the table below.

Brand A	Brand B	Brand C	Brand D
24	23	22	32
21	30	26	22
24	28	26	32
31	28	40	35
30	30	23	20

1. Estimate the model effects: (4 pts)

$\hat{\mu}$	$\hat{\tau}_a$	$\hat{\tau}_b$	$\hat{\tau}_c$	$\hat{\tau}_d$

2. Compute the ANOVA table (give the p-value to four decimal places): (5 pts)

Source	DF	Sum of Squares	Mean Square	F	P-Value
Model					
Error				xxxxx	xxxxx
Total			xxxxx	xxxxx	xxxxx

3. Does the test reject the null hypothesis of equal nutrient effectiveness for all brands at the 0.1 level? (1 pts)

D.2 Distribution Free Models

Name: ____ Circle One: Pre-Test Post-Test

Part 1

It is known that the weights of newborn male moose are symmetrically distributed. Environmental protection officials in the New Hampshire White Mountains took a random sample of ten newborn male moose weights. They want to obtain a point estimate of the population median weight and a confidence interval of level as close to 0.95 as possible.

1. What is the data type? (Circle One)

   Single population, Arbitrary shape

   Single population, Symmetric

   Two populations, Same shape

   Multiple populations, Same shape

   X-Y data

2. What will the inference be about? (Circle One)

   Median

   Location Shift

   Population Effects

   Monotone Association

3. What test should they use? (Circle One)

   Sign test

   Wilcoxon Signed Rank test

   Wilcoxon Rank Sum test

   Kruskal-Wallis test
   Spearman correlation test

TURN OVER for Part 2 →

Part 2

The distribution of incomes is heavily right skewed. The administration of a school district wishes to estimate the median income of its teachers and has decided to use the Sign test. They also want to compute a 95% confidence interval based on the Sign Test. The administration took a random sample of 15 teachers in the district. The results where:

51000	45000	50000	46000	65000	55000	52000	60000
48000	49000	55000	49000	46000	48000	49000	

1. Obtain a point estimate of the median income.

The point estimate is: _______________

2. The formula for finding a 95% confidence interval is:

$$(Y_{(n-k+1)}, Y_{(k)})$$

What is the value of k for this problem?

The value of k is:_______________

Name: ____ Circle One: Pre-Test Post-Test

Part 1

It is known that the distribution of male college students' heights is left skewed. A group of WPI students wish to estimate the median height of male students and obtain a confidence interval as close to the 0.90 level as possible.

1. What is the data type? (Circle One)
   - Single population, Arbitrary shape
   - Single population, Symmetric
   - Two populations, Same shape
   - Multiple populations, Same shape
   - X-Y data

2. What will the inference be about? (Circle One)
   - Median
   - Location Shift
   - Population Effects
   - Monotone Association

3. What test should they use? (Circle One)
   - Sign test
   - Wilcoxon Signed Rank test
   - Wilcoxon Rank Sum test
   - Kruskal-Wallis test
   - Spearman correlation test

TURN OVER for Part 2 →

Part 2

A researcher wants to compare the weights of newborn male moose and newborn female moose in the White Mountains. The researcher has decided to use the Wilcoxon Rank Sum test to check whether or not the median weights differ. The researcher collected weights for 5 newborn male moose and 5 newborn female moose. The researcher then calculated the **differences between male weight and female weight.** The differences are listed below.

100	150	200	100	150
50	100	150	50	100
75	125	175	75	125
0	50	100	0	50
150	200	250	6	200

1. Obtain a point estimate of the difference in median weight between newborn male moose and newborn female moose.

The point estimate is:_______________

2. Obtain the level of confidence as close to the 0.95 level as possible.

The level of confidence is: _______________

E. Survey

1. Which section of MA2612 were you registered for?

	Response Percent	Response Count
B01 (8AM Lab)	46.2%	6
B02 (9AM Lab)	**53.8%**	**7**

2. Did you notice that there were hints and help questions available during one of the WeBWork assignments?

	Response Percent	Response Count
Yes	**76.9%**	**10**
No	23.1%	3

3. Did you use the hints?

Yes	30.8%	4
No	**69.2%**	**9**

4. Did you use the help questions?

Yes	46.2%	6
No	**53.8%**	**7**

5. Please rank the following statements from 1 (Disagree Strongly) to 5 (Agree Strongly)							
	1	**2**	**3**	**4**	**5**	**N/A**	**Response Count**
When I used the hints, I thought they were helpful.	0.0% (0)	0.0% (0)	8.3% (1)	33.3% (4)	8.3% (1)	**50.0% (6)**	12
It was easy to navigate the problem.	0.0% (0)	8.3% (1)	0.0% (0)	**41.7% (5)**	**41.7% (5)**	8.3% (1)	12
The help questions were useful.	0.0% (0)	0.0% (0)	8.3% (1)	25.0% (3)	16.7% (2)	**50.0% (6)**	12
I got frustrated working through the problems.	16.7% (2)	**33.3% (4)**	**33.3% (4)**	8.3% (1)	0.0% (0)	8.3% (1)	12
I thought problems with help took too long.	0.0% (0)	16.7% (2)	33.3% (4)	8.3% (1)	0.0% (0)	**41.7% (5)**	12
I understood the material better after using the hints and help questions.	0.0% (0)	8.3% (1)	16.7% (2)	25.0% (3)	0.0% (0)	**50.0% (6)**	12
I would like it if more classes used hints and help questions within WeBWorK.	0.0% (0)	0.0% (0)	16.7% (2)	33.3% (4)	**41.7% (5)**	8.3% (1)	12

6. Please let us know here if you have any additional comments or suggestions:

If the lab were an hour later, I probably would have noticed the help section

Annotated Bibliography

Ahlgren, A., & Garfield, J. (1988). Difficulties in Learning Basic Concepts in Probability and Statistics: Implications for Research. *Journal for Research in Mathematics Education, 19*(1), pp. 44-63. Retrieved from http://www.jstor.org/stable/749110.
This paper examines how pre-college and college students have difficulty understanding/learning statistics due to misconceptions, etc. It speaks about research related to this idea, and suggests other areas for research.

American Federation of Teachers. (2010). Assigning Effective Homework. Classroom Tips. *American Federation of Teachers.*
A brochure outlining the American Federation of Teacher's research and position on what makes an effective homework assignment.

Atkinson, R. K., Renkl, A., & Margaret Merrill, M. (2003). Transitioning from Studying Examples to Solving problems: Effects of Self-Explanation Prompts and Fading Worked-Out Steps. *Journal of Educational Psychology, 95*(4), 774-783. doi:10.1037/0022-0663.95.4.774.
The authors combined fading with the introduction of prompts designed to encourage learners to identify the underlying principle illustrated in each worked-out solution step. This combination produced medium to large effects on learning without requiring additional time on task.

Ben-Zvi, D., & Garfield, J. (2008). Introducing the Emerging Discipline of Statistics Education. *School Science and Mathematics, 108,* 355. Retrieved from http://go.galegroup.com/ps/i.do?id=GALE%7CA191854789&v=2.1&u=mlin_c_worpoly&it=r&p=AONE&sw=w.
The article discusses the importance of learning statistics and the challenges students have when learning the subject. The article talks about the differences between mathematics and statistics, and how statistics is a "bridge between mathematics and science".

Ben-Zvi, D., & Garfield, J. (2007). How Students Learn Statistics Revisited: A Current Review of Research on Teaching and Learning Statistics. *International Statistical Review, 75*(3), 372-396. doi:10.1111/j.1751-5823.2007.00029.x.
The article discusses research on how statistics is taught and how students learn the subject. It mentions difficulties/misconceptions students have, how teachers can fix these problems, and suggests "principles" to help students learn statistics more effectively.

Bressoud, D. M. (2009). Launchings: WeBWorK. Retrieved September 7, 2011, Retrieved from http://www.maa.org/columns/launchings/launchings_04_09.html
Discusses WeBWorK, students perceptions of online homework, and provides evidence of how it is effective.

Brewer, S. (2009). Using Online Homework in Traditional College Math Classes or How to Grade 45,000 Homework Problems and Still Keep Smiling. Retrieved August 31, 2011, from Teaching with Technology Idea Exchange: http://ttix.org/archives/2009-

sessions/using-online-homework-in-traditional-college-math-classes-or-how-to-grade-45000-homework-problems-and-still-keep-smiling/
The purpose of this presentation is to present research-based results regarding the effectiveness of using online homework as part of a college math class.

Broers, N. (2009). Using Propositions for the Assessment of Structural Knowledge. *Journal of Statistics Education, 17*(2).
In this article they discuss a new method for assessing a student's structural knowledge and compare it with other traditional methods. They say that in order to fully understand Statistics, students need to have a conceptual understanding of the underlying theory.

Bull, K. S., Shuler, P., Overton, R., Kimball, S., Boykin, C., & Griffin, J. (1999). Processes for Developing Scaffolding in a Computer Mediated Learning Environment. Retrieved from http://www.eric.ed.gov/ERICWebPortal/contentdelivery/servlet/ERICServlet?accno=ED429765; http://www.eric.ed.gov/ERICWebPortal/detail?accno=ED429765.
This article discusses scaffolding and the use of online scaffolding. It provides pointers on appropriate uses of scaffolding and discusses how scaffolding can help students learn.

Burnham, N. (2005). How Online Homework Reduces the Grading Load and Improves Instructor Feedback. Retrieved August 31, 2011, from WPI: http://www.wpi.edu/Academics/ATC/Collaboratory/Stories/burnham.html
Discusses how online homework reduces the amount of grading for professors and improves the feedback professors receive.

Catrambone, R. (1998). The Subgoal Learning Model: Creating Better Examples So That Students Can Solve Novel Problems. *Journal of Experimental Psychology: General, 127*(4), 355-376. doi:10.1037/0096-3445.127.4.355.
Subgoals guide problem solving by helping learners focus on the steps to modify in novel problems that involve the same subgoals but require new steps to achieve them.

Chi, M. (1996). Constructing Self-Explanations and Scaffolded Explanations in Tutoring. *Applied Cognitive Psychology, 10*, S33-S49.
One-on-one tutoring is a form of instruction that requires interaction between a tutor and a tutee. The effectiveness of tutoring is examined from the perspectives of the tutor's actions, the tutee's actions, and successive interactions. A tutor's actions that may not lead to successive interactions consist of asking an initiating question, providing feedback, and asking a comprehension-gauging question. It is suggested that these types of actions can lead to the learning of an ideal template of solution procedures for solving problems, but may not lead to deep understanding.

Cooper, H. (2008). *A Brief History of Homework in the United States. Research Brief* National Council of Teachers of Mathematics. 1906 Association Drive, Reston, VA 20191-1502. Tel: 800-235-7566; Tel: 703-620-3702; Fax: 703-476-2970; e-mail: orders@nctm.org; Web site: http://www.nctm.org/publications/. Retrieved from http://www.eric.ed.gov/ERICWebPortal/detail?accno=ED505965
This brief presents a history of homework, in which media-fueled outcries for more or less

homework occur cyclically, about fifteen to twenty years apart. It describes homework practices today and the beneficial and negative effects it can have beyond achievement on young children.

Cooper, H. (2008). *Effective Homework Assignments. Research Brief* National Council of Teachers of Mathematics. Retrieved from: http://www.eric.ed.gov/ERICWebPortal/detail?accno=ED505964
Discusses effective homework, interspersing easy and hard problems throughout an assignment, individualization of assignments, etc.

Emathzone. (2008). *Importance of Statistics in Different Fields.* Retrieved September 11, 2011, from http://www.emathzone.com/tutorials/basic-statistics/importance-of-statistics-in-different-fields.html
Discusses important fields which commonly use statistics.

Gagnon, J. (2011, September 23). Visiting Assistant Professor of Mathematics, Worcester Polytechnic Institute. Interview

Gal, I., & Garfield, J. (1999). Assessment and Statistics Education: Current Challenges and Directions. *International Statistical Review / Revue Internationale De Statistique, 67*(1), pp. 1-12. Retrieved from http://www.jstor.org/stable/1403562.
The article lists the main goals of statistics classes and what skills students should take away from the class. It also discusses new ways to assess a student's understanding and competence in statistics.

Garfield, J. (1995). How Students Learn Statistics. *International Statistical Review / Revue Internationale De Statistique, 63*(1, A Collection of Papers on Statistical Education), pp. 25-34. Retrieved from http://www.jstor.org/stable/1403775.
The article discusses common mistakes people make when dealing with statistics and explains methods that help students understand concepts better. The author also suggests certain "principles" to help students learn statistics more efficiently.

Goos, M. (2004). Learning Mathematics in a Classroom Community of Inquiry. *Journal for Research in Mathematics Education, 35*(4), pp. 258-291. Retrieved from http://www.jstor.org/stable/30034810.
Conclusions are drawn from observations about how different students learn when a teacher uses a method based on scaffolding and the zone of proximal development. Detailed explanations of the learning process are provided.

Heffernan, C. (2011). *ASSISTments: About.*, Retrieved September 6, 2011, Retrieved from http://teacherwiki.assistment.org/wiki/About.
Discusses the free web-based platform, ASSISTments.

Heffernan, N. (2011, September 15). Associate Professor of Computer Science, Worcester Polytechnic Institute. Interview.

Jonassen, D., Rebello, S., Wexler, C., Hrepic, Z., & Triplett, G. (2007). Learning to Solve Problems by Scaffolding Analogical Encoding. *American Society for Engineering Education*.
In order to teach students how to solve problems in most science and engineering courses, instructors demonstrate how to solve a problem and then ask learners to apply that method to a new transfer problem. Transferring learning from a single example to a new problem requires that learners induce a schema for that kind of problem from that single example and then apply that schema to a new, contextually varied problem. This single-example approach to teaching problem solving usually results in students attempting to mimic the process for solving the problem while ignoring the semantic, structural characteristics of the problem.

Knowledge@Wharton. (2008). *The Use -- And Misuse -- of Statistics: How and Why Numbers Are So Easily Manipulated.* Retrieved September 11, 2011, from http://knowledge.wharton.upenn.edu/article.cfm?articleid=1928
Mentions common uses of statistics in everyday life. Also mentions how statistics can be misused.

Kodippili, A., & Senaratne, D. (2008). Is Computer-Generated Interactive Mathematics Homework More Effective than Traditional Instructor-Graded Homework? *British Journal of Educational Technology, 39*(5), 928-932. doi:10.1111/j.1467-8535.2007.00794.x.
A study of 70 students which compared the academic performance of students assigned computer-generated homework and those assigned traditional homework. While the evidence was not conclusive, the higher number of passing grades in the group assigned computer homework is enough to call for further study.

Le, M., & Roungrojkarnranan, V. (2011). *Designing WeBWorK Problems For Applied Statistics II*. Retrieved from http://www.wpi.edu/Pubs/E-project/Available/E-project-042811-152813/unrestricted/final_report.pdf.
IQP project at WPI that created and tested new WeBWorK homework problems for Applied Statistics II.

learning style. (n.d.). *Dictionary.com's 21st Century Lexicon*. Retrieved October 3, 2011, from Dictionary.com website: http://dictionary.reference.com/browse/learning+style

Li, Z., & Xia, Y. (2011). *WeBWorK Problems For Statistics I* Retrieved from http://users.wpi.edu/~jdp/webwork_iqp/Final_Report_March_1.pdf
IQP project at WPI that accomplished the following:
Created a Perl package for WeBWorK that allowed multi-part problems with hints and more interaction; Developed a new instructional model called the Forward Fading Model (FFM); Used new package to implemented the FFM into Statistics I homework questions; Tested the effectiveness of the new model.

Lohr, S. (2009, August 9, For Today's Graduate, Just One Word: Statistics. *New York Times.* Retrieved from http://www.nytimes.com/2009/08/06/technology/06stats.html
Discusses the importance of statistics in many different fields.

McLoughlin, C. (1999). Scaffolding: Applications to Learning in Technology Supported Environments. Retrieved from http://www.eric.ed.gov/ERICWebPortal/contentdelivery/servlet/ERICServlet?accno=ED446740; http://www.eric.ed.gov/ERICWebPortal/detail?accno=ED446740.
This article discusses scaffolding, its common forms, and some history. It also calls for a reconsideration of the nature of scaffolding and for the alignment of theory with practice.

Merrill, M. M., Atkinson, R. K., & Catrambone, R. (2003). Aiding Transfer in Statistics: Examining the Use of Conceptually Oriented Equations and Elaborations During Subgoal Learning. *Journal of Educational Psychology, 95*(4), 762-773.
In this study, authors hypothesized that learners who studied statistics examples with conceptually oriented equations would transfer more successfully to novel problems compared with learners who studied examples using computationally oriented equations. The study discusses the authors' experiments to test this hypothesis, as well as their results.

Palocsay, S. W., & Stevens, S. P. (2008). A Study of the Effectiveness of Web-based Homework in Teaching Undergraduate Business Statistics. *Decision Sciences Journal of Innovative Education, 6*(2), 213-232. doi:10.1111/j.1540-4609.2008.00167.x.
This article compares traditional homework with three web-based homework systems for business statistics courses at the college level. The authors find no advantage to using online homework systems and discuss the systems' limitations.

Petruccelli, J.P. (2011, September 15). Professor of Mathematics, Worcester Polytechnic Institute. Interview.

Porter, T. S., & Riley, T. M. (1996). The Effectiveness of Computer Exercises in Introductory Statistics. *The Journal of Economic Education, 27*(4), pp. 291-299. Retrieved from http://www.jstor.org/stable/1183236.
One introductory statistics class was given homework involving exercises generated by computer software, and a second class was assigned problems from the textbook. The software provided an explanation if students' answers were incorrect, but those using textbook problems had no such feedback. The class using the software typically completed the assignments in less time and had higher scores on homework-related exam questions.

Schneiter, K. (2008). Two Applets for Teaching Hypothesis Testing. *Journal of Statistics Education, 16*(3).
This article introduces two applets used in class to demonstrate hypothesis testing. They have the ability to enhance statistics teaching with multiple representations of new concepts and allowing the students to experiment.

Section 3: Course Descriptions, Mathematical Sciences. (2011). *WPI undergraduate catalog 2011-2012* (pp. 155) Retrieved from http://www.wpi.edu/Images/CMS/Pubs-Catalogs-

Ugrad/FINAL_undergrad-cat_2011-2012.pdf
Gives class descriptions of Applied Statistics I and Applied Statistics II.

Selvanathan, S., & Cybinski, P. (2005). Learning experience and learning effectiveness in undergraduate statistics: Modeling performance in traditional and flexible learning environments. *Decision Sciences Journal of Innovative Education, 3*(2), 251-271.
This project investigates the relationship between student performance, past mathematics experience, and perceptions of statistics education for two groups of university students studying statistics in different learning environments.

Smith, M. (2005). *Careers Involving Probability and Statistics.* Retrieved September 11, 2011, from http://www.ma.utexas.edu/users/mks/statund/careers.html
Lists career areas that use statistics.

statistics. (n.d.). *Dictionary.com Unabridged*. Retrieved September 11, 2011, from Dictionary.com website: http://dictionary.reference.com/browse/statistics

Tashjian, G. (2011, September 27). Adjunct Assistant Professor of Mathematics, Worcester Polytechnic Institute. Interview.

Vance, E. (2010). Improving Self-efficacy in Statistics: Role of Self-Explanation and Feedback. *Journal of Statistics Education, 18*(3).
Novice problem solvers often fail to recognize structural similarities between problems they know and a new problem because they are more concerned with the surface features rather than the structural features of the problem. The surface features are the story line of the problem whereas the structural features involve the relationships between objects in the problem. We used an online technology to investigate whether students' self-explanations and reception of feedback influenced recognition of similarities between surface features and structural features of statistical problems.

Wallis, Claudia (09/04/2006). "The Myth about Homework". *Time international (Asia ed.) (1064-0304)*, (35), p. 44.
Discusses negative effects of homework.

Wu, Z. (2011, September 13). Assistant Professor of Mathematics, Worcester Polytechnic Institute. Interview.