Science and Mathematics Education Centre

Evaluation of an Instructional Unit Utilizing the Worst Case Method in Improving Students' Understanding of Uncertainty Analysis and Propagation of Error

Paul Alan Haugen

This thesis is presented for the Degree of

Doctor of Science Education

of

Curtin University

September 2010

Declaration

This thesis contains no material which has been accepted for the award of any other degree or diploma in any university. To the best of my knowledge and belief, this thesis contains no material previously published by any person except where due acknowledgement has been made.

Signature: ……………………………………………

Date: …June 21, 2010…………..

Abstract

The objective of most physics laboratory exercises is to investigate the validity of a physical law or theory. Students compare predictions, based on theoretical grounds, with experimental results and are often confronted with a discrepancy between these two. Instead of submitting an analysis or a conclusion that incorporates uncertainty analysis, students will often resort to a list of excuses to explain the difference, such as equipment malfunctions or human error. They fail to recognize that their results may support the theory, even without perfect correlation.

Physics teachers are challenged to provide instruction on uncertainty analysis rigorous enough to analyze laboratory data while, at the same time, understandable to entry-level students. This study focused on evaluating the effects of an algebra-based instruction unit on student understanding of uncertainty analysis and propagation of error. A comparison of scores on a pretest and posttest showed a statistically significant improvement in scores. In Phase Two of the study, student laboratory assignments were evaluated for changes in the level of understanding. Students demonstrated improved ability to incorporate uncertainty analysis and propagation of error in their laboratory reports, but most did not obtain an in-depth level of understanding. In a similar manner, conceptual change was evident at the lower level of assimilation, but few students achieved a complete conceptual change regarding uncertainty analysis.

Keywords: uncertainty analysis, propagation of error, conceptual change

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

CHAPTER 1

INTRODUCTION

One question that people frequently ask one another is, “What time is it?” In response, most people look at their watch and may say 12:05 without considering how accurate this is. Is it 12:05 and 10 seconds or 12:05 and 40 seconds, or is the actual time 12:06? After stating the time, it is no longer that time anymore; therefore, how accurate is the time that is stated? A student standing in the bookstore might answer the time question above with a simple statement like, “It is 12:05,” or they might even say “lunch time” without giving the statement a second thought. Furthermore, people rarely calibrate their watches to the Greenwich Mean Time, which is considered to be the world standard. What does the word “mean” imply in Greenwich Mean Time? How important is it to know the exact time at any given second? Is it acceptable to state an estimate of the time along with a range of say plus or minus three minutes? My own personal experience tells me yes. In fact, if the general public were to adopt this criteria and always include the uncertainty in the time, then according to the rules of uncertainty analysis one would record the time as (12:05 ± 03) minutes. This kind of reasoning is at the heart of uncertainty analysis. Moreover if you accept this example as typical, it is safe to say that everyone has used uncertainty analysis at least once in their lifetime. Even Alan Greenspan, the previous Chairman of the Federal Reserve in the United States and the person who adjusted interest rates in fractions of a percent, stated “It is better to be roughly right than precisely wrong.”

An introductory physics student standing in the bookstore might answer “lunch time” to the question about what time it is, but this same student would probably react very differently in a physics laboratory when asked the same question. The instructor would likely observe the student calculating the time to the millisecond or beyond, at least as far as his or her calculator has spaces for the decimals. After all it is crucial to give the most accurate, correct answer. Students are well programmed from their previous educational experience, which is dominated by true/false and multiple choice test questions, to believe that there is always one correct answer. They enter the physics laboratory and are informed that there is uncertainty in their measurements and that a range of values is more important than any one value no matter how careful it is determined. This places students in immediate conflict with their previous way of thinking. Measurements obtained in laboratories contain *uncertainty* in their values, and it is important that students understand the meaning of this uncertainty and how to incorporate it in laboratory reports. According to Taylor (1997) regardless of how careful one is in obtaining a measurement it can never be completely free of some uncertainty. Although students recognize that any instrument used for measurement has limitations in its accuracy and precision, they lack the knowledge and techniques to quantify those limitations and uncertainties. They also struggle with the concept of a range of results as an acceptable answer in lieu of one correct answer. Furthermore, students are asked to use these measurements with uncertainty in calculations to derive other values, otherwise known as *propagation of error*.

An overview and background of the position of educators and researchers on the importance of including uncertainty analysis in curricula follows, along with a review of the mathematical models utilized in teaching this concept. This chapter concludes with a description of the purpose of the study and its significance to the field of science education along with a statement of the research questions.

Background

Clarifying the Terminology

Uncertainty analysis is a term that is employed inconsistently in the science community, in part, because there are multiple related and confounding terms. Some of these terms include error analysis, human error, propagation of error, random error, systematic error tolerance, precision and accuracy. "Error" is present in many of these terms, but uncertainty analysis does not imply that there are mistakes. Swartz (1993) points out that, "As we shall emphasize, error represents uncertainty and has nothing to do with mistakes or sloppiness" (p. 1). Taylor (1997) also supports this position when he explains, "In science the word error does not carry the usual connotation of the terms mistake or blunder" (p. 3). This confusion may have contributed to students' persistent use of the term "human error" when explaining discrepancies in laboratory results. In this study, the terms, uncertainty analysis and propagation of uncertainty are frequently discussed together, but they are not referring to mistakes or errors in measurement. The term uncertainty in a measurement will simply refer to that defined in Swartz (p. 9) where the uncertainty is determined by the experimenter as the extreme possible boundaries of a measurement limited by the instrument, also known as a scale uncertainty. Propagation of uncertainty, or sometimes called propagation of error, is that aspect of uncertainty analysis dealing with the effect the uncertainty in a measurement has on derived quantities based on the uncertainty in a that measurement. The important point is, when the term uncertainty analysis is used in this study, it is referring to the propagation of uncertainty and or the scale uncertainty, both of which are considered different aspects of the more general term of uncertainty analysis. A more thorough explanation of uncertainty analysis and propagation of error is included in the Methodology Chapter (Chapter 3).

Professional organizations in science education have emphasized the importance of including uncertainty analysis in introductory science curricula. In 1998, the American Association of Physics Teachers (AAPT) identified five fundamental goals of introductory physics laboratories. Instruction in measurement uncertainty is a component of "Goal II: Experimental and Analytical Skills". The AAPT identified the importance of students' ability to analyze data and demonstrate an understanding of the relationship between laboratory results and mathematical interpretation of results at varying levels of sophistication.

Students should understand the uncertainty associated with measurement and the distinction between experimental uncertainties and mistakes in reading or recording information. Students should learn enough about uncertainties to understand the inherent limitations of measurement processes (p. 483).

In addition to professional organizations, individual researchers and educators advocate including uncertainty analysis in the curriculum. Phillips (1972) stated, "I maintain that the calculation of errors is an integral part of all experimental science and that answers obtained without it are not worth the paper they are written on" (p. 383). Students are often frustrated with the concept of uncertainty analysis, and the following familiar phrase illustrates this: “If it’s green or it wiggles, it’s biology; if it stinks, it’s chemistry; if it doesn’t work, it’s physics” (Roberts, 1983, p. 155). Many students are not able to analyze data in terms of the associated uncertainty. Students often say that the experiment did not work, even when their data are valid. The results may be valid if they incorporate uncertainty, but rather than recognize this, students try to rationalize their inconsistent results as resulting from *human error*. They usually continue to look for one correct answer. As Allie, Buffler, Campbell and Lubben (2003) stated:

> For many students, the ideal is to perform a single perfect measurement with the utmost care. When presented with data that are dispersed, they often attempt to choose the “correct” value (for example, the recurring value) from amongst the values in the ensemble. (p. 394)

If students enter the classroom with the idea of finding a correct answer, then it is easy to understand why educators often find that the topic of uncertainty analysis wreaks havoc in the classroom when first presented to entry-level physics students. According to Deardorff (2001), novices tend to assign more importance to the absolute difference between results and ignore the uncertainty in the values. It is common to see students report answers on their written assignments with as many as 13 decimals, when the measurement instruments may only be accurate to 1 or 2 decimal spaces. As Fairbrother and Hackling (1997) noted, students who look for the right or wrong answer are unable to see that “an experiment which works is similar to a car which runs properly – all the parts fit together, it functions and gives *an* answer which can be defended” (p. 891).

Mathematical Models of Instruction on Propagation of Uncertainty

Educators have attempted to teach students about uncertainty analysis, but there are many challenges associated with instruction on this topic. Some of these are complicated by the previous knowledge of introductory level students which is limited by their level of mathematical background (Blasiak, 1983). Therefore, educators and researchers have incorporated instruction on uncertainty analysis in introductory physics laboratories utilizing a variety of strategies and difficulty levels. Uncertainty analysis can be taught at several levels of difficulty: statistics and calculus, the algebra-based Worst Case Method, and percent difference with significant figures. The method employed in this study is the Worst Case Method for Propagating Uncertainties. The term algebra-based is used here to emphasize the fact that this method can be utilized by students possessing only a rudimentary knowledge of algebra. This method is explained in detail in the research design section in this chapter. Additionally, the Organization for Standardization developed a framework for addressing uncertainty in measurements, which is known as the GUM (Guidelines to the Expression of Uncertainty in Measurement) method (Pillay, Buffler, Lubben, & Allie, 2008). An overview of the three most common methods follows.

At the most advanced level, uncertainty analysis employs the use of statistics and calculus, which provide the most powerful analytical tools for performing data analysis. Entry level students usually do not possess the mathematical skills required for the full, rigorous treatment of statistical analysis. Moreover, according to Roberts (1983) and St. John (1980), the attempt to educate entry-level students about the rigors of statistics does more harm than good. These methods – statistics and calculus - are most appropriate for data obtained from multiple measurements, which are more suited for statistical analysis.

The International Organization for Standardization (ISO) issued GUM in 1995 in response to the fragmented way of applying measurement reporting in the science disciplines. This guide to calculating and reporting measurements and uncertainties has been adopted by "all international standards organizations including the International Union of Pure and Applied Physics and the National Institute of Standards and Technology (Pillay et al., 2008).

The GUM method utilizes probability density functions (pdf) associated with a measurement and assigns a standard uncertainty based on the form of the pdf. While this method is superior to determining uncertainties in a measurement based on scale, as used in this study, the Pillay et al. study did not include evaluation of propagation of error of these uncertainties. In 2008 a workbook was developed by Allie and

Buffler incorporating all components of the GUM method. The format followed a user friendly version appropriate for beginning science students. Had this workbook been available at the time this study, in would have contributed to some aspect of the research design. The GUM method for propagating uncertainty is again based on the more advanced framework using calculus and statistics. More on this in the literature review.

At the other end of the spectrum of difficulty are the rules that define the use of percent difference and significant figures for analyzing uncertainty. Bauman and Swartz (1984) point out that the application of uncertainty analysis in most science laboratories is limited to the calculation of percent error and the use of significant figures. They also conclude that these techniques are too limited for calculating the uncertainty in laboratory data and do not address propagation of error.

A third technique of analyzing uncertainty in data is a method that is less complex than the statistical method but sophisticated enough to analyze physics laboratory data.. The Worst Case Method, first employed by Gordon, Pickering and Bisson (1984), is the method used in this study. This method focuses on calculating the minimum and maximum values, hence "worst case" of a measurement rather than the traditional approach derived from calculus and statistics. In addition this method yields an absolute uncertainty and bypasses the need to determine relative uncertainties which students can find confusing. In a study conducted at Princeton University, Gordon, et al. concluded that upper division chemistry and physics students preferred the Worst Case Method over the use of the traditional method as a tool for analyzing uncertainty. It is a fact that any instrument used to measure a quantity has a finite limit of precision and the fluctuations in values are due to the uncertainties in the measurements. All measurement uncertainties have an effect on equations and on the final results of the experiment (Bevington, 1969). For example, if someone is taking a trip and wants to estimate how long it will take to get there, he or she divides the distance by speed. The distance and speed are of course not known exactly. Assume the distance is approximately 100 miles away, give or take 5 miles, and your speed is 50 mph give or take 10 mph. In the notation of uncertainty analysis the "give or take amount" is defined as the uncertainty in the distance and, by convention, is written as $\delta = \pm 5$ miles for distance and $\delta = \pm 10$ mph for speed. The δ symbol stands for "the uncertainty in". The distance along with its uncertainty is 100 ± 10 miles and the speed is 50 ± 10 mph. The question now is how does the uncertainty in the distance and speed affect the estimate in the time? The impact that uncertainties have on mathematical operations is called propagation of error and the Worst Case Method provides the student with techniques to calculate this impact. This is the method employed in the current study and a more in depth example is discussed in the Research Methodology chapter (Chapter 3).

The Worst Case Method provides a way to calculate uncertainty that is consistent with the mathematical background of the introductory physics students. It is important for educators to evaluate prerequisite knowledge for comprehending concepts. Entry-level physics students are not equipped with the mathematical techniques that dictate how uncertainties in the data propagate into calculations and formulas (Blasiak, 1983). As mentioned earlier, statistics and calculus do encompass the methods to analyze uncertainty in data, but entry-level students do not have these skills (Roberts, 1983). Few instructors teach the full breath of this content, including propagation of error. Most instructors attempt to provide some instruction and explanation on the topic of propagation of uncertainty when discussing uncertainty analysis, but they usually do not present a comprehensive and consistent picture (Roberts, 1983).

The previous section discussed the background to the problems encountered with teaching and understanding uncertainty analysis and propagation of error. It also described the mathematical models used by educators, including an overview of the Worst Case Method. The next section covers the motivation for the study and gives an explanation of the overall design. The purpose, significance and research questions are discussed.

Motivation for the Study

As an instructor in introductory physics courses, I have incorporated content on uncertainty analysis in the course materials, both in response to professional recommendations and in response to the idea that it is essential in facilitating student understanding of their laboratory data. One of the most challenging experiences I have encountered is helping students to understand these concepts. Students' difficulty in understanding and applying these concepts increases their frustration and inability to correctly analyze laboratory data. When students learn about a physics concept and calculate predicted values from mathematical formulas, they expect to obtain exactly the same results from measurements which they make in the actual laboratory. When their results are different from what is predicted they believe that they made an error and become frustrated, especially when identical measurements yield a range of data. After recording the data that they believe is not as accurate as it should be, they must then use these data in calculations for their laboratory written assignment. Students' lack of understanding of uncertainty analysis and propagation of error makes it difficult for them to interpret the data, and they become confused and disappointed in the results.

Instruction on uncertainty analysis has been a topic in the literature and research among physics teachers. When including this content in the curriculum, success has been noted in an improvement in students' ability to record measurements with an associated uncertainty. Little success has been noted with improving student understanding of the concept. These concepts are generally taught in calculus, and many entry-level physics students have taken algebra as their most advanced mathematics course. Therefore, the students do not have the prerequisite knowledge to understand propagation of error.

I began teaching uncertainty analysis and propagation of error in the fall of 2001, utilizing the Worst Case Method. I recognized an immediate improvement in the quality of data analysis in laboratory reports. Students began to quantify the margin of error associated with an experimental result and the theoretical prediction. They also realized that an experiment could support a theory even though the laboratory results may not be an exact match to the predicted theoretical values. In contrast to the lack of understanding of the concept in studies, it appeared that students understood the concepts. I heard more “Oh, I see” comments with smiles on students’ faces. My previous experience was that I would show a student how to calculate uncertainty in one problem, but the student would return with similar questions as if they had no understanding of how to calculate the uncertainty. After instruction using the Worst Case Method, I noted less return visits to my office. Students seemed to understand how to apply the concepts to new but similar problems. They were quick to spot and correct their minor mistakes, with little assistance. In addition, I saw a significant increase in the confidence students demonstrated while participating in a laboratory exercise.

These observations led me to believe that the concepts of uncertainty analysis are essential to understanding and analyzing data in the physics laboratory and that it could be taught using only simple mathematics along with some algebra. In fact, students appeared to grasp and understand these concepts with the benefit of the simpler, more consistent tools offered by this level of mathematics. Students appeared to be able to expand on their understanding of uncertainty to the more advanced concept of propagation of error. As a result of my experiences and observations, I decided to conduct a research study to determine if my observations were supported in a more scientific investigation. Consequently, I decided to evaluate an algebra based instructional model on teaching uncertainty analysis and its effectiveness in increasing the level of understanding of these concepts.

Purpose and Design of the Study

The purpose of this study was to evaluate the effectiveness of an educational intervention based on the Worst Case Method for teaching uncertainty analysis and propagation of error. The intervention consisted of classroom instruction and the use of a manual, developed by this researcher, based on the Worst Case Method which provides the student with a theoretical background, along with many examples on how to use the Worst Case Method. (Further details are in Chapter 3 Research Methodology, and the manual is in Appendix A.)

Phase One of the study comprised a pretest and posttest designed to identify misconceptions which students had about measurements and how they interpret data. More specifically the goal was to categorize how students think about measurements in terms of either point or set reasoning as defined by Allie and Buffler (2003). This was accomplished through the use of a 12 item two-tier questionnaire based on similar instruments developed by Treagust (1988, 1995), Odom and Barrow (1995), and Tan, Goh, Chia, and Treagust (2002). The development of the two-tier questionnaire is discussed in greater detail in chapter three.

Phase Two involved the analysis of laboratory work submitted by students over the course of a normal semester. The purpose of Phase Two was to evaluate how students progressed in their understanding of the application of the methods provided in the intervention. The level of understanding was evaluated utilizing Wiggins and McTighe's (2005) six facets of understanding. Laboratory assignments were also evaluated for the students' level of conceptual change.

Research Questions

The research questions for Phase One are:

1. To what extent can the concepts of uncertainty analysis and propagation of uncertainty be successfully taught to entry-level students with a minimal background in algebra?
2. Over the course of a typical college physics laboratory class, how do first year students' concepts of uncertainty compare from the beginning to the end?

The research questions for Phase Two are:

1. How successful are students in applying the concepts of uncertainty analysis while participating in a traditional physics laboratory course?
2. How do students' concepts of uncertainty analysis develop or change while participating in a traditional physics laboratory course?

Significance of the Study

The literature supports the claim that there exists little agreement among educators on the best way to teach propagation of error, but there is considerable agreement on the importance of the subject in the laboratory setting. The methods available for students to analyze and propagate uncertainty, at the freshman level, are inconsistent in their mathematical background. Although the use of significant figures provides a first approximation to uncertainty, it is limited in application to elementary mathematical functions such as addition, subtraction, multiplication and division. It is also not appropriate to use with more complicated relations often used to relate variables in laboratory exercises such as exponential, trigonometric and logarithmic functions. In contrast, the more advanced methods, such as the total differential and statistics, are often beyond the scope of mathematical ability of the average freshman or high school student. The results of this study I believe will fill the gap between the ‘too easy method’ and the ‘too hard method’. This is significant in that many teachers are frustrated with the current inconsistencies that exist among the methods and therefore may tend to avoid the subject altogether. As pointed out by Thompson (1997), terms such as systematic and random errors along with accuracy and precision are not used consistently even in physics publications. A consistent technique that is applicable to all the mathematical functions and that is easy to teach and learn will, I believe, result in this important tool finding wide acceptance in the secondary and college laboratory science classrooms. Secondly, the laboratory experience will be more satisfying for students because they will have a reliable method to evaluate the difference between what is predicted by the physics and their actual results, rather than attempting to rationalize what they do not understand by labeling it as human error or equipment problems.

Limitations

One limitation was that there was no guarantee that student’s would actually read the manual. One assumption was that the majority of students enrolled in a college physics course is typically very motivated and are eager to apply any additional materials to assist them in attaining the highest grade possible. The best evidence that students were consulting the manual was the many questions asked about specific content areas throughout the semester. This fact suggests an additional limitation in that the manual should have been field tested before its use in this study. This may have clarified some of the terminology and notation that students found confusing.

Overview of the Chapter

In summary, this chapter introduces the reader to the research questions and the challenges of teaching uncertainty analysis to beginning science students with a limited mathematical background. Moreover students' misconceptions and preconceived ideas, learned during their previous education, complicate the process of accepting a new concept. The belief that there is only the one correct answer or that human error is the cause behind unacceptable results make it very difficult for students to accept the idea that a range of values is a valid way to describe data. The various methods for calculating uncertainties, described in the literature, are also discussed, including the most relevant method utilized in this study - the Worst Case Method. The importance of applying uncertainty analysis is well documented and there is considerable agreement among educators that the advanced nature of the calculus-based methods is beyond the ability of most entry level students. In response to this need for a simpler approach to calculating uncertainties, a manual was developed utilizing the Worst Case Method. A brief description of the Worst Case Method and the concept of propagation of error were also presented.

CHAPTER 2

REVIEW OF LITERATURE

The importance of uncertainty analysis and propagation of error in science education has been emphasized for many years, and researchers have investigated methods to successfully teach entry level students about measurement error. Additionally, researchers have pursued a better understanding of how students learn and how conceptual change affects the educational experience. The ability to understand uncertainty analysis and propagation of error requires students to look at data differently. Students arrive in the classroom with previous knowledge and misconceptions about measurement, and they must incorporate this knowledge with the new material to develop new concepts about analyzing data. Uncertainty analysis and propagation of error requires students to transition from looking for one right and most accurate answer to looking at data in terms of ranges with uncertainty around them. Some researchers have differentiated between looking for one answer to looking for ranges in data to point and set reasoning. These studies generally reported positive changes in students' ability to mechanically measure and record uncertainty, but were less positive in improving students' understanding of uncertainty. Understanding uncertainty analysis and propagation of error requires a conceptual change, which is far more complex than the ability to record uncertainty. This chapter discusses literature on conceptual change in terms of the process and how it applies to uncertainty analysis, along with the research on point and set reasoning. The chapter concludes with an overview of one method of evaluating student understanding that is employed in this study.

Conceptual Change

Conceptual change "refers to cognitive restructuring different from what is evidenced in conceptual growth" (Duit & Confrey, 1996, pp. 80-81). The student learns through a process of restructuring, which can vary from mild adjustment to an actual change in thinking. The amount of restructuring is often mitigated by the student's previous knowledge and preconceptions. Duit and Treagust (2003) discussed transition in understanding conceptual change since the 1980's. They reviewed research and literature spanning from the classical view in the 1980's and 1990's to viewing changes in concepts about science as incorporating a much broader range of factors including the student's learning environment and their beliefs about science and about the real world. Conceptual change is a complicated process that involves many factors. Some of the concepts from constructivism, and more recently the research that looked at broader concepts such as ontological beliefs, epistemological beliefs and factors that favor students adopting new concepts, are helpful in understanding the process of students experiencing conceptual change as they learn about uncertainty analysis. In the following section, the literature on constructivism is covered followed by an overview or pre-instructional, epistemological, and ontological beliefs, as they relate to conceptual change.

Constructivism

Treagust et al. (1996) explain that during the constructive process students do not simply transfer the teacher's ideas into their heads; rather they incorporate the knowledge into their own ideas. There is often much variance from one student to the other, as this process is influenced by students' pre-instructional knowledge and subjective views of the world. Assimilation and accommodation are two concepts that are a component of the constructivist theory of conceptual change, and they are rooted in Piaget's theory of cognitive development. According to Harrison and Treagust (2000), assimilation has been called "weak knowledge restructuring", and accommodation has been referred to as "strong/radical knowledge restructuring" or conceptual change (p. 672). Students can be at various stages of assimilation or accommodation of new information, but those who experience conceptual change have a new construct of the concept. Students are confronted with new knowledge or new situations, and they must adapt and re-organize how they think. Hewson (1996) discussed conceptual change as acquiring new conceptions or exchanging existing conceptions for the new ones. Hewson points out that students learn by making connections between new ideas and those that they already have. This is referred to as assimilation or conceptual capture, and it generally occurs when students' current ideas are consistent with the new material they are learning. Hewson also discusses the higher level of conceptual change which is similar to accommodation. The student who has assimilated knowledge on uncertainty analysis may at times incorporate measurement uncertainty in their answers, but the same student may revert to the old way of thinking, such as including 13 decimal points in an answer claiming more accuracy just because the calculator is tabbed to that many places. The student who has accommodated the concept of uncertainty analysis sees data differently. The student stops looking for the one right answer and sees data as a range of values and understands the reasons behind the ranges along with the impact of using these data in further calculations. The development of a whole new concept is rare, and the deeply engrained ideas involving a student's pre-instructional, epistemological and ontological beliefs are some of the factors influencing the process.

Preinstructional, Epistemological, and Ontological Beliefs

Treagust, Duit, and Fraser (1996) indicate that research shows that pre-instructional conceptions are difficult to change and that students strongly resist change in these beliefs. Duit and Treagust (1998) point out that students enter the classroom with pre-existing ideas that may be in conflict with the material presented. For example, if a student believes that there is one right answer, then that belief is in conflict with the idea of accepting a range of values in an answer as more scientific. As students progress through their elementary education, they generally learn that there is one right answer to questions and problems, and this idea can be very resistant to change.

The idea of one right answer is more involved that pre-existing knowledge learned in school and it extends into a student's beliefs about science and the very nature of reality. These types of ideas are referred to as epistemological and ontological. As Treagust and Duit (2008) point out, "conceptual change from an epistemological and an ontological perspective refers to students' personal views, on the nature of coming to know – what we refer to as epistemological – and on the nature of reality – what we refer to as ontological" (p. 299). Chinn and Brewer (1993) state that ontological beliefs refer to "the fundamental categories and properties of the world" (p. 17). Cohen and Manion (1989) state that epistemology involves "the very bases of knowledge –its nature and forms, how it can be acquired, and how its communicated to other human beings" (p. 6). When students collect data in a laboratory and obtain different results with each subsequent measurement, they often attribute it to human error. It is difficult to say what leads students to conclude that the data has errors or that human error contributed to the problem or that it is their fault without probing the thoughts of each student as they approach the issue. It is likely that the ideas of right answers in the world and ideas about science and mathematics being accurate and precise come into play. Clinging to these beliefs prevents students from viewing data in a new way. Venville and Treagust (1997) interviewed students in science classes to evaluate their ideas and beliefs about science concepts over the course of a semester. Their findings are very revealing on the process of conceptual change. They found that student's ideas about genes transitioned from the idea of a gene as a 'passive particle' to the idea that a gene was active and contained instructions that influenced characteristics. It is important to note that all students did not achieve the same level of understanding. Their beliefs about science and about the world are influencing their thinking as they are learning about new concepts, including uncertainty analysis.

Treagust and Duit (2008) comment on the challenges of teaching concepts in a manner that facilitates a conceptual change. They point out that "research has shown that students come to science classes with pre-instructional conceptions and ideas about the phenomena and concepts to be learned that are not in harmony with science views" (p. 298). They continue to state that these beliefs are often very difficult to change, and pre-instructional beliefs should not be under estimated when teaching new concepts. Duit and Treagust (2003) describe these beliefs as a set of goggles that interpret how students see all the material presented in class. For example, as demonstrated in the Venville and Treagust (1997) study, if a student considers genes to be passive particles, it is very difficult to teach students about the complex biochemical process involved in genetics. Chinn and Brewer (1993) point out that when students are presented with information about the world that conflict with their ideas, they will usually cling to their pre-instructional beliefs. Ideas are harder to change when they are "deeply embedded in a network of other beliefs" (p. 15). One embedded belief in the minds of many entry level physics students is that human error is impossible to eliminate and is an acceptable explanation for why results from an experiment do not support the theory.

Process of Conceptual Change

The previous section illustrates some of the difficulties found in changing students pre-existing concepts about the world and science, and that when a new way to look at the ideas is introduced, there is often little initial change in student thinking. In fact, as Duit and Treagust (1998) discuss in a review of literature, research studies have demonstrated that progress to learning new concepts is often very limited. They go on to report that they did not review any study where the student's original concepts were "completely extinguished and then replaced by the science view" (p. 673). They continue to state that the original concepts frequently remain in some form.

Researchers have investigated the process of conceptual change in an effort to better understand the process. Many have incorporated the idea of a new concept leading to some sort of dissatisfaction in the student when new ideas are presented. The work of Chinn and Brewer (1993) focused on anomalous data or "presenting students with evidence that contradicts their pre-instructional beliefs" (p. 2). When presented with this information, students may accept the data and change their ideas or discount the data and revert to their pre-instructional way of thinking. They continue with their discussion and further elaborate on a continuum of responses. Students may ignore the data and not even explain it. Another possible outcome is that students may reject the idea. Subsequently, there may be some explanation for the rejection, but they do not change their way of thinking. Other possibilities are to exclude the data as outside of the current focus of discussion, to procrastinate contemplating the issue and put it off to another time, or to reinterpret the data. The seventh possibility is the only option that is actually conceptual change, which is for students to accept the data and change their thinking. Few students reach that point (Chinn & Brewer, 1993). Hewson and Hewson (1984) also focused on the idea of some feelings of dissatisfaction that is similar to the idea of anomalous data. They stated that without dissatisfaction, the new concept may be added to the old one, and they exist side by side. When dissatisfaction occurs, there are a couple of possibilities. If the new idea is valued more highly, it may be adopted and accommodation occurs. If the old idea is more highly valued then conceptual change and accommodation do not occur at that time, but they point out that the concept is still there and may be incorporated at another time. Therefore, instruction on science concepts such as uncertainty analysis may not result in immediate changes in thinking, but they may plant the seeds for more understanding at a later time.

Hewson (1992) also discusses the fact that there are criteria that impact whether or not students adopt new ideas and concepts, which include whether the idea is "intelligible (knowing what it means), plausible (believing it to be true), and fruitful (finding it useful)" (p. 8). Students can demonstrate if the concept is intelligible to them by describing it in their own words. To be plausible, the concept must first be intelligible, and it also must present a viewpoint that the student believes is possible in the world. It has to fit in with other ideas. To be fruitful, a concept should first be intelligible and plausible, but it also has to be helpful and show a better way of doing or explaining things.

When applying these ideas to uncertainty analysis, it is evident that students must first have some way of understanding the concept. One problem with this has been in terms of whether or not they have the mathematical tools to incorporate the ideas. As mentioned previously, entry-level students may not have the mathematical background to understand uncertainty analysis and propagation of error when the instructional methods are calculus based. The plausibility of uncertainty analysis may depend on a student's ability to let go of the idea of one right answer, and or human error. The new idea of a range of answers may not fit with his or her view of the world. Finally, to be fruitful, the student must see this as a better way of explaining experimental data. If these criteria are present, then conceptual change could occur in students when taught the concepts of uncertainty analysis.

Entry-level physics students come to the classroom from a wide variety of backgrounds and beliefs. Their previous knowledge may not include the mathematical skills required to understand uncertainty analysis. Their preconceptions may include beliefs that contradict accepting the idea that a range of values can be a correct. Their view of science and about the laws of nature may interfere with the concepts of uncertainty in measurements and mathematical calculations. These beliefs and previous knowledge are crucial for educators to explore when teaching this topic. Even when an educator addresses these ideas, conceptual change may be a slow process. Although some students will transition to a new idea and have a new concept of the science topic, many students are at varying places in the transition. They may at times appear to grasp an idea but quickly revert to some of the old ideas. It is not uncommon for new and old concepts to co-exist. Educators may find some consolation to their frustration when teaching uncertainty analysis in realizing that they may be planting the seeds for future conceptual change, even when they do not see clear evidence of it in a student in their classroom.

Point and Set Reasoning

One body of research that is closely related to conceptual change includes the studies on point and set reasoning. The studies on point and set reasoning included in this research began with the work of Allie, Buffler, Kaunda, Campbell and Lubben (1998), who described the way students reason about measurement data as point or set reasoning, and expands to the studies of Volkwyn, Allie, and Buffler (2008), who evaluated changes in reasoning in entry level physics students. The earlier work by Allie et al. provided insight into students' understanding about measurements in terms of how students judge the quality of data and how those judgments affect experimental procedures. Students who were identified in the study as 'perfecters' expressed the need to keep practicing in an attempt to either eliminate random and or systematic errors or to keep measuring until the same number kept appearing. Obviously interpreting results in this fashion influences the experimental procedure. These same students would later be identified as point reasoners by Lubben, Campbell, Buffler, and Allie (2001). In the Lubben et al. study many more students expressed the need to establish a mean. Interestingly a small number of these students also believed it was important to consider the spread and uncertainty in the measurements. Lubben et al. later characterized these students as set reasoners. Although these studies did establish two fundamental categories of how students think about experimental data, no evidence was presented in either study on how to teach uncertainty analysis.

Most educators and researchers reported an improvement in students' ability to calculate measurement uncertainty after instruction, but few reported an improvement in understanding. In an effort to provide a framework for teaching measurement uncertainty, Lubben et al. (2001) classified how students reason about the reliability and validity of experimental data into two broad categories: point reasoning and set reasoning. Point reasoning is characterized by the underlying notion that each measurement could in principle be the true value. As a consequence each measurement is independent of the others, and the individual measurements are not combined in any way. This is similar to the previously addressed preconceptions that there is one right answer.

On the other hand, set reasoning is characterized by the idea that the data must be considered as a collective set of measurements that randomly deviate around the true value. Consequently, the best estimate of the true value is obtained by combining the measurements such as the mean and standard deviation. Lubben et al. (2001) noted that these students used point and set reasoning in a "fragmented way" (p. 325), and they noted contradiction between how students seem to reason about data and their actions. For example, they may have stated that they are taking repeat measurements to find an average but they then chose a recurring value as an answer rather than the mean value. Overall Lubben et al. report that a 'full understanding' is not apparent in these entry-level students.

Rollnick, Lubben, Lotz, and Diamini (2002) examined the effect of instruction on students' understanding about the quality of measurements at two South African universities. More specifically these studies probed students' knowledge about the quality of data and classified their perception about the quality of the data collected. Students who considered measurements as independent of each other and represented data from single attempts as a true value fell into the category of point paradigm. The more advanced student recognized the importance of the mean and spread of the data and fell into the category of set paradigm. Students at both universities made considerable gains after instruction in moving from the point paradigm to the set paradigm thinking.

Buffler, Allie, Lubben, and Campbell (2003) evaluated the effectiveness of a research based curriculum for teaching measurement in the first year physics laboratory. The framework for this investigation was also based on categorizing responses according to the point and set paradigms. The evaluation of this new course used diagnostic testing before and after the course. This diagnostic instrument was administered to 106 freshmen before the course began and then again after completion of the class. Before the course only 1% of students exhibited set paradigm thinking compared with 89% after the course. Although these results show that the new course was successful in facilitating students' transition from point paradigm to set paradigm reasoning, no attempt was made to evaluate the impact of this new curriculum on students' understanding of propagation of error.

The effectiveness of the Guidelines to the Expression of Uncertainty in Measurement (GUM) method for teaching in an introductory physics laboratory comprised of 76 first year students was evaluated by Pillay, Buffler, Lubben and Allie (2008). The results were compared to a control group that received instruction in a more traditional curriculum for teaching measurement. Students were evaluated in terms of their point and set reasoning both pre and post-instruction. Those who received instruction with the GUM framework demonstrated an improved "understanding of a measurement based on a single observation" (p. 657).

Volkwyn, et al. (2008) expanded upon previous research by investigating point and set reasoning of 53 students who were physics majors. Prior to this time, the samples comprised introductory students who were not physics majors. The assumption was that physics majors would enter the introductory laboratory with a more advanced knowledge of science and mathematics required to understand uncertainty analysis. These students were enrolled in a physics laboratory course with 12 three-hour laboratory sessions, and the initial sessions covered experimental skills, including measurement and uncertainty. A majority of students demonstrated an improvement in the mechanics of measurement and the inclusion of uncertainty in their reported values. When compared to the 40% of non-physics majors who demonstrated set reasoning in the Lubben, et al. (2001) study, over 90% utilized set reasoning in this study. However, the analysis of the level of understanding did not result in a significant conceptual improvement, with only 20% of the physics majors demonstrating an improved understanding of the concept required to analyze and compare data. Although the researchers expected more improvement in level of understanding with these more advanced students, the results were comparable to the

Lubben, et al. study which found that 19% of the physics non majors demonstrated a higher level of understanding.

Buffler, Allie and Lubben (2008) recently presented to the community of science educators a summary of a workbook designed according to the ISO publication, *GUM.* The GUM method introduces the student to the *measurand*. The measurand is a term generally used to identify anything undergoing the measurement process. The objective of the measurement process is to obtain as much information about the measurand by identifying any and all possible sources of uncertainty associated with the measurand. By increasing the precision of the measurement and decreasing the amount of uncertainty one can know more about the measurand but never its true value as that would take infinite precision and zero uncertainty.

The methods are described in terms of what is called the probabilistic approach, using *probability density functions* as a framework for identifying uncertainties. The GUM method utilizes probability density functions (pdf) associated with a measurement and assigns a standard uncertainty based on the form of the pdf. There are three probability density functions employed by the workbook; flat, triangular and Gaussian. The flat, or rectangular, is used for a single digital reading, the triangular is used for a single analog reading and the Gaussian is employed for repeated or scattered data. The workbook reduces the inherently advanced mathematical nature of the more rigorous methods to a more user-friendly version. The introduction outlines the frustrations that educators are faced with when confronted with the desire to teach uncertainty analysis and the inconsistencies that exist among the different methods accumulated over the past.

The majority of the workbook focuses on how to interpret data and how to assign uncertainties to a measurement. A small portion of the workbook focuses on the propagation of uncertainties. The authors present a general formula, derived from calculus, that is the standard for calculating the uncertainty in all mathematical forms, but as is noted in the workbook, can be a complicated process. Also presented are simplified forms of the equation for determining uncertainties for addition, subtraction and functions involving powers but not for others, such as trigonometric and logarithmic functions. The Worst Case Method, outlined in

Appendix A, focuses on how to propagate uncertainties and can be used for all mathematical expressions. The GUM workbook does outline a very consistent methodology for the assignment of uncertainty for both single and multiple measurements compliant with the recommendations of ISO at a level appropriate for freshmen science students. More on this in the discussion section.

When reviewing these South African studies, it is apparent that instruction was helpful in teaching students about quantifying uncertainty and in measurement factors, but most did not demonstrate a significant change in their level of understanding. Some students continued to look for one right answer and cling to the point paradigm type of reasoning, and other students reported the importance or ranges and means but reverted to point reasoning at other times. Although instruction demonstrated improvement in terms of transition to set reasoning, most researchers did not find a change in depth of understanding. Students tended to persist in their pre instructional ways of thinking to some degree. Their pre-conceptions about accurate answers seemed to persist in some fashion. Few students looked at data and viewed it in terms of sets and ranges and then incorporated this new way of looking at the data into a new concept that included the idea of uncertainty and propagation of error in their written work. When comparing these results to the literature on conceptual change, these results are not surprising.

Facets of Understanding

Studies on point and set reasoning describe changes in level of understanding. Wiggins and McTighe (2005) have studied the meaning of understanding and how to evaluate it. The word understanding has many meanings, and the criteria to evaluate student understanding of concepts is very complex. Understanding is more than knowledge, and it involves the meaning of concepts and the inclusion of theory in the understanding of concepts. Understanding is knowing why, and it involves a higher level of sophistication than knowledge (p. 38). Wiggins and McTighe developed the six facets of understanding: explanation, interpretation, application, perspective, empathy, and self-knowledge (pp. 85-100), to evaluate level of understanding. The facet of explanation is utilized to evaluate understanding in this study, and a more in-depth discussion is included in Chapter 3: Research Methodologies.

Overview of the Chapter

This chapter has focused on some of the background research and literature that impacts upon the current study. In summary, the following areas have been discussed in Chapter 2:

- Uncertainty analysis has been a difficult concept for entry-level physics students to learn and understand
- Conceptual change is necessary for students to fully understand these concepts, and the literature on constructivism and pre-instructional beliefs provides insight into the challenges in teaching these concepts.
- Point and Set Reasoning is one body of research dedicated to identifying student misconceptions about the quality of measurements made in the laboratory. Instruction on uncertainty analysis showed some improvement in understanding how to interpret data in the laboratory.
- Understanding of concepts can be evaluated by the Facets of Understanding developed by Wiggins and McTighe (2005).
- Student interviews have also provided a means to evaluate student understanding and the process of conceptual change.
- No studies were found that evaluated student understanding of uncertainty analysis and propagation of error over the course of a semester by evaluating weekly laboratory assignments especially in the areas of epistemological and ontological beliefs

CHAPTER 3

RESEARCH METHODOLOGY

The research design was developed to measure a change in entry level physics students' understanding of uncertainty analysis before and after an educational intervention. The intervention was an instructional manual on uncertainty analysis. The design included two phases. The first phase evaluated the effectiveness of the educational intervention in improving students' understanding of uncertainty analysis as measured by a pre-and posttest design (Cohen & Manion, 1989). At the time when this research was implemented, there were no available instruments in the form of a two-tier design to assess uncertainty analysis so a pilot study was completed to gather information from students to develop a pre and posttest for this first phase. The second phase evaluated the change in students' level of understanding in the application of knowledge about uncertainty analysis expressed in their written laboratory reports.

This chapter is divided into eight sections which describe the research methodology of this study. The first section describes the purpose of the study, and it is followed by the research questions. Next, the research setting is identified, and the demographics of the sample are discussed. The research design follows, with a description of the development of the instructional manual and pilot study prior to the main study, including an example of the mathematical method used in the manual. This section also discusses the two phases of the study, including the two-tier pretest and posttest and the evaluation of laboratory assignments for students' level of understanding and conceptual change. An overview of the data analysis procedure, the limitations, and the ethical issues involved in the study conclude the chapter.

Purpose of the Study

The purpose of this study was to evaluate the effectiveness of a self-study manual based on the Worst Case Method to teach uncertainty analysis to entry level physics students possessing only knowledge of elementary algebra (see Appendix A) .

The Worst Case Method, first developed by Gordon, Pickering and Bisson (1984), was used in a study where students were instructed on uncertainty using the more traditional approach and the Worst Case Method. The results indicated students preferred the Worst Case Method over the traditional approach. No prerequisite knowledge of calculus or statistics is required.

Research Questions

Phase One: The Pretest and Posttest

The research questions for Phase One are:

1. To what extent can the concepts of uncertainty analysis and propagation of uncertainty be successfully taught to entry-level students with minimal background in algebra?
2. Over the course of a typical college physics laboratory class, how do first year students' concepts of uncertainty compare from the beginning to the end?

Phase Two: Evaluation of Student Laboratory Assignments

The following research questions pertain to Phase Two:

1. How successful are students in applying the concepts of uncertainty analysis while participating in a traditional physics laboratory course?
2. How do students' concepts of uncertainty analysis develop or change while participating in a traditional physics laboratory course?

Research Setting

The study was conducted in a community college in a suburb of a major metropolitan area in the southwestern United States. Data from the pilot study were collected during the summer semester of 2004, and data for the main study were collected during the fall semester of 2004.

Sample

All students who were enrolled in a college level physics class ranged in age from 19 to 36 years old. The sample size for the pilot study was 25 students enrolled during the summer semester of 2004. The sample used for both Phase One and Phase Two consisted of 36 students enrolled in fall semester of 2004. The initial enrollment was 53, with 17 students withdrawing.

An initial survey revealed that students' proposed majors included engineering, pre-health professions, liberal arts, and technology majors. A majority of these students planned to transfer to the university to pursue a baccalaureate degree. Several students entered the class with previously earned baccalaureate degrees, and their goal was to advance their knowledge of science or physics. The students included in this study have selected majors that are consistent with students in the larger population (Institutional Resource Office, 2004).

Research Design

This study includes two phases. The first phase comprised a pretest and posttest to identify increased understanding of uncertainty analysis and to evaluate the effectiveness of the manual. Students completed a pretest. Following the pretest the manual on uncertainty analysis was distributed to the students. The students were instructed to use the manual as a reference when calculating uncertainties needed for their laboratory work during the semester. The researcher provided very little verbal instruction on uncertainty analysis. At the end of the semester the students completed a posttest.

The second phase comprised an evaluation of student laboratory assignments for the application of uncertainty analysis concepts and for examining students' level of understanding of these concepts. Five laboratory assignments were included in the study, and student level of understanding was evaluated as students progressed through the semester. The level of understanding was evaluated by utilizing Wiggins and McTighe's Facets of Understanding (2005).

There was no control group for either Phase One or Two. According to Salkind (1997), this fact presents some limitations. Although use of a control and experimental group is ideal, it would have resulted in two problems in this study. The first is that information would be withheld from the control group, and it is valuable information for completing laboratory assignments. As an educator, the researcher did not want to withhold information. It is also doubtful that the information could have been restricted to the experimental group. In a small community college, with a class sample size of 36, there was a high probability of the groups talking to each other and sharing the information. The other problem is that with a total sample size of 36, the separation of 36 students into two groups would have resulted in an extremely small sample size. The results of the pilot study indicated most students had a low level of knowledge regarding uncertainty analysis. This has also been the experience of this researcher over the last twenty years of teaching freshman level physics laboratory courses.

The literature also supports the fact that this problem still exists in entry- level physics students continuing to state their frustrations with the interpretation of laboratory results. There is little support for the idea that students will learn this concept indirectly from completing laboratory assignments. Therefore, exposure to the intervention that instructs students using algebra is likely to be correlated with a change in test scores.

Design of the Manual for the Intervention

Prior to completing Phase One of this study it was necessary to develop the intervention in the form of a manual and to develop the pretest and posttest. These are discussed first, prior to discussing the design of the two phases of the study. The researcher developed a manual with the goal of teaching uncertainty analysis based on algebra and arithmetic concepts rather than on calculus concepts. The title of the manual is *The Worst Case Method Applied to the Propagation of Uncertainties: Supplement to Introductory Physics Laboratory* (A copy is provided in Appendix A). The example that follows illustrates the Worst Case Method.

Pendulum Example of the Worst Case Method

The pendulum laboratory provides an example of the issues facing students and teachers when applying uncertainty analysis to data. Also it is a good example of the application of the Worst Case Method in determining agreement between predicted and experimental values. In this example, predicted value refers to values that are obtained using formulas derived from physics. The experimental value refers to values obtained through measurements in the laboratory setting. A mass connected to a string swings back and forth, and the time for this mass to swing back and forth is the period of a pendulum. Students calculate the predicted time (T) using the formula, and then they measure the time directly, using a stopwatch. The student then compares the two values and determines if the values agree with each other. The formula to predict the period of the swing (*T)* is given as:

$$T = 2\pi\sqrt{\frac{l}{g}}$$

In the above formula, l is the length of the string and g is the constant for the acceleration of surface gravity. Students measure the l (length) of the string using a meter stick. They then insert l into the above formula and obtain the predicted value of T (period). This is the predicted value for T. Students then use a stopwatch to measure T, the time it takes for the of one swing of the pendulum. They compare the predicted T to the experimental T and determine if they are in agreement. There is generally some discrepancy and students usually respond by making excuses, blaming the equipment, or simply stating the old standby, “the values are close enough; the discrepancy is due to human error”. Students must make a judgment as to whether or not these two values agree, and in order to make a scientific conclusion they must first determine the uncertainty in each quantity. This is precisely where the Worst Case Method provides a simple way to compare values quantitatively.

The next section explains how the Worst Case Method is used to determine the uncertainty in the predicated period, using the formula, followed by the identification of the experimental period and its uncertainty.

Calculation of the predicted period (T). In the above formula, l is the length of the string, and g is the acceleration of surface gravity. Obtaining the predicted value requires the students to measure l, the length of the swing. For example, the student may measure the length as 50.6 cm. Some uncertainty in l must be assigned, as it is impossible to obtain an exact value for l. This is based on the fact that the meter stick is only accurate to a certain point. There are not an infinite number of lines on the meter stick, and the eye can only resolve the lines to a certain point. The uncertainty is determined by the instrument specifications and by the experience of the experimenter. In this example, the uncertainty might be $\pm$ 0.5 cm or by convention, $\delta l = \pm$ 0.5 cm. The Greek letter lower case delta, δ , is used to represent the uncertainty in any quantity. Therefore, if a student measures l = 50.6 cm, this measurement has an uncertainty of $\pm$ 0.5 cm. The measurement should include a range based on the uncertainty or l = 50.6 cm $\pm$ 0.5 cm. This describes a range of values from 50.1 cm to 51.1 cm. The range implies that upon repeated measurements of the length, the values will lie between 50.1 cm and 51.1 cm.

Therefore, any physical quantity capable of being measured actually consists of three values. The first is called the best estimate or the measure and as defined by ISO (Pillay et al., 2008). This is the value determined by the instrument and the experience of the person making the measurement to be the best estimate of the true value, and in the above example it is 50.6cm. The other two values are obtained by adding and subtracting the uncertainty from the best estimate. In the above example, the range is 50.1 cm to 51.1 cm. When there are three different values for the length, a reasonable question at this point is which value to enter into the formula to obtain the predicted period (*T).* The short answer is all three, and this is precisely where the Worst Case Method provides a simple way to propagate uncertainty into a formula. When there are three values for the length, there will be three values for the predicted period (*T*) generated from the formula: the best estimate and the minimum and maximum values. This method involves four simple steps. The first one determines the best estimate. The next two steps simply determine the range of the period (*T),* while the last step identifies the uncertainty in the standard form of $\pm \delta T$. The four steps that follow illustrate the process of propagation of error as demonstrated by the Worst Case Method.

1. Using the formula for the predicted period enter the best estimate for the length or *l*= 50.6 cm. This derives the best estimate of the period or T_{BE} :

$$T_{BE} = 2\pi\sqrt{\frac{(50.6)\ cm}{981\ cm/s^2}} = 1.419921\ s$$

2. Using the same formula for the period enter in the maximum possible value of the length or *l*=(50.6 + 0.5)cm = 51.1 cm. This gives a maximum period or T_{max} :

$$T_{\max} = 2\pi\sqrt{\frac{(51.1)\ cm}{981\ cm/s^2}} = 1.4340\ s$$

3. Using the same formula for the period plug in the minimum possible value of the length or *l*=(50.6 – 0.5)cm = 50.1 cm. This gives a minimum possible period or $T_{\min}$:

$$T_{\min} = 2\pi\sqrt{\frac{(50.1)\ cm}{981\ cm/s^2}} = 1.426\ s$$

We now have a range for the period, *T,* based on the range of the length. The range of possible values for the predicted period is 1.426 s to 1.4340 s. It is understood that upon repeated measurements of the length, the calculated period will lie between 1.426 seconds and 1.434 seconds. Note that the number of decimal places in all three values is different. The last step determines how many decimal places are justified in the final answer and also the uncertainty in the period.

4. To determine the uncertainty in the period, identified as δT, we simply subtract the minimum from the maximum and divide the result by two:

$$\delta T = \frac{1.434 - 1.426}{2} = 0.008\ s$$

This is the uncertainty in the best estimate for the predicted period. According to Taylor (1997), all uncertainties should be rounded to one significant figure. To clarify this point, consider an uncertainly of ± 0.0082 seconds. The value is uncertain in the thousandths place by ± 0.008, so how can one know anything about the ten thousandths place? The best estimate of the period, based on the calculator, was 1.419921 seconds, but the uncertainty of 0.008 seconds limits the best estimate to three decimals. The predicted period is reported as the best estimate along with its uncertainty. The answer and its uncertainty must have the same number of decimal places Taylor (1997). The best estimate of the predicted period along with its uncertainty is: 1.419 ± 0.008 seconds. We say the uncertainty in the length of $\delta l = \pm$ 0.5 cm has propagated into an uncertainty of $\delta T = \pm 0.008$ s into the theoretical period.

It should be noted that since the production of this thesis a new technique on how to determine the uncertainty, in a predicted value based on a formula, has been developed by this researcher that requires only half the steps outlined above.

Experimental Period and its Uncertainty

When compared to calculating the predicted value, this determination is simple. This is a direct measurement of the period (*T)* using a stopwatch. This uncertainty is not the result of propagation of error and no formulas are required to obtain it. The uncertainty in a digital stopwatch is provided directly from the manufacture and experience of the user and in this case is $\delta T \pm 0.007$. In this example the student makes a measurement of the period and reports its value as 1.431 seconds. The experimental best estimate of the period (*T)*, along with its uncertainty, is 1.431± 0.007 seconds.

Both the predicted and experimental periods and their associated uncertainties have been determined and are presented in Table 3.1.

Table 3.1

Predicted and Experimental Periods

	Period (seconds)	Uncertainty in period (seconds)
Predicted (formula)	1.419	0.008
Experimental (stopwatch)	1.431	0.007

Taylor (1997) states that two values are in agreement if their uncertainties overlap. In other words, if the best estimate of the experiential value is found to be inside the range of the predicted value the quantities are said to agree within uncertainty. When this occurs the student can be confident in his or her conclusion concerning agreement between the values. The minimum experimental value for (*T)* is 1.419 - 0.007= 1.424 seconds. The maximum predicted value of (*T)* is 1.419 + 0.008 = 1.427 seconds. These two values overlap, and there is agreement between the predicted and experimental *T.* The use of error bars helps illustrate this as shown in figure 3.1. The horizontal axis is the time scale in seconds. The lines, extending to the left and right sides of the data marker, represent the plus and minus values of 0.007 and 0.008 seconds of uncertainty on each side of the experimental and predicted values of the period.

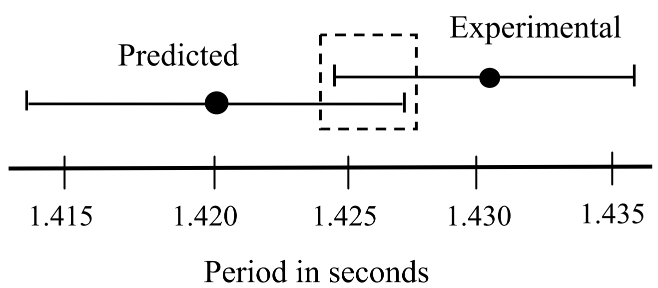


Figure 3.1 The dashed rectangular box indicates the degree of overlap between the upper error bar of the predicted period and the lower error bar of the experimental period.

In summary, the Worst Case Method provides a way to determine the quality of a measurement by assigning limitations on its implied accuracy. More specifically the uncertainty in a value determines the number of significant decimal places to retain.

Once the uncertainties have been assigned to the values, measured or calculated, the Worst Case Method provides the student a way to determine if the values are in agreement. This is perhaps the most significant benefit of this method because the student can now present a conclusion based on analysis rather than on opinion. The student no longer feels compelled to invoke the human error excuse because, in essence, the Worst Case Method quantifies the limitations humans have in their ability to measure. The entire structure and content of the manual is similar in presentation to this example and is provided in its entirety in Appendix A.

Development of the Pretest and Posttest

A review of the literature revealed that there was no specific instrument available to use in the study to assess knowledge about uncertainty analysis. Subsequently, a pilot study was completed to assess student knowledge of measurement uncertainty, concepts of averages, and misconceptions in student reasoning and to use these responses to develop a pre and posttest for the study. The pilot study questionnaire was structured to gather information from students that would aid in the construction of a two-tier design questionnaire. The two-tier design requires the student to select an answer from a list of responses and then to provide a reason for choosing that answer. Students were asked questions about how to interpret measurements. The students then selected an answer from a list. The second part of the question asked students for the reason behind their choice. Instead of selecting a reason that was provided, this section was open-ended, thereby eliciting a broad range or responses. The use of a 12 item two-tier questionnaire is based on similar instruments developed by Treagust (1988, 1995), Odom and Barrow (1995), and Tan, Goh, Chia, and Treagust (2002). This two-tier design was a modification of the standard multiple choice diagnostic test and consists of two parts. The first part incorporates a content question based on propositional knowledge followed by two or three answers. The second part is what differentiates the two-tier design from the standard format. This part provided four multiple choice selections that supply the correct answer, distracter items, and the wrong reason if necessary. Treagust (1988) discusses the development of diagnostic tests to be used in the classroom as an effective means for identifying misconceptions held by students in the specific content areas such as covalent bonding and structure, and photosynthesis and respiration. He goes on to explain that although significant research exists regarding students' understanding and/or misconceptions, the application of these findings is not easily integrated in the classroom. Through the use of diagnostic testing, more specific information about the nature of what students actually understand can be realized. The benefit of this approach is that the teacher can now modify the curriculum to focus on correcting the misconceptions identified from the tests. Odom and Barrow incorporated the two-tier design in a study intended to explore students' misconceptions regarding diffusion and osmosis. The findings suggested that this design was an effective tool in assessing students' understanding of diffusion and osmosis and in exposing their misconceptions even after instruction. In addition, Odom and Barrow also recommended that interviews be conducted with students who have taken the test in an effort to judge the clarity of the terminology of each question. The purpose would be to revise the first tier prior to collecting information on the free response questionnaire. In a similar study by Tan et al. a two-tier diagnostic instrument was developed to assess high school students' understanding of inorganic chemistry qualitative analysis. The development of this instrument also involved a free response pilot study and interviews. The final test included alternative concepts based on the pilot study.

Consistent with these researchers a pilot study was conducted for the two-tier test developed for this thesis to identify common alternative reasons based on misconceptions that students had about data analysis and uncertainty in the form of free response questions. Interviews were not conducted as part of the development of this current study, and this is a limitation as interviews would have proved beneficial. A significant number of students included in this study attended night school and had very limited time for anything other than class time and laboratory exercises. The final version of the instrument consisted of 12 questions which appeared to be an acceptable length of these instruments among most researchers; an exception is the 19-item instrument developed by Tan et al. (2002).

Consider the following example shown in Figure 3.2 taken from the pilot study questionnaire. The responses from two different students to the same question are presented.

3) Two students calculate the velocity of a rolling ball by dividing distance by time. The distance traveled is 3.1 meters and the time is 1.3 seconds. The calculator displays the answer to seven decimal places. Student A reports a value of 2.4 m/s while student B reports a value of 2.384615 m/s.

a) The value 2.384615 m/s is more accurate than 2.4 m/s
b) The values are equally accurate.

The reason I selected this answer is because:

~~[illegible]~~
Student B's value is not more accurate because it has more decimal places. Student A's answer accounts for precision of instruments by maintaining significant figures.

Figure 3.2 Student selects the correct answer and provides a reason acknowledging the limitations of combining measurements.

A different student responded to the same question, as illustrated in Figure 3.3.

a) The value 2.384615 m/s is more accurate than 2.4 m/s
b) The values are equally accurate.

The reason I selected this answer is because:

Most everything in Life is rounded & to have 7 decimal places is overdoing it. I would write 2.4.

Figure 3.3 Student selects the correct answer but the reason is more of an opinion.

The responses are at extremes to each other. The first student recognizes the importance of both significant figures and that these are crucial due to the limitations of the instruments. The second student simply gives his opinion based on common sense and on what is typical of his experience with no reference to quantitative thought.

All the students' rationales were categorized based on their level of understanding. The numbers following the category refer to questions from the test. The Pilot Study Test is included in Appendix B. The student responses fell into three fundamental categories:

1. Responses that relate to students' misconceptions about human error (questions 1, 6, 7, 9, 12).

2. Responses that address students' understanding of repeated measurements, average, and margin of error (1, 3, 4, 5, 7, 8, 10, 12).
3. Responses that reflect students' quantitative ability (2, 3, 7, 8, 10, 11).

In general a majority of students responded according to point reasoning vs. set reasoning. Lubben, Campbell, Buffler, and Allie (2001) describe the student with point reasoning as one who thinks of data as being independent of each other and believes that any single measurement could in fact be the true value. In this same study these researchers identify students who exhibit set reasoning as those who interpret a single measurement as an estimate of the true value and any deviation is considered random. After analyzing students' responses and collating them in terms of understanding, multiple choice questions were developed to assess students' rationale in terms of level of understanding. A few students demonstrated a higher level of understanding and provided reasons that were useful in writing answers for the multiple choice answers. Many students gave incorrect reasons, but these were valuable distracters in the final questionnaire. Special attention is given to those responses that occurred frequently and were of similar content. These responses provided a baseline for identifying misconceptions about uncertainty analysis.

Phase One: The Pretest and Posttest

A 12 item pretest and posttest was developed from the Pilot Study utilizing the two-tier design and is defined as a One-Group Pretest Posttest format (Salkind, 1997). As mentioned earlier, the two-tier design requires the student to select an answer from a list of responses and then to select a reason for choosing that answer. The choices are derived from the results of the pilot study as shown in Figures 3.2 and 3.3. In other words, the purpose of each question is to solicit a response from the students that will identify their knowledge and also provide insight into their reasoning when selecting an answer. This type of assessment is recommended by Treagust (1995). The Pretest and posttest is included in its entirety in Appendix C. A sample question is provided in Figure 3.4 for clarity. In responding to this question, the student selects one answer and then selects a reason for the choice.

1. A mass, attached to a string, swings back and forth as a pendulum. The student makes seven measurements of the time for the mass to swing out and back. The seven measurements are: 1.53, 1.53, 1.60, 1.53, 1.56, 1.53, 1.49 seconds. The true swing time is probably closest to:

   a) 1.53 seconds
   b) the average of all the times

The reason I selected this answer is because:

   1. No one value is more important than the others are as they each have equal merit.
   2. 1.53 seconds occurs four out of seven times.
   3. Averaging is the best method to use with a collection of numbers.
   4. 1.53 seconds, as the other values were probably affected by outside influences.

Figure 3.4 Sample Question from Pre Posttest

On the first day of their first laboratory exercise, students were asked to complete the two-tier pretest. No instruction on uncertainty analysis was provided prior to the test. All students completed the test without a time limit, and all finished within an hour. After completing the pretest, students received the manual on uncertainty analysis with instructions for use. As an incentive, students were advised of the ability to earn more points on laboratory scores if they demonstrated and included an understanding of uncertainty analysis in their written work. Although the primary method of instruction on uncertainty analysis was through the Manual, the researcher included verbal instruction and demonstrations during laboratories throughout the semester. At the end of the semester, students completed the same test as a posttest. The scores were compared with the pretest scores and a t-test was used to determine the significance, if any, for dependent means. The two-tier instrument was also analyzed for validity and reliability.

Phase Two: Evaluation of Student Laboratory Assignments

Phase two evaluated students' application of uncertainty analysis in written laboratory assignments. A total of five laboratory assignments were included in the study: Laboratory A, B, C, D, and E, and a list of these laboratories is included in Appendix D. Each laboratory assignment comprises the following sections: Objective, Procedure, Data Analysis, Results and Conclusions. The latter two sections are evaluated for the application of concepts of uncertainty analysis and for the sophistication of the students' level of understanding.

Phase two consist of two parts; part one evaluates the laboratory reports according to Six Facets of Understanding developed by Wiggins and McTighe and part two examines the reports in terms of conceptual change based on the research by Hewson (1996) ,and Venville and Treagust (1997).

Part One: Facets of Understanding

The method for assessing the level of understanding is based on the Rubric for the Six Facets of Understanding developed by Wiggins and McTighe (1998). The levels are assigned a value from 1 to 5, with a score of 5 demonstrating more advanced understanding. The Facets of Explanation is utilized for Phase Two, as this category is the most objective and quantitative method to assess the level of understanding. Wiggins and McTighe's rubric of Explanations is utilized to evaluate students' level of understanding in the laboratory reports. The categories of this rubric of understanding along with a description of categories are listed in Table 3.2.

Faculty members in physics, chemistry, and mathematics at the college developed the following evaluation criteria based on areas determined to be most important in demonstrating students' understanding of data analysis and uncertainty in the science laboratory. The four criteria are: a) identifies uncertainty accurately, b) performs mathematics correctly, c) reports uncertainty in the correct form, and d) uses uncertainty to support conclusions. Each of these criteria was evaluated according to Wiggins and McTighe's rubric for students' level of understanding and was scored with a value from 1 to 5. A score of '1' corresponds to "naïve", and a score of '2' with "developed", a score of '3' with "in-depth", a score of '4' with "systematic", and a score of '5' corresponds to "sophisticated and comprehensive", in accordance with the rubric. A score of '0' indicates that the student did not include any information on uncertainty analysis.

Table 3.2

Level of Understanding of the Facet of Explanation

Level of Understanding	Description
Sophisticated and	"an unusually thorough, elegant, and inventive account (model, Comprehensive theory, or explanation); fully supported, verified, and justified; "an unusually thorough, elegant, and inventive account (model, theory, explanation); fully supported, verified, justified; the deep and broad; goes well beyond the information given"
Systematic	"an atypical and revealing account, going beyond what is obvious or what was explicitly taught; makes subtle connections; well supported by argument and evidence; novel thinking displayed"
In-Depth	"an account that reflects some in-depth and personalized ideas; the student is making the work his own, going beyond the given; there is supported theory here, but insufficient or inadequate evidence and argument"
Developed	"an incomplete account but with apt and insightful ideas; extends and deepens some of what was learned; some reading between the lines; account has limited support/argument/data or sweeping generalizations; there is a theory with limited testing or evidence"
Naïve	"account; more descriptive than analytical or creative; a fragmentary or sketchy account of facts/ideas or glib generalizations; a black-and-white account; less theory than an unexamined hunch or borrowed idea"

Wiggins & McTighe, 2005, pp. 178-179.

An example of how the first laboratory assignment was scored for the first five students is shown in Table 3.3. Thirty-six students submitted five laboratory reports during the semester. As discussed earlier, each student received a score from 1 to 5 for each of the criteria evaluating level of understanding. For example, student number 1 received a score of 1 (naïve) for "identifies uncertainty accurately", a score of 4 (in-depth) for "performs mathematics correctly", a score of 3 (developed) for "reports uncertainty in the correct form", and a score of 3 (developed) for "uses uncertainty analysis to support conclusion". This same system was used for each of the 36 students for each of the five laboratories selected for the study.

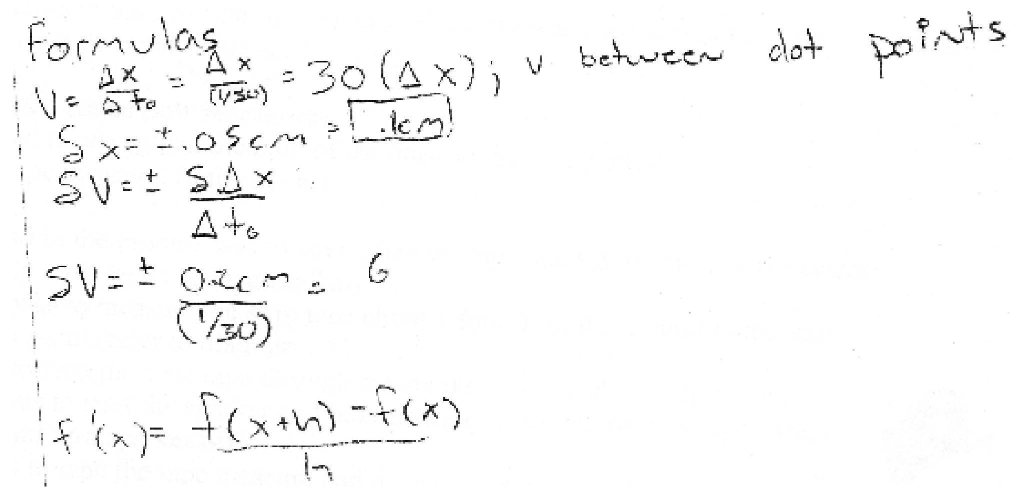
Table 3.3

Sample of Student Scores on Five Criteria of Understanding on Laboratory A

Five Criteria		Student/Scores				
		1	2	3	4	5
a)	Identifies uncertainty accurately	1	5	0	1	5
b)	Performs mathematics correctly	4	2	0	4	4
c)	Reports uncertainty in the correct form	3	2	0	3	5
d)	Uses uncertainty analysis to support conclusion	3	2	0	3	4

The first criterion assesses a student's level of understanding and application of uncertainty concepts in identifying uncertainty accurately. At the naïve level, the student reports an uncertainty out of context with the laboratory. The student does not connect the number with the data or the laboratory and seems to include the value just because he or she thinks it is supposed to be included. At a more sophisticated level of understanding, the uncertainty value is connected with a variable and the student comments correctly on its relevance. Finally, the student clearly connects the uncertainty value with the variable, its measurement, and the laboratory as a whole.

The following examples illustrate this distinction. The student in the first example scored a '1' in this category, and the student work is shown in Figure 3.5. The second example of student work is illustrated in Figure 3.5, and this student scored a '5'.



Formulas

$V = \frac{\Delta x}{\Delta t_o} = \frac{\Delta x}{(1/30)} = 30(\Delta x)$; v between dot points

$\delta x = \pm .05 cm \approx \boxed{.1cm}$

$\delta V = \pm \frac{\delta \Delta x}{\Delta t_o}$

$\delta V = \pm \frac{0.2cm}{(1/30)} = 6$

$f'(x) = \frac{f(x+h) - f(x)}{h}$

Figure 3.5 Example Category One. This work received a '1' on the scale.

The student's work in Figure 3.5 presents a calculation as a task identified as formulas without any reference to its purpose. The uncertainty is not identified. The value of 6, reported as the uncertainty in the velocity, has no units or indication of the range that should be reported as ± 6 cm/s.

The student's work in Figure 3.6 on the other hand defines the table as uncertainty analysis. The equations are presented in an ordered fashion with units and each step is provided. The units are reported along with the ± notation needed to indicate maximum and minimum values.

Derivation of δv (uncertainty of calculated velocity between tape marks)
$v = \Delta x \div \Delta t$ cm/s
$\Delta v_{max} = (\Delta x_n + \delta\Delta x) \div \Delta t_o$
$\Delta v_{min} = (\Delta x_n - \delta\Delta x) \div \Delta t_o$
$\delta\Delta v = \pm(\Delta v_{max} - \Delta v_{min}) \div 2 = \pm(\delta\Delta x \div \Delta t_o)$
$\delta\Delta x = \pm.2$ cm ; $\Delta t_o = 1/30$ s
$\delta\Delta v$ **= ±6 cm/s**

Figure 3.6 Example Category One. This would score a '5' on the scale.

The second criterion is to perform mathematics correctly in terms of calculating uncertainty. At the naïve level, students may omit the calculation or list a number without showing work. As students progress, the equations are presented with units and in the correct sequence. At the sophisticated level, definitions, comments, and explanations are also provided.

The third criterion is to report the uncertainty in the correct form. Students progress from listing a value with no units and inconsistency in decimal place to accurately listing a value with units and rounding decimals to the correct value. Students also progress to reporting values with accurate significant figures.

The fourth and last category is possibly the most important. The ability to use uncertainty analysis to support conclusions is very important in assessing students' ability to apply uncertainty concepts. At the naïve level of understanding, there is no mention of uncertainly in the analysis. At higher levels of understanding, students mention uncertainty as another part of the assignment but without application to the laboratory results. They fail to recognize the uncertainty in justifying their conclusions. At the more sophisticated levels of understanding, the student employs uncertainty analysis to justify and support a conclusion about the laboratory results. Again the following examples illustrate this category with excerpts from one of the laboratories completed by the students. The purpose of this laboratory exercise was to verify conservation of total mechanical energy or T.M.E. The first student simply stated an opinion with no support from the data or uncertainty analysis. The student completed the uncertainty analysis calculations but failed to employ the values to support the conclusions as indicated by phrases like "fairly accurate" and "didn't line up with others" as illustrated in the following response:

> After seeing the results I believe that the T.M.E. was conserved throughout the pendulum swing. The data received was fairly accurate with the exception of one data point which didn't line up with the others. But with the information found I noticed that when K.E. went up P.E. went down and visa versa. We showed the results from 6 data points to find the uncertainty including the data point that didn't correlate with the other information.

This conclusion scores a '1' on the Wiggins and McTighe scale. Compare the above response with a more sophisticated response that follows.

> The prediction of this lab was that the total mechanical energy of the pendulum would be conserved. The initial T.E. was 0.198 Joules and the final T.E. was 0.193 Joules. This alone was not enough. An uncertainty in the calculation of the total energy had to be calculated to allow room for error. The uncertainty in the P.E. = ± 0.001 J and the uncertainty in the K.E. = ± 0.003 J. This was the average taken from the calculations of every third point. In using these uncertainties in the K.E. and P.E. the uncertainty in the Total energy could now be found. The average was taken from the uncertainties from the same eight points and it produced an uncertainty in the total energy of 0.008 J. The uncertainty of T.E. was applied to the average T.E. which came out to be 0.195J. Since T.E. is 0.195 ± 0.008 J, the initial and final T.E. fall within that range. The experiment agrees with the prediction that the total mechanical energy is conserved.

This student identified each uncertainty and reported each value. The analysis includes the final statement in support of the prediction that total mechanical energy is conserved providing the actual range of allowed values based on uncertainty analysis. This conclusion would receive a '5'on the Wiggins and McTighe scale.

Over the course of the semester, students submitted written laboratory assignments on a weekly basis as part of their coursework, and the researcher assessed each laboratory report for the sophistication of students' application of uncertainty concepts as described above. Therefore, changes in level of understanding are indicated in comparing the weekly scores during the semester. Although ten laboratory reports were collected over the semester only five were used in this study due to the high student drop-out rate. In other words, only five reports were completed by all 36 students who comprise the sample size.

Part Two: Conceptual Change

The final section of the study evaluated the level of conceptual change demonstrated in the same five laboratory assignments mentioned above. The method to evaluate conceptual change was based on the model used by Venville and Treagust (1997). Student comments in the results and conclusion section of the laboratory assignments were evaluated for the level of conceptual change. Responses were categorized as intelligible, plausible, and fruitful, with fruitful being the highest level obtained.

Data Analysis Procedures

Phase One: Pretest and Posttest

The means of students' scores on the pretest and posttest were analyzed for statistical significance. An item analysis of the questions was also completed and evaluated for any changes from point to set reasoning. The pretest and posttest was also analyzed for internal consistency.

Phase Two: Evaluation of Laboratory Assignments

Five laboratory assignments completed during the semester were analyzed for level of understanding using the rubric of Facets of Understanding developed by Wiggins and McTighe (2005). These assignments were also evaluated for conceptual change using Hewson's (1996) criteria of intelligible, plausible, and fruitful.

Limitations

One limitation of the study is the sample size of 36. It is difficult to generalize the results to a broader population with the smaller sample size. Another limitation in this study is that the pretest and posttest were written by the researcher for this study. There was no test available that was previously tested for internal and external validity. The two-tier design of the questionnaire, however, is a previously utilized method for assessing student responses (Treagust, 2006). The level of understanding section is based on the research of Wiggins and McTighe (1998). Therefore, even though the pretest and posttest is a new design, the basis of the design of questions is derived from validated methods.

Difficulty in assessing entry level knowledge is a third limitation. Although one assumption of the study is that students enter the class with minimal to no knowledge of uncertainty analysis, there were some students who entered with some knowledge.

Finally, the limitation of the teacher as researcher is evident. Although this is not the ideal situation, the methods to assess students provided more objectivity to the results. The first part of the pretest and posttest assess factual knowledge about uncertainty analysis. The students either know the answer or they do not. The rationale section is based on an established method of determining level of understanding. This same method was also employed in Phase Two of the study to assess students' application of their knowledge.

Ethical Issues

The researcher advised all students of the purpose of the study and of the confidentiality of the data. Each student received a letter in regards to this and signed and returned the letter, which incorporated a permission to participate form. A copy of the letter and permission are included in Appendix E. Student scores remain anonymous and all individual data were kept confidential.

Overview of the Chapter

This chapter covered the methodology and research design of the study. The different methods for teaching uncertainty are outlined, ranging from the more advanced techniques utilizing calculus and statistics to the use of percent difference and significant figures. Supporting arguments in favor of using the Worst Case Method over these other methods and an example demonstrating the simplicity of this method is provided. An overview of the development of the manual using the Worst Case Method follows, followed by a description of the development of the pre and posttest for Phase One. The pilot study for the pre and posttest is explained along with the research concerning two-tier questionnaires.

Phase Two incorporated the facets of understanding as defined by Wiggins and McTighe (2005) and used as the rubric for scoring laboratory reports. The five levels of understanding were discussed, including student examples of the levels. Finally, the methods utilized for data analysis and the limitations and ethical issues are presented.

CHAPTER 4

RESULTS

Chapter four presents the data for both phases of the study. Phase One comprises the analysis of the two-tier questionnaire. The pretest was administered at the beginning of the semester followed by the posttest given on the last day of class. Phase Two evaluates student understanding of uncertainty analysis based on laboratory work. Students submitted laboratory assignments throughout the semester, and these assignments are analyzed for level of understanding uncertainty analysis and for conceptual change. The following outline provides a more specific description of the contents of Phase One and Phase Two.

Phase One: Pretest and Posttest

Part One includes an analysis of student performance on the two-tier questionnaire. Mean student scores were compared, and the statistical significance for dependent means was determined.

Part Two includes an item analysis of each test question. Each of the 12 items on the pretest and posttest was analyzed for change following the instructional intervention. Student understanding is analyzed for a movement from point to set reasoning as defined by Lubben, Campbell, Buffler, and Allie (2001).

Phase Two: Evaluation of Student Laboratory Assignments

Part One: Laboratory assignments were examined for the mathematical ability of students to implement concepts in the manual for propagating uncertainties. These were scored from 1 – 5 utilizing Wiggins and McTighe’s facets of understanding (1998).

Part Two: Laboratory reports were evaluated for evidence of conceptual change according to the criteria established by Hewson (1992): intelligible, plausible, or fruitful. This method was employed by Venville and Treagust (1997) to analyze conceptual change of science students.

Phase One

The purpose of Phase One was to measure students' knowledge regarding topics of measurement, human error, and data manipulation following instruction on the Worst Case Method used for calculating uncertainty and propagation of error. A pretest was administered on day one of the semester followed by a posttest at the end of the semester. Phase One was guided by the following research questions:

1. To what extent can the concepts of uncertainty analysis be successfully taught to entry-level students possessing only a background in elementary algebra.
2. Over the course of a typical college physics laboratory class, do first year students' concepts of uncertainty compare from the beginning to the end?

Twelve questions comprised the two-tier questionnaire and each item consisted of two components. The first component required students to select the correct answer, while the second component required students to provide a reason for their answer. Phase One was a one group pretest and posttest design. According to Salkind (1997), this type of design involves administering a pretest, employing an intervention, and following with posttest. Both the pretest and posttest was given to the same group of 35 students who completed both the pretest and posttest. Students had 45 minutes to complete each test, and they were allowed to use calculators. Data from the pretest and posttest were analyzed according to two different parameters. In both parameters a response was considered correct only if the student answered both parts correctly. The response to a question was incorrect if either tier was incorrect.

Part One: Student Performance on Pretest and Posttest

Pretest and posttest scores were compared by determining the mean score for the 35 students. The results are included in Table 4.1.

Table 4.1

Part One: Mean Student Scores answered correctly out of 12 questions.

Group	n	M	SD
Pretest	35	4.6	2.0
Posttest	35	6.5	2.3

The pretest mean was 4.6 questions answered correctly, and the posttest mean was 6.5 correct answers. In relative terms, the pretest average was 38.3%, and the posttest average was 54.2% with a relative percentage gain of 41.3% (54.17 – 38.33)/ 38.33) x 100). A one tailed *t* -test for dependent means (with df = 34) of 7.94 showed that the means are statistically significantly different at the p = 0.0001 confidence level.

Figure 4.1 shows the individual scores for each of the 35 students as a percentage of the total possible scores on both the pretest and posttest. The values are plotted in descending order starting with the largest gain of posttest over pretest ending with smallest gain. A total of 52 students participated in the pretest, and each student was originally assigned a number from 1 to 52. Seventeen students dropped the course before the posttest was given, leaving 35 who completed both parts of the test. Students kept their original number assignment, therefore there are missing numbers on the graph and numbers higher than 35, but the total number of students remained at 35.

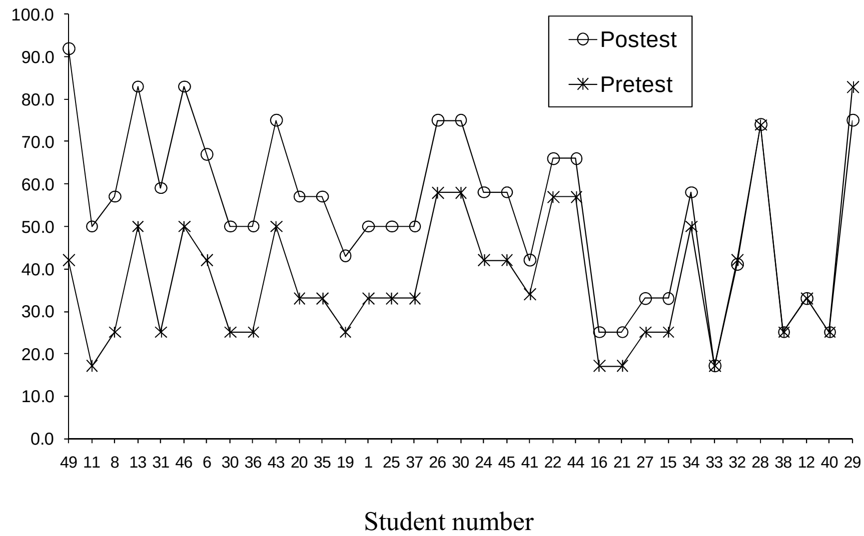


Figure 4.1 Student Pretest and Posttest Scores as a Percent of the Total vs. Student Number

Student 49 showed the largest gain in performance after receiving the intervention, with an increase from approximately 40 to 90 percentage points. The next eleven students, 11-35, showed a significant increase of about 35 percentage points. Following this, seven students, 19-24, showed a modest increase of about 18 percentage points. The scores for students' numbers 16, 21, and 27 improved slightly. The scores for students near the end of the graph, 33 through 29, were the same before and after the intervention. The last two students, 40 and 29, showed an eight percentage point decrease in performance.

The results of the questionnaire were also tested for internal consistency, which produced a Cronbach Alpha Coefficient of 0.28 for the pretest and 0.48 for the post test. It is generally accepted that scores below 0.70 indicate the instrument is not measuring a single concept (Salkind, 1997). Although this is not uncommon for a new instrument, further evaluation of instruments in future research could improve the internal validity. As shown in Table 4.1, the mean student scores answered correctly is less than 50% on the pretest and a little over 50% on the posttest, so the there is a large percentage of students who have the incorrect responses. However, as would be desired there was increased consistency on the posttest compared to the prestest. The low percentages correct alone will lead to a low reliability value combined with a low number of participants. This level of internal validity makes it difficult to generalize the findings. A more in-depth discussion of the validity is included in the discussion section of this study, along with recommendations for further research.

Part Two: Item Analysis of Test Questions

Part Two examines how students performed on each of the 12 questions on the pretest and posttest. Each item is analyzed for change following the instructional intervention. A question was considered correct only if both tiers were answered correctly: the answer component and the reason component. A question was considered incorrect if the student answered either tier incorrectly. The purpose was to rank each question in terms of pretest and posttest gains

The questions were originally catagorized according to the following three different content areas and the results are presented in Table 4.2.

1. Responses that relate to students' misconceptions about human error (questions 1,6,7,9,12).
2. Responses that address students' understanding of repeated measurements, average, and margin of error (1,3,4,5,7,8,10,12).
3. Responses that reflect students' quantitative ability (2,3,7,8,10,11).

Table 4.2

Total Number of Correct Responses of the 35 Students in the Three Categories

	Human Error	Repeated Measurement	Quantitative Ability
Pretest	59	126	77
Posttest	84	172	102
Percent Change	42.4	36.5	32.5

The total scores were higher for the categories of repeated measurements and qualitative ability, because these categories contained eight and six questions respectively versus only five for the human error category. The percentage gain for the category involving human error was 42.4 % and for the category of repeated measurements was 36.5 %, followed by 32.5 % for the quantitative ability category. The percent gain for all 12 questions without considering categories was 33.9 %. When compared to the test as a whole the performance in each of the three categories is essentially the same. This is not surprising, because several questions overlapped into more than one content area. After closer examination, the majority of questions were identified as either point or set paradigm type items as characterized by Lubben et al. (2001). Therefore the items were analyzed according to how the questions were answered in terms of point or set reasoning and not on the three categories.

On a typical multiple choice test, consisting of four to five questions, satisfactory understanding is achieved if 75 % of students answer a question correctly (Gilbert 1977). The chances of guessing the correct answer on a standard four item multiple choice test is 25% and only 12.5% on the two-tier design when both components are considered. The standard of 75 % was used in this study to judge each question. Table 4.3 summarizes the results of the pretest and posttest.

Table 4.3

Percentage of Students Who Answered Both the Answer and Reason Components Correctly For Each Question.

Item number	Pretest	Posttest	Percentage Point Gain
1	51.4	82.8	31.4
4	31.4	60	28.6
8	22.9	42.9	20
5	71.4	88.6	17.2
2	54.3	71.4	17.1
3	22.9	37.1	14.2
6	5.7	17.1	11.4
12	25.7	34.3	8.6
10	71.4	74.3	2.9
7	8.6	11.4	2.8
9	40	40	0
11	20	20	0

The pretest scores ranged from 8.6% to 71.4% with no scores of 75% or over. For the posttest, the lowest score is 11.4% and the highest is 88.6% with only two questions scoring at or above the minimum score of understanding of 75%. The largest gains appeared in items 1, 4, 5 and 8. Items 2, 3 and 12 indicate an average gain of 12.8%. The pretest and postest average scores for items 6 and 7 are less than 12.5% . Scores in this range are most likely the result of quessing the answer. Items 9,10, and 11 saw no appreciable gains. Figure 4.2 shows the results of Table 4.3 in terms of the percentage point gains made for each question in decreasing order.

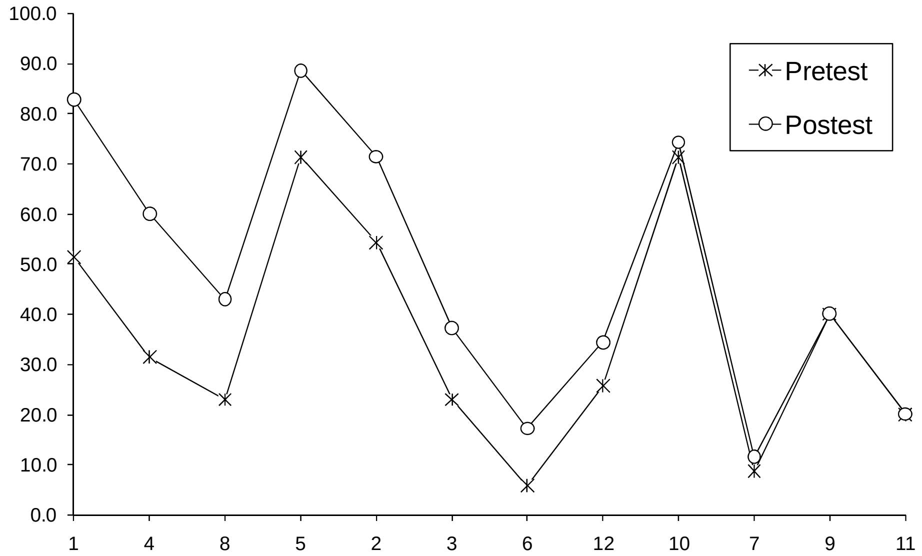


Figure 4.2 Percentage Point Gains in Decreasing Order

Each item was examined in terms of how students think about data. The purpose was to identify students as either point thinkers or set thinkers according the definitions outlined by Lubben et al. (2001). In order to facilitate the analysis, each item is presented in its entirety in the same order they appear in Figure 4.2: 1, 4, 8, 5, 2, 3, 6, 12, 10, 7, 9, and 11.

Question 1

A mass, attached to a string, swings back and forth as a pendulum. The student makes seven measurements of the time for the mass to swing out and back. The seven measurements are: 1.53, 1.53, 1.60, 1.53, 1.56, 1.53, 1.49, seconds. The true swing time is probably closest to:

a) 1.53 seconds
b) the average of all the times

The reason I selected this answer is because:

1. No one value is more important than the others are as they all have equal merit.

2. 1.53 seconds occurs four out of seven times.
3. Averaging is the best method to use with a collection of numbers.
4. 1.53 seconds, as the other values were probably affected by outside influences.

The correct answer is ‘b’ for content and ‘3’ for the reason. For convenience I will refer to this combination as b3, which was utilized by Tan (2002) and continue in this manner when discussing answers to questions. In the pretest, only 3 (8.6%) of the 35 students answered the question correctly, and in the posttest 7 (20%) answered it correctly. Answer b2 is close and represents set paradigm thinking; therefore the correct answer was expanded to include either b2 or b3. Under this criterion the pretest score was 18 (51.4%), and the posttest score was 29 (82.9%), which indicated a posttest gain of 31.4 percentage points. In the pretest 16 (45.7%) students selected the distracter combinations a2 and a4, while in the posttest 6 (17.1 %) students chose the same distracters. Based on these data it appears that 10 (28.6 %) students selected the answer based more on set reasoning rather than point reasoning.

Question 4

Two students calculate the velocity of a rolling ball by dividing distance by time. The distance traveled is 3.1 meters and the time is 1.3 seconds. Using these numbers, the calculator displays the answer to seven decimal places. Student A reports a value of 2.4 m/s while student B reports a value of 2.384615 m/s. The value 2.384615 m/s is more accurate than 2.4 m/s

a) I agree with this last statement
b) I do not agree with this last statement.

The reason I selected this answer is because:

1. More decimals imply higher degree of accuracy.
2. 2.4 m/s is the most accurate. The remaining decimals are not based on physical measurement.

3. The value of 2.384615 m/s rounds off to 2.4 m/s.

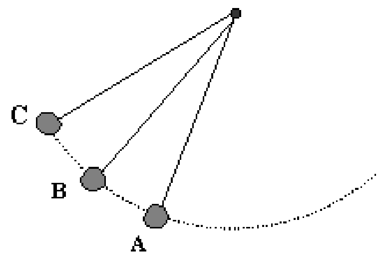
4. 2.384615 m/s has more significant digits.

The best answer to question 4 is b2. The selection of b3 might also be considered correct but the pilot study showed students who selected this did so out of procedural knowledge quoting that: “rounding off is always the right thing to do”. The purpose of this question was to expose the misconception that more decimal places always mean greater accuracy regardless of the accuracy of the arguments. Students selecting b2 recognized the limitation placed on the answer was more related to the accuracy of the values and not on the power of the calculator. The pretest and posttest scores were 11 (31.4%) and 21 (60%), respectively. The distracter combinations a1 and a4 were selected by 23 (65.7) students on the pretest and 7 (20%) on the posttest. These 23 students indicated in the pretest they believed that more decimal places in the answer were more accurate regardless of the accuracy of the initial values stated in the question. Judging single values in this manner, without considering all the data, was consistent with the point paradigm model (Lubben et al. 2001; Allie & Buffler 2003). The posttest results showed that 10 (28.6%) students changed to set thinkers after the instructional intervention.

Question 8

Referring to the same experiment above, the student now wonders if the swing time depends on how far back the mass is before it’s released. Three different positions are tested. The results are shown below. The margin of error in the stopwatch is ± 0.02 seconds.

Position	**Swing time (seconds)**
A	1.27
B	1.29
C	1.31



a) The swing time does depend on where the mass is released.

b) The swing time does not depend on where the mass is released.

The reason I selected this answer is because:

1. The overall trend is in the correct direction of increasing swing time
2. The increase in the time is the same between A and B and B and C
3. All three numbers round-off to the same value of 1.30 seconds.
4. There is some overlap of the margin of error for each time.

The correct answer to item 8 is b4. Although positions A and C only touch each other and do not actually overlap, selecting 4 as the reason is still the best answer in that it states there is *some* overlap of the margin of error for each time, indicating some knowledge of uncertainty. Eight students (22.8%) responded correctly in the pretest and 11 (31%) in the posttest. The most significant distracter is item a1. Selecting this item illustrated the misconception that students have about trends in data. Each measured value of time has an associated uncertainty of ± 0.02 seconds, resulting in significant overlap among the three times. Sixteen (45.7%) students chose this in the pretest suggesting that these students considered each time as a point independent of any associated range. This same distracter was selected by only 5 (14.3%) students. It appears that after the intervention 11 (31.4%) of these 16 students think of these data as a set related to each other through uncertainty, indicating set point thinking

Question 5

It is a fact that the acceleration of gravity is 9.81 m/s^2. A student, using very good equipment, performs an experiment four separate times. The results are as follows: (9.79, 9.82, 9.84, 9.78) m/s^2. The student should keep trying until he obtains the accepted value of 9.81 m/s^2

a.	I agree with this statement

b.	I do not agree with this statement.

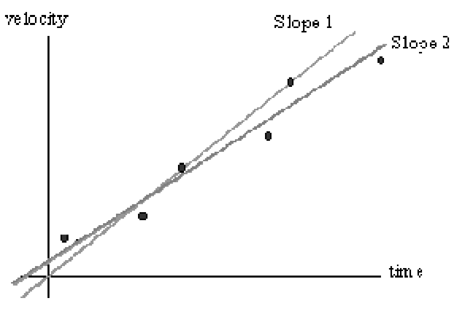
The reason I selected this answer is because:

1. It is not necessary to determine the exact value because all measurements have uncertainty no matter how careful you are.
2. The second value, 9.82 m/s^2, is very close to the accepted one.
3. The student needs to determine what is wrong and correct the problem.
4. Gravity is very constant and the experiment should produce the accepted value.

The correct response for this item is b1. In general students performed well on this item with 25 (71.4%) on the pretest and 31 (88.6 %) on the posttest, being correct, illustrating set point thinking. According to Gilbert (1977), a score > 75 % indicates a satisfactory level of understanding. In the pretest 6 (17.1 %) students selected the alternative combinations b2 and a3, indicating point paradigm reasoning followed by 3(8.6%) students on the posttest.

Question 2

Data is collected on the velocity and time of an accelerating car. The data is plotted on the graph below. The slope is the acceleration of the car. Two possible slope lines are drawn in. Which slope has the smaller margin of error.



a.	slope 1

b.	slope 2

The reason I selected this answer is because:

1. slope 1 goes through more actual data points.
2. slope 1 goes through the origin
3. slope 2 has equal number of points above and below the line
4. slope 2 is closest to most of the points

The correct answer is b4. The pretest and posttest results showed no appreciable difference for this item. Of the 35 students, 31.4 % answered correctly on each part of the test. The next most popular combination, b3, was closely related to the correct answer b4 in that each of these describe the data as they relate to each other without assigning too much weight to any single data point. This combination, b3, was selected by 9 students (25.7%) on the pretest and 15 students (42.8 %) on the posttest. These students seemed to recognize the significance of the data as a set with no one value having more relevance over another. In contrast to b3 and b4 the combination, a1 describes the line as intersecting only two points challenging students with the misconception that all points should line up perfectly on the line. Still twelve students (34%) viewed this response as the closest comparison to what they believed. Selecting a1 indicates point thinking. Only 6 students (17%) selected A1 in the posttest suggesting that these students may be thinking about measurements as part of a set.

Question 3

A student estimates the length of a room by walking from one side to the other. The distance between each step is one meter ± 2 %. She reports the length as 10 m. The margin of error in the length of the room is

a. ± 20 %
b. still ± 2 %

The reason I selected this answer is because:

1. the $\pm$ 2 % margin of error is constant for the whole distance measured.
2. the total distance of 10 meters has a margin of error of $\pm$ 2 % which equals $\pm$ 20 %
3. each step has an error of $\pm$ 2 % and therefore limits the error in the length to $\pm$ 2 %
4. the maximum error results from compounding the error in each step.

This question attempts to identify those students who have some experience in calculating uncertainties based on the Worst Case Method. The correct answer is a4. This assumes that the 2% uncertainty is added for each step taken. Statistically this is not what would actually happen but is one of the assumptions made with the Worst Case Method employed by this study. In the pretest, 8 (22.8 %) answered correctly and 11 (31.4 %) answered correctly in the posttest. This question tests whether a student is able to propagate uncertainty correctly. This aspect remains a challenge for students in that the posttest score showed only three additional students were able to correctly propagate the uncertainty over the pretest.

Question 6

Human error is inherent in all scientific endeavors and impossible to eliminate.

a) I agree with this statement
b) I do not agree with this statement.

The reason I selected this answer is because:

1. use of extremely accurate instruments can eliminate human error.

2. Humans are not perfect and sometimes make errors that affect the results.
3. Eliminating mistakes will eliminate human error.
4. Usually there is a discrepancy between experimental and accepted values. This is evidence of human error.

The correct response is b3. Only two students selected this on the pretest and six students on the posttest. Over half the students (57.1%) selected either a2 or a4 indicating that human error is an acceptable explanation as to why experimental results may not support the theory. This misconception is also acknowledged by Buffler et al. (2009) in their workbook in which they comment on the frequent use of human error to explain results.

The alternate combinations a1 and b1 express the misconception of connecting human error and inaccurate instruments. Approximately 34 % of students selecting these answers exhibited the notion that some perfect exact value must exist, provided the instrument was more precise; this response again exhibits point reasoning. This misconception persisted in the posttest showing no appreciable change.

Question 12

A student runs an experiment designed to determine the density of water. Performing the same experiment twice the results are: 1.11 gm/cm^3 and 0.93 gm/cm^3. The accepted value for the density of water is 1.00 gm/cm^3. Assuming no mistakes were made, I believe

a. the equipment used is probably **not** good enough to determine the density of water
b. the equipment is probably good enough to determine the density of water

The reason I selected this answer is because:

1. the values are too different from each other
2. good enough- the difference is just human error

3. must make many more runs and average the values
4. Probably good enough if the margin of error was reported.

For this item, 9 (25.7 %) of the students selected the correct response, b4, followed by 12 (34%) students in the posttest. The most frequent response was a1 with 11 (31.4%) students selecting this choice in the pretest and only 5 (14.3 %) students on the posttest. These students tended to judge the merit of a measurement based on how close repeated measurements are to each other consistent with point reasoning.

Question 10

The table below shows polling data for two presidential candidates A and B. Based on the table of data, select one of the following conclusions:

Candidate	Percent of the vote	Margin of error
A	50 %	□ 2 %
B	47 %	□ 2 %

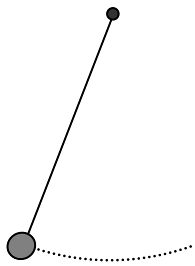
a. Candidate A has a significant lead over B
b. The two candidates are tied

The reason I selected this answer is because:

1. it is more important to compare the values for percent of the vote than the margin of error.
2. Candidate A has a 3-percentage point lead over candidate B, which is larger than the margin of error.
3. The margins of error overlap
4. The margin of error for each candidate is the same.

Students generally performed well on this item with 25 (71.4%) students choosing the correct answer, b3, on the pretest and 27 (77.1%) students on the posttest. Question eight described a similar situation; however, only 22.8 % of students identified overlapping uncertainties as the criteria for agreement between two values. Almost 36 % chose to focus on the trend shown by the numbers rather than the uncertainty or margin of error. This may suggest students recognize the importance of uncertainty when presented in more familiar context verses a laboratory environment.

Question 7



A popular physics laboratory exercise is to measure the time required for a mass, attached to a string, to swing back and forth. Let's assume the theory time is exactly 1.30 seconds and the experiment produced a value of 1.27 ± 0.02 seconds. The percent difference between the theory and the experimental time is 2 %. The student must decide if the lab was successful or not. Which of the following do you agree with?

a. these results are acceptable
b. these results are unacceptable

The reason I selected this answer is because:

1. This percent difference is probably due to human error.
2. Physics is an exact science and the results should agree at least to the same decimal place.
3. The experimental value, rounded off to the same decimal place does agree with the theory value.
4. The accuracy of the instruments is probably not high enough.

This question ranked last in terms of gains. Only 3 (8.6%) chose b4, the correct answer, followed by 4 (11.4 %) students in the posttest. Once again 19 (54.3 %) students chose the combination a1 or "human error" as an explanation for the discrepancy between the experimental and theoretical values. This misconception persisted in the posttest with 8 (22.8 %) students responding the same.

Question 9

A student, using a meter stick, measures the length of a playing card ten times. The student understands that the uncertainty in the meter stick is ± ½ millimeters. This value refers to:

a. The impact of human error on measurements.
b. The greatest difference between a measurement and the true value.

The reason I selected this answer is because:

1. No reason why the meter stick should reveal different values for the same object being measured.
2. All humans are not perfect and will most likely make mistakes when making measurements.
3. All instruments have limited precision.
4. The meter stick is not precise enough to determine the exact length of the card.

For this item the results for the pretest and posttest are identical. Fourteen (40 %) students selected the correct response, b3, before and after the intervention. Only 6 (17.1 %) students in the pretest, and 4 (11.4%) in the posttest selected the alternative combination, a2, that offers human error as an answer and a reason. The actual definition of a measurement uncertainty: *The greatest difference between a measurement and the true value* appears in this question for the first time as answer b. It may be that this alternative explanation was more tangible when compared to the misconceptions associated with the old standby, human error.

Question 11

A student uses a meter stick that is accurate to within ± 2 % to calculate the volume of a perfect cube of side 2.00-cm. The volume is therefore 8 cm. The calculated volume is only accurate to within

a. 2 %

b. 8 %

c. 6 %

The reason I selected this answer is because:

1. There are three sides. Each has an error of 2 % so 2 x 3% = 6 %
2. The volume is calculated by cubing the $(2.00\ cm)^3 = 8\ cm^3$ .The error is determined the same way or $(2\ \%)^3 = 8\ \%$.
3. Only one side is measured and it's a perfect cube so the error is 2 %.

This question has a calculation error. Under the reason component, number one reads: "There are three sides. Each has an error of 2 %, so 2 x 3% = 6 %." The correct mathematical operation should be 3 x 2% = 6 %. This may be the cause of the poor performance on this item since there is really no correct answer. In addition, this is the only question with three possible answer selections and three possible choices for the reason part. The intention of this question was to assess the computational skill of how uncertainty propagates into formulae. The correct response was c1 and even with the error, 7 (20 %) students selected this in both the pretest and posttest. The majority of students, about 57 %, selected a3. The combination a3 assumes the uncertainty in a volume is equivalent to that in any single linear measurement made on the body, in this case 2% of one side. This is consistent with how students report uncertainty in their laboratory work.

Phase Two

Phase Two consists of an analysis of laboratory reports submitted by students over the course of one semester. These reports are the basis for addressing the following research questions of this study.

1. How successful are students in applying the concepts of uncertainty analysis while participating in a traditional physics laboratory course?

2. How do students' concepts of uncertainty analysis develop or change while participating in a traditional physics laboratory course?

The review of laboratory assignments provided the opportunity to explore the progression of ideas about uncertainty analysis over a sixteen week semester. Students initially received instruction on the content and then were asked to apply it to actual laboratory experiments. Their ability to comprehend and apply these concepts was evaluated for their level of understanding and for conceptual change.

Description of Laboratory Reports

A total of eight laboratory exercises were conducted during the semester, and five reports were selected for this study. Thirty-six students completed the semester, but they did not consistently complete each of the eight laboratories. Twenty-three students did complete the same five laboratories, and these 23 students and 5 laboratories were included in this study. This selection was made to maintain uniformity and consistency of the data.

The titles of the five reports:

A. Graphical analysis and the determination of the value of Pi.
B. Determination of the value for the acceleration of gravity
C. Modeling the force of friction
D. Conservation of energy
E. Spring constant

Appendix D includes the outline and format of laboratory assignments.
Each laboratory assignment included an in class laboratory experiment where students set up an apparatus designed to provide data about a particular physics concept. Students also calculated an anticipated value based on physics theories. After completing the laboratory, students submitted a written analysis of the lab. Each written laboratory assignment comprises the following sections: Objective, Procedure, Data and Analysis, Results and Conclusions.

Part One: Facets of Understanding According to Wiggins and McTighe

The reports were examined for technical ability in executing the mathematics of uncertainty analysis according to four specific criteria previously discussed in the methods section. The four criteria are: (a) Identifies uncertainty accurately, (b) Performs mathematics correctly, (c) Reports uncertainty in the correct form, and (d) Uses uncertainty analysis to support conclusion. Student laboratory assignments were evaluated according to these four criteria. Table 1 illustrates the results of the scores. The scores for each category were averaged for all 23 students for all five laboratory exercises. The lowest score was 1.6 out of 5 for laboratory E under the category of "Employs uncertainty to support conclusion". The highest score was 4.4 out of 5 for laboratory C under the category of "Performs mathematics correctly."

Table 4.4

Average Score of all 23 Students for Each of the Four Categories for Laboratory Reports A, B. C, D and E

Laboratory	Correctly Identifies uncertainty.	Performs mathematics correctly	Reports uncertainty in correct form	Employs uncertainty to support conclusion	Total Average
A	3.3	3.4	2.5	2.9	3.1
B	3.2	3.3	3.2	2.3	2.9
C	3.9	4.4	3.1	3.1	3.6
D	3.7	3.5	3.0	2.7	3.2
E	2.4	2.8	1.9	1.6	2.2
Average	3.3	3.5	2.74	2.52	3.0

Note : Laboratory A = Graphical analysis and the determination of the value of Pi. Laboratory B= Determination of the value for the acceleration of gravity. Laboratory C= Modeling the force of friction. Laboratory D = Conservation of energy. Laboratory E = Spring constant

Students were provided with the manual along with instruction on uncertainty analysis. As the results in the table show, there was no significant trend showing an increase in student understanding of uncertainty analysis over the course of the semester. Laboratories A, B, C, and D were all similar in that each required the students to compare the experimental values with the accepted value. Laboratory E, or The Spring Constant, was different in that it did not involve an accepted value. Students were instructed to simply plot the distance the spring stretched vs. the force and determine the constant for the spring. The constants for these particular springs were unknown. It is interesting to note that the scores for this laboratory, among the four categories, were the lowest among all of the laboratories even though this was the last laboratory exercise completed by the students. Moreover this same laboratory received a score of 1.6 in the category: *Employs uncertainty to support conclusion.* This is the lowest of all the scores. Here it seems as though, without a theoretical value to compare to, students were less confident in reporting their findings and attributed the less than perfect results to human error and mistakes. Among the highest scores reported were those for the category of: *Performs mathematics correctly.* This supports the findings of Allie, et al. (2003) in that students, in general, are able to carry out the mechanics for calculating uncertainties. The low average scores in the category of: *Employs uncertainty to support conclusion*, indicated that they are still unable or unwilling to use these calculated values to validate their findings. Some examples from each category are presented here, for comparison. Included are examples illustrating the lowest and highest scores on the facets of understanding according to Wiggins and McTighe (2005). The highest score is a five and the lowest is a one.

Student response:

Category a): Correctly identifies uncertainty: High score of 5.
Laboratory A: Graphical analysis and the determination of the value of Pi

Student response:

> *Next we determine the possible range of values of Pi based on the uncertainties in the circumference and diameter. These include $\delta C = \pm 0.05$ cm and $\delta D = \pm 0.05$ cm".*

This response received a score of 5 out of 5. This student clearly identifies a range of values for the circumference and diameter and includes the units and the correct notation for uncertainty, namely the lower case delta symbol "δ" and the "±".

Category a): Correctly identifies uncertainty: Low score of 1.
Laboratory B: Determination of the value for the acceleration of gravity

Student response:

> *Given the slope, the percent error is calculated out to be 0.61 cm/s*^{2} *± 0.05 m/s*^{2}*. Knowing that the percent error was 0.61 cm/s*^{2} *± 0.05 m/s*^{2} *makes the experiment and its measurements and the uncertainty very small.*

This student incorrectly identifies the uncertainty as a percent difference. The value stated above includes units with an absolute uncertainty added to this percent value which is inconsistent with the expression of percent. Furthermore, the student confirms the misconception in the last sentence that uncertainty is simply the percent difference.

Category b): Performs mathematics correctly: High score of 5.
Laboratory: Determination of the value for the acceleration of gravity.

Student response:

Measurand:

$$g = \frac{\Delta V}{\Delta t} = \frac{V_f - V_0}{\Delta t} = \frac{381.6 - 60.9}{11/30 - 1/30} = \frac{320.7}{10/30} = 962.1\ \mathrm{cm/s^2}$$

$$g_{max} = \frac{(381.6 + 6) - (60.9 - 6)}{(11/30 - 0) - (1/30 + 0)} = \frac{332.7}{10/30} = 998.1\ \mathrm{cm/s^2}$$

$$g_{min} = \frac{(381.6 - 6) - (60.9 + 6)}{(11/30 + 0) - (1/30 - 0)} = \frac{308.7}{10/30} = 926.1\ \mathrm{cm/s^2}$$

$$\delta g = \frac{g_{max} - g_{min}}{2} \rightarrow \frac{998.1 - 926.1}{2} = 36\ \mathrm{cm/s^2}$$

$$\delta g = \pm 36\ \mathrm{cm/s^2}$$

This analysis is complete. Beginning with the equation for the acceleration of gravity the student goes through each step maintaining the correct form and showing the actual values inserted including the units along the way. Finally the uncertainty in g is reported including units and indicating the range with the ± notation. It is note worthy to recognize, at the top of the example, this student is one of few to correctly identify the acceleration of gravity as the measurand.

Category b): Performs mathematics correctly: Low score of 1.
Laboratory: B): Determination of the value for the acceleration of gravity.

Student response:

$$\%\text{error} = \frac{980-971}{980} \times 100 = .9\%\ \text{error or uncertainty}$$

The calculation shown here relies on the standard definition of percent difference to express the uncertainty in the value of the acceleration of gravity. The percent difference is defined as the difference between the accepted value for the acceleration of gravity of, 980 cm/s2, and the value obtained from experiment of 971 cm/s2 divided by the accepted value. The result is then multiplied by 100. This is the wrong mathematical approach for calculating the uncertainty using the Worst Case Method. The student also mixes the terms error and uncertainty together adding confusion to the meaning of uncertainty.

Category c): Reports uncertainty in correct form: High score of 5.
Laboratory B): Determination of the value for the acceleration of gravity.

Student response:

Variable	Lab Values	Theoretical Value or Limit	Theoretical Agrees with Lab Range
g	971 ± 30 cm/s^2	981 cm/s^2	Yes
v_0	38 ± 8 cm/s	32.7 cm/s	Yes

This table includes all the relevant information including the measured along with the units and its associated uncertainty expressed to the correct number of decimal places. A comparison is made and a conclusion is presented stating agreement between the theoretical and laboratory values.

Category c): Reports uncertainty in correct form: Low score of 1.
Laboratory B): Determination of the value for the acceleration of gravity

Student response:

> *In this case the slope represents gravity and my calculation was – 975.52 m/s^2 . And the actual gravity in this world is – 981 m/s^2. So my conclusion of g ± δg > <u>g ± δ 6.52</u>*

The correct way to report this acceleration and its uncertainty is - 975 ± 6 m/s^2. The value of gravity is first reported as – 975.52 m/s^2 which contains two more decimal places allowed by the uncertainty. The symbol δ is shown, but there is no equal sign to refer to the value. The value itself of 6.62 is missing units and should contain only one significant figure.

Category d: Employs uncertainty to support conclusion: High score of 5.
Laboratory C: Modeling the force of friction

Student response:

> *We calculated the acceleration down the track as 2.86 ±0.75 m/s^2. The experimental value we got was 2.88 ± 75 m/s^2. Why are these values acceptable and correct? If you take a look at the uncertainty you will notice that the calculated value for acceleration completely falls within the experimental values. Therefore we can say these values agree.*

This student is very direct about the conclusion. The student reports the values with their uncertainties in the first sentence. Moreover he refers directly to the associated uncertainties as the criteria for agreement between the values. Finally a judgment is made, based on the uncertainty analysis, which verifies that the values are in agreement.

Category d: Employs uncertainty to support conclusion: Low score of 1.
Laboratory C: Modeling the force of friction

Student response:

> *This shows that our results are really close; in fact our percent difference between the actual and predicted values is only 4.26%. We had a lot of confidence in these numbers and we had taken really careful measurements so we could obtain the most accurate numbers.*

This student calculated the uncertainty correctly as $\delta a = \_\pm 0.08$ m/s^2 but never even reports it in the conclusion. Instead he refers to the care taken in obtaining the measurements and the discrepancy in the values of 4.26% as the criteria for agreement between theory and experiment. After grading this report it turns out the experimental value for acceleration was 2.80 ±0.08 m/s^2 and the theoretical value was 2.75 m/s^2. Because these values overlap, the experiment supports the theory within the uncertainty but the student failed to recognize this fact.

Part Two: Conceptual Change

The qualitative exploration involves the text portion of the laboratory reports. The Data and Analysis and the Results and Conclusions sections contain students' written discussions of the results. It is in this section that students discuss the variation between their mathematically calculated values and their data obtained in the laboratory. It is in this explanation that students apply their concepts of uncertainty analysis. Their comments reveal how students are thinking about uncertainty analysis in terms of conceptual change.

Student comments on laboratory reports were also analyzed in a study completed by Campbell et al. (2000). In this qualitative study individual student comments revealed their underlying thought process, knowledge base, and ability to assimilate information. Their results revealed the difficulty many students had in accommodating new information. Although the results supported the researchers' premise that instruction on laboratory concepts and report writing are essential in the science curriculum, the results also revealed a wealth of information about student ideas and thoughts about concepts.

Treagust and Duit (2008) interviewed students and evaluated their thinking about concepts in terms of the process of conceptual change. In the same manner, student comments in these laboratory assignments provide insight into the process of conceptual change in relation to uncertainty analysis. The criteria utilized in the study of Treagust and Duit are employed in this study. They include Intelligible, Plausible, and Fruitful. An additional category of "None" was added to describe responses absent of intelligible, plausible, or fruitful comments in relation to uncertainty analysis.

When a student made no mention of uncertainty and discussed the discrepancy between the predicted outcome and the experimental results in terms of human error, the response was considered "None". When discussing the difference in calculating π and the actual measurement, one student stated, "We know that the exact value of π is 3.14 so; this difference comes from human error." Another student commented, "Also my uncertainties could be due to the thickness of the measuring tape that I used in the experiment. The thickness of the tape measure could have thrown off my calculations in the precision of my measurements. Maybe next time I could incorporate the thickness of the tape measure into my calculations, which could cut down on the uncertainty of my measurements and calculations." Other students use percent differences to refute their results or calculate answers with multiple decimal places, which demonstrate a lack of understanding of uncertainty analysis.

Intelligible responses are defined by Treagust and Duit (2008) as "sensible if it is non-contradictory and its meaning is understood by the student" (p. 299). They also utilized the criteria of Hewson (1992) to define intelligible as evident when the student seems to "know what the concept means" (p. 8). Students can describe the concept in their own words. In this study, responses were considered intelligible when the student discussed uncertainty as an important aspect of the analysis.

Students described the concept, but they did not take the step of incorporating uncertainty in explaining the variance between their data and the calculated results. One student calculated the percent difference and then went on to state, “Therefore, I find our results to be well within the uncertainty of the accepted value ...” The student demonstrated some understanding of a range of acceptable values. Another student discussed percent error in the calculation and measurement of circumference and also discussed the measurements falling between bars on a graph which indicated an understanding of uncertainty. In the spring constant laboratory a student stated, “connecting points 1 and 9, I can see some of the other points are below, some are above and some are right on the slope line, thus this tells me my spring constant is what it is suppose to be. . .” Another student discussed the data as “within the range of the data points” which demonstrates an intelligible response.

A plausible response is “considered believable in addition to the student knowing what the concept means” (Treagust & Duit, 2008, p. 299). Treagust and Duit described plausible as a student “believing that this is how the world actually is” (p. 304). The student makes a judgment about the data that is based on the concept of uncertainty analysis. For example, in comparing the laboratory data and calculations of circumference and diameter, one student stated, “I conclude that the results do agree with the theory. The obtained slope is within the allowable for Pi. It may not have been exactly 3.14, but the chances of getting that were just as possible as getting 3.12, which was obtained. The student went beyond defining what uncertainty is to applying it to understanding the data and conveyed believability in the answer. Another student stated, “So I have a 5% discrepancy between my calculated and theoretical values, but the two values do overlap in uncertainties, so it is an acceptable discrepancy.” This student also supported the data using uncertainty.

Fruitful responses are both intelligible and plausible, but the student also incorporates the concept to “solve other problems or suggests new research directions” (Treagust & Duit, 2008, p. 299). The student sees the new concept as a “better way of explaining things” (p. 304). In this study, students conveyed that they believed in uncertainty and grasped it at a level where they may even imply its application outside the individual laboratory assignment. The student makes repeated references to the uncertainty as the rationalization for data and recognizes that uncertainty represents the limitations of both the instrument and human beings. Using the uncertainty as a measure of the quality of the data, the student may extend and explore others sources not accounted for the experiment. For example, this student's comment demonstrated the application of the data results to the overall theory of conservation of energy. "We found.... that this experiment agrees with the conservation of mechanical energy. This is known by looking at the Total energy plot in the above graph. There you will see that the average Potential Energy intersects and overlaps with the uncertainty error bars of .004m. If energy was not conserved the PE plot would have escaped those error bars." Another student commented on the range of data and uncertainty and then goes on to state, "With this agreement, I can say my analysis is almost identical between actual and predictive acceleration and therefore proves that Newton's Law is correct."

Responses were categorized and the laboratory assignments were evaluated for conceptual change. Little progression in conceptual change was noted when looking at all of the students in chronological order across the semester. In fact, there were a high number of students who did not incorporate uncertainty analysis in their conclusions on the final laboratory assignment. A discussion of this laboratory and how it impacted the results is included in Chapter 5 of this study. Table 4.5 indicates the category of student responses as None, Intelligible, Plausible, and Fruitful for the five laboratories: A, B, C, D and E over the course of the semester.

Table 4.5

Conceptual Change in Student Responses on Laboratory Assignments

Level of Conceptual Change	Number of Students				
	Laboratory				
	A	B	C	D	E
None	10	10	8	8	11
Intelligible	3	4	1	4	8
Plausible	5	3	8	4	3
Fruitful	5	6	6	7	1
Total	23	23	23	23	23

When looking at all 23 students a progression toward the plausible and fruitful levels of conceptual change is not indicated above, especially when looking at the results of Laboratory E. Nevertheless, the progression of conceptual change, in relation to uncertainty analysis, is apparent when looking at some of the individual student examples. Concepts discussed in relation to conceptual change such as the persistence of preconceived ideas, the idea of the hybrid levels of change where there is a mixture of thinking, and the frequent return to old ideas is clear in many student comments. It is also clear that there is much individual variability among students in terms of the attainment of conceptual change.

Two students demonstrated a progression in their conceptual change from Laboratory A to Laboratory E and three students progressed in their thinking between Laboratory A and Laboratory D, but then they revert to a low level of conceptual change in Laboratory E. Examples of their comments provided insight into the process of conceptual change as it relates to uncertainty analysis.

Student 27 was able to describe uncertainty on the intelligible level in Laboratory A by minimally stating that the results were "within the uncertainty of the accepted value". There was no mention of uncertainty in terms of supporting the results in Laboratory B and C. In Laboratory D, the student drew a range of uncertainty on a graph and indicated data points within the range. These comments were considered intelligible. On the final assignment, Laboratory E, the student achieved the plausible level when discussing the data. This student discussed uncertainty as a range and used the results to support the data falling on a linear regression line.

Student 45 never addressed uncertainty until the final laboratory when the response was scored as plausible. This student began with indicating that measurement error explained the discrepancy in Laboratory A, to stating that the data were an exact match to the theoretical values in B, to no discussion of results in C. In Laboratory D, this student mentioned the idea of uncertainty values but did not apply it accurately. In Laboratory E, the student incorporated uncertainty in a graph and discussed it in the results as supporting the equation. "This equation works." These two students were the only students in the 23 to progress to plausible from lower levels of conceptual change.

Some students showed progression in their thinking through Laboratory D and then reverted to lower level in Laboratory E. Their comments on the first four laboratories provided examples of conceptual change. Student 32 explained results in Laboratory A as being attributed to human error after stating that "if we wrap the tape around nothing the circum cannot be zero". Laboratory B did not address uncertainty. In Laboratories C and D, the student jumped to the plausible level of conceptual change. The student compared actual to theory results and used a drawing indicating a range. He utilized these results to support agreement of the lab data with theory. After a discussion how uncertainty of one variable affected the uncertainty in another in Laboratory D, this student stated, "so uncertainty is like a chair, and every part is under influence of the past part". These comments demonstrated an understanding and acceptance of uncertainty and its use to explain results.

Student 49 progressed in a step-wise manner from none to fruitful from Laboratory A to Laboratory D, and returned to a plausible understanding in Laboratory E. In Laboratory A, the student talked about having made a mistake, because the data did not correspond to the theoretical values. In Laboratory B, this student talked about uncertainty in values, and then mentioned uncertainty as a possible contributor to the difference in values: "a larger uncertainty in the value of X than that which was recorded." In Laboratory C, the student stated, "so I have a 5% discrepancy between my calculated and theoretical values, but the two values do overlap in uncertainties, so it is an acceptable discrepancy. He then went on to state that the experiment supported the theory being tested, and this response was considered plausible. In Laboratory E, the student discussed uncertainties in depth and concluded that "based on the data from table 1 and the uncertainty of TME (total mechanical energy), it can be concluded that TME was conserved. This response was considered fruitful because of the in-depth discussion of uncertainty and the use of uncertainty to support the conservation of total mechanical energy. This student was one of few to reach the plausible level on Laboratory E. He addressed uncertainty and used it to support answers in a believable manner.

Although few students progressed in the manner of Student 49, their comments in the laboratory assignments clearly described concepts of conceptual change.

Preconceived ideas about the ability to get a right answer were very apparent. When data did not match the theoretical values, students explained this as human error or making a mistake. One of the most apparent ideas was the hybrid comment, where old and new ideas are mixed together. It was common to see new ideas about uncertainty analysis in one sentence, followed by comments about human error in the next. It is also evident that students progressed to a higher level of conceptual change at one point and then quickly reverted to old ideas. No students demonstrated consistently fruitful comments about uncertainty analysis. The following list contains examples of student responses as examples of this mixture of conceptual change.

> *Student 23: This value is in the range, crediting the process of uncertainty analysis and acceleration of gravity. "It is fair to conclude errors that exceeded uncertainty were to blame."*

In the above example, the student has hybrid thinking that contains both the new ideas about uncertainty and old ideas about error.

> *Student 40: "Even though the uncertainties are already accounted for in my measurements, I could have been more accurate in my data collected."*

This response also illustrated an understanding of uncertainty at one moment while reverting to the idea of human error at the same time. In this case the data supported the theory, so the comment about error also demonstrated a lack of understanding of the concept at the same time as demonstrating an understanding.

> *Student 49: The uncertainty value of ±0.05cm does not include the maximum measured value for t, so it confuses me and I believe I made some kind of mistake in the experiment."*

This student completed the experiment correctly and discussed uncertainty but returned to the concept of human error to explain data.

Reverting to old ways of thinking was clearly evident in Laboratory E, when 11 students did not include uncertainty analysis in their conclusions. The number of students in the “None” category was higher in this laboratory than it was in the first two laboratories.

Overview of the Chapter

This chapter presented the results of Phase One and Phase Two and correlated the results with the research questions. Phase One included an analysis of student scores on a pre and posttest and also an item analysis of each of the test questions. The results demonstrated an improvement in scores from the pre to the posttest. For Phase One the posttest scores were significantly higher at the $p = 0.05$ level over the pretest results. The pretest scores ranged from 8.6 % to 71 %, and the posttest scores ranged from 11.4% to 88.6 %. In general about one half of the students made some progress in advancing from point thinkers to set thinkers. Four of the 35 students performed above the 65 % level on both tests showing little change from pretest to posttest. These students appeared to think about data consistent with the set paradigm on both tests.

Phase Two included the results from a review of laboratory assignments during the semester and an analysis of change in the level of student understanding. A basic description of the laboratory exercises was also provided. Based on the five facets of understanding, it was apparent that although most students demonstrated the ability to perform the needed mathematical operations, very few students recognized the importance of applying the results of the calculations in the conclusion section of the report.

The results in this chapter showed that in general students failed to connect the values obtained using the Worst Case Method with the measured results of the laboratory exercise. When the values predicted by the theory did not exactly match the experimental data, students continued to site human error and or equipment problems as the cause for the discrepancy even up to and including the last report. As pointed out in this chapter, only a small percentage of students experienced the kind of conceptual change needed to realize that uncertainty is a way of accepting a range of values as a valid conclusion.

CHAPTER 5

DISCUSSION

This chapter includes a discussion of the results presented in Chapter 4 in terms of their support for the research questions in Phase One and Phase Two of this study. These questions are:

Phase One

1. To what extent can the concepts of uncertainty analysis and propagation of uncertainty be successfully taught to entry-level students using algebra?
2. Over the course of a typical college physics laboratory class, how do first year students' concepts of uncertainty compare from the beginning to the end?

Phase Two

1. How successful are students in applying the concepts of uncertainty analysis while participating in a traditional physics laboratory course?
2. How do students' concepts of uncertainty analysis develop or change while participating in a traditional physics course?

Each phase of the study includes two parts, and these are discussed sequentially along with the associated research question. Following the discussion of both phases, there is a section on suggestions for further research.

Phase One: Pretest and Posttest

Part One: Student Performance on Pretest and Posttest

As indicated by the results presented in Chapter 4, the pretest and posttest did support the first research question that uncertainty analysis and the propagation of uncertainty can be taught to, and learned by, students with a minimal background in algebra.

The group as a whole had a gain of 41.3% from pretest to posttest, even when both tiers needed to be answered correctly. These results are statistically significant and support the use of the Worst Case Method in teaching uncertainty analysis. The size of the sample was small, which creates difficulties in generalizing the results to a larger population without further research. Additionally, the Cronbach Alpha Coefficient was low, which also renders the results difficult to generalize to a larger population. Therefore, in spite of the positive results, more studies are needed to validate these findings.

Part Two: Item Analysis of Test Questions

This section of the study includes an item analysis of each of the questions on the pretest and posttest. The research question addressed with these results is: Over the course of a typical college physics laboratory class, how do first year students' concepts of uncertainty compare from the beginning to the end. The main focus of this part of the study was to evaluate students' understanding of uncertainty analysis by examining each question in terms of the students' ability to see data as points or as sets of data.

The results of the item analysis support the results of previous studies on point and set reasoning, which demonstrated that students do benefit some from instruction on data analysis. Some individual students were able to transition from point reasoning to set reasoning indicated by the significant increase in performance on the posttest questions 1, 4 an 8, all of which had to do with how students look at data. Students were able to select the correct answer when they looked at data as a set rather than as a point. For example, on question 1 involving the swinging pendulum, "no one value is more important than the others as they all have equal merit" requires a student to see the data as a set. The correct answer in question 8 is that "there is some overlap of the margin of error for each time" also indicates that students are looking at the data as a set rather than one right answer.

Lower scores were obtained for other questions, especially for question six. The pretest average of correct responses was 5.7% and 17.4 % on the posttest. Students had much difficulty transitioning from point to set reasoning on this question, and some of the information on conceptual change may help to explain their difficulties. This question targets the idea of human error in explaining results, and the pre-instructional belief that human error is at the heart of all the discrepancies in scientific data is very difficult to change. Students generally got this question wrong, because they continue to select the answer that they agree with "human error inherent in all scientific endeavors and impossible to eliminate."

Phase One of the study demonstrated that instruction on uncertainty analysis was beneficial in teaching students about how to interpret a set of data. Their scores improved after the instructional unit to a significant degree. However, their ability to understand data better in terms of uncertainty analysis is not as definitive. Although many students progressed from evaluating data in terms of points to sets, others showed little progress. It also appears that certain questions have content that involved pre-instructional ideas that are very difficult to change, especially those addressing human error.

Phase Two: Evaluation of Student Laboratory Assignments

Phase Two included the analysis of five laboratory assignments during the semester. The five laboratories are A, B, C, D, E, and the general outline for all assignments is included in Appendix D. In the first part of this section, the assignments were evaluated for the students' level of understanding, utilizing the rubric for facets of understanding developed by Wiggins and McTighe (2005). In the second part of this section, student assignments were evaluated for conceptual change. This section also provides insight into the process of conceptual change in select students.

Part One: Facets of Understanding

The facets of understanding rubric was utilized to evaluate the first research question in this phase: "How successful are students in applying the concepts of uncertainty analysis while participating in a traditional physics laboratory course?" The results demonstrated that students were able to incorporate the mathematics of uncertainty analysis and to carry out the procedures described in the manual. These skills comprised the first three categories of the Facets. Student scores decrease as they moved into category 4, which included criteria evaluating the students' grasping the whole picture. Little change was noted as students progressed from one laboratory to the next, so time and repeated exposure did not seem to be a factor. It is important to remember that students received instruction on uncertainty analysis and a copy of the manual prior to writing any assignments. The initial benefit of instruction discussed in Phase One did not seem to expand and improve understanding as time went on. The lowest scores were in the area of using uncertainty analysis to support conclusions. Students overwhelmingly resorted to explaining differences in theoretical and laboratory data as a function of human error. These preconceived ideas are discussed further in the following section on conceptual change.

Part Two: Conceptual Change

The final section of the study evaluated the research question of "how students' concepts of uncertainty analysis develop or change while participating in a traditional physics laboratory course." Many science teachers' goals are that students will achieve conceptual change as a result of their educational experience. Many teachers measure their success as educators with facilitating students in their attainment of a new way of thinking or comprehending and understanding material that is over and above the rote presentation of ideas followed by testing students' ability to recall facts. Students' difficulties with conceptual change can be one of the most frustrating experiences that teachers confront. In the physics laboratory, uncertainty analysis has often interfered with both the teacher's success and with student's confidence in understanding of scientific data. In Phase Two of this study, the ability of students to understand this concept and the ability to acquire a conceptual change was evaluated. Student laboratory assignments were evaluated over the course of a semester in terms of their level of understanding and conceptual change.

Conceptual change has been a challenging concept both to achieve in the classroom and to evaluate in student work. In this study, it was anticipated that instruction, through the use of a self-study manual utilizing the Worst Case Method would provide students with the means to understand how uncertainty analysis affects laboratory results. Although, as discussed in Phase One, students did demonstrate the ability to utilize uncertainty analysis as measured by a pretest and posttest, the results of improving their understanding or achieving conceptual change were not as positive. In fact, little change was noted when evaluating all 23 students over the course of the semester. Rather than adopting the new way of thinking, most students clung to their preconceived ideas. The impact of these ideas, especially when many are rooted in ontological and epistemological beliefs about the world and reality, was profound. Nevertheless, the results provide excellent examples of conceptual change in process and demonstrate that it is rarely a linear progression. The findings do provide valuable information for future successful instruction on this topic.

Although a few students demonstrated a chronological progression to a higher level of conceptual change as the semester progressed, most vacillated back and forth. The highest number of students who did not include uncertainty analysis in their results, or who reverted to the old ways of thinking was the highest in the last laboratory of the semester. It was expected that this laboratory would have the highest scores because Laboratory E was the only laboratory where students were not asked to compare their data to a theoretical value. It is possible that this fact contributed to the low level of conceptual change noted in the laboratory reports.

Students also scored the lowest Wiggins and McTighe's level of facets of understanding for this laboratory, so it is possible it was just a more difficult laboratory for them to understand. Nevertheless, in spite of this possible explanation, a continued progression to higher levels of understanding was not evident in the first four laboratories either. There was some improvement noted for many students but it was not significant.

Student comments in the laboratories did provide valuable insight into the process of conceptual change. Many students did reach the fruitful level of conceptual change at some point and were able to use uncertainty analysis to explain the physics of laboratory. It was apparent, however, that this change could disappear the following week. Few students reached a fruitful or plausible level of conceptual change and then continued to demonstrate this at the same level each week.

Student comments do provide a wealth of information about how students think and the impact it can have on sustaining a high level of conceptual change. Pre conceived ideas have been discussed in the conceptual change literature for many years. These can be very difficult to change. It is possible that these beliefs are embedded in other beliefs about science and about the world, and they are very difficult to change. In more recent literature, the challenges of these ideas, especially when they are linked to epistemological and ontological beliefs about the world, is emphasized. The findings in this study support this literature. The idea of finding a right answer is engrained in student thinking before they enter the physics laboratory. It is difficult to say if it is an idea from early childhood or if it is emphasized in elementary schools, but students arrive in the introductory physics laboratory with a preconceived idea of one right answer, and this thinking is almost a direct contradiction to the concept of uncertainty analysis which indicates that a range of data is acceptable. In the pretest and posttest section, students are able to learn what uncertainty analysis is and even how to calculate uncertainty, but their beliefs about right answers is a very difficult one for them to unlearn. They frequently revert to a discussion focused on finding the right answer, which is most often apparent when they talk about human error. As noted in Pillay et al (2008), when students are confronted with a range of data, they may attribute the discrepancy to human error and the idea of having made a mistake. This is frustrating for both the student and the teacher.

When conceptual change occurs, the results in this study support those of many authors and researchers. Conceptual change does not occur in a linear manner with a new engrained way of thinking that is apparent in a short period of time. Mixed thinking is very apparent and this was evident in this study. In one paragraph, students would include uncertainty analysis in their discussion and then substitute human error as the cause for the variance. It was also evident that students would appear to grasp the concept on one laboratory and then revert to old ways of thinking the next week. Researchers have found this to be typical of student thinking during conceptual change.

In conclusion, an effective method to teach uncertainty analysis to entry level students addresses a long standing problem with data interpretation in introductory physics laboratory courses. Conceptual change is very difficult to accomplish with certain concepts, and uncertainty analysis seems to be one of them. This method to calculate uncertainty was helpful, but it did not result in significant conceptual change. Time is a factor, and it may take much longer than expected to make a conceptual change. It is essential to evaluate the preconceived ideas of students and how these relate to their epistemological and ontological beliefs. The impact of these ideas cannot be under estimated. If it is any consolation to educators, they may very well be planting the seeds for future conceptual change in their efforts. The timing of conceptual change is not pre-determined and it will vary from student to student. In this study, there were many examples of this, but one stands out. A student whose laboratories were consistently evaluated as "none" in terms of conceptual change has now progressed to being admitted to a doctoral program in physics. I still have contact with this student, and I can attest to the fact that he has now grasped the complexities of uncertainty analysis. I would like to think that some of my efforts at least planted a few seeds to help him develop in his understanding of this concept.

Suggestions for Future Research

Teaching uncertainty analysis to beginning science students is a complex task. It may even be said that teaching uncertainty analysis is more difficult than teaching physics. The teachers are challenged by the limited mathematical skills of their students, complicated by the fact that the traditional methods for teaching uncertainty analysis are inconsistent and often unfamiliar to both students and teachers. The majority of the data included in this study support the fact that students enter the science classroom as point thinkers and that only a small fraction progress to thinking about data in terms of sets. Therefore students are challenged by the persistence of old ideas learned throughout their high school education. As this and other studies have shown, science students find it very difficult to assimilate a new concept believable enough to replace the old standby of insisting on finding the "one right answer" and invoking human error as the cause for any discrepancies encountered in the data. Additionally, as indicated in the literature review, the techniques often presented to students for handling uncertainties are laborious, incoherent and sometimes appear even contradictory. Students, exposed for the first time, to this kind of unintelligible alternative cling to their misconceptions and very little conceptual change take place. Studies have shown that the Worst Case Method for determining uncertainties and propagating errors is the preferred method among students due to its simple calculations and is conceptually easy to understand and remember. This is also the method used in this study and is described on page 29 using the pendulum example. Although determining uncertainties using this method is simple, the calculations can be quite lengthy and therefore students can quickly lose interest and sometimes even forget what they are calculating.

One of the outcomes of this research study was the development of a modified version of the Worst Case Method that resulted in obtaining the same results with only approximately half of the calculations and thereby half the time. Students have responded very positively to the new manual, and it would be beneficial to complete a new study with this revised manual. It would be interesting to evaluate the students' transition from point and set reasoning and overall conceptual change when using the revised and simplified method. I plan to publish this new method in a future article.

Research has clearly documented that students arrive in the classroom with preconceived ideas. This has been repeatedly demonstrated in the literature. Future research into the origin of some of these ideas such as the one right answer and human error may be helpful in addressing these ideas when students arrive in the entry level physics laboratory.

One source of preconceived ideas are the teachers themselves. If secondary education teachers are unfamiliar with uncertainty analysis and propagation of error, then they may be conveying these erroneous concepts to their students. Future research could focus on the outcome of teacher education on uncertainty analysis in very much the same manner that students were instructed on uncertainty analysis in this study. Once again, I am very optimistic that the revised manual may help with this education also.

Early in this study a quote taken from Roberts (1983) states that: "If it's green or it wiggles, it's biology; if it stinks, it's chemistry; if it doesn't work, it's physics". Most students would agree with the last phrase: "if it doesn't work, it's physics". If students are relying on their previous beliefs and preconceived ideas about how to interpret data then it is no surprise that this is their conclusion. Phase Two of this study clearly illustrates, even after instruction on uncertainty analysis, that the majority of students fail to recognize the value of applying uncertainty analysis in their conclusions. Instead of using uncertainty analysis to account for the discrepancy that may exist between what the theories predict and what the experiment actually produced, most students fall back on human error and or equipment problems as excuses for the discrepancy. This researcher and others agree that the most important reason for calculating uncertainties is to qualify the results presented in the conclusion of a laboratory report. The conceptual change necessary for students to realize the full benefit of uncertainty analysis in this important application was not apparent in this study. Further research into this specific area is needed. Maybe this would result in a revision of the above statement to "if it works within the uncertainty then its physics".

Although I am very optimistic about the new manual, there is one result from this study that has been apparent in many previous studies involving conceptual change. It is a very complex process, and there is no simple answer. Students change their thinking at a very individual pace and fluctuate back and forth. I have learned a great deal about conceptual change through this research study. I might venture to say that I have undergone a conceptual change at the fruitful level in terms of expectations on teaching uncertainty analysis. I am now able to appreciate the depth of the challenge and will be able to see small changes in a more positive manner rather than feeling frustrated when the term "human error" reappears along with other indications of improved student understanding.

REFERENCES

Allie, S, Buffler, A., Kaunda, L, Campbell, B., & and Lubben, F. (1998). First year students' perceptions of the quality of experimental measurements. *International Journal of Science Education,* 20, 447-459.

Allie, S., Buffler, A., Campbell, B., & Lubben, F. (2003). Teaching measurements in the introductory physics laboratory. *The Physics Teacher,* 41, 394-401.

American Association of Physics Teachers (AAPT) (1998). Goals of the introductory physics laboratory, *American Journal of Physics,* 66, June 1998, 483-485.

Bauman , R. & Swartz, C. (1984). Treatment of errors and significant figures in the laboratory manuals. *The Physics Teacher, 22,* 235-236.

Bevington, P.R. (1969). *Data reduction and error analysis for the physical sciences* New York: McGraw-Hill Book Company.

Blasiak, W. (1983). Errors and uncertainty in physical measurement. *Physics Education, 18*, 290-294.

Buffler, A., & Allie, S. (2001). The development of first year physics students' ideas about measurement in terms of point and set paradigms. *International Journal of Science Education,* 23, 1137-1156.

Buffler, A., Allie, S., Lubben, F., & Campbell, B. (2003). Evaluation of a research-based curriculum for teaching measurement in the first year physics laboratory. *International Journal of Science Education, 23,* 1137-1156.

Buffler, A., Allie, S., & Lubben, F. (2008). Teaching measurement and uncertainty the GUM. way. *The Physics Teacher, 46,* 539-543.

Buffler, A., Allie, S., Lubben, F., & Campbell, B. (2009). *Introduction to measurement in the physics laboratory: A probabilistic approach.* South Africa: University of Cape Town.

Campbell, B., Kaunda, L, Allie, S, Buffler, A, & Lubben, F. (2000). The communication of laboratory investigation by university entrants. *Journal of Research in Science Teaching,* 37, 839-853.

Chinn, C. & Brewer, W. (1993). The role of anomalous data in knowledge acquisition: A theoretical framework and implications for science education. *Review of Educational Research,* 63, 1-49.

Cohen, L., & Manion, L. (1989). *Research methods in education* (3rd ed.). London: Routledge.

Deardorff, D.L. (2001). *Introductory physics students' treatment of measurement uncertainty.* Unpublished doctoral dissertation, North Carolina State University, Raleigh.

Duit, R. & Confrey, J. (1996). Reorganizing the curriculum and teaching to improve learning in science and mathematics. In D.F. Treagust, R. Duit, & B.J. Fraser (Eds.), *Improving teaching and learning in science and mathematics* (pp. 79-93). New York: Teachers College.

Duit, R. & Treagust, D. (2003). Conceptual change: A powerful framework for improving science teaching and learning. *International Journal of Science Education*, 25, 671-688.

Fairbrother, R. & Hackling, M. (1997). Is this the right answer? *International Journal of Science Education,* 19, 887-894.

Gilbert, T.K. (1977). The study of student misunderstandings in the physical sciences. *Research in Science Education, 7,* 165-171.

Gordon, R., Pickering, M., & Bisson, D. (1984). Uncertainty analysis by the "Worst Case" Method. *Journal of Chemical Education, 61,* 780-781.

Harrison, A.G. & Treagust, D.F. (2000). Learning about atoms, molecules and chemical bonds: A case-study of multiple model use in grade-11 chemistry. *Science Education,* 84, 352-381.

Hewson, P.W. (1992, June). *Conceptual change in science teaching and teacher education.* Paper presented at a meeting on "Research and Curriculum Development in Science Teaching" under the auspices of the National Center for Educational Research, Documentation, and Assessment, Ministry for Education and Science, Madrid, Spain.

Hewson, P.W. (1996). Teaching for conceptual change. In D.F. Treagust, R. Duit, & B.J. Fraser (Eds.). *Improving teaching and learning in science and mathematics* (pp. 131-140). New York: Teachers College.

Hewson, P.W., & Hewson, M.G., (1984). The role of conceptual conflict in conceptual change and the design of scientific instruction. *Instructional Science,* 13, 1-13.

Hewson , P.W., & Hewson, M.G. (1992). The status of students' conceptions. In R. Duit, F. Goldberg, and H. Niedderer (Eds.), *Research in physics learning: Theoretical issues and empirical studies* (pp. 59-73). Kiel: IPN.

Institutional Research Office (2004). Fall 2004 semester 45 day report. Tempe, AZ: Maricopa Community Colleges.

Lubben, F., Campbell, B., Buffler, A., & Allie, S. (2001). Point and set reasoning in practical science measurement by entering university freshmen. *Science Education,*85, 311-327.

Odom, A.L. & Barrow, L.H. (1995). Development and application of a two-tier diagnostic test measuring college biology students' understanding of diffusion and osmosis after a course of instruction. *Journal of Research in Science Education, 32,* 45-61.

Phillips, M.D. (1972). A simple approach to experimental errors. *Physics Education, 7*, 383-388.

Pillay, S. Buffler, A, Lubben, F. & Allie, S. (2008). Effectiveness of a GUM-compliant course for teaching measurement in the introductory physics laboratory. *European Journal of Physics,* 29, pp. 647-659.

Roberts, D.(1983). Errors, discrepancies, and the nature of physics. *The Physics Teacher, March*, 155-160.

Rollnick, M., Lubben, F., Lotz, S, & Diamini, B. (2002). What do underprepared students learn about measurement from introductory laboratory work? *Research in Science Education, 32,* 1-18.

Salkind, N.J. (1997). Experimental research methods. *Exploring research* (3rd ed.). Upper Saddle River, NJ: Prentice Hall, 233-248.

St. John (1980). Thinking like a physicist in the laboratory. *The Physics Teacher, 18,* 436-442.

Swartz , C.E. (1993) *Used math for the first two years of college science.* College Park, MD: American Association of Physics Teachers.

Tan, K.C.D., Goh, N.K., Chia, L.S., & Treagust, D.F., (2002). Development and application of a two-tier multiple choice diagnostic instrument to assess high school students' understanding of inorganic chemistry qualitative analysis. *Journal of Research in Science Teaching, 39,* 283-301.

Taylor, J.R.(1997). *Error analysis: The study of uncertainties in physical measurement,* (2nd ed.). Sausalito, CA: University Science Books.

Treagust, D. F. (1995). Diagnostic assessment of students' science concepts. In S. Glynn & R. Duit (Eds.). *Learning science in the schools: Research reforming practice* (pp. 327-346). Hillsdale NJ: Erlbaum.

Treagust, D.F. & Duit, R. (2008). Conceptual change: A discussion of theoretical, Methodological and practical challenges for science education. *Cultural Studies of Science Education,* 3, 297- 328.

Treagust, D.F., Duit, R., & Fraser, B.J. (1996). Overview: Research on students' preinstructional conceptions - The driving force for improving teaching and learning in science and mathematics. In D.F. Treagust, D. Duit, & B.J. Fraser (Eds.), *Improving teaching and learning in science and mathematics* (pp. 1-14). New York: Teachers College.

Venville, G.J. & Treagust, D.F. (1997). Exploring conceptual change in genetics using a multidimensional interpretive framework. *Journal of Research in Science Education,* 35, 1031-1055.

Volkwyn, T.S., Allie, S., & Buffler, A. (2008). Impact of a conventional introductory Laboratory course on the understanding of measurement. *Physical Review Special Topics,* 4, 1-10.

Wiggins, G. & McTighe, J. (2005). *Understanding by design* (Expanded 2nd edition). Alexandria, VA: Association for Supervision and Curriculum Development.

Appendix A

WORST CASE METHOD INSTRUCTIONAL MANUAL

Worst Case Method Applied to the Propagation of Uncertainties: Supplement to Introductory Physics Laboratory

Table of Contents

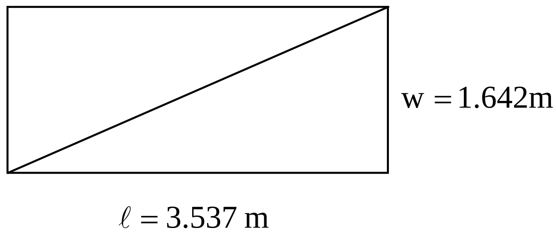
Uncertainty Analysis and Propagation of Uncertainty are concepts that are essential to data analysis in the physics laboratory. Many introductory physics students have not mastered this content, as they have not taken a calculus class. This manual is designed to supplement introductory physics lectures on uncertainty analysis and propagation of uncertainties without a pre-requisite knowledge of calculus. The process outlined in this manual for compounding uncertainties will be referred to as the **Worst Case Method**. This method can be applied to both simple and complex math functions. The first section of this manual demonstrates how uncertainties in measured quantities propagate into simple operations such as addition, subtraction, multiplication and division. The second section applies the worst case method to more complex mathematical operations, including powers and exponents, logarithms, trigonometric function, and graphical parameters such as slope and intercept.

Section 1

Simple Math Functions

The following example is discussed throughout this section on simple math functions.

Consider the perfect rectangle shown. It has sides ℓ = 3.537 m and w = 1.642 m along with a diagonal line. The measured values of ℓ and w each have an uncertainty of $\pm$ 0.005 m. Three geometric aspects of the rectangle are: a) the perimeter b) the area and c) the slope of the diagonal line.



The equations for each are: the perimeter, $P = 2\ell + 2w$; the area $A = \ell \cdot w$; and the slope, $S = w/\ell$. These three equations represent three different operations on the length and width of the rectangle namely addition, multiplication and division.

Before calculating values for the perimeter, area and slope, the following questions about the quality of the calculations should be addressed:

A) How does the uncertainty of $\pm$ 0.005 m in ℓ and w, affect the precision of the perimeter, area and slope?

B) How many decimal places should I include in the final values for the perimeter, area and slope?

C) Does the uncertainty of $\pm$ 0.005 m in ℓ and w have the same impact on the precision of the perimeter, area and slope?

These and other questions are answered in the next section, which quantifies the impact of the uncertainty in ℓ and w on the precision of the perimeter, area and slope. The mathematical steps used for quantifying the effect of the uncertainty in

ℓ and w on the three quantities are the same, but the results in the final answers are different.

1.1 Addition and Subtraction

Although the methods described below are used to find the uncertainty in the perimeter of a rectangle, it is important to recognize that these techniques apply to any calculation based on addition or subtraction.

A student wants to determine the perimeter of the perfect rectangle shown in

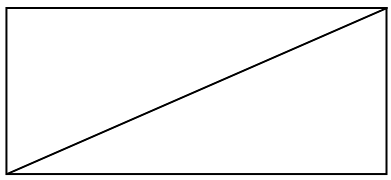


Fig. 1

Figure 1.

Using a meter stick the student measures the length as $\ell = 3.537\,\text{m}$ and the width as $w = 1.642\,\text{m}$. We will refer to these values as the **measurands.**

The uncertainty in the meter-stick is ± 0.5 cm or ± 0.005 m. This suggests that the maximum measurement of the length could be as large as

$\ell_{max} = 3.537 + 0.005\text{m} = 3.542\,\text{m}$ and the minimum measurement as small as

$\ell_{min} = 3.537 - 0.005\text{m} = 3.532\,\text{m}$ **.**

One can be very confident that the true length is somewhere between 3.532 m and 3.542m. Similarly the range for the width is $w_{max} = 1.642 + 0.005\text{m} = 1.647\,\text{m}$ and $w_{min} = 1.642 - 0.005\text{m} = 1.637\,\text{m}$. One can be very confident that the true width lies somewhere between 1.637 and 1.647 meters.

The correct interpretation of these results is, that any person performing the same measurements of the length and width of the rectangle would expect to report a value of the length no smaller than 3.532 m and no larger than 3.542 m. The width would be no smaller than 1.637 m and not greater than 1.647m.

A more convenient way of reporting the measurements of the length and width is to include the measurand and its corresponding uncertainty in a single expression. Following this idea we would report the measurements as $\ell = 3.537 \pm 0.005\,\mathrm{m}$ and the width as $\mathrm{w} = 1.642 \pm 0.005\,\mathrm{m}$

The question is, if the length and width are not exactly known, how does this effect our confidence in the precision of the calculation of the perimeter of the rectangle. Let's first calculate the measurand of the perimeter. Referring to Fig 1, the perimeter is given by the formula $\mathrm{P} = 2\ell + 2\mathrm{w}$. To find the measurand for the perimeter we just plug in the measrurands of the length and the width into this formula or $\mathrm{P} = 2(3.537) + 2(1.642) = 10.358\,\mathrm{m}$

Is this value reported for the perimeter really accurate to three decimals? Does it also have an uncertainty? The length and width have an uncertainty, therefore, the perimeter must also have an associated uncertainty. To find the uncertainty in the perimeter we first find the largest and smallest value of the perimeter based on the range reported above for the length and the width.

The maximum value is

$$\mathrm{P}_{\max} = 2\ell_{\max} + 2\mathrm{w}_{\max}$$

$$\boxed{\mathrm{P}_{\max} = 2(3.542) + 2(1.647) = 10.378\,\mathrm{m}}.$$

The minimum value is

$$P_{min} = 2\ell_{min} + 2w_{min}$$

$$\boxed{P_{min} = 2(3.532) + 2(1.637) = 10.338\,\text{m}}$$

One can be very confident that the true perimeter lies somewhere between 10.338 and 10.378 meters. One could include all three values in reporting the result .The measurand of the perimeter is 10.358m with a range of 10.338m to 10.378 meters. This is somewhat cumbersome in that you have to report all three values. Note that the measurand lies exactly in the middle of the range. The difference between the measurand and the upper limit is 0.02m and the difference between the measurand and lower limit is - 0.02 m. The point is that the absolute difference is the same, 0.02 meters. The variance of ±0.02 m is defined as the uncertainty in the perimeter or symbolically $\delta P = \pm 0.02\text{m}$. A more convenient format is to include the uncertainty along with the measurand. Report the measurand along with this plus + and minus- value that is added and subtracted to the measurand to identify the lower and upper limits. If you add 0.02 m to the measurand we get the upper limit and subtracting 0.02 m we get the lower limit.

An equivalent method that produces the same result is to find the difference between the max and min values and divide by two. The difference is just 10.378 - 10.338 = 0.04 m. If we divide this by two, we split the difference and obtain the same uncertainty in the perimeter as follows:

$$\delta P = \frac{10.378 - 10.338}{2} = \pm\ 0.02\text{m}\ .$$

This is the uncertainty in the measurand of the perimeter. This quantity carries the same units as the measurand and is therefore more specifically called an *absolute uncertainty*.

We are now ready to report the final answer for the measurement of the preimeter. You might be tempted to report the perimeter as $10.358 \pm 0.02\text{m}$ **BUT THIS WOULD BE WRONG.** Consider the perimeter value of 10.358 m. The number eight, in the thousnths place, is not physically significant. According to the uncertainty in the perimeter, the hundreths place is already uncertain to $\pm 0.02\,\text{m}$. It does not make sense to report a value for the perimeter accurate to more decimal places than the uncertainty allows. **ALWAYS REPORT THE MEASURAND AND ITS UNCERTAINTY TO THE SAME DECIMAL PLACE.**

As a result, we must drop the eight and not use it to round off the value. The final best answer is $\boxed{10.35\,\text{m} \pm 0.02\text{m}}$

There is a shortcut that allows one to obtain the uncertainty diretly using just one equation. It applies to any equation that involves the sum of quantities. In this case the short cut is

$$\delta P = 2(\delta \ell) + 2(\delta w)$$

Plugging the numbers in we get

$$\boxed{\delta P = 2(0.005\text{m}) + 2(0.005\text{m}) = \pm 0.02\,\text{m}}$$

The results are identical to the worst case method outlined above. For a derivation of the short cut see appendix A)

1.2 Multiplication

Although the methods described below are used to find the uncertainty in the area of a rectangle, it is important to recognize that these techniques apply to **any quantity** that is the result of multiplication.

The same student wants to now determine the area of the same rectangle and the uncertainty in the area. The area is given by the formula $A = \ell w$ and therefore gives

$$A = (3.537\text{m})(1.642\text{m}) = 5.8078\text{m}^2$$

You should ask, " how many decimals should be included when reporting the area?" The answer depends on the uncertainty in the area. Let's calculate the uncertainty in the area using the same technique described above. First we find the max and min values for the area. The maximum value is

$$A_{max} = (\ell_{max})(w_{max})$$

$$\boxed{A_{max} = (3.542\text{ m})(1.647\text{ m}) = 5.834\text{m}^2}$$

The minimum value is

$$A_{min} = (\ell_{min})(w_{min})$$

$$\boxed{A_{min} = (3.532\text{ m})(1.637\text{ m}) = 5.782\text{m}^2}$$

Using the same equation for the uncertainty outlined above, we can determine the uncertainty in the area. Plugging in the numbers we get

$$\delta A = \frac{A_{max} - A_{min}}{2} = \frac{(5.834 - 5.782)\,m^2}{2} = \pm\,0.026\,m^2\,.$$

This is the *absolute* uncertainty in the area of the rectangle.

How many decimals should you include when reporting an uncertainty? The answer is: **"ALWAYS ROUND OFF ALL UNCERTAINTIES TO ONE SIGNIFICANT DIGIT"!**

The uncertainty would be rounded to $\delta A = \pm\,0.03\,m^2$. We can now answer the first question "how many decimals should I include when reporting the measurand, or the area?" The answer is **TWO** decimal places because we must, **"ALWAYS REPORT THE MEASURAND AND ITS UNCERTAINTY TO THE SAME DECIMAL PLACE".**

We must conclude that the area and its uncertainty is

$$\boxed{5.81 \pm 0.03 \ \text{m}^2}$$

We can now state with confidence that the true area is definitely somewhere between 5.78 and 5.84 m^2 .

There is a shortcut that allows one to obtain the uncertainty in the area diretly using just one equation. It applies to any equation that involves the product of quantities. In this case the short cut is

$$\delta A = \left[\ell(\delta w) + w(\delta \ell)\right]$$

Plugging the numbers in we get

$$\boxed{\delta A = \left[3.537\text{m}(0.005\text{m}) + (1.642\text{m})0.005\text{m}\right] = 0.0259\text{m}^2}$$

Rounded to one significant figure the answer is

$$\boxed{\delta A = \pm 0.03 \ \text{m}^2}$$

The results are identical to the method outlined above. For a derivation of this short cut see the appendix B)

1.3 Division

Although the methods described below are used to find the uncertainty in the slope of a rectangle, it is important to recognize that these techniques apply to any quantity calculated that is based on **division.**

Consider the diagonal line drawn through the corners of the same rectangle in Figure 1. The student wants to calculate the slope of this line. Assuming the lower point goes through the exact corner of the rectangle, the slope is calculated by dividing the width by the length. So we have:

$$\boxed{S = \frac{w}{\ell} = \frac{1.642\text{m}}{3.537\text{m}} = 0.464}$$

This is the measurand of the slope.

To determine the uncertainty in the slope we employ the same procedure used above for the perimeter and area calculations. We first find the max and min values for the slope based on the min and max values of the length and width just like before. There are, however, a c

$$S_{max} = \frac{w_{max}}{\ell_{min}}$$

Here we want the quotient to be a **maximum** so we must use w_{max} in the numerator and ℓ_{min} as a denominator

$$\boxed{S_{max} = \frac{1.647\text{m}}{3.532\text{m}} = 0.4663}$$

$$S_{min} = \frac{w_{min}}{\ell_{max}}$$

Here we want the quotient to be a **minimum** so we must use w_{min} in the numerator and ℓ_{max} as a denominator

$$\boxed{S_{min} = \frac{1.637\text{m}}{3.542\text{m}} = 0.462}$$

The uncertainty in the slope ie, δS, is found the same way as before.

$$\delta S = \frac{S_{max} - S_{min}}{2} = \frac{(0.4663 - 0.462)}{2} = \pm\, 0.00215$$

Remember to **"ALWAYS ROUND OFF ALL UNCERTAINTIES TO ONE SIGNIFICANT DIGIT"!** The uncertainty in the slope, reported to one significant figure is

$$\boxed{\delta S = \pm\, 0.002}$$

In this case, the uncertainty allows 3 decimal places for the slope. Recall that **THE MEASURAND AND ITS UNCERTAINTY SHOULD AGREE TO THE SAME DECIMAL PLACE".** The slope, along with its uncertainty, is reported as

$$\boxed{0.464 \pm 0.002}$$

Summary

Measurements are made on a single or on several variables. Physics provides the relationship, in the form of an equation, which defines a specific operation between the measured variables. These numbers are then substituted into the equation, which produces a result called the measurand. Inherent in all measurements is some degree of uncertainty. The worst case method provides a way to quantify the impact of these measurement uncertainties on the precision of the measurand. In the case of the rectangle we began with an uncertainty in the measured quantities of the length and width defined as $\delta\ell = \pm 0.005\,\mathrm{m}$ and $\delta w = \pm 0.005\,\mathrm{m}$. These values were used to calculate the measurands of the perimeter, area and slope. As a result of the uncertainty in the length and width, the perimeter, area and slope each acquired a related uncertainty.

Table 1 summarizes the results obtained thus far. Included in the table are the measurands and the related uncertainty. The uncertainty is reported as both an *absolute* uncertainty, with units, and in the form of percent uncertainty. The percent uncertainty is calculated by the following equation

$$\text{Percent Uncertainty} = \frac{\text{uncertainty in measurand}}{\text{measurand}} \times 100$$

Uncertainty Results

Variable	Measurand	Uncertainty	Percent uncertainty
Length	3.537 m	± 0.005 m	0.1
Width	1.642m	± 0.005 m	0.3
Perimeter	10.35 m	± 0.02 m	0.2
Area	5.81 m^2	± 0.03 m^2	0.5
Slope	0.464	± 0.002	0.4

Note that the percent uncertainty in the area is over two times geater than the perimeter uncertainty. The point is that the size of the uncertainty of the measurand depends on the mathemeatical relationship between the other measuremnts, in this case ℓ and w. Three important observations concerning the reported values are:

1) The measurand and its absolute uncertainty are reported to the **same number of decimal places.**
2) The uncertainty values are all rounded to **one significant figure.**
3) The uncertainty values have the **same units** as the measurand.

Once the uncertainties in the measurements have been identified, the worst case method consists of three simple steps.

1) Add these uncertainties to the measurement values and plug these into the equation to obtain the **maximum** value.
2) Subtract these uncertainties from the measurement values and plug the results into the equation to obtain the **minimum** value.
3) Determine the difference between the max and min values and divide by two. This same operation can be employed to calculate the uncertainty in any equation.

Other mathematical forms that define the relationship between many physical quantities include: exponential, logarithmic and trigonometric functions. The next section provides examples of the **Worst case method** to these other important functions with answers. The detailed steps have been omitted. The student should check each answer.

Section 2

Complex Math Functions

2.1 Exponents and Powers

Example. According to theory, the total power output of a light bulb is given as $P = kT^4R^2$ where R = radius = 0.045 m and T is the temperature = 373 ^{o} . The constant "k" $= 7.125x10^{-7} \frac{w}{^o k^4 m^2}$.

The uncertainties in the radius and the temperature are: $\delta R = \pm 0.003\,m$ and $\delta T = \pm\ 2^o K$.

Calculate the power and the uncertainty in the power. According to the calculator, P = 27.92834 ± 4.6751 watts. Of course we must report the uncertainty to one significant figure and the measurand to the same decimal place. The final correct answer in the form of $P \pm \delta P$ is

$$\boxed{P = 27 \pm\ 5 \text{ watts}}$$

There is a shortcut that allows one to obtain the uncertainty diretly using only one equation. It applies to any equation that involves powers. In this case the short cut is

$$\boxed{\frac{\delta P}{P} = \frac{4\,\delta T}{T} + \frac{2\,\delta R}{R}}$$

Plugging in the numbers we get $\frac{\delta P}{P} = \frac{4(\,2^o\,)}{373^o} + \frac{2\,(0.003m)}{0.045m} = 0.155$ or simply

$$\boxed{\frac{\delta P}{P} = 0.155}$$

This is the relative uncertainty. Multiplying this by 100 gives the percent uncertainty or 15.5 %. To get the absolute uncertainty, solve for δP in the above equation

$$\delta P = P\,(0.155) = 27.9W(0.155) = \pm\ 4.32 \text{ watts}$$

This agrees very well with the direct substitution method. For a derivation of this short cut see the appendix D).

2.2 Logarithms

Example. The human ear correlates sound intensity with loudness according to the decibel scale. The decibel scale is related to the intensity of sound by $dB = 4.3\ \ln(I_r)$ where dB is in decibels, I_r = relative sound intensity and 4.3 is a constant. $I_r = 50000 \pm 250$ watts/m^2

Calculate the loudness in decibels and the uncertainty in decibels. The calculator gives 46.525 dB ± 0.02152 dB. Of course we must report the uncertainty to one significant figure and the measurand to the same decimal place. The final correct answer in the form of $dB \pm \delta dB$ is

$$\boxed{46.5 \text{ dB } \pm\ 0.02\text{dB}}$$

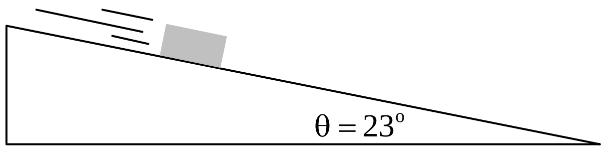
There is a shortcut that allows one to obtain the uncertainty diretly using just one equation. It applies to similar equations that involve logarithums . In this case the short cut is

$$\boxed{\delta dB = 4.3\left[\frac{\delta I_r}{I_r}\right]} \quad \text{or} \quad \delta dB = 4.3\left[\frac{250}{50000}\right] = 0.0215$$

The results are identical to the direct substitution method. For a derivation of this short cut see the appendix E).

2.3 Trigonometric functions

Example. A popular physics experiment is to determine the acceleration of a



frictionless cart moving down an incline and compare this with the acceleration predicted by theory.

Using a spark timer, measurements are taken on the position and time as the cart slides down the incline. The acceleration and its uncertainty is calculated to be: $a = 3.51\,m/s^2 \pm 0.07\ m/s^2$. This is the experimental value.

From theory the acceleration is given as $a = g\sin\theta$. The angle is measured to be $\theta = 23 \pm 1^o$ ($\delta\theta = \pm 1^o$), and the acceleration of gravity is g = 9.81 m/s^2.

Determine the theoretical acceleration and its uncertainty. Using the method of direct substitution, the answer in the form of $a \pm \delta a$, is

$$3.833 \pm 0.1576\ m/s^2$$

This is the theoretical acceleration.

Again we must report the uncertainty to one significant digit. and the measurand to the same decimal place. The final correct answer is

$$\boxed{3.8 \pm 0.02\ m/s^2}$$

There is a shortcut that allows one to obtain the uncertainty in the theoretical acceleration diretly using just one equation. It applies to any equation that involves the sin function . In this case the short cut is

$$\boxed{\delta a = g\cos\theta\delta\theta}$$

The uncertainty in the theoretical acceleration is $\delta a = g\cos\theta\delta\theta$ where $\delta\theta$ must be in radians. As stated above $\delta\theta = \pm 1^{o}$ and converting to radians, we have

$$\delta\theta = \pm 1^{o} \times \frac{\pi\,\mathrm{rad}}{180^{o}} = 0.0174\,\mathrm{rad}$$

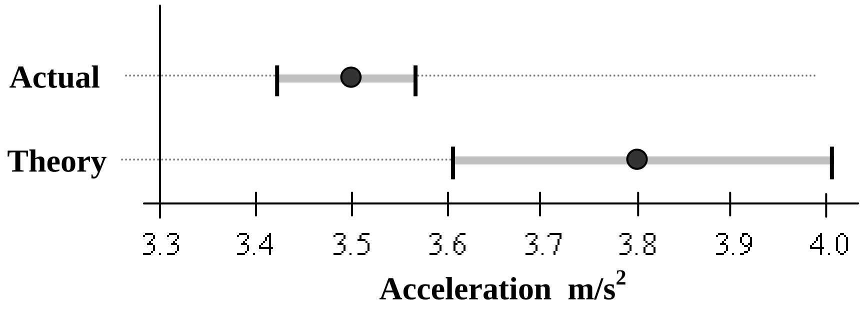
The uncertainty in the theoretical acceleration is

$$\delta a = 9.81 \mathrm{m/s^2}(\cos 23^{o})0.0174 = 0.157 \mathrm{m/s^2}$$

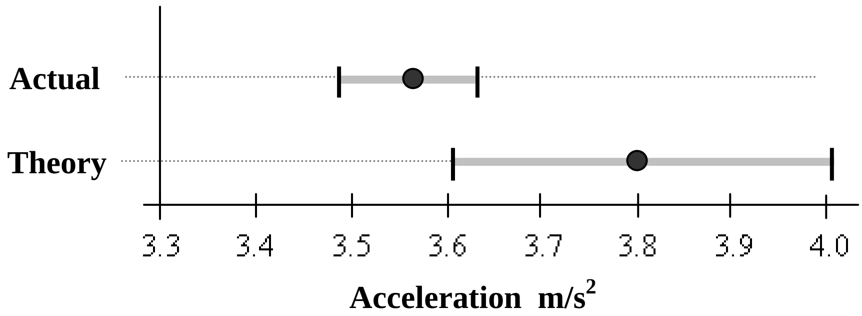
The results are identical to the direct substitution method. For a derivation of this short cut see the appendix F)

Comparison between theoretical and experimental values of acceleration.

The theoretical acceleration is reported as $a = 3.8 \pm 0.2\ \mathrm{m/s^2}$. From above the experimental acceleration is $a = 3.51 \pm 0.07\ \mathrm{m/s^2}$. One must now ask, is this discrepancy between the two values a result of measurement uncertainty or something else?. When these values are plotted on a number line, as seen below, the uncertainty bars **do not overlap**. This fact indicates that the discrepancy between the two values is due to something other than measurement uncertainty, such as the presence of friction that the theory does not account for.



After the track was cleaned, more measurements were obtained and plotted below



As seen above, even though there is still a discrepancy between the two values, the left and right uncertainty bars overlap indicating agreement between theory and experimental values.

Section three

Graphical Parameters

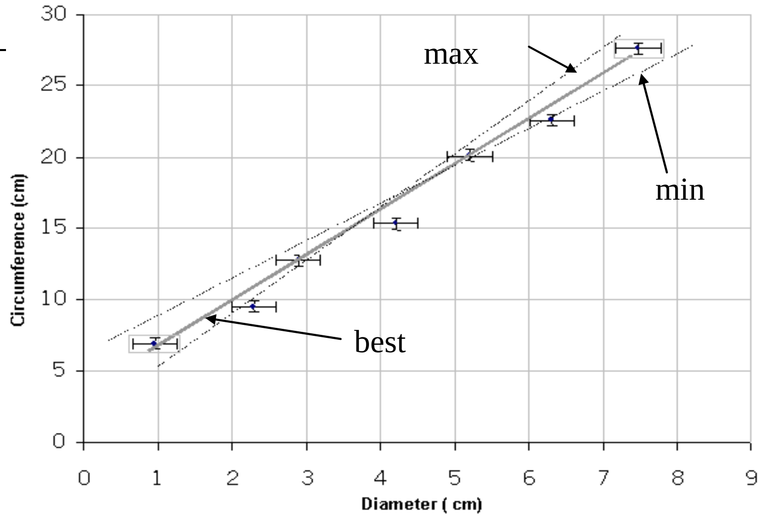
3.1 Slope

Most physics laboratory exercises require students to plot data in graphical form and determine the slope and intercept. The objective of this section is to demonstrate how the uncertainty in measured quantities can affect the accuracy and reliability of graphical parameters such as slope and intercept. The accuracy and reliability of the slope and intercept cannot be greater than the data used to calculate each.

Example. In a physics lab, a student obtains measurements of the diameter and circumference of several round objects. The data is plotted on a graph of diameter vs. circumference.

The uncertainty in each measurement is included in the table.

Diameter (cm)	Circumference (cm)
$\delta D = \pm\ 0.005$	$\delta C = \pm\ 0.05$
1.120	4.51
2.305	6.81
2.920	10.05
4.215	12.62
5.235	17.46
6.320	19.95
7.395	24.11



The bars on each data point, displayed in the graph, are included with each data point to illustrate the uncertainty of each measurement of the circumference and diameter. Consider the blow up of the data point (D_1, C_1) or 1.120 cm, 4.51 cm in the

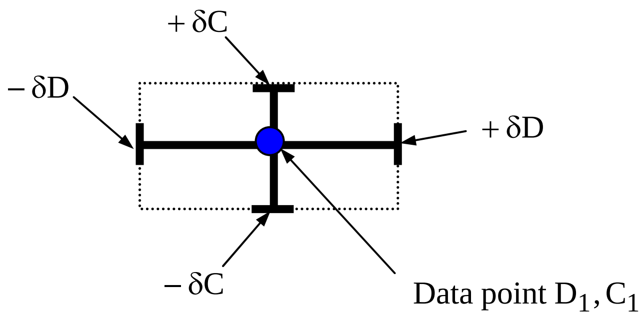


diagram. The uncertainty bars indicate the

uncertianty in both the X and Y varibles i.e. D (diameter) and C (circumference). The data point (D_1, C_1) , highlighted in gray in the table, could have occurred with equal probabilty anywhere in the area outlined by the dotted rectangular box.

Calculate the slope.

The slope is calculated using two points on the line, and each point has an uncertainty, therefore the slope must also be uncertain to some degeree. In other words, the slope must have a range of possible values just like the parameters of the rectangle analyzed earlier. This range is bounded by the **maximum** and **minumum** slope lines shown above. We will find the max and min values later.

First calculate the measurand for Pi (π). The value of Pi will be determined from the data used to plot the graph. The value of Pi can be calculated from the formula $\pi = \frac{C}{D}$ (assuming zero intercept) . This can be rearranged as $C = \pi D$. This formula is in the same format as the standard equation of a straight line or $y = mx + b$ where m is the slope and b is the intercept. So the linear equation of the graph above, in terms of C and D, is identified as $\boxed{C = \pi D + b}$. There is a y intercept in this example, so this is indicated by adding b to the formula.

If this equation is correct, then the slope of the graph should be π . According to the standard slope equation for a straight line the slope is given by:

$$\pi = \frac{\Delta C}{\Delta D} .$$

When expanded, we have

$$\pi = \frac{C_2 - C_1}{D_2 - D_1} .$$

First lets calculate the slope, or Pi, based on the best-fit line shown above. Always try to select data points **that actually fall on the line** as these will more accurately reflect the slope.

We will use the shaded data points in the table above or (D_1, C_1) or (1.120 cm, 4.51 cm) and (D_2, C_2) or (7.395 cm. , 24.11cm) . Plugging in to the formula we have:

$$\boxed{\pi = \frac{C_2 - C_1}{D_2 - D_1} = \frac{24.11 - 4.51}{7.395 - 1.120} = 3.123505976}$$

Rounded off to three significant figures π = 3.12. This is the measurand or the best etimate of Pi. We now must determine the range of possible values of Pi based on the known uncertainties in the circunference and diameter. These uncertainties are $\delta C = \pm 0.05$ cm and $\delta D = \pm 0.005$ cm . Lets first determne the max value of the slope, or Pi. Reffering to figure 2, look at the max slope line in the graph. The lower part of the line goes through the lower right corner of the box and the upper part of

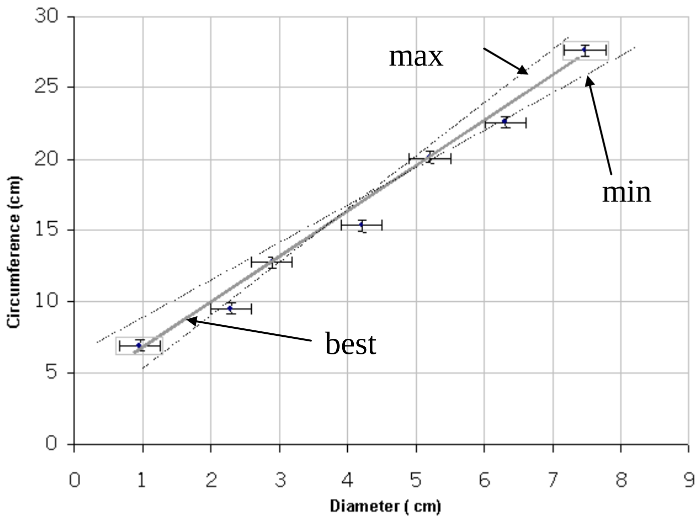


Fig 2

the line goes through the upper left corner of box.

A blow up of the first data point (D1, C1) on the graph is illustrated by figure 3.

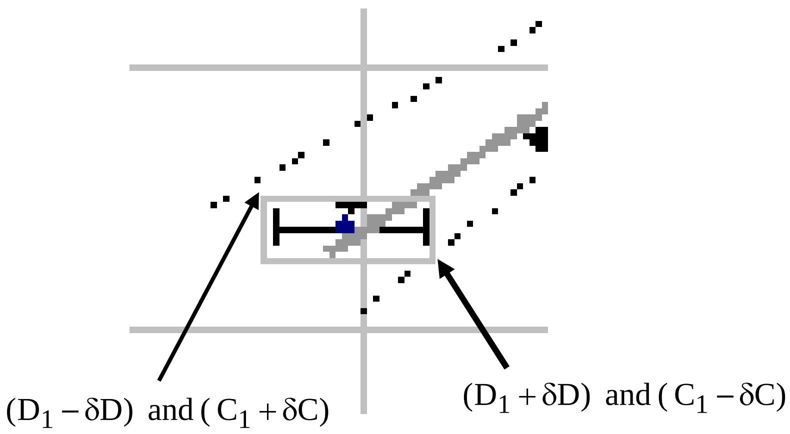


Fig. 3

From this we see that D_1 could as large as $(D_1 + \delta D)$ and C_1 could be as small as $(C_1 - \delta C)$ The second data point , (D_2, C_2),is treated similarly. D_2 could have a minimum of $D_2 - \delta D$ and C_2 could have a maximum of $C_2 + \delta C$. The maximum slope is

$$\pi_{max} = \frac{C_{2\,max} - C_{1\,min}}{D_{2\,min} - D_{1\,max}}$$

This arrangemnt maximizees the numerator and minimizese the denominator which results in the largest possible value of Pi ! Following the same logic, we simply switch the max min values for each data point and we generate the minimum value for Pi :

$$\pi_{min} = \frac{C_{2\,min} - C_{1\,max}}{D_{2\,max} - D_{1\,min}}$$

The **maximum** value is:

$$\pi_{max} = \frac{(C_2 + \delta C) - (C_1 - \delta C)}{(D_2 - \delta D) - (D_1 + \delta D)} = \frac{(24.11 + 0.05) - (4.51 - 0.05)}{(7.395 - 0.005) - (1.120 + 0.005)} = 3.144$$

and the **minimum** is:

$$\pi_{min} = \frac{(C_2 - \delta C) - (C_1 + \delta C)}{(D_2 + \delta D) - (D_1 - \delta D)} = \frac{(24.11 - 0.05) - (4.51 + 0.05)}{(7.395 + 0.005) - (1.120 - 0.005)} = 3.103$$

The range of values is 3.103 to 3.144 . As demonstrated in the previous examples the "±" uncertainty in Pi is:

$$\boxed{\delta\pi = \frac{\pi_{max} - \pi_{min}}{2} = \frac{3.144 - 3.103}{2} = \pm\, 0.0205}$$

"ALWAYS ROUND OFF ALL UNCERTAINTIES TO ONE SIGNIFICANT DIGIT"!

Rounded off to one significant figure:

$\boxed{\delta\pi = \pm\ 0.02\ .}$

Using the calculator, the measurand for Pi is 3.123505976. We now understand that all digits past the hundredth place are not significant.

THE MEASURAND AND ITS UNCERTAINTY SHOULD AGREE TO THE SAME DECIMAL PLACE".

We report our results in the standard form of $\pi \pm \delta\pi$ as:

$$\boxed{3.12 \pm 0.02.}$$

This result suggests that if the experiment is repeated, the value of Pi will fall somewhere between 3.12 ± 0.02. One may conclude, that although the experiment did not produce the exact accepted value of π , the accepted value of π =3.14 falls within this range. The answer, including the uncertainty, implies the technique used for determining the value of π is at least reasonable and supports the proposed relationship. It may be possible to obtain closer agreement if one could reduce the uncertainty in the measurements of both D and C.

As you can see, the direct substitution method to obtain the uncertainty in the slope is very time consuming and vulnerable to mistakes. The following shortcut is much simpler and less likely to result in mistakes. It can be applied to determine the uncertainty in any slope if the uncertainties in the x and y variables are known. In this case the short cut is :

$$\boxed{\frac{\delta\pi}{\pi} = \frac{2(\delta D)}{\Delta D} + \frac{2\,(\delta C)}{\Delta C}}$$

Here ΔC is just the **rise** in the slope and ΔD is just the **run** in the slope. Plugging in he numbers we get

$$\frac{\delta\pi}{\pi} = \frac{2(0.005)}{(7.395 - 1.120)} + \frac{2\,(0.05)}{(24.11 - 4.51)} = 0.0067 \text{ or}$$

$$\frac{\delta\pi}{\pi} = 0.0067$$

Multiplying 0.0067 by π we get $\delta\pi = 3.12\,(0.0067) = 0.02089$ or

$$\boxed{\delta\pi = \pm\,0.02}$$

The results are identical to the direct substitution method but much easier. For a derivation of this short cut see the appendix G)

3.2 Intercept

Calculate the Intercept. Since the intercept depends on the slope and the slope has an uncertainty, then the intercept will have uncertainty as well. First lets calculate the intercept from the data. If $\boxed{C = \pi D + b}$ then $\boxed{b = \overline{C} - \pi\overline{D}}$ The bar above the C and D indicate the average of D and C. Then

$$b = 13.64 - (3.12)4.216 = 0.486\text{cm} \quad .$$

To find the range of the intercept values we use the uncertainties again.

$$b_{max} = (\overline{C} + \delta C) - \pi_{min}(\overline{D} - \delta D)$$

$$\boxed{b_{max} = (13.64 + 0.05) - 3.10(4.216 - 0.005) = 0.64\,\text{cm}}$$

and

$$b_{min} = (\overline{C} - \delta C) - \pi_{max}(\overline{D} + \delta D)$$

$$\boxed{b_{min} = (13.64 - 0.05) - 3.14(4.216 + 0.005) = 0.336\text{cm}}.$$

This gives $b_{max} = 0.64\,\text{cm}$ and $\boxed{b_{min} = 0.336\text{cm}}$. Again we report the uncertainty in the intercept as $\pm\,\delta b$

$$\boxed{\delta b = \frac{b_{max} - b_{min}}{2} = \pm\frac{0.64 - 0.336}{2} = \pm\,0.152\,\text{cm}}. \text{ saguaro}$$

Rounding this value off to one significant figure our final value for the intercept is reported as:

$$\boxed{0.5 \pm 0\,.1\ \text{cm}}$$

In the conclusion section of a laboratory report one should report the final results as

3.12 ± 0.02 for Pi and 0.5 ± 0.1 cm for the intercept.

Remember that in physics experiments, the slope and intercept are not just numbers but always relate to some physical aspect of the experiment.

Appendices

A) Sums and Differences

Consider the function, $f = ax + by$, defined with constants *a* and *b*. Example physics equations are $v_f = at + v_o$, $F = ma$, $x_f = vt + x_o$ The independent variables *x* and y have individual uncertainties of δx and δy. Assuming the uncertainties are random , then the upper limit of uncertainty for *x* is $(x + \delta x)$, and for *y* is$(y + \delta y)$. Substituting these maximum values for *x* and *y* results in a largest possible value of *f* given as $(f + \delta f)$ We want to find uncertainty in the function *f* , i.e. δf .

$$f = ax + by \tag{1}$$

If I add the uncertainties δx and δy to *x* and *y* then *f* is increased by δf ,

$$f + \delta f = a(x + \delta x) + b(y + \delta y) \tag{2}$$

Solving now for δf results in the following:

$$\delta f = ax + a\delta x + by + b\delta y - f \tag{3}$$

Substituting $(ax + by)$back in for f results in:

$$\delta f = (a\delta x + b\delta y) \tag{4}$$

To find the minimum value of *f*, I need to subtract the values of δx and δy and get $f - \delta f = a(x - \delta x) + b(y - \delta y)$. Following the same steps above the result is the same or

$\delta f = \pm(a\delta x + b\delta y)$. So the "±" allows for both maximum and minimum values.

This result shows the addition rule for the propagation of uncertainty. Given two or more variables, *x* and *y*, each with an uncertainty δx and δy, then the uncertainty in the sum of these parameters is the sum of the individual uncertainties multiplied by the respective constants *a* and *b*. The same result is obtained when the function is the difference between the same parameters, i.e. $f = ax - by$. The uncertainties still add, and we get $\delta f = \pm(a\delta x + b\delta y)$

B) Multiplication

Consider the function below where k is some constant.

$$f = kxy \qquad (5)$$

Example physics equations are $V = IR$, $F = kx$, $P = mv$.

If I add the uncertainties δx and δy to *x* and , then the this adds δf to the function *f* . The goal is again to solve for δf.

$$f + \delta f = k(x + \delta x)(y + \delta y) \qquad (6)$$

Solving for δf we get

$$\delta f = k(xy + x\delta y + \delta xy + \delta x\delta y) - f \qquad (7)$$

Consider the cross term $\delta x \delta y$ in equation (3). If δx and δy are both small than $\delta x \delta y \cong 0$ and one can neglect its contribution, then we get:

$$\delta f = k(x\delta y + y\delta x) \qquad (8)$$

This result shows the product rule for the propagation of uncertainties and carries the same units as the function.

Dividing equation (8) by f = kxy we get

$$\frac{\delta f}{f} = \frac{\delta y}{y} + \frac{\delta x}{x} \qquad (9)$$

Equation (9) gives the relative uncertainty and is used for calculating percent uncertainty

C) Division

Consider the function below where k is some constant.

$$f = k\frac{y}{x} \qquad (10)$$

Example physics equations are $R = V / R$ and $a = F / m$.

If I rewrite this as

$$y = \frac{1}{k} f \; x$$

and Apply the product rule defined in the last section we get:

$$\frac{\delta y}{y} = \frac{\delta f}{f} + \frac{\delta x}{x}.$$

Solving this equation for δf/f we get:

$$\frac{\delta f}{f} = \frac{\delta y}{y} - \frac{\delta x}{x}$$

Since δx can be either ± we will assume the worst case and add the two so we get

$$\frac{\delta f}{f} = \frac{\delta y}{y} + \frac{\delta x}{x} \qquad (11)$$

This result shows the quotient rule for the propagation of uncertainties. It is identical to the product rule as seen in equation (9).

D) Powers and Exponents

Consider the function below where k is some constant

$$f = kx^n y^m \qquad (12)$$

Example physics equations are $F = kq^2 r^{-2}$, $T = 2\pi g^{-\frac{1}{2}} \ell^{\frac{1}{2}}$

If I add the uncertainties δx and δy to *x* and *y* then the this adds δf to the function *f*

$$f + \delta f = k(x + \delta x)^n (y + \delta y)^m$$

Factoring out the x and y we get

$$f + \delta f = k\left[x^n(1+\frac{\delta x}{x})\right]^n \left[y^m(1+\frac{\delta y}{y})\right]^m \quad = \quad kx^n y^m\left[(1+\frac{\delta x}{x})^n(1+\frac{\delta y}{y})^m\right] \tag{13}$$

The terms in parenthesis can be approximated by using the binomial expansion provided that $\delta x/x$ and $\delta y/y$ are small compared to 1 . The expanded terms are therefore:

$$(1+\frac{\delta x}{x})^n = 1 + n\frac{\delta x}{x} \quad \text{and} \quad (1+\frac{\delta y}{y})^m = 1 + m\frac{\delta y}{y}$$

Substituting these results back into equation (13) we get

$$f + \delta f = kx^n y^m\left[(1+n\frac{\delta x}{x})(1+m\frac{\delta y}{y})\right] \tag{14}$$

After multiplying the new terms we get

$$f + \delta f = kx^n y^m\left[1 + m\frac{\delta y}{y} + n\frac{\delta x}{x} + \left(nm\frac{\delta x\delta y}{xy}\right)\right] \tag{15}$$

The last term in equation (15) is $\left(nm\frac{\delta x\delta y}{xy}\right)$. Assuming the quotient $\frac{\delta x\delta y}{xy}$ is very small this last term is negligible and we get

$$f + \delta f = kx^n y^m\left[1 + m\frac{\delta y}{y} + n\frac{\delta x}{x}\right]$$

Now dividing both sides by $f = kx^n y^m$ we get

$$1+\frac{\delta f}{f}=\frac{kx^n y^m}{f}\left[1+m\frac{\delta y}{y}+n\frac{\delta x}{x}\right]$$

$$1+\frac{\delta f}{f}=\left[1+m\frac{\delta y}{y}+n\frac{\delta x}{x}\right]$$

Finally solving for $\frac{\delta f}{f}$ we get

$$\frac{\delta f}{f}=\frac{m\,\delta y}{y}+\frac{n\,\delta x}{x} \qquad (16)$$

This result shows the power rule for the propagation of uncertainties. This is again a relative uncertainty and is used for calculating the percent uncertainty.

E) Logarithms

Consider the function below where k is some constant.

$$f = k\ln x \qquad (17)$$

If I add the uncertainty in *x*, i.e. δx , to x then this adds δf to the function and we get

$$f+\delta f = k\ln(x+\delta x)$$

We want to find an equation for the uncertainty in the function *f* , i.e. δf . We begin by factoring out *x* and we get

$$f + \delta f = k \ln\left[x(1 + \frac{\delta x}{x}) \right]$$

Applying the product rule for logarithms we get

$$f + \delta f = k\left[\ln x + \ln(1 + \frac{\delta x}{x}) \right]$$

Consider the term $\ln(1 + \delta x / x)$. If $(\delta x / x) << 1$ then $\ln(1 + \delta x / x) \cong \delta x / x$ so we get

$$f + \delta f = k\left[\ln x + \frac{\delta x}{x} \right]$$

Substituting $f = k \ln x$ in for f we have

$$\delta f = k \frac{\delta x}{x} \tag{18}$$

This result shows the logarithm rule for the propagation of uncertainties.

F) Trigonometric Functions

Consider the function below where k is some constant.

$$f = k\sin\theta$$

If the angle has uncertainty of $\delta\theta$ then adding $\delta\theta$ and θ adds δf to the function f

$$f + \delta f = k\sin(\theta + \delta\theta) = k(\sin\theta\cos\delta\theta + \cos\theta\sin\delta\theta) \quad (\text{ trig identity })$$

Now $\cos\delta\theta \approx 1$ for small $\delta\theta$ i.e. like $\delta\theta = \pm\, 0.5$ degree.

And $\sin\delta\theta \approx \delta\theta$ (in radians) for small $\delta\theta$ i.e. like $\delta\theta = \pm\, 0.5$ degree

$$f + \delta f = k(\sin\theta(1) + \cos\theta(\delta\theta))$$

And solving for δf we get

$$\delta f = k(\sin\theta + \cos\theta(\delta\theta)) - f$$

$$\boxed{\delta f = k\cos\theta(\delta\theta)}$$

This result shows the trigonometric rule for the propagation of uncertainties.

G) Slope and Intercept

We begin by defining the slope of a straight line as:

$$m = \frac{\Delta y}{\Delta x} \tag{19}$$

Applying the quotient rule to equation (19) for uncertainty propagation gives

$$\frac{\delta m}{m} = \frac{\delta(\Delta y)}{\Delta y} + \frac{\delta(\Delta x)}{\Delta x}$$

Now Δx and Δy are defined as

$$\Delta x = x_f - x_i \quad \text{and} \quad \Delta y = y_f - y_i$$

Applying the difference rule to each for uncertainty propagation we get

$$\delta(\Delta x) = \delta x_f + \delta x_i = 2\delta x$$

$$\delta(\Delta y) = \delta y_f + \delta y_i = 2\delta y$$

This assumes that δx and δy are the same for the initial and final values. Upon substituting, we obtain the relative uncertainty for the slope m:

$$\frac{\delta m}{m} = \frac{2\delta y}{\Delta y} + \frac{2\delta x}{\Delta x} \tag{20}$$

This gives the relative uncertainty in the slope of a line.

Intercept

Now we can also determine how the uncertainty propagates into the *y* intercept b. The intercept is given by $b = y - mx$. To find the uncertainty in the intercept we can use the difference rule for uncertainty propagation and we get $\delta b = \delta y + \delta(mx)$. The second term, *mx*, is a product. Applying the product rule for uncertainty propagation gives

$$\frac{\delta(mx)}{mx} = \frac{\delta m}{m} + \frac{\delta x}{x} \tag{21}$$

Upon solving for $\delta(mx)$ we get

$$\delta(mx) = mx\left(\frac{\delta m}{m} + \frac{\delta x}{x}\right)$$

Substituting this back into equation (21) and simplifying gives

$$\delta b = \delta y + x\delta m + m\delta x \tag{22}$$

This gives the absolute uncertainty in the in the *y* intercept of a line and carries the same units as the *y* axis. Note how the uncertainty in the intercept, δb, depends on the uncertainty in the slope or , δm.

Appendix B

PILOT STUDY QUESTIONNAIRE

1) The main objective of any physics experiment is to try to obtain a result that agrees with the exact value predicted by the theory.

   a) I agree with this statement
   b) I do not agree with this statement.
   c)

The reason I selected this answer is because:

2) A mass attached to a string swings back and forth as a pendulum, The student makes seven measurements of the time for the mass to swing out and back. The seven measurements are 1.53, 1.53, 1.60, 1.53, 1.56, 1.53, 1.49, seconds. The true swing time is probably closest to:

   a) 1.53 seconds
   b) the average of all the times

The reason I selected this answer is because:

3) Two students calculate the velocity of a rolling ball by dividing distance by time. The distance traveled is 3.1 meters and the time is 1.3 seconds. The calculator displays the answer to seven decimal places. Student A reports a value of 2.4 m/s while student B reports a value of 2.384615 m/s.

   a) The value 2.384615 m/s is more accurate than 2.4 m/s
   b) The values are equally accurate.

The reason I selected this answer is because:

4) An instructor informs students that a meter stick is accurate to within ± 0.05 cm. This is defined as the uncertainty in the meter stick. A student ,using the meter stick , reports the length of a toothpick as
7.3 cm.

   a) The student used the meter stick correctly
   b) The student did not use the meter stick correctly.

The reason I selected this answer is because:

5) A student uses a meter stick that is accurate to within ± 2 % to calculate the volume of a perfect cube of side 2.00 cm. The volume is therefore 8 cm^3 .The calculated volume is only accurate to within

   a) 2 %
   b) 8 %
   c) 6 %

The reason I selected this answer is because:

6) A student estimates the length of a room by walking from one side to the other. The distance between each step is about one meter ± 2 %. She reports the length as 10 m. The margin of error in the length of the room is

   a) not greater ± 20 %
   b) still ± 2 %
   c) exactly ± 5 %

The reason I selected this answer is because:

7) The length of the room in #7) above was found to be exactly 11.5 meters. This means that the method of stepping off a room to obtain its length is

   a) not accurate enough
   b) Maybe accurate enough

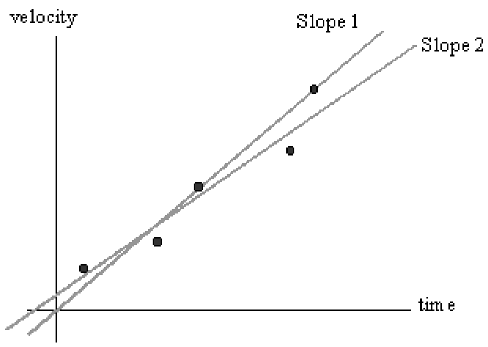
The reason I selected this answer is because:

8) Using a stopwatch accurate to three students record the time for a ball to fall a certain distance. The results, including the margin of error for each, are listed below. If the stopwatch is accurate to ± 0.05 seconds which answer is consistent with its margin of error.

   a) 1.32 ± 0.05 seconds
   b) 1.326 ± 0.05 seconds
   c) 1.3 ± 0.05 seconds

The reason I selected this answer is because:

9) Data is collected on the velocity and time of an accelerating car. The data is plotted on the graph below. The slope is the acceleration of the car. Based on the two slope lines draw on the graph which slope has the smallest margin of error.
   a) slope 1
   b) slope 2
   c) the two slopes have the same margin of error.



The reason I selected this answer is because:

10) A student performs an experiment that determines the acceleration of gravity as (975 ± 8) cm/s^2 . The accepted value of the acceleration of gravity is 981 cm/s^2.

a) There is a discrepancy between the two values.
b) There is not a discrepancy between the two values

The reason I selected this answer is because:

11) In the experiment above, 975 does not equal 981 cm/s^2 exactly, The measurements are different from each other due to human error. Which of the following is correct concerning human error

a) Human error is an acceptable characteristic of all scientific measurements
b) Human error can be reduced to zero
c) Human error helps explain why experimental and accepted values rarely agree

The reason I selected this answer is because:

12) A student runs an experiment designed to determine the density of water. Performing the same experiment twice the results are: 1.11 gm/cm^3 and 0.93 gm/cm^3 . The accepted value for the density of water is 1.00 gm/cm^3. Assuming no mistakes were made, I believe

a) the equipment used is probably **not** good enough to determine the density of water
b) the equipment is probably good enough to determine the density of water

The reason I selected this answer is because:

Appendix C

GENERAL LAB FORMAT

In general I will grade not on the quantity but on the quality of the content of your report. I will be looking for logical development and continuity. Each lab is worth 25 pts each. How these points are distributed is outlined below.

OBJECTIVE : (2 points) Present a clear descriptive statement about the purpose of the lab exercise. What is it you are investigating? What is predicted by the physics you are to examine?

PROCEDURE: (3 points) Here describe the method on how the data were collected. Consider the following points as a guide for writing about the procedure.

1. Write a statement about what you did. Describe the physical apparatus using diagrams and or drawings.

2. Discuss the equipment used and your technique on how it was used.

3. Include here any important measurement details i.e. if you had to put two meter sticks together to measure a length say so.

DATA AND ANALYSIS : (10 points) Your data are the numbers you obtained from the measurements taken. The data represent the relationship that may exist between variables in its simplest form. All data should be presented in neat legible tables. DON'T FORGET UNITS AND DATA TITLES. Consider the following points as a guide for reporting data and analysis.

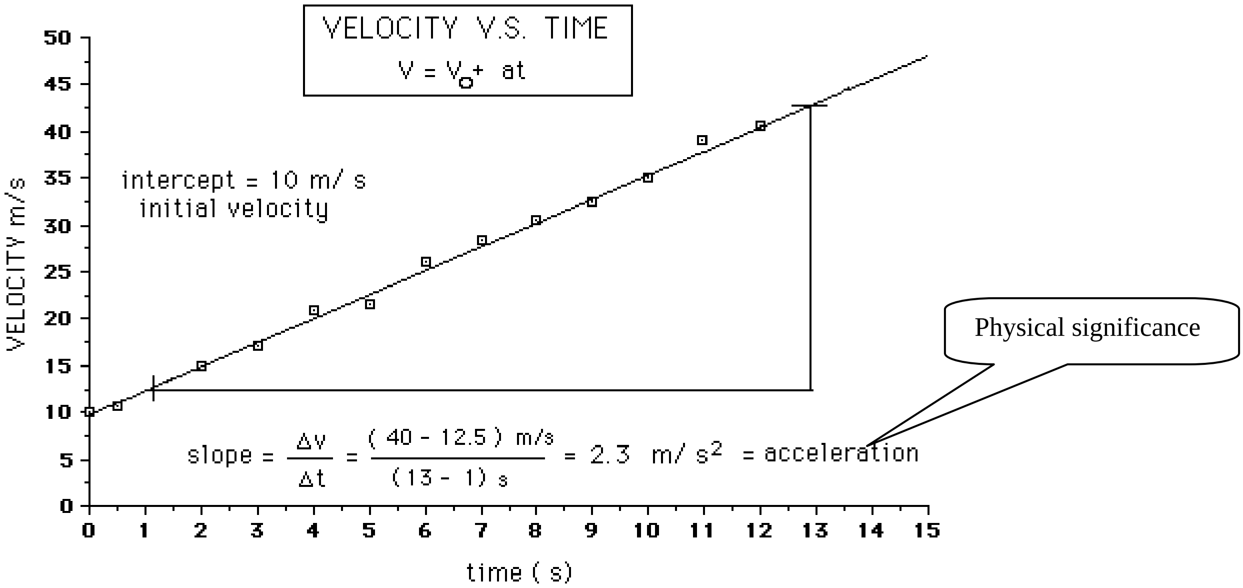
1. Include any equations and or calculations you perform on the data, (don’t just plug into calculator and show results, write down the calculation with units).

2. Be sure to comment on the accuracy of the equipment used, i.e. a voltmeter may be accurate to $\pm$ 0.05 volts or a meter stick to $\pm$ 0.05cm. Don’t report a length of 3.213 cm if the meter stick is only accurate to $\pm$ 0.05 cm.

3. Usually the analysis generates important numbers and results that are at the heart of the lab. Present these results in a table or some other clear format and DON'T FORGET UNITS . A good way to lose points is to report numbers without telling me where they came from and without units.

4. Include all the pertinent equations and how they relate to the lab exercise. Don't forget to identify all the variables i.e. V= voltage, a = acceleration etc.

5. Include all graphs and the calculations related to any graph i.e. the slope and intercept. Remember the numbers you get from calculating the slope and intercept HAVE UNITS and will help you to identify the **physical significance** of the slope and intercept of a graph.. Below is an example of a complete graph.



RESULTS AND CONCLUSIONS: (10 points) This is where you interpret the results from the previous section and identify any relationships that may exist between variables studied. Compare your conclusions about the results with the accepted physics. You must show logical development from one idea to the next. Consider the following points as a guide for writing a conclusion

1. Do the results support or refute any theory or predictions about the experiment. Be sure to identify what physics is making what predictions e.g. Newton's 2nd law predicts that a mass will accelerate directly proportional to the force applied etc.

2. Discuss any relationship the data may suggest. Most of the time the results will not reflect a perfect agreement with the physics being investigated i.e. there will be some % difference between the experimental and predicted values. You should include these calculated values of % difference and explain the origin of such discrepancies. If the experiment produced an acceleration of gravity of 9.72 $m/s^2 \pm 0.05\ m/s^2$ then report it here and calculate % difference from accepted value of 9.81 m/s^2 using this standard form below

$$\boxed{\%\Delta = \frac{\text{experimental value} - \text{accepted value}}{\text{accepted value}} \times 100}$$

3. You should backup any statements and/or conclusions with evidence from the data e.g. if Newton's second law predicts a linear relationship between force and acceleration then you must use the data and graphs to support or refute the theory in your conclusions. It is best to report the numeric values obtained in the analysis section to support any statements made in this section. Don't just talk about it – use the numbers obtained during the lab to support your argument.

Appendix D

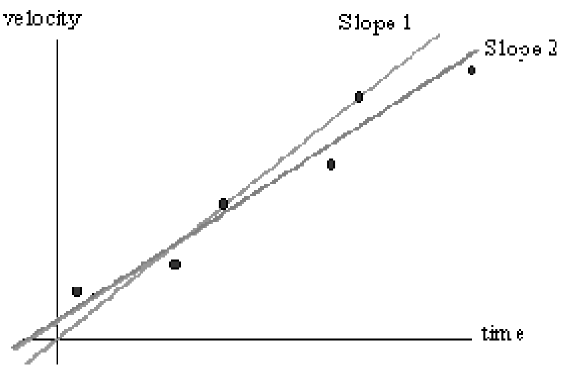
PRETEST AND POSTTEST

1) A mass, attached to a string, swings back and forth as a pendulum. The student makes seven measurements of the time for the mass to swing out and back. The seven measurements are 1.53, 1.53, 1.60, 1.53, 1.56, 1.53, 1.49, seconds. The true swing time is probably closest to:
2)
    a) 1.53 seconds
    b) the average of all the times

The reason I selected this answer is because:

a) No one value is more important than the others are as they all have equal merit.
b) 1.53 seconds occurs four out of seven times.
c) Averaging is the best method to use with a collection of numbers.
d) 1.53 seconds, as the other values were probably affected by outside influences.

3) Data is collected on the velocity and time of an accelerating car. The data is plotted on the graph below. The slope is the acceleration of the car. Two possible slope lines are drawn in. Which slope has the smaller margin of error.



a) slope 1
b) slope 2

The reason I selected this answer is because:

a) slope 1 goes through more actual data points.
b) slope 1 goes through the origin
c) slope 2 has equal number of points above and below the line
d) slope 2 is closest to most of the points

4) A student estimates the length of a room by walking from one side to the other. The distance between each step is one meter ± 2 %. She reports the length as 10 m. The margin of error in the length of the room is

a) ± 20 %
b) still ± 2 %

The reason I selected this answer is because:

a) the ± 2 % margin of error is constant for the whole distance measured.
b) the total distance of 10 meters has a margin of error of ± 2 % which equals ± 20 %

c) each step has an error of ± 2 % and therefore limits the error in the length to ± 2 %
d) the maximum error results from compounding the error in each step.

5) Two students calculate the velocity of a rolling ball by dividing distance by time. The distance traveled is 3.1 meters and the time is 1.3 seconds. Using these numbers, the calculator displays the answer to seven decimal places. Student A reports a value of 2.4 m/s while student B reports a value of 2.384615 m/s. The value 2.384615 m/s is more accurate than 2.4 m/s

a) I agree with this last statement
b) I do not agree with this last statement.

The reason I selected this answer is because:

a) More decimals imply higher degree of accuracy.
b) 2.4 m/s is the most accurate. The remaining decimals are not based on physical measurement.
c) the value of 2.384615 m/s rounds off to 2.4 m/s.
d) 2.384615 m/s has more significant digits.

6) It is a fact that the acceleration of gravity is 9.81 m/s^2. A student, using very good equipment, performs an experiment four separate times. The results are as follows: (**9.79, 9.82, 9.84, 9.78**) m/s^2
The student should keep trying until he obtains the accepted value of 9.81 m/s^2

a) I agree with this statement
b) I do not agree with this statement.

The reason I selected this answer is because:

a) It is not necessary to determine the exact value because all measurements have error no matter how careful you are.
b) The second value, 9.82 m/s^2, is very close to the accepted one.
c) The student needs to determine what is wrong and correct the problem.
d) Gravity is very constant and the experiment should produce the accepted value.

7) Human error is inherent in all scientific endeavors and impossible to eliminate.

a) I agree with this statement
b) I do not agree with this statement.

The reason I selected this answer is because:

a) use of extremely accurate instruments can eliminate human error.
b) Humans are not perfect and sometimes make errors that effect the results.
c) Eliminating mistakes will eliminate human error.
a) Usually there is a discrepancy between Experimental and accepted values. This is evidence of Human error.

8) A popular physics laboratory exercise is to measure the time required for a mass, attached to a string, to swing back and forth. Let's assume the theory time is exactly 1.30 seconds and the experiment produced a value of 1.27 ± 0.02 seconds. The percent difference between the theory and the experimental time is 2 %. The student must decide if the lab was successful or not. Which of the following do you agree with?

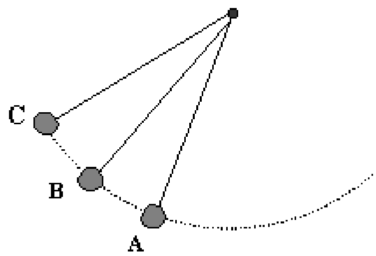
a) these results are acceptable
b) these results are unacceptable

The reason I selected this answer is because:

a) This percent difference is probably due to human error.
b) Physics is an exact science and the results should agree at least to the same decimal place.
c) The experimental value, rounded off to the same decimal place does agree with the theory value.
d) The precision of the instruments is probably not high enough.

9) Referring to the same experiment above, the student now wonders if the swing time depends on how far back the mass is before its released. Three different positions are tested. The results are shown below. The margin of error in the stopwatch is ± 0.02 seconds.

Position	**Swing time (seconds)**
A	1.27
B	1.29
C	1.31



a) The swing time does depend on where the mass is released.
b) The swing time does not depend on where the mass is released.

The reason I selected this answer is because:

a) the overall trend is in the correct direction of increasing swing time
b) the increase in the time is the same between A and B and B and C
c) all three numbers round-off to the same value of 1.30 seconds.
d) There is some overlap of the margin of error for each time.

10) A student, using a meter stick, measures the length of a playing card ten times. The student understands that the uncertainty in the meter stick is ± ½ millimeter. This value refers to _____________ .

a) the impact of human error on measurements.
b) The greatest difference between a measurement and the true value.

The reason I selected this answer is because:

a) no reason why the meter stick should reveal different values for the same object being measured.
b) all humans are not perfect and will most likely make mistakes when making measurements.
c) All instruments have limited precision.
d) The meter stick is not precise enough to determine the exact length of the card.

11) The table below shows poll data for two presidential candidates A and B. Based on the table of data, select one of the following conclusions:

CANDIDATE	PERCENT OF THE VOTE	MARGIN OF ERROR
A	50 %	$\pm$ 2 %
B	47 %	$\pm$ 2 %

a) Candidate A has a significant lead over B
b) The two candidates are tied

The reason I selected this answer is because:

a) it is more important to compare the values for **PERCENT OF THE VOTE** than the margin of error.
b) Candidate A has a 3-percentage point lead over candidate B, which is larger than the margin of error.
c) The margins of error overlap
d) The margin of error for each candidate is the same.

12) A student uses a meter stick that is accurate to within ± 2 % to calculate the volume of a perfect cube of side 2.00-cm. The volume is therefore 8 cm^3 .The calculated volume is only accurate to within

a) 2 %
b) 8 %
c) 6 %

The reason I selected this answer is because:

a) There are three sides. Each has an error of 2 % so 2 x 3% = 6 %
b) The volume is calculated by cubing the $(2.00\ cm)^3$ = 8 cm^3 .The error is determined the same way or
$(2\ \%)^3$ = 8 %.
c) Only one side is measured and it's a perfect cube so the error is 2 %.

13) A student runs an experiment designed to determine the density of water. Performing the same experiment twice the results are: 1.11 gm/cm^3 and 0.93 gm/cm^3. The accepted value for the density of water is 1.00 gm/cm^3. Assuming no mistakes were made, I believe

c) the equipment used is probably **not** good enough to determine the density of water
d) the equipment is probably good enough to determine the density of water

The reason I selected this answer is because:

a) the values are too different from each other
b) good enough- the difference is just human error
c) must make many more runs and average the values
d) Probably good enough if the margin of error was reported.

Appendix E

STUDENT PERMISSION AND LETTER

Curtin University of Technology
Science and Mathematics Education Centre.
Division of Engineering Science and Computing

I agree to participate in the research project being conducted by Paul Haugen. I understand that in addition to participating in the regular coursework, I may be asked to complete some questionnaires. I also understand that some aspects of my laboratory reports may be examined for research input and that I may be asked to participate in a follow-up interview.

It is my understanding that my name or any other identifying information will not be used in any written reports; that, if individuals are identified pseudonyms will be employed. I am assured that the data collected will be secured electronically for a period of five years after which they will be destroyed. I understand that my grade in this course is in no way related to or dependent on any part of this study. I also understand that I may unconditionally withdraw from this study at anytime.

Date: September 2, 2004

____
Name

____
Signature

Note: If you are under the age of 18 your parents' signature is required.