UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

PRE-SERVICE TEACHERS AND STATISTICS: INTERRELATIONSHIPS BETWEEN CONTENT CONFIDENCE, KNOWLEDGE, AND ATTITUDES; PEDAGOGICAL CONTENT KNOWLEDGE; AND TEACHER INTEREST IN PROFESSIONAL DEVELOPMENT IN STATISTICS

A DISSERTATION

SUBMITTED TO THE GRADUATE FACUTLY

in partial fulfillment of the requirements for the

Degree of

DOCTOR OF PHILOSOPHY

By

STEPHEN MAX LANCASTER
Norman, Oklahoma
2007

UMI Number: 3264594

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

PRE-SERVICE TEACHERS AND STATISTICS:INTERRELATIONSHIPS BETWEEN CONTENT CONFIDENCE, KNOWLEDGE, AND ATTITUDES; PEDAGOGICAL BELIEFS; CLASSROOM PRACTICES; AND TEACHER INTEREST IN PROFESSIONAL DEVELOPMENT IN STATISTICS.

A DISSERTATION APPROVED FOR THE
DEPARTMENT OF MATHEMATICS

BY

Curtis McKnight, Chair

Teri J Murphy, Chair

Krishnan Shankar

Marilyn Breen

Randa Shehab

Acknowledgements

I would like to thank my co-advisors, Dr. McKnight and Dr. Murphy for their confidence in me. I especially thank Dr. Murphy for the unreasonable amount of reading and editing that she was willing to invest in this paper. I would like to thank all of the members of my advisory committee for their efforts in scheduling meetings and for their support during my graduate work in general. I also would like to thank my committee members for their support during my work on this paper. I thank the members of the RISE institute for their assistance when I was first learning to do research and write papers. In particular, I would like to thank my co-authors from the paper on office hours and female engineering students; Dr. Murphy, Dr. Trytten, and Dr. Walden.

I would like to thank my wife, Marla, who has sacrificed much to help me pursue my ambitions. I thank Luke and John for their patience while I spent many evenings working. I also would like to thank my dad and my uncle who paved the way for me when they completed their graduate degrees those many years ago.

Table of Contents

Abstract

The purpose of this study was to identify key variables that exist during preservice teacher training and that could affect teacher interest in pursuing continuing professional development in statistics. A study of the current literature was conducted. This body of literature included the growing importance of statistics in the K-12 popular curriculum, the attitudes that undergraduate students tend to possess toward statistics as a subject to learn, the importance of beliefs and content knowledge for improved K-12 classroom learning, and the growing importance of continuing professional development to help improve teacher beliefs and content knowledge levels. An important result from the literature is past studies indicated that there is little correlation between teacher knowledge and teacher attitudes toward statistics.

Instruments were used to gather data about the defined variables. Two pre-existing statistics knowledge instruments were used as well as three pre-existing attitudes scales. One of the three attitudes scales measured two separate subcategories of attitudes. Two new instruments were constructed for the study. A short Likert-type instrument was designed to investigate participant attitudes toward continuing professional development in statistics. A grading project was developed to qualitatively investigate preservice teacher knowledge of statistics content as applied to teaching.

Results show that, for preservice teachers, many affective variables do not correlate with statistics knowledge levels. This corresponded to the existing literature. Yet there were moderate correlations between some of the affective variables, such as

attitudes toward statistics as a field, and the knowledge levels of preservice teacher groups in the study. A key to this new result was the inclusion of more specific attitude measures than previously used in studies.

There were other relevant results. Preservice primary teachers taking their first required (general introductory) mathematics course had significantly higher self-efficacy to use their current statistics knowledge than did other majors in the same general mathematics course. Preservice primary teachers taking their first required mathematics course had significantly higher self-efficacy to learn statistics in the future that preservice primary teachers who were taking their last (4^{th}) required mathematics course. Yet preservice primary teachers finishing their last required mathematics course had significantly higher attitudes toward statistics as a field than did the participants from the first required mathematics course for preservice teachers.

These results are suggestive rather than conclusive. Future research should focus on preservice teacher self-efficacy to learn statistics in the future among other issues. More work must be done for preservice mathematics teacher educators to extensively utilize the results from this study. However, we now know more about the factors affecting preservice teacher attitudes toward statistics than we did before this study.

Chapter 1 Introduction

The purpose of this study was to research those factors that may affect both primary and secondary level teachers' interest in supplementing their ability to teach statistics through professional development once they are practicing teachers. In this chapter, I defend the stance that this work is important to mathematics departments and that statistics is an important specific mathematical topic to investigate. I then present the objectives of the study.

The Call for Improved Statistics Education K-12

Research and various mathematics education standards indicate that statistics is gaining an increasing role in the K-12 curriculum (Mathematical Sciences Education Board 1996, 1997; National Center for Improving Student Learning and Achievement in Mathematics , 2004; National Center for Research in Mathematical Sciences Education, 1994a, 1994b; National Council for Accreditation of Teacher Education, 2003; National Council of Teachers of Mathematics, 1989, 1992, 2000). This new emphasis has led to a need to increase the statistical content knowledge and statistical pedagogical content knowledge of all teachers. As early as 1988, Garfield and Ahlgren recommended that researchers should design studies to investigate how K-12 students think statistically and probabilistically, and to find methods to help these students to learn stochastics in better ways (Garfield & Ahlgren, 1988). Makar and Confrey (2004) suggest that "Statistics and data analysis are becoming increasingly important in our society for a literate citizenry". (p. 354)

Reference to Both Statistical and general Results in Mathematics education Research

In this paper, the terms *statistical content, or statistics,* refer to concepts within the areas of probability, descriptive statistics, and inferential statistics. Thus the phrases statistical content, and statistics, are used to describe material often referred to as stochastics (Canada, 2004). This conformity was set for the instructional protocols for the preservice primary teachers. The goal was to reduce the chance that the preservice primary teachers would not recognize the terminology during the instructions and on the instruments. The conformity is maintained in the paper to allow uniformity between the language in the paper and the language in the instruments used in the study.

The primary mathematics topics in this study were statistical. Many of the literature references in this paper refer to statistics education research results. However, some areas of concern in statistics education had not been researched in a statistical framework. To provide insight into these issues, results from mathematics education research that is not topic specific were referenced.

The Relevance of the Study to Mathematics Departments

Mathematics departments serve many purposes. Some of these purposes vary from department to department. Many mathematics departments provide service courses for other departments within the college or university. One set of courses that most mathematics departments provide includes mathematics content-for-teaching courses for preservice teachers. Those members of a mathematics department faculty who are actively involved in the education of preservice teachers serve a special role.

For the purposes of this paper, these faculty members are referred to as *preservice mathematics teacher educators* (PMTEs).

General issues that mathematics departments must address to provide preservice teacher education. There are a variety of decisions that affect, on various levels, the service a mathematics department provides for preservice teachers. A decision may be made to increase the number of required courses in K-12 mathematics content for teaching. In such a case the mathematics department accrues an increased responsibility to these preservice teachers. Other decisions that must commonly be made include the selection of new texts and curricula to adopt in order to better fit the needs of preservice teachers based on indications from recent research.

Specific issues concerning the meanings that preservice teachers acquire due to their experiences in the preparation program. Much of the current research indicates that the way in which preservice teachers are taught affects their attitudes and beliefs as well as the content knowledge they accumulate. In the Mathematical Sciences Education Board letter report *The Preparation of Teachers of Mathematics: Considerations and Challenges*, (MSEB, 1996, p. 6) the MSEB recommends that "It's not just the mathematics. Knowing mathematics does not ensure the effectiveness of prospective teachers. How they come to know their mathematics matters as well." According to Llinares (2002), how teachers come to know their mathematics can affect teacher beliefs concerning mathematics teaching. In fact, beliefs can play at least two roles in the preservice teacher undergraduate classroom. Beliefs are entities that preservice teachers carry into the classroom; beliefs are also a potential target for learning (Llinares, 2002).

Attitudes and beliefs can be affected by addressing important non-content issues in the learning environment. The development of habits such as *reflective practices* can help teachers to appreciate the opinions they develop toward a mathematics topic as they work in a course to improve their content knowledge of that topic (Cooney, 1998; Llinares, 2002). A goal of Llinares was to present course material in a social context designed to allow the teachers to experience the material in much the way in which they will need to apply the content when they are teaching. Exposure to learning in a social context can affect teacher attitudes about teaching in a social context.

Beyond beliefs and attitudes, there are other factors affected by preservice teacher mathematical education. Development of pedagogical knowledge, providing exposure to alternative ways of learning mathematics, creating sensitivity to the feelings students have toward mathematics, and promoting reflection are all goals of preservice teacher education (Philippou & Christou, 2002; Zaslovsky & Leikia, 1999). Prospective teachers should be trained to be learners of what they need to know. To try to teach them everything that they may need to learn in the context of a single course or program is not realistic (Heaton, 2000). This philosophy also governs many of the premises that underlie this study.

Certainly teacher accumulation of content knowledge is a primary focus in mathematics methods courses. But preservice teachers also develop attitudes toward subject material and toward teaching methods during any preservice teacher course. Thus it is important for mathematics departments to carefully consider how these preservice teacher content methods courses are developed and taught.

Potential consequences of the beliefs and attitudes that preservice teachers keep or acquire during their preparation program. In chapter 2, a case is built for the importance of *continuing professional development for teachers* (CPD) in the process of improving mathematics learning in the kindergarten through twelfth grade (K-12) classroom. This situation places two burdens on mathematics departments. First, mathematics departments are likely to contribute to the CPD itself. Second, PMTEs have the potential to take responsibility for motivating teachers to return for CPD.

One question posited within this paper is "What correlation exists between (1) teachers' beliefs and attitudes about pursuing professional development related to a specific mathematics topic, and (2) other specific preservice teacher characteristics related to that topic." Some of these specific characteristics are teachers':

- beliefs concerning the topic's importance and difficulty,
- attitudes toward the learning and teaching of that topic,
- self-efficacy to learn that topic, and
- knowledge levels pertaining to that topic.

Each of these specific characteristics is also a construct that preservice teachers are likely to develop or solidify during their preparation program.

There are two primary situations under which PMTEs have the opportunity to influence teacher interest in CPD. First, they can make decisions affecting the preservice teacher preparation program. This affects teachers before they enter the field. Second, they can attempt to affect teachers once they are in the field. Certainly PMTEs have greater contact with teachers during the years of their preservice teacher preparation. If PMTEs are interested in increasing the likelihood that teachers will participate in CPD, then it is reasonable to consider possible factors that influence preservice teachers' interest in CPD and that may be addressed during the preparation program.

Objectives of the Study

One objective of this study was to develop a framework within which a more complete understanding of the attitudes that preservice teachers have towards statistics can take place. As far as the literature indicates, there has not yet been a study completed in which both attitudes toward statistics in general and, more specifically, self-efficacy towards understanding and learning statistics, were measured on preservice teachers. Although many studies have addressed either teacher (preservice teacher) attitudes toward the teaching of statistics or teacher attitudes toward statistics as a mathematical tool to use, neither of these types of teacher attitudes were the focus of this study. This study emphasized the attitudes and efficacy of preservice teachers to learn about statistics and to take part in developmental activities either related to the learning of statistics or related to learning about the teaching of statistics.

Another objective of this study was to search for indicators that may provide insight to factors (such as level of content knowledge) that positively affect preservice teachers' interest in pursuing professional development in the future. Similarly, the study attempted to identify factors that negatively affect such interest. Preservice teacher educators could then seek ways to avoid the development of these preservice teacher attributes.

The question for the study. Much of Chapters 1 and 2 is devoted to arguing the cause of CPD. However, the study did not directly investigate CPD. The focus of this dissertation was on what mathematics teacher educators in mathematics departments can do now, within preservice teacher programs to encourage and facilitate participation in statistics-related professional development. Specific focus was placed within the courses taught in the mathematics department. The goal was to increase the possibility that preservice teachers will participate in CPD. Specifically, this study was seeking answers to the question "Are there indicators from preservice teacher attitudes toward, and knowledge of, statistics that might assist PMTE efforts to increase the possibility that these preservice teachers will pursue CPD in statistics?"

A residual objective of this study was to identify correlations that may exist, either by quantifiable measures or by qualitative suggestive patterns, among the following factors:

- self-efficacy to use statistics
- self-efficacy to learn statistics
- attitudes toward statistics
- appropriate knowledge of statistics for the level to be taught
- ability to accurately grade and evaluate student work in statistics

A second residual objective was to compare two instruments that are each intended to measure the level of statistics knowledge in undergraduate students. These two instruments had not been administered to the same set of students in prior studies. In this aspect, this study served a unique role separate from its primary focus.

Structure of the Dissertation

In chapter two, I provide the theoretical basis for this study. In it I defend the need for CPD for teachers once they have completed their initial degree and have accumulated experience teaching K-12 students. Because successful development of improved K-12 educational opportunities is complicated – and involves factors that are dependent on teacher experience, beliefs, attitudes, content knowledge, and pedagogical content knowledge – it is necessary to provide a complete picture of these interactions.

In chapter three, I present the methodological structure of the study. In it I describe the setting of the study, including the participants, the data collection methods, and the specific surveys utilized. I also provide a rationale for choices made in the selection of participants approached for the study, the surveys chosen and created, other data collection methods utilized, and the types of data analyses used.

In chapter four I describe the data collected. In this chapter I discuss the analyses performed on the original data. The analyzed data come from both quantitative and qualitative analyses. Where exploratory data analyses occurred, explanations are provide for the patterns identified. Where statistical analyses occurred, the process by which the data was derived is provided. Patterns, relationships, correlations, and interactions identified are discussed with respect to practicality and potential extrapolation.

In chapter five I present the discussion of the results. In it I elaborate on any decisions concerning practicality of analyses from chapter four. The data analyses are related to the broader context of their implications for preservice teacher educators with respect to choices to be made concerning curricular and pedagogical issues. The hypotheses and the objectives of the study are re-evaluated. Answers are provided to the study question, "What interventions, if any, can be taken within the mathematics department preservice teacher program courses to encourage and facilitate future participation in professional development activities that focus on improving the teaching of statistics topics?" Within this section, suggestions for potential interventions in current preservice teacher content methods courses based on the results from chapter four are discussed. The weaknesses of the study are discussed. This includes suggestions for how the design of the study could have been improved to better answer the question of the study. There is also a discussion of the limitations that may exist with respect to the implementation of the suggestions. Suggestions for future research in this area are then presented.

Chapter 2 Literature Review and Rationale

The goal of the study was to identify factors within preservice teacher programs that will increase teacher interest in pursuing continuing professional development (CPD) once they are in the field, specifically with respect to statistics. The fundamental premise of the study was that CPD plays an important role in K-12 student learning. This makes preservice teacher interest in CPD important to preservice mathematics teacher educators (PMTEs) since the promotion of K-12 student learning is am important goal of all PMTEs. A key goal is for this chapter to defend the stance that CPD can affect K-12 student learning in ways that preservice teacher programs cannot.

In this chapter, I argue the importance of CPD in statistics education by showing that CPD can affect multiple teacher characteristics, each of which either directly or indirectly affects K-12 student learning of statistics. This goal is accomplished in two ways. First, evidence of connections between CPD and K-12 student learning from mathematics education is presented as relevant to statistics education via shared characteristics between mathematics and statistics learning. Evidence from statistics education is also presented. Second, I present issues specific to the process that teachers traverse as they progress from learning statistics to learning to teach statistics to teaching statistics. This process is referred to as *teacher preparation to teach statistics* (TPTS). The issues tied to this process shape preservice teacher training. They also affect CPD in the sense that some of these issues cannot be adequately addressed during preservice teacher training.

Outline for Chapter Two

To accomplish the key goal of this chapter, it is necessary to discuss individually important factors that shape TPTS. These factors include:

- the emergence of more effective methods of delivery of mathematics,

  In this section I present the emergence of constructivist learning methodologies, discovery learning, and cognitively guided instruction.

- teacher needs for implementing successful teaching strategies,

  In this section I address the three needs for teachers to implement successful teaching strategies; teachers must know the content well, teachers must know their students and how those students learn mathematics and statistics, and teachers must possess certain beliefs and attitudes toward the learning of and teaching of mathematics and statistics. I then present indications of dependency among these three needs.

- evidence that CPD can positively affect these needs,

  In this section, I present specific cases in which the literature calls for CPD to provide unique, critical support for teacher development.

- K-12 statistics education issues,

  In this section I address three primary focus issues in K-12 statistics education; calls for increased statistical content in the K-12 curriculum, the state of teacher preparation to teach statistics, and the role teacher attitudes toward learning statistics have in teacher development.

- the role of attitudes and beliefs on teacher acquisition of content knowledge and specific applications to the learning and teaching of statistics,

  In this section I develop an argument for the important role that attitudes toward learning a topic has in the learning of that content material. To accomplish this, I present work from studies on general mathematical attitudes and beliefs. I discuss attitudes toward statistics of undergraduates in general. To put the attitudes and the learning in context, I present evidence of teachers' statistical knowledge levels. I present results that focus on teachers as the learners of the mathematics and how their attitudes affect their learning. I then discuss teachers as the teacher rather than as the learner to help to clarify the differences between these two categories of attitudes, attitudes toward learning statistics and attitudes toward teaching statistics.

- professional development programs for statistics education.

  In this section I present samples of programs that have been implemented to improve K-12 statistics education.

During the discussion addressing all of these factors, I present evidence that many of these factors are dependent on each other. The dependence among these factors makes it difficult to consider either teacher attitudes toward learning specific content such as statistics, or continuing professional development, in isolation from the other factors. An outline of the chapter structure is provided in table 2.1.

Table 2.1. *Outline of Chapter 2 Topics*

The Responsibilities PMTEs Owe to Teachers and K-12 Students
The Emergence of More Effective Methods of Delivery of Mathematics in General
Teacher Needs for Implementing Effective Mathematics Teaching Strategies
Concerns about teachers' mathematical content knowledge
The importance of teachers' pedagogical content knowledge
The importance of teachers' beliefs and attitudes
Interactions between teacher content knowledge, teacher knowledge of students, and teacher beliefs

Table 2.1. *continued*

Evidence that CPD Can Play a Unique Role in Changing Teacher Content Knowledge, Knowledge of Students, and Beliefs
Comprehensive professional development.
Examples of CPD programs with agendas beyond the scope of preservice teacher education
K-12 Statistics Education Issues
Calls for increased statistical content in the K-12 curriculum
Statistical thinking and learning of K-12 students
Programs supporting K-12 statistics education
The Role of Attitudes and Beliefs on Teacher Acquisition of Content Knowledge and Specific Applications to the Learning and Teaching of Statistics
Goldin's work on affect and mathematical belief structures
Attitudes and other concerns in undergraduate statistics education
Evidence of teachers' statistical content knowledge levels
The effects of attitudes and self-efficacy toward the learning of mathematics and statistics on teacher success learning mathematics and statistics
Teacher attitudes toward the teaching of statistics
Professional Development Programs for Statistics
Chapter Summary

The Responsibilities PMTEs Owe to Teachers and K-12 Students

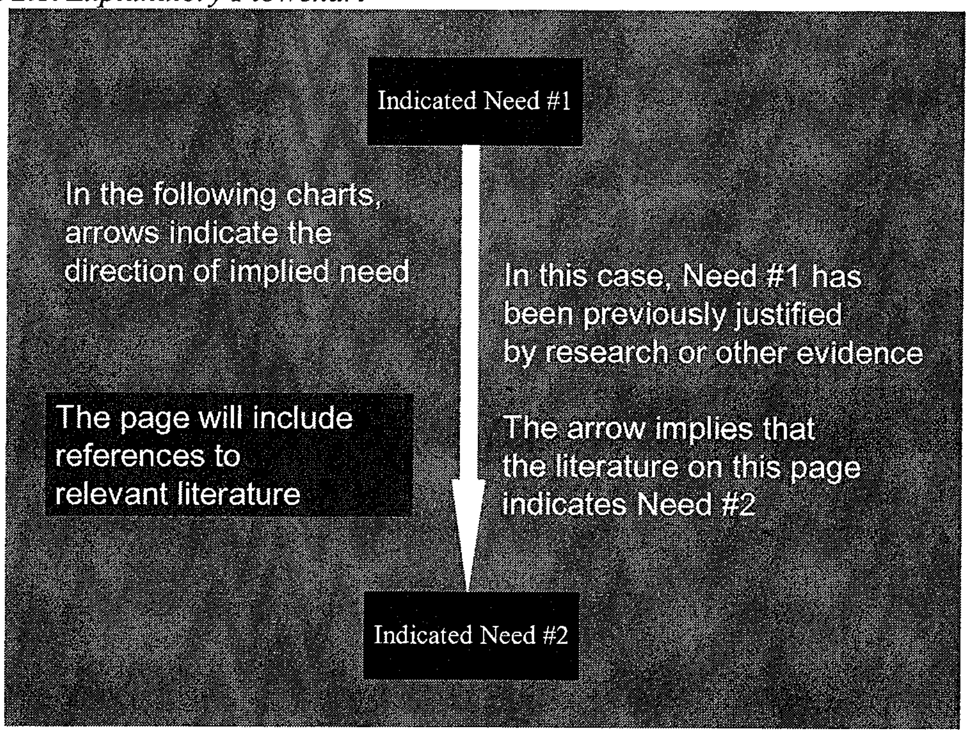
Preservice teacher educators sit at the top of a metaphorical pyramid. They influence and instruct future teachers who then influence and instruct K-12 students.

There is a "trickle-down" effect that they initiate. To appropriately debate decisions preservice teacher educators make concerning how to teach preservice teachers, it is necessary to start at the bottom of this pyramid and follow the progression of needs up to the preservice teacher preparation program.

The literature indicates three distinct paths of influence that connect teacher education programs to K-12 student outcomes. The first path addresses the flow of content knowledge from its acquisition by teachers in preservice teacher preparation programs (PTPPs), and continuing professional development programs (CPDPs), to K-12 students through K-12 classroom events. The second path addresses a system that starts with learning methods and content knowledge acquired by teachers from PTPPs and CPDPs. The learning methods used in PTPPs and CPDPs and the content knowledge acquired in these programs affect teacher attitudes and beliefs toward the teaching of specific content. The attitudes and beliefs teachers possess then affect the methods used to present material in the K-12 classroom. The third path addresses a system that starts with the attitudes teachers bring to PTPPs. Teacher attitudes toward leaning may be affected during PTPPs. These attitudes are likely to affect teacher interest to participate in CPDPs. When the three paths are combined, we get a picture of the complex interactions that affect K-12 student learning through (1) direct content and pedagogical knowledge of the teacher, (2) teaching methods used by the teacher, and (3) improvements in (1) and (2) through teacher participation in CPDPs.

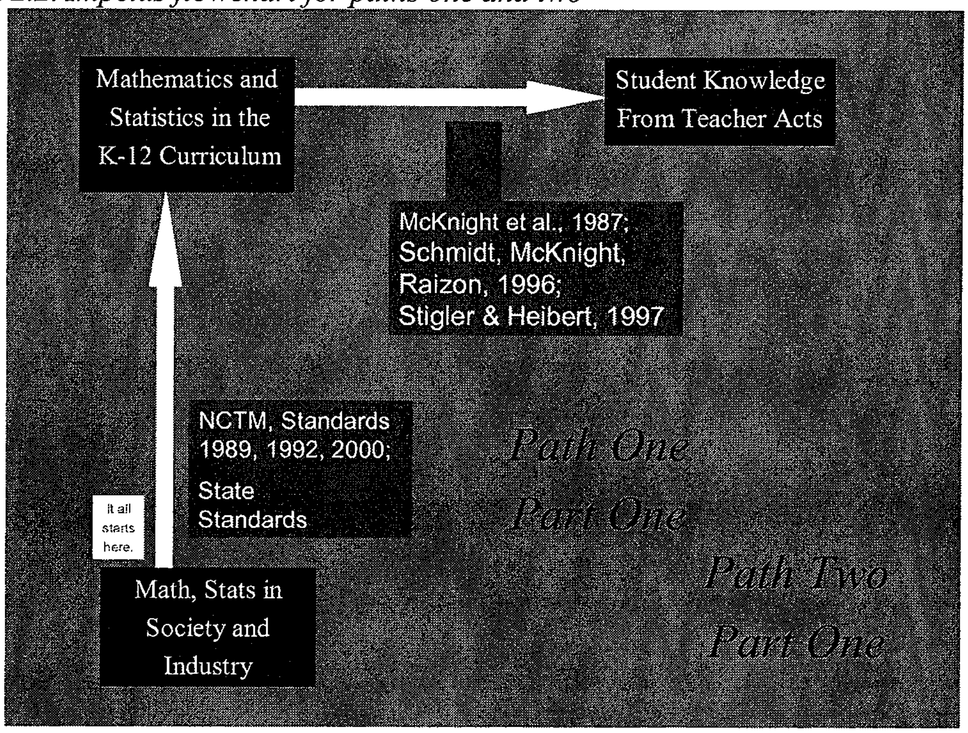
The following flowcharts provide visual cues for this "trickle-down" effect. The flowcharts are not designed in the direction of influence from PMTEs to teachers to K-12 students. Instead, the flows of these charts are in the reverse order. This order indicates the direction of dependence based on the literature. Each flowchart box represents a factor in the "trickle-down" effect. Each connecting arrow represents an indicated dependence based on research finding and other relevant publications such as curricular standards. An explanatory flowchart is presented in chart 2.1.

Chart 2.1. *Explanatory Flowchart*



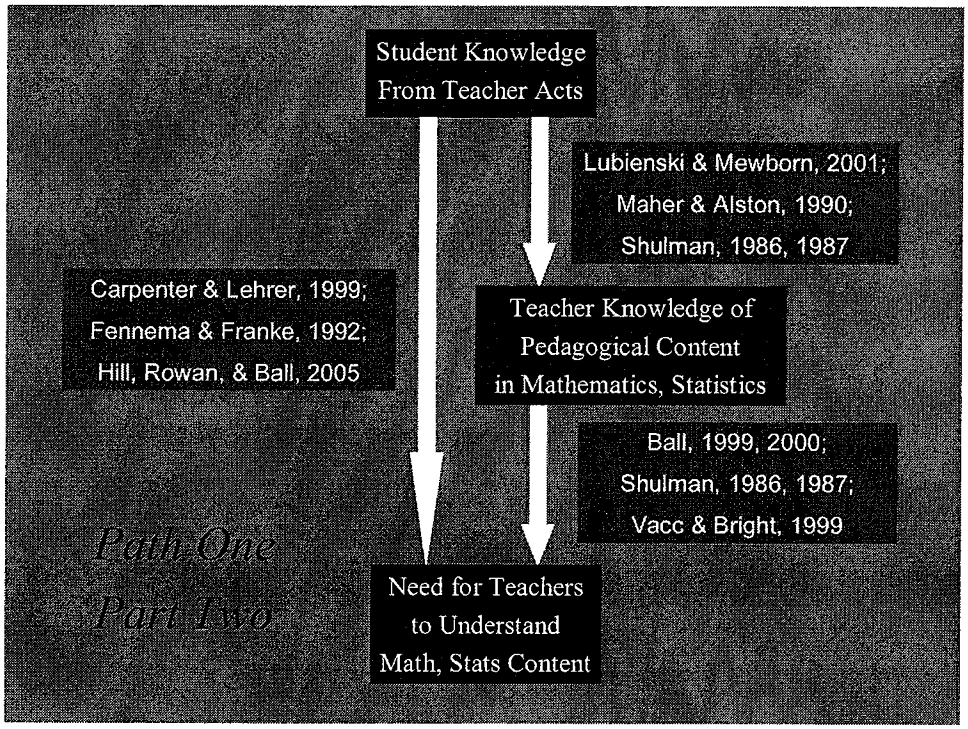
The impetus for paths one and two: Calls for better K-12 mathematical learning. These paths begin with the importance of mathematics and statistics in society and industry. This importance is manifest in curricular standards (NCTM, 1989, 1992, 2000). Evidence exists that indicates that K-12 mathematical and statistical learning needs to improve (McKnight et al., 1987; Raizon, 1996; Schmidt, McKnight, Stigler & Heibert, 1997). The literature also indicates two key elements of improved K-12 mathematical and statistical learning: (1) teacher knowledge of content and of students, and (2) the methods through which the material is experienced by the K-12 students. These two elements generate the first two flowchart paths. The following chart, chart 2.2, provides the impetus flowchart for paths one and two.

Chart 2.2. *Impetus flowchart for paths one and two*



Path one, the system of dependency based on teacher knowledge. It has been shown that two key types of teacher knowledge are necessary to facilitate mathematics understanding in K-12 students. Teachers need content knowledge beyond superficial understandings (Carpenter & Lehrer, 1999; Fennema & Franke, 1992; Hill, Rowan, & Ball, 2005). Teachers also need pedagogical content knowledge (Ball, Lubienski & Mewborn, 2001; Maher & Alston, 1990; Shulman, 1986, 1987). In fact, teachers must possess deep content knowledge to successfully develop their pedagogical content knowledge (Ball, 1999, 2000; Shulman, 1986, 1987; Vacc & Bright, 1999). These interactions are presented in the flowchart on chart 2.3.

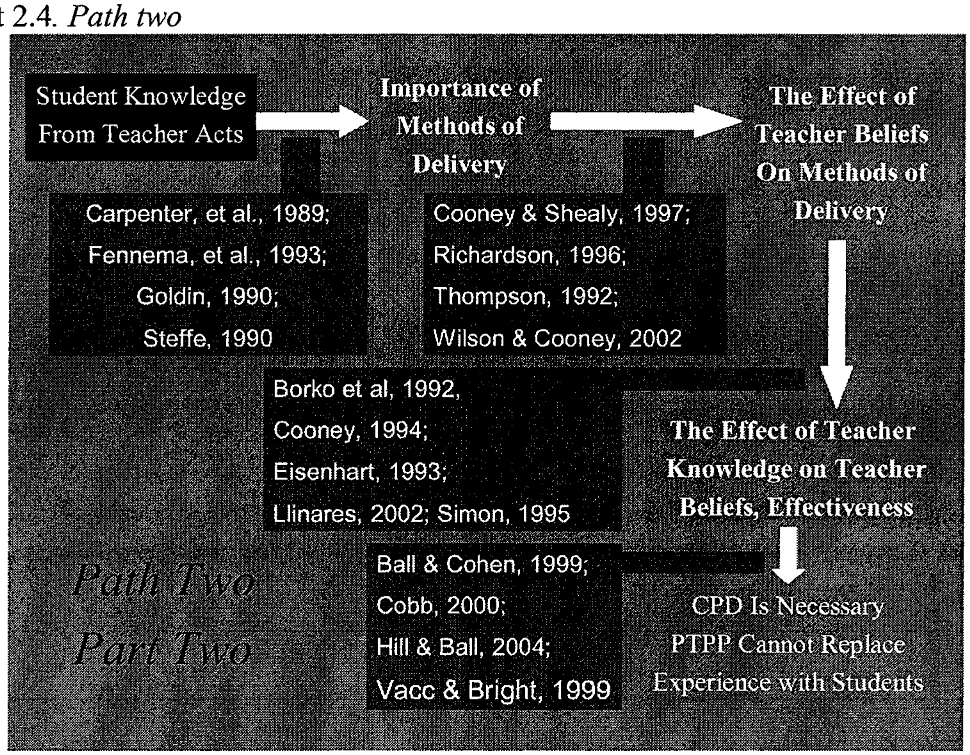
Chart 2.3. *Path one*



Path two, the system of dependency based on teacher beliefs and the use of beneficial learning methods. Certain types of methods for presenting mathematical material are more effective than others (Carpenter, T. P., Fennema, E., Peterson, P. L., Chiang, C.-P., & Loef, M., 1989; Fennema, E., Franke, M. L., Carpenter, T. P. & Carey, 1993; Goldin, 1990; Steffe, 1990). Both the willingness of teachers to utilize certain beneficial teaching methods and the success teachers have using these methods depend on the beliefs the teachers possess concerning mathematical learning (Cooney & Shealy, 1997; Richardson, 1996; Thompson, 1992; Wilson & Cooney, 2002). Other research evidence indicates that teachers have difficulty embracing the appropriate beliefs of mathematical learning if they fail to posses the necessary content knowledge (Borko, H., Eisenhart, M., Brown, C. A., Underhill, R. G., Jones, D. & Agard, P. C., 1992; Cooney, 1994; Eisenhart et al., 1993; Llinares, 2002; Simon, 1995). This evidence links path two to path one. However, there is more involved.

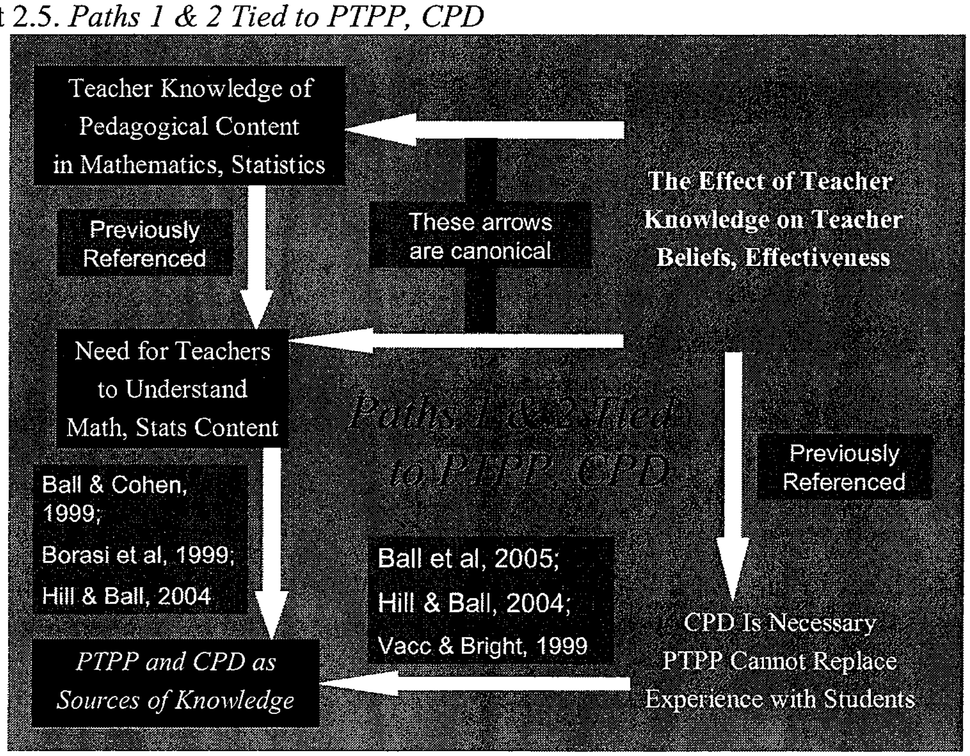
Studies indicate that certain types of knowledge are difficult to achieve during only the preparation of teachers through the PTPP (Ball & Cohen, 1999; Cobb, 2000; Hill & Ball, 2004; Vacc & Bright, 1999). These authors pointed out that participation in CPD can play an important role in both teacher beliefs and teacher use of more beneficial teaching methods. More details are presented in subsequent sections with regarding this important conclusion. The flowchart for path two is on chart 2.4.

Chart 2.4. *Path two*



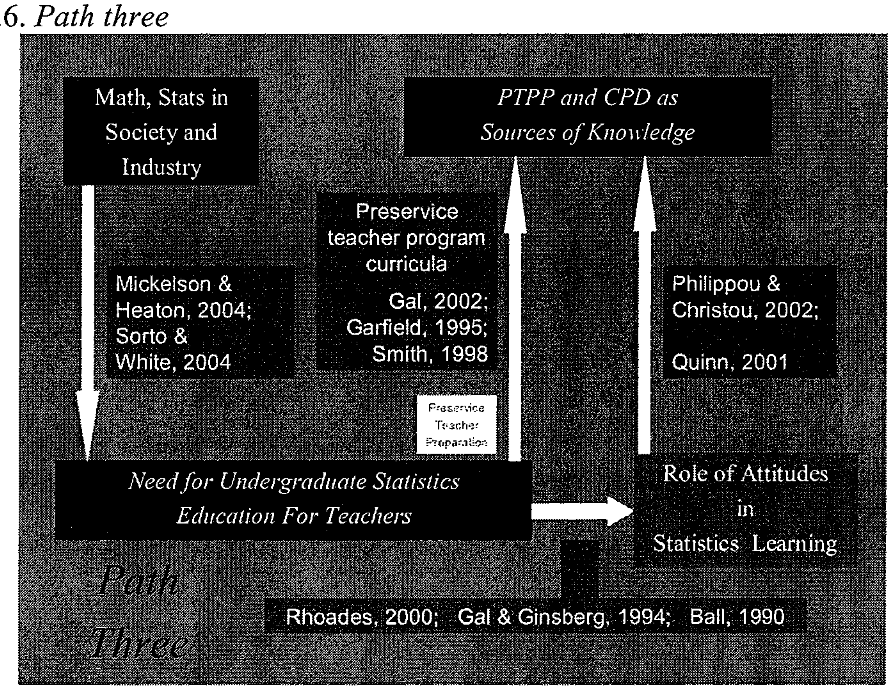
Paths one and two both lead to PTPP and CPD as sources of content knowledge. Teachers can learn much of what they need to know regarding content knowledge and pedagogical content knowledge through appropriate PTPP (Ball & Cohen, 1999; Borasi, R., Fonzi, J., Smith, C. F. & Rose, B. J., 1999; Hill & Ball, 2004). This knowledge affects both teacher ability to directly communicate mathematics to K-12 students and teacher ability to embrace appropriate beliefs for accessing more beneficial teaching methods. Teachers can gain types of knowledge during CPD that are difficult to obtain during PTPP (Ball, et al., 2005; Hill & Ball, 2004; Vacc & Bright, 1999). The PTPP and CPD links to paths one and two are presented in chart 2.5.

Chart 2.5. *Paths 1 & 2 Tied to PTPP, CPD*



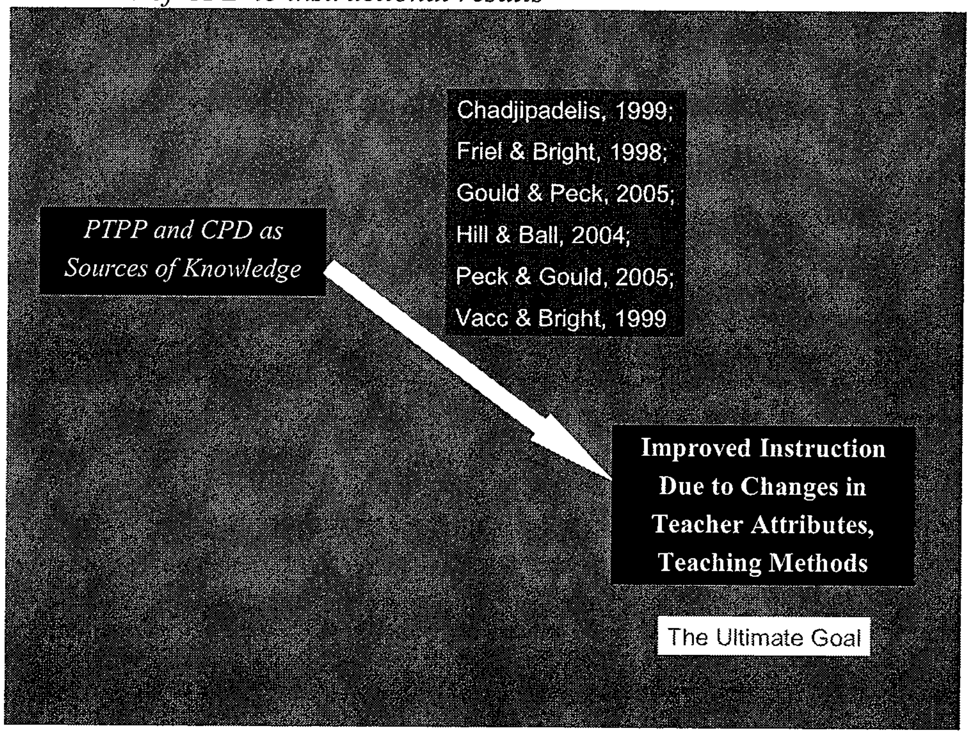
Path three, the flow of dependency based on teacher attitudes to learn specific content and the need for CPD. Teachers have been shown to need better statistical knowledge for K-12 teaching (Mickelson & Heaton, 2004; Sorto & White, 2004). While few studies have provided evidence of clear methods for solving this problem, I propose that the core for improving teacher statistical knowledge is from PTPP and CPD. Many people, including teachers, possess negative attitudes toward the learning of statistical content (Ball, 1990; Gal & Ginsberg, 1994; Rhoades, 2000). There is some evidence that PTPP and CPD can help teacher statistical knowledge (Gal, 2002; Garfield, 1995; Smith, 1998). There is also evidence that PTPP and CPD can affect teacher attitudes toward statistics (Goldin, 2002). However, whether or not PTPP or CPD can affect teacher attitudes toward the learning of statistical content has not been investigated. The flowchart for path three is on chart 2.6.

Chart 2.6. *Path three*



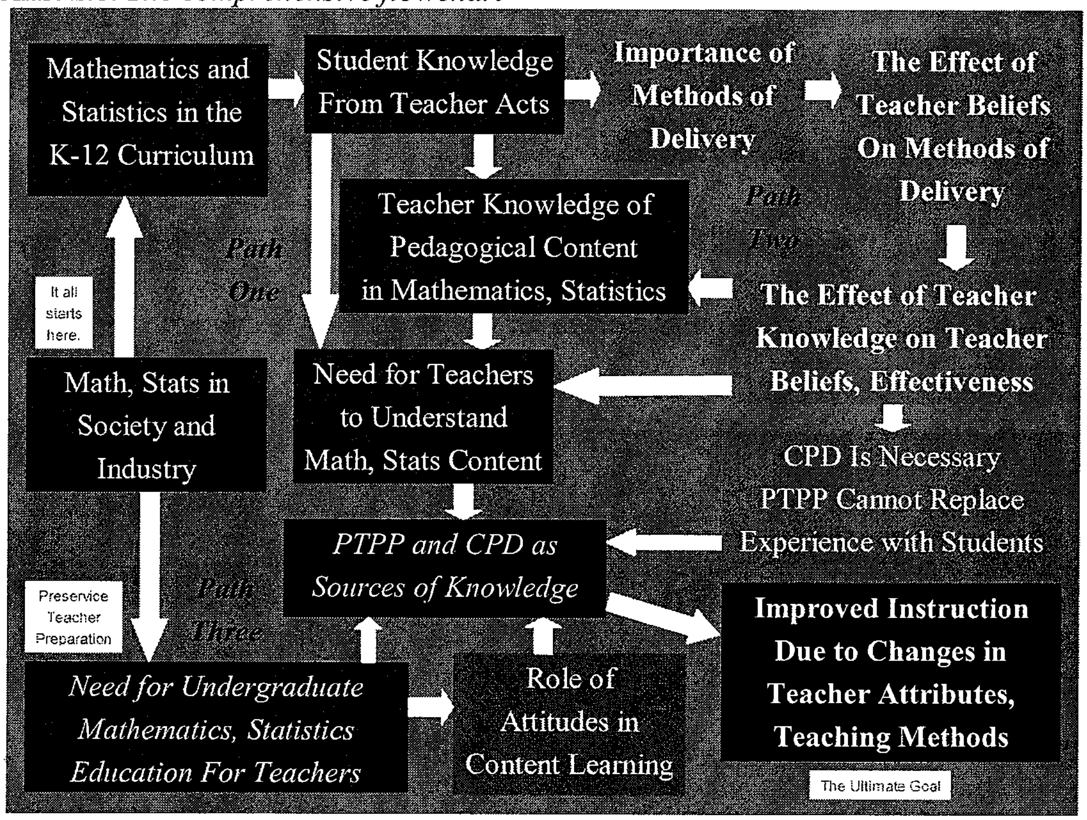
It is important to find ways to attract teachers back for CPD in statistics. CPD been shown to provide certain content knowledge and pedagogical content knowledge that are difficult to acquire during PTPPs. But even more importantly, CPD has been shown to directly improve K-12 statistics instruction (Chadjipadelis, 1999; Friel & Bright, 1998; Gould & Peck, 2005; Peck & Gould, 2005; Vacc & Bright, 1999). This flowchart is on chart 2.7.

Chart 2.7. *Value of CPD to instructional results*



We can consider (1) the three paths, (2) the links between the paths at PTPP and CPD, and (3) the implications of good CPD, to create a comprehensive flowchart. This flowchart is presented in chart 2.8. There may be evidence of direct dependency between some factors on this chart that do not have connecting arrows. Still, this flowchart provides an overview of how various aspects of teacher preparation affect each other. It also provides some evidence of the importance of CPD to improved K-12 mathematics and statistics education.

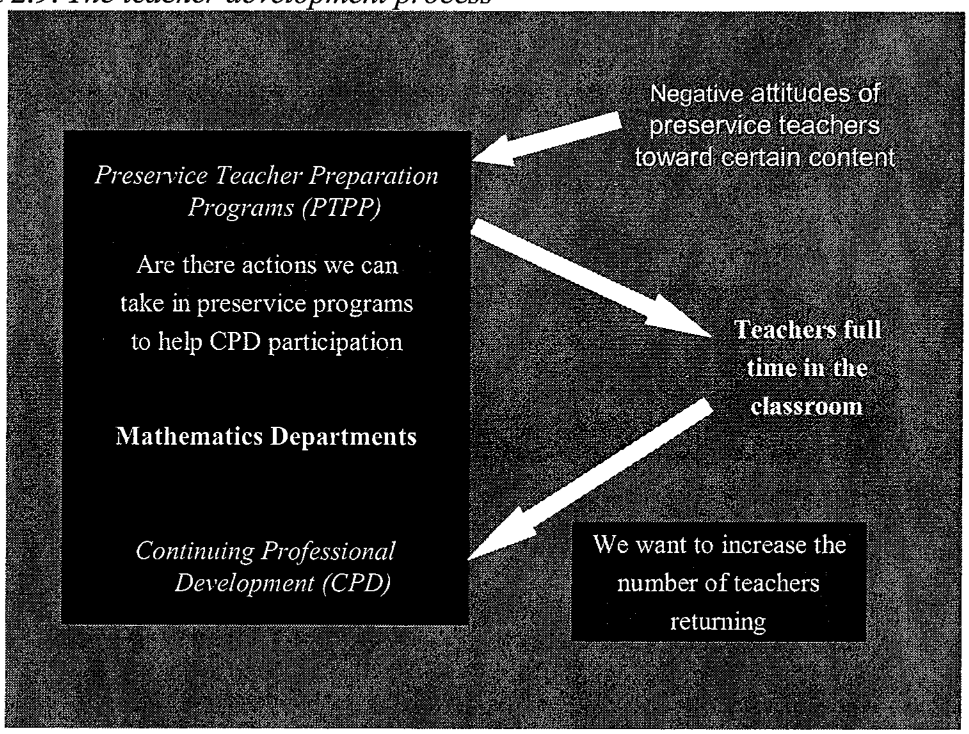
Chart 2.8. *The comprehensive flowchart*



More detailed discussions of some of the points highlighted in the flowcharts occur throughout the remainder of this chapter. The chapter continues to build the argument that some delivery methods can only be mastered by adequate preservice training followed by a certain minimum level of CPD once the teacher has experience interacting with students. The importance of CPD places a burden on the preservice teacher educator to find ways to motivate preservice teachers to be willing to seek out

CPD once they are practicing professionals. This burden is presented in the context of teacher development in the flowchart on chart 2.9. This becomes an additional challenge when specific CPD programs need to address a content subject that has historically created negative affective responses from preservice teachers. Statistics is shown to be one of these content subjects.

Chart 2.9. *The teacher development process*



The Emergence of More Effective Methods of Delivery of Mathematics in General

Over the past few decades, observed student misconceptions about mathematical ideas have lead researchers to focus on the methodological issues surrounding mathematics education (Maher & Alston, 1990). Heibert (1999) indicated that in the traditional classroom, student skills and knowledge are fragile due to the superficial level at which the content is investigated, with a lack of focus on conceptual understanding. Heibert claimed that, "If students have more opportunity to construct mathematical understandings, they will construct them more often and more deeply." (p. 14)

In response to such reports, studies were conducted to identify methods of mathematic instruction that would utilize the ways in which people cognitively assimilate and organize mathematics. Several successful instructional designs have been developed that have been shown to improve the levels of understanding and the problem solving skills of K-12 students (Goldin, 1990; Carpenter, T. P., Fennema, E., Peterson, P. L., & Carey, D. A., 1988, Carpenter et al., 1989, Carpenter, T.P., Fennema, E., & Franke, M.L., 1996; Fennema & Franke 1992; Fennema et al., 1993; Fennema & Romberg, 1999). These various designs are generally referred to as *progressive methods* or *progressive methodologies* throughout this chapter.

To appreciate the state of research concerning progressive methodologies of mathematics instruction, consider that Schoenfeld (1994) argued that in the not so distant past the debate concerning the future of mathematics education in the United States was treated as a comparison between the "traditional" proven approaches and the new "experimental" approaches. (pp. 57, 64) Although few researchers in mathematics education would review such research in those terms today, this attitude is far from eliminated in the philosophies of high school teachers and collegiate professors. In this light, it is understandable that many students entering preservice teacher education programs today still have attitudes toward the teaching of mathematics that reflect such skepticism of progressive teaching methods.

One prominent movement in general over the past 30 years has been the constructivist philosophy of mathematics learning (Cobb & Moore, 1997; Simon, 1995). Teaching philosophies that have developed based on the constructivist learning philosophy include discovery learning and cognitively guided instruction. Discovery Learning is based on the premise that for large numbers of students at all levels of mathematics education, the old methods of presenting rules and procedures for students to use verbatim are less successful than methods involving mathematical discovery (Goldin, 1990). Among those whose work corresponds to that of Goldin is Steffe. Steffe (1990) proposed that students "invent" mathematical concepts as they are exposed to experiences.

Cognitively guided instruction is designed to help teachers understand children's mathematical thinking by helping them understand the development of children's thinking in well-defined content domains (Carpenter, 1985; Carpenter & Fennema, 1992; Carpenter, Fennema, & Franke, 1996; Carpenter et al., 1989; Fennema et al., 1993). Although CGI focuses on primary level mathematics teaching, its success has ramifications at all levels of mathematics education. The impact of CGI on mathematics education research is apparent in the number of separate studies that utilize CGI as the underlying methodology.

Among the studies that have utilized CGI are those of Cooney and Shealy (1997), Lubinsky & Jaberg (1997), and Franke, Fennema, & Carpenter (1997). A number of studies that involve CGI were discussed in Wilson and Berne (1999). Vacc and Bright (1999) also utilized CGI in a study in which they measured the changing beliefs of elementary preservice teachers' during a two year sequence of courses designed to improve teacher beliefs and perceptions about mathematics. This study is revisited during the discussion about teacher beliefs.

Teacher Needs for Implementing Effective Mathematics Teaching Strategies

Although progressive methods of instruction have been developed, the use of such methods is still far from widespread. According to McKnight et al. (1987), contemporary mathematics teaching could almost universally be characterized as formal, symbolic presentations of mathematical rules or procedures in lecture formats. In fact, traditional United States mathematics curriculum is relatively repetitive, unfocused, and undemanding (Schmidt, McKnight, & Raizon, 1996). As recently as 1997, it was being reported that 96% of the time, students do seatwork in which they practice procedures they have been shown to do (Stigler & Heibert, 1997). This is consistent with traditional mathematics teaching methods.

If progressive methods are known to be more effective, and if there is evidence that these methods are not being widely utilized, then there is a need to identify the reasons that such methods are not being adopted by teachers. Most of teachers' mathematical knowledge is situated in well-defined problems from the culture of the classroom since this is where they have learned the majority of their mathematics. This includes teachers' learning of advanced mathematics (Fennema & Franke, 1992). Fennema and Franke contrasted this with the more effective method of taking complex mathematics concepts and breaking them down into constructs that are understandable for students. They argued that this process requires deeper understanding of the concepts than traditional methods demand. They also argued that this process requires understanding of students and how students learn mathematical concepts to recognize the best fundamental constructs in which to represent the concepts.

Carpenter and Lehrer (1999) indicated that students should have opportunities to discuss their ideas and that the overarching emphasis within classrooms should be to develop understanding. Both Fennema and Romberg (1999) and Carpenter and Lehrer (1999) asserted that without requisite understanding of mathematics and of students, teachers will be relegated to the routine presentation of ideas designed by others and not explicitly created to meet the mathematical thinking of the teachers' specific students. They asserted that without requisite understanding of mathematics and of students, teachers cannot appropriately engage students in productive discussions of alternative strategies because student responses that fail to fit within specific protocols will not be readily recognized by the teacher.

Concerns about teachers' mathematical content knowledge. Preservice teachers bring certain mathematical understandings with them to teacher education from pre-college and college mathematics. The understandings that many preservice teachers bring are "rule-bound and thin" (Ball, 1990, p. 449). Ball asserted that teachers should first have correct knowledge of concepts and procedures; she also asserts that teachers must have an understanding of the underlying principles and meanings.

Ball (2000) stated, "It is not just what mathematics teachers know, but how they know it and what they are able to mobilize mathematically in the course of teaching" (p. 243). Although CGI and other progressive methods of instruction had been recognized for over ten years at the time, Ball still lamented that "many teachers are unable to hear students flexibly…, and think about things in ways other than their own." (p. 243). Hence the lack of content understanding was affecting teacher ability to listen to students' thinking. This affects teacher ability to acquire what is known as pedagogical content knowledge.

The importance of teachers' pedagogical content knowledge. Shulman (1986) introduced the concept of "pedagogical content knowledge". Pedagogical content knowledge describes a type of teacher knowledge that links content and pedagogy. Shulman explains that teachers not only need to know both the content of the material that they teach and general pedagogical issues; they also need to know such things as which representations are most useful for teaching a specific topic and what topics children find more interesting.

Many of the tenets of CGI utilize results from Shulman's work in pedagogical content knowledge. According to Vacc and Bright (1999), teachers who use CGI principles when teaching meet four criteria. First, they "believe that their understanding of children's thinking is a critical component of instructional planning." Second, they "facilitate children's problem solving and discussions of children's thinking." Third, they "listen to their children and question them until the students' thinking becomes clearer." Fourth, they "are willing and able to make instructional decisions that are appropriate to the mathematics needs of their students." (p. 90)

Although these four criteria were presented by Vacc and Bright in the context of teacher beliefs, it stands to reason that teachers must also be able to facilitate these activities. For example, in the first criteria, teachers not only need to believe that their understanding of children's thinking is important; teachers need to be able to develop such understanding of their children's thinking. This is an example of pedagogical content knowledge. In fact, all four of the criteria discussed by Vacc and Bright involve the application of pedagogical content knowledge.

The importance of teachers' beliefs and attitudes. Studies have linked the type of strategies utilized by teachers with teacher beliefs (Richardson, V., Anders, P., Tidwell, D., & Lloyd, C., 1991). Teachers' beliefs have been linked to their ability both to change and to develop teaching practices consistent with progressive teaching strategies (Wilson & Cooney, 2002). This appears to be a two-way relationship. Teachers who utilize reform methods are more likely to have a broader view of mathematics and their beliefs are likely to be consistent with this view (Cooney & Shealy, 1997; Thompson, 1984, 1992).

Pajares (1996) noted that efficacy beliefs of teachers were related to factors such as professional enthusiasm, instructional experimentation, implementation of progressive and innovative methods, and students' positive outlook. However, Pajares also noted that these relationships exist only within a very specific focus. Hence efficacy beliefs, and the subsequent factors such as implementation of progressive teaching methods, are likely to vary from topic to topic, even within a single content subject such as geometry or statistics.

Interactions between teacher content knowledge, teacher knowledge of students, and teacher beliefs. Pedagogical content knowledge, by definition, is dependent on content knowledge (Shulman, 1986). A teacher cannot understand how their students think about a mathematical topic if the teacher does not have a firm understanding of the topic. Teacher content knowledge has been shown to affect the way in which teachers view the study of mathematics. When a teacher's mathematical knowledge is procedural and sparsely connected, then according to Simon (1993), the teacher is likely to view the study of mathematics as the acquisition of particular computational procedures. Teachers who possess such a view are unlikely to foster a learning atmosphere that supports deep understanding of the content.

Teacher content knowledge and teacher beliefs are so intertwined that studies have cautioned researchers whom might try to separate these traits. According to Wilson and Cooney (2002), it is unlikely that teachers can change their teaching methods in fundamental ways without experiencing fundamental changes in their associated beliefs. These authors claim that it is a tricky business to try to separate knowledge levels and beliefs and that when separation occurs in research, the descriptions are often incomplete. They also note that participation in reform-oriented instructional settings enhance teachers' beliefs about implementing progressive teaching strategies.

It has been shown that content knowledge and beliefs are intertwined. However, it has also been shown that professed beliefs can be difficult to implement in application when a practice teacher lacks appropriate subject matter or pedagogical content knowledge (Borko et al., 1992). In this case, a preservice teacher professed belief in progressive teaching methods. The preservice teacher lacked the necessary content knowledge for a lesson presented, and subsequently had difficulty executing a lesson that matched the professed beliefs. The authors added that without both appropriate subject matter knowledge and pedagogical content knowledge, the student teacher is likely to never achieve the effective teaching styles to which the beliefs correspond.

Evidence that CPD Can play a Unique Role in Changing Teacher Content Knowledge, Knowledge of Students, and Beliefs.

Guskey (1986) stated that professional development programs attempt to bring change in teachers' practices and in their beliefs and attitudes. Specific to mathematics, the Mathematical Sciences Education Board states that, "we recommend...a national focus on professional development" (MSEB, 1997, p. 4). The primary traits that professional development is designed to change are those that have been discussed thus far: teacher content knowledge, teacher pedagogical content knowledge, and teacher beliefs.

Comprehensive professional development. Leaders in mathematics teacher education propose a comprehensive approach to teachers' professional development, encompassing preservice education and ongoing professional development for teachers in the field (Ball & Cohen, 1999; Ball et al., 2005). Preservice teacher education is well-positioned to emphasize core content knowledge and to introduce progressive teaching methods (Philipou & Christou, 2002; Zaslowsky & Leiken, 1999). However, Ball and Cohen (1999) asserted that, "preservice teacher education offers a weak antidote to the powerful socialization into teaching that occurs in teachers' own prior experience as students" (p. 5). To address this weakness, among other issues, ongoing teacher professional development opportunities can offer iterative experience with students and with the content and ideas about teaching practice (Ball & Cohen, 2004).

Kagan (1992) described a process in which novice teachers can begin to step back from their personal beliefs and images of teaching and acknowledge when these beliefs and images are incorrect or inappropriate. Kagan emphasizes that this process takes place as novice teachers gain experience interacting with pupils. It may be difficult to simulate adequately such experience within the confines of a preservice teacher education program.

Examples of CPD programs with agendas beyond the scope of preservice teacher education. A particular professional development program reported by Borasi et al. (1999) focused on engaging teachers as learners in mathematics learning experiences, providing supported field experiences, providing teachers with multiple opportunities to reflect, and encouraging the participation of teams of teachers. Only the first of these goals can be appropriately addressed in preservice teacher courses. The other goals are much more likely to be effective once teachers have experience with students in the classroom.

The Mathematics Professional Development Institutes (MPDIs) in California were a heavily funded set of institutes designed to improve teacher subject matter knowledge in mathematics as well as English language content (Hill and Ball, 2004).

These institutes are designed to utilize teacher familiarity with classroom mathematics topics to motivate the activities in the professional development program. In these institutes, effort was made to utilize both education specialists within mathematics departments and mathematics specialists within education departments to provide content training that is mathematically accurate and guides teachers toward those types of understandings that are needed to implement progressive teaching methodologies. Hill and Ball determined, based on analyses of the MPDIs, that primary level teachers indeed can learn the mathematics they need in the context of a single CPD program.

Teachers, like children, learn in social contexts in which they can interact and make sense of their experiences (Cobb, 2000; Mahar & Ahlston, 1990). This is a central theme behind work by Krainer as well as an important study by Llinares. Krainer (1999) focused a study on a continuing professional development program that emphasized the use of both reflection and social interaction between teachers to enhance teaching quality. Such social themes can be developed between preservice teachers. However, in such an event, none of the participants can bring with them a depth of experience with students from authentic learning situations. This is an important factor in the reflection portion of this professional development.

K-12 Statistics Education Issues

In the past thirty years, two important developments changed the role that statistics plays in the K-12 curriculum. The advent of computer technology has created many opportunities for statistical inquiry at all levels of education and society. (Lajoie, 1998) Tukey's new approaches to statistics created more accessibility to statistics for those who do not have strong mathematical backgrounds (Tukey, 1977). An example of such accessible approaches is the box-and-whisker plot that helped make median and quartiles, and thus measures of center, more visual and less numerical to students at all levels. These developments have helped to increase the amount of statistics in curricular recommendations, increase the amount of research on K-12 student understanding of statistics, and increase the number of programs that support K-12 statistics education.

Calls for increased statistical content in the K-12 curriculum. There are many reasons to include statistics and probability in both primary and secondary curricula (Gal, 2002, 2004; Wild & Pfankuch, 1999). This is manifest in the steady increase of statistical content in curricular recommendations in the past twenty years (NCTM, 1989, 1992, 2000; American Mathematical Society, 2001). The NCTM (1992) published an addendum series to supplement the 1989 recommendations by clarifying and illustrating the new statistics standards introduced in 1989. In 2000, the NCTM recommended that "students need to know about data analysis and related aspects of probability in order to reason statistically – skills necessary to becoming informed citizens and intelligent consumers." (NCTM, 2000, p 47) In these recommendations, the NCTM called for yet another increase in statistics and probability in the elementary curriculum. The National Center for Research in Mathematical Sciences Education (NCRMSE, 1994a) states that "Students encounter numerous statistical claims in their daily lives. Data are collected, summarized, analyzed, and transformed in most of this country's media, work places, and homes. The collecting, representing, and processing of data are assuming major importance in most nations. While statistics was once taught primarily to college students pursuing professional or academic careers, it is now becoming a part of the school mathematics curriculum." (p 1)

Statistical thinking and learning of K-12 students. A number of studies have focused on specific aspects of student thinking when investigating or otherwise learning statistics concepts (Batanero, C., Merino, B., & Diaz, C., 2003; Batanero & Serrano, 1999; Ben Zvi & Arcavi, 2001; Jones, G., Langrall, C., Thornton, C., & Mogill, A. T., 1999; Lakoma, 2000; Lehrer & Schauble, 2000; Mokros & Russell, 1995; Watson, 2002; Watson, J., Kelly, B., Callingham, R., Shaughnessy, J. M., 2003). Some of these have focused on student learning of specific statistics concepts. Ben Zvi and Arcavi (2001) studied global views of data representation constructed by junior high students. Batanero and Serrano (1999) performed a study on secondary students to measure their understanding of the meaning of randomness.

Garfield and Chance (2000) reported on the importance of assessment in K-12 statistics education. They argue that the traditional assessment methods of cookbook computations and wrote memory lead students to believe that routine skills and memorized formulas are what teachers view as important. The authors suggest that teachers must incorporate newer assessment methods and assess students in ways that not only inform the teacher but that also provide feedback to the students. This includes assessing in ways that focus on students' reasoning and on more authentic tasks. Such studies serve two key purposes. One, they emphasize the importance of specific teacher attributes such as deep understanding of statistical concepts. Two, these studies infer concerns that should be addressed during preservice programs and CPD.

Programs supporting K-12 statistics education. There have been multiple studies performed that report on previous, ongoing, or planned programs oriented towards improvement of K-12 statistics education. The reports from these studies provide valuable information for preservice teacher educators and researchers at multiple levels. First, they provide generic evidence of the growing importance of statistics education for K-12 students. Second, they provide suggestions for preservice teacher educators as to programs that may be worth consideration for adoption into preservice teacher programs. Third, they provide evidence and arguments in defense of these programs. This evidence can be debated among preservice teacher educator researchers in journals and conferences. Such debate helps promote the development of research in this field of study. It also helps to shape the direction in which future research will aim.

In the National Center for Research in Mathematical Sciences Education report on the studies of V. R. Jacobs & S. Lajoie, and of S. Lajoie & N. C. Lavigne (NCRMSE , 1994b), it is noted that two activities were particularly useful in promoting discussions that enabled students to learn statistical content. One was student-generated tests and the other was student-directed classes. Suzanne Lajoie (1999) supported the concept of using collaborative methodologies to enhance statistics learning. To implement such methodologies in K-12 learning, she created the Authentic Statistics Project (ASP). The ASP is a two week program for 8th grade students. During this program, the students work in groups of three at a computer station where they can utilize statistics programs to develop understandings of statistical concepts. Lajoie reports that in the ASP, students work to master basic concepts and skills. They then practice these skills. They are thus able to extend understanding by applying this knowledge to student-designed experiments.

Lajoie (1999) stated, "Classrooms with a problem-solving focus help students learn about statistics as well as what to do with them." (p 118) A central theme in the ASP is to situate the learning in tasks that are meaningful to the learner. This is known as "authentic" learning. The use of natural social interactions through dialogues about the statistics helps to create statistical understanding. Teachers must help the students by modeling and coaching students in constructing statistical relationships. They do this with the goal that when students develop an understanding of the relationships, then they will be able to apply this knowledge to the projects they (the students) design.

There are other programs that have been reported. Cobb and McClain (2004) developed a set of design principles for developing statistical reasoning at the elementary level. Two of the primary design principles presented are the use of Tukey's exploratory data analysis (EDA) and the process of generating data. In fact, Cobb and Moore (1997) proposed that EDA is a necessary precursor to statistical inference.

As with many of the statistics issues addressed in this chapter, several of the studies reported in this section were conducted outside of the United States. This emphasizes the global focus on the improvement of statistics education. The Ministry of Education in China is pushing for reform to include statistics and probability in its elementary and secondary curriculum. This is due to the influence of the worldwide movement to introduce elements of statistics and probability into school curricula (Li, 2004). Li investigated two studies performed in China concerning statistics and probability education K-12. Some important results of the studies are: (1) Students' understanding of probability does not improve naturally with age – and teaching plays an important role; (2) Students' understanding of probability can be improved after instruction; (3) Introducing probability using the experimental approach or using the theoretical approach cannot replace each other (i.e. students need both probability theory with the concepts and formulas, and probabilistic applications with experiments such as flipping coins and pulling numbered cards out of a hat); and (4) Students' cognitive development in frequentist probability is slow, if no direct instruction is given.

Carvalho & Cesar (2002) declared that in Portugal, "as in most western countries", the 1970's and 1980's witnessed a growth in concerns related to finding solutions for the problematic issues generated by the necessary inclusion of statistics into the compulsory curricula. They expatiate that this comes from the fact that the role of statistics is becoming increasingly important in today's society. They believe that the emphasis of teaching must not be place in a so-called knowledge acquisition, but on "knowing in action". They investigated the value of collaborative work to enhance statistics learning based on the social-learning paradigms introduced by

Vygotsky (Vygotsky, 1962, 1978). They found that collaborative work facilitates better performances on behalf of the pupils, regardless of whether they have had difficulties in related subject knowledge.

The implementation of such collaborative methodologies concerning statistics puts teachers in a situation similar to that which they are in when implementing constructivist strategies such as CGI. Of course, the tenets of CGI can be applied to statistics lessons just as they can to any mathematics lesson. In both cases, (1) teacher utilization of CGI on a typical mathematics topic such as division, and (2) teacher to utilization of collaborative methodologies on a statistics topic; the teacher must have both a deep understanding of the content and a deep understanding of how students (specifically those students that they are teaching) learn under the conditions the teacher intends to create.

With the existence of programs to improve statistics education K-12 comes a need for K-12 teachers to be able to implement and utilize these programs. Such programs are critically important for improved student understanding of statistics, but these programs put great pressure on teachers to understand content and pedagogical content for statistics. It is likely that such programs will continue to be developed and improved. These conditions force preservice and practicing teachers to, (1) learn about and understand these programs and understand why they are valuable, and (2) prepare to overcome the deficiencies that prove to hinder efforts to successfully execute such programs.

The Role of Attitudes and Beliefs on Teacher Acquisition of Content Knowledge and Specific Applications to the Learning and Teaching of Statistics

For teachers to teach statistics, they must know statistics content. They acquire this knowledge in various ways. Opportunities for teachers to have learned statistics include their primary grade level education, their secondary grade level education, collegiate statistics courses, preservice mathematics methods courses, and research-oriented undergraduate courses. However, it is common for teachers to not have statistics experience from many of the possibilities listed.

Just as teachers need pedagogical content knowledge to teach any mathematical topic, they need pedagogical content knowledge, focused on statistics learning, to teach statistics. Just as progressive methods have been designed to address mathematics teaching in general, there are progressive methods for teaching statistics content. Just as teachers need to have appropriate beliefs to effectively administer progressive teaching methods in general, teachers need to appropriate beliefs to effectively administer progressive statistics teaching methods.

First, it is appropriate to address teacher acquisition of statistics content knowledge. In this section, I present evidence that the acquisition of statistics content knowledge is dependent on factors such as learner beliefs, and learner attitudes, toward the acquisition of statistics content knowledge. How learner attitudes affect acquisition of knowledge is particularly relevant when applied to preservice teachers as the learners of statistics content. To address statistics learning, it is beneficial to look at work that has been reported concerning the ways in which attitudes and beliefs affect learners of mathematics in general.

Goldin's work on affect and mathematical belief structures. Goldin (2002) provided theoretical perspectives on mathematical beliefs drawn from analyses of the affective domain. He specifically focused on the interplay between meta-affect and belief structure in sustaining each other in the individual. Goldin's work provided valuable information concerning those feelings and attitudes that preservice teachers have developed, or may eventually develop, with respect to a mathematical field of study. Goldin stated that, "My main assertion is that the stability of beliefs in individuals has much to do with the interaction of belief structures not only with affect (feelings), but with meta-affect (feelings about feelings) – that through their psychological interplay, meta-affect and belief structures sustain each other." (p. 59) Though his work encompasses the affective domain, much of Goldin's work centered on the notion of teacher beliefs. Goldin defined beliefs as "multiply-encoded cognitive/affective configurations to which the holder attributes some kind of truth value (e.g. empirical truth, validity, or applicability)." (p. 59)

Categories of affective representations. Goldin (2002) treated beliefs as one of four fundamental categories of affective representations. The following descriptions come from Goldin's work. Emotions are rapidly changing states of feeling that tend to be local or exist within some context. Attitudes are moderately stable inclinations toward ways of feeling in various classes of situations. Attitudes involve a balance between affect and cognition. Beliefs are representations that hold truth, validity, or applicability for the individual. They are relatively stable, highly cognitive, and can be highly structured. Values, ethics, and morals, are deeply-held preferences that could be characterized as personal truths. These are highly affective as well as cognitive, and are stable.

Some of the types of beliefs that Goldin acknowledged are: (1) beliefs about the nature of mathematics (and branches thereof) including philosophy and foundations, (2) beliefs about the individuals' mathematical (statistical) ability and how this manifests itself, (3) "beliefs about the learning of mathematics, the teaching of mathematics, and the psychology of doing mathematics", and (4) "beliefs about oneself in relation to mathematics, including one's ability, emotions, history, integrity, motivations, self-concept, stature in the eyes of others, etc." (p. 68)

Goldin described affective constructs that he has found to be especially important to mathematics. Some of these include vulnerability, personal caring, private experience, and possible creative expression. These are characterized under the term *mathematical intimacy*. There are other constructs such as recognition and acceptance of inadequate understanding. This construct fits into a category characterized under the term *mathematical integrity.*

Schoenfeld (1985) noted that beliefs and affective structures are important for understanding individual's mathematical problem solving heuristics and strategies. A specific representation that Goldin elucidates is fear of mathematics. Fear, and the meta-affects and beliefs that are affected by fear, can both be generalized across a wide range of mathematics topics (even applied to all of mathematics) and applied independently to specific topics. Hence a student (or a teacher) can feel intimidated by any statistical topic if they have generalized a number of specific statistics-related experiences to the broader field of statistics.

Affective constructs generated by a socio-cultural environment. Goldin (2002) addressed the socio-cultural environment that exists external to an individual with respect to mathematics. This environment generates "remarkably consistent feedback" concerning (1) shared emotions, (2) prevailing or acceptable attitudes, (3) belief systems that exist across the culture, and (4) values, ethics and morals communicated through peer groups, schooling, and examples from adult family members. I believe one potential result due to a socio-cultural environment is that preservice teachers can develop a fear of statistics due to exposure to a preservice teacher and undergraduate student environment that propagates a general fear of statistics.

Attitudes and other concerns in undergraduate statistics education. Undergraduate statistics education has been the focus of three types of research that are addressed here. There have been calls for improved undergraduate statistics education. Attitudes and beliefs toward statistics have been shown to be negative for many undergraduate students. Effective methods for presenting statistical material during lesson have been identified. The American Statistical Association has endorsed a set of proposals called the GAISE reports (ASA, 2005). The GAISE reports (Guidelines for Assessment and Instruction in Statistics Education) were focused on improving statistics education in both the K-12 setting and the undergraduate setting. These reports consisted of a curricular framework for preK-12 statistics education and specific recommendations for the improvement of introductory college statistics courses (GAISE project, 2005a, 2005b, 2005c, 2005d; Landwehr, 2004; Wilson, 2004).

There is reason to believe that many undergraduate students are intimidated by statistics (Hirabayashi, 1999; Mvududu & Nyaradzo, 2003; Rhoades, 2000). These students are likely to have trouble with statistics due to their negative attitudes or beliefs toward the subject (Gal & Ginsberg, 1994). This trouble often has no cognitive basis. Preservice primary teachers are known to have increased fear of mathematics relative to undergraduates in STEM majors such as secondary mathematics majors (Ball, 1990). Preservice elementary teachers are not as likely as STEM majors to take a complete course in undergraduate statistics since most elementary education programs do not require such a course and most STEM majors do.

Effective teaching methods have been identified for undergraduate statistics education. Gal (2002) constructed a model of statistical literacy that indicates factors such as statistical knowledge, mathematical knowledge, critical questions, and beliefs and attitudes. From such constructs, Gal and others have found evidence of teaching devices that help statistics students at both K-12 and undergraduate (K-16) levels. These teaching devices include activity based courses; the use of small groups and cooperative learning to force students to argue their case and to improve attitudes toward statistics; testing students and providing consistent feedback on their misconceptions; the use of more open-ended problems while using fewer goal-specific problems; and incorporating active-learning strategies such as designing studies, collecting data, analyzing results, preparing written reports, and presenting results orally (Gal, 2002; Garfield, 1995; Smith, 1998)

If these specific teaching methods benefit undergraduate students, it seems possible that the same methods would benefit both primary and secondary level preservice teachers. Furthermore, since many of these devices have also been specifically identified as beneficial for K-12 students, teaching the teachers using these methods may increase the likelihood that the teachers will use these effective devices when teaching statistics to their K-12 students.

Evidence of teachers' statistical content knowledge levels. The statistical knowledge levels of teachers have been shown to be relevant to K-12 statistical learning. Three primary issues each indicate that teachers need strong statistical understanding.

Teachers need statistical knowledge so that they can adequately implement curricula (Friel & Bright, 1998). Teachers also need to be able to respond to surprise questions, unanticipated responses, and unintended outcomes during statistical investigation (Mickelson & Heaton, 2004). These situations require teachers to "think on their feet" statistically. (p 327) This is similar to the notion of "on line" actions of teachers in the mathematics classroom coined by Schoenfeld (1998). (p 1) Schoenfeld argues that teacher goals, beliefs, and knowledge interact to shape the moment-to-moment actions of teachers in the classroom. This becomes especially relevant when situations occur such as those suggested by Mickelson and Heaton. There is also evidence that teachers need not just statistical knowledge, but the ability to reason with data by using statistics (Makar & Confrey, 2004).

It is clear that statistical knowledge and reasoning are needed. Unfortunately, the evidence indicates that there are concerns about the statistical knowledge levels that teachers possess. Teachers tend to have weak statistical concept skills, and even when understanding does exist, it is often fragmented (Mickelson & Heaton, 2004). Types of reasoning skills can be developed within one content area and yet not exist in other content areas. Even when statistical reasoning skills are present, teachers tend to perform better in the domain of pure statistical knowledge than the domain where they have to apply this knowledge to teaching (Sorto & White, 2004).

Explanations have been presented for these conditions. Makar & Confrey (2002) note that teachers receive mixed messages concerning statistics in the United States, and many other countries, due to the dichotomous nature of the increased calling for statistical literacy as a vital skill for citizens, versus the overemphasis of misused, high-stakes accountability systems that have the potential of retarding the growth of progressive teaching methodologies. Peck and Gould (2005) noted that in the United States, the majority of secondary school teachers of statistics have backgrounds in mathematics with little training in statistics; they emphasize that while statistics is a mathematical science, it is fundamentally different from mathematics.

The statistical understanding of teachers is becoming so relevant that within the past three years at least three dissertations have focused on teacher knowledge of statistics within specific parameters. These parameters have included preservice teachers' understanding of variation, teachers' development of statistical inquiry as applied to the analysis of high-stakes testing results, and middle school teachers' knowledge of data analysis as applied to teaching (Canada, 2004; Makar, 2004, Sorto, 2004). More than one of these studies indicated that teachers need to improve their understanding of statistical concepts.

The effects of attitudes and self-efficacy toward the learning of mathematics and statistics on teacher success learning mathematics and statistics. Ball and Cohen (1999) believed that to embrace progressive teaching methodologies, teachers must develop and expand their ideas about learning. This includes contemplating what it means to learn. Such behavior is not necessarily easy to promote in preservice teachers.

Borko et al. (1992) noted that a student teacher who was unsuccessful presenting a lesson due to a lack of content knowledge failed to have either the impetus or desire to seek a correct way to present the material. Due to the fact that the student teacher was involved in a study, the teacher had interactions with researchers immediately after the lesson. Three times over a period of several days after the lesson, the researchers asked the student teacher if there was an effective way to present a specific problem that had impacted the original lesson. Even on the third try, the student teacher had no valid response to the question. This indicates a situation in which the teacher failed to see the importance of correcting the deficiency in their teaching knowledge.

This teacher's attitudes toward the correction of their knowledge levels affected their ability to learn the material. A teacher is certainly unlikely to improve their knowledge levels when they are unmotivated to make any effort to correct known deficiencies. I believe that teacher beliefs about learning are the internal guides that motivate them to seek or not seek understanding when confronted with deficiencies in their knowledge levels.

Such results are disconcerting when combined with the indications that the teaching profession has been unsuccessful in attracting prospective primary teachers who have positive attitudes toward mathematics (Philippou & Christou, 1998). It may very well be that many primary teachers lack the motivation to correct their deficiencies in mathematical knowledge due to attitudes and beliefs that do not support reflection and self-correction.

Teacher attitudes toward the teaching of statistics. When the literature is paired down to studies that address primary level teacher attitudes toward the teaching of statistics, conflicting results appear. Teachers tend to have little experience learning statistics before they begin to teach the subject (Friel & Bright, 1998). This experience is often not a positive one (Friel & Bright, 1998). Material that is unfamiliar to teachers tends to be intimidating to them (Bright & Friel, 1993). Yet primary level teachers have indicated that they are relatively secure in their ability to teach statistics (Begg & Edwards, 1999). In the same study, the teachers showed a weak understanding of several statistical topics including mean, median, and mode; probability; and stem-and-leaf plots. These teachers unanimously believed that it is important for teachers to encourage children to explore their own mathematical ideas rather than for teachers to simply impart their own understanding to the students (Begg & Edwards, 1999). They seem to not understand the connections between teacher ability to teach in this manner and deep understanding of content.

Despite weak reasoning skills in specific topics, and despite low levels of statistical training, the teachers in Begg and Edwards' study indicated that, for these primary level teachers, continuing professional development in statistics would either not be a priority or would only be of low priority (Begg & Edwards, 1999). The teachers did not necessarily see their lack of statistical knowledge as a problem for their teaching. In fact, most of the teachers felt that a person need not be highly skilled in mathematics to be able to grasp concepts such as average. When asked about the types of possible in-service courses they would prefer concerning statistics, nearly all said that they would want more ideas and activities for use in their classroom. Only two said they would want the work to specifically improve their statistical understanding. Begg concludes by suggesting that it will be difficult for teachers to enable their students to take possession of the content if the teachers have not taken possession of the content themselves.

Professional Development Programs for Statistics

Numerous cases exist to support the argument that both preservice training and continuing professional development are important for improved K-12 statistics education (Callingham, 1995; Cobb & McClain, 2004; McClain, McGatha, & Hodge, 2000; Mickelson & Heaton, 2004; NCISLAMS, 2004). The rationale for such cases varies from the increase in the use of inferential statistics in society, to the value of using open discussion to enhance learner statistical reasoning, to the differences between the reasoning required to teach statistics and the reasoning required to directly use statistical concepts, to creating lessons better suited for the ways in which students think and learn about statistics.

Batanero and Godino (2004) found that students' strategies for solving probability problems are quite different from those of professional statisticians. This emphasizes a need to provide statistical training for teachers that is different from that which would be provided in a collegiate course designed primarily to train students for professional application of statistics. In fact, Batanero and Godino suggest that it is urgent to offer better training in statistics during preservice training as well as continuous support from University departments and research groups.

International studies have reflected this pattern of recommendations. Greer and Ritson (1993) state that in Northern Ireland, teachers at all levels need in-service training to upgrade their understanding of probability and statistics to be aware of appropriate teaching methodologies. This has lead to the development of specific CPD programs to assist teachers. Chadjipadelis (1999) developed a CPD program in Greece in which teachers were exposed to several hours each of statistics theory and laboratory group work. The teachers also organized a project in a school class during these activities. Some of the concepts focused upon were: the use of excel to analyze data, the making of tables, the understanding of mean and median, and the writing of essays to formulate and communicate their thought processes.

Domestically, three major studies have provided valuable information concerning what works and what does not work in CPD for teacher statistics reasoning and classroom application. An example of one CPD program in statistics is the Teach-Stat program developed and implemented by Friel and Bright. This program was designed to help grades 1-6 teachers to learn more about statistics and integrate teaching about and teaching with statistics in their instruction (Friel & Bright, 1998).

The program incorporated teacher administration of statistics lessons while supervised by teacher educators during 2- or 3-week long workshops. The researchers found that teachers improved their statistical understanding but even after practicing the teaching of statistics for at least a year after the workshop, many teachers still had holes in their understanding of statistics content. Possibly more importantly is the result that teachers who participated in the program reported that they felt they had improved their effectiveness at questioning students and encouraging inquiry. The teachers also felt that being involved in the program had caused a positive change in the way they viewed teaching and on their students' responses to their teaching.

Studies on professional development statistics programs can provide insight into teacher needs. A study involving professional development for preservice teachers was performed at the University of Nebraska (Heaton & Mickelson, 2002). The program consisted of a one semester course in which students learned to use statistical investigation by posing appropriate questions, identifying variables, performing data collection, summarizing data, and reporting findings. The course goals were to help preservice teachers to develop perspective, insight, and enthusiasm for statistical investigation.

The researchers found that the preservice teachers resorted to simple graphs to explain many of the analyses, a tool with which they were familiar. This supports the suggestion that preservice teachers need to have a level of statistical training beyond that which is practical within the confines of a preservice teacher program. Even the one semester course provided for preservice teachers at the University of Nebraska is far beyond the level of statistical training most programs are willing or able to expect from preservice primary teachers. Heaton and Mickelson found that although the program appeared to have failed to provide the desired depth of statistical content knowledge, it did provide a context for learning statistical investigation that they found interesting and relevant.

Another example of a CPD program in statistics is the INSPIRE project developed by Gould and Peck (Gould & Peck, 2005; Peck & Gould, 2005). In the project, they combined a weeklong workshop with a nine-month online course. In line with research findings, they incorporated teacher educator to teacher interactions (via the workshop and online communications) with teacher to student interactions in the classroom. The program provided success for teacher growth in statistics reasoning. However, Gould and Peck noted a disappointing aspect of the first course of the project; there was a lower than anticipated level of student-to-student interaction. This may reflect the tendency for teachers to separate themselves as a community of professionals and for individual teachers to become isolated in their classrooms.

Chapter Summary

The rationale for what is presented in chapters three and four follows the results of what has been discussed in chapter two. Teachers need to utilize CPD due to factors such as exposure to students. Professional development is necessary to help teachers to successfully implement progressive teaching strategies. Progressive teaching strategies are known to provide more effective ways for students to learn.

If PMTEs intend to improve the likelihood that preservice teachers will eventually embrace CPD in a specific content area such as statistics, then it is essential to identify those factors on which attitudes toward CPD are dependent. Although some dependence, or lack thereof, has been shown for isolated factors such as general attitudes toward statistics and statistics knowledge on students in non-mathematically oriented majors (Templaar, 2003), no study has investigated general attitudes toward statistics, self-efficacy to learn statistics, self-efficacy to use statistics, statistical knowledge, and pedagogical statistical knowledge all in one study. Such studies with more specific attitudes measures have been recommended (Finney & Schraw, 2002). This study not only accomplished all of those investigations, it was also conducted with a principal focus on preservice teachers.

Chapter 3; Methodology

The goal of the study was to answer the research question, "Are there indicators from preservice teacher attitudes toward, and knowledge of, statistics that might assist the efforts of preservice mathematics teacher educators (PMTEs) to increase the possibility that these preservice teachers will pursue continuing professional development (CPD) in statistics?" To seek answers to this question, I determined to collect data about both primary and preservice secondary teachers. There were seven primary constructs to be measured.

During the discussion of these constructs a number of acronyms are defined and used. These acronyms represent lengthy titles that are referred to repeatedly. Table 3.1 outlines these acronyms. The constructs and the methods used to collect data for each construct are listed in table 3.2.

Table 3.1 *List of Acronyms Used*

Acronym	Phrase
PMTE	Preservice Mathematics Teacher Educators
CPD	Continuing Professional Development
ATS	Attitudes Toward Statistics
CSSE	Current Statistical Self-Efficacy
SELS	Self-Efficacy to Learn Statistics
SCI	Statistics Concepts Inventory
ARTIST	Assessment Resource Tools for Improving Statistical Thinking
ATPDS	Attitudes Toward Professional Development in Statistics

Table 3.1 *continued*

Acronym	Phrase
IRB	Institutional Review Board
SPSS™	A Marketed Software Package That Can Perform A Variety of Statistical Analyses

Table 3.2 *Constructs to Be Measured and Methods of Measurement*

Construct to be measured	Method of measurement
• Attitudes toward statistics in general	• ATS, Wise (1985)
• Self efficacy toward current ability to perform statistical procedures	• CSSE, Finney & Schraw (2003)
• Self-efficacy toward potential ability to learn statistical reasoning	• SELS, Finney & Schraw (2003)
• Statistical knowledge and statistical reasoning	• ARTIST scales and SCI instrument
• Ability to grade student work involving statistical reasoning	• Grading project, Appendix C
• Background in statistical training	• Demographic questionnaire, SCI
• Interest in pursuing CPD in statistics (ATPDS)	• Instrument, Appendix B

Defining the Variables and Abbreviated References

Tables 3.3 to 3.9 provide the variables and the abbreviated references used in the analyses with descriptions of the variables and the references that have been abbreviated. For example, for the qualitative analyses the participants are identified via a short identifier that does not jeopardize the identity of each. SPSS™ truncates each variable to eight characters or fewer (e.g., the variable ATScourse is presented as ATScours on each SPSS™ output).

Table 3.3 *Variables Representing Instrument Scores from Phase 1*

Variable	description
ATSfield	Wise's ATS instrument – scores for the "field of statistics" subcategory
ATScourse	Wise's ATS instrument – scores for the "current statistics course" subcategory
ATSsum	Total of the field and course subcategories for the ATS instrument
CSSE	Current Statistics Self-Efficacy instrument scores
SELS	Self-Efficacy to Learn Statistics instrument scores
SCI	SCI scores

Table 3.4 *ARTIST Scales Variables*

Variable	description
SCgraphs	ARTIST scale scores for understanding of graphical representations
SCcenter	ARTIST scale scores for understanding of measures of center
SCspread	ARTIST scale scores for understanding of measures of spread
SCprob	ARTIST scale scores for understanding of introductory probability
SCsum4	Sum of the first four ARTIST scales
SCdacoll	ARTIST scale scores for understanding of data collection
SCcnfint	ARTIST scale scores for understanding of confidence intervals
SCsmpvar	ARTIST scale scores for understanding of sample variation
SCsum7	Sum of the first seven ARTIST scales

Table 3.5 *Interest in CPD in Statistics Variables (from the ATPDS Instrument)*

Variable	description
WKen	Response to the question about participant belief that they would enjoy a CPD workshop in statistics
WKfr	Response to the question about participant belief that they would attend a CPD workshop in statistics if colleagues who were friends attended
WKhp	Response to the question about the value of CPD workshops in statistics for the participant's classroom teaching
WKpd	Response to the question about participant belief that they would attend a CPD workshop in statistics if the cost of the workshop was covered
WKin	Response to the question about participant belief that they would attend a CPD workshop in statistics if a small stipend was provided
WKls	Response to the question about participant belief that they would attend a CPD workshop in statistics if the participant had to pay for the workshop
WKtotal	Total score from the six WK responses

Table 3.6 *Squares of Certain Attitude Variables for Quadratic Models*

Variable	description
ATSsumSq	The square of the ATSsum value for each participant
ATSfldSq	The square of the ATSfield value for each participant
CCSEsqrd	The square of the CSSE value for each participant
SELSsqrd	The square of the SELS value for each participant
ATScsSqd	The square of the ATScourse value for each participant

Table 3.7 *Qualitative Variables used in the Quantitative Analyses*

Variable	description
SEMESTER	0 = fall 1 = spring
EDMAJOR	0 = elementary education major or secondary mathematics ed. major 1 = otherwise
COURSE	1 = 1473 the introductory undergraduate mathematics course 2 = 3213 the second math content-for-teaching course for elementary Education majors 3 = 432 the mathematics content-for-teaching course for secondary Education majors
COLLSTAT	0 = the participant has indicated that they have not completed a full-semester course in statistics 1 = the participant has indicated that they have completed a full-semester course in statistics

Table 3.8 *Participant Group References*

Abbreviation	description
1473	All 81 participants from the introductory mathematics course
1473ed	The 12 participants from 1473 who are elementary education majors
1473not	The 69 participants from 1473 who are not elementary education majors
3213fa	The 22 participants from the fall semester of the second primary mathematics content course
3213sp	The 22 participants from the spring semester of the second primary mathematics content-for-teaching course
3213sp21	The 3213sp data set with participant #18 removed due to no WK scores (so n=21)

Table 3.8 *continued*

Variable	description
3213sp19	The 3213sp data set with participants #1, #17, and #18 removed due to a lack of complete ARTIST scores (so n=19)
3213all	The combined participants from 3213fa and 3213sp
432n6	The six participants from the secondary mathematics content-for-teaching course[watch that these participants are not identifiable]
432n4	The four participants from the 432n6 data set with WK scores [watch that these participants are not identifiable]

Table 3.9 *Qualitative Analyses References*

Abbreviation	description
El	3213sp participant #1, a primary grading project participant
Cy	3213sp participant #21, a primary grading project participant

The terms survey and instrument. In this chapter, I used the terms *survey* and *instrument.* The term instrument refers to any single device that was designed to measure a specific construct such as self-efficacy toward current ability to perform statistical procedures. The term survey refers to the union of all the instruments as presented to a particular set of participants. Throughout the study, each instrument stayed the same. However, the survey changed across differing groups of participants due to variation in the specific set of instruments that comprised the survey.

Categories of data collection. I used a combination of quantitative and qualitative data collection techniques, including Likert-style surveys and open-ended instruments. Knowledge levels were measured using multiple choice instruments. Various forms of the survey were administered to measure the constructs listed in table 3.1. A qualitative grading project was completed by a small subgroup of the participants. This instrument was open-ended.

Development of the methodology. The resulting methodology developed in two distinct phases. By the end of the summer before data collection was to commence, I was prepared to gather data for the preservice teachers concerning (1) general attitudes toward statistics, (2) self-efficacy to use statistics, (3) self-efficacy to learn statistics, (4) knowledge of statistics, and (5) pedagogical content knowledge as applied to grading student work. The goal was to investigate teacher attitudes and knowledge of statistics in a more complete manner than had previously been attempted. The planned timeline was to gather data in both the fall (semester 1) and subsequent spring (semester 2) semesters of that academic year. IRB approval was secured during that summer and gave permission to gather data during the subsequent fall and spring semesters.

Near the end of semester 1, I determined to add an instrument to the survey to gather information about the teachers' attitudes toward continuing professional development in statistics. The instrument is referred to as the *attitudes toward continuing professional development in statistics* (ATPDS) instrument. This instrument was added to the initial data gathering structures. Because this instrument expanded the research question, there was substantial difference between the data collected in semesters 1 and 2. Hence the plan for the study entering semester 1 is referred to as phase 1 and the new plan for the study implemented between semesters 1 and 2 is referred to as phase 2.

Design Development: Phase 1

The existing research indicated that there were no meaningful correlations between undergraduate students' content knowledge of statistics and attitudes toward statistics (Templaar, 2003). Statistics reasoning skills were most closely related to mathematics outcomes, even when compared with statistics outcomes. These results were found using a general attitudes-toward-statistics instrument from Schau, Stevens, Dauphinee, and DelVecchio (1995) and a statistics reasoning instrument from Garfield (1998).

I believed that there was potential for more complete studies to offer additional details on the teacher attributes "attitudes toward statistics" and "knowledge of statistics". This idea was supported by recommendations from the literature (Finney & Shraw, 2003). Thus the research question at that time was "Can a study that incorporates more complete data gathering reveal that there are interactions between teacher attitudes toward statistics and teacher content knowledge of statistics?"

One significant difference developed between my initial study design and the design of the studies used in the existing literature. My initial study was designed to investigate the attitudes of preservice teachers to learn statistics. The previous studies had focused on either general attitudes of the preservice teachers toward statistics or preservice teacher's attitudes toward the teaching of statistics (Begg & Edwards, 1996; Callingham, 1995; Hoy, 2000). In fact, my initial study design was to investigate both teachers' general attitudes toward statistics and teachers' self-efficacy to learn statistics. My rationale for this decision was indications that the confidence a person has in their ability to accomplish a task is the best known behavioral predictor of whether or not the person will accomplish the task (Bandura, 1986; Pajares, 1996).

During semester 1, the initial research question guided the study design and data collection. During the course of the semester, as data were gathered, I continued to investigate what I considered to be the relevant literature. I gradually began to visualize the complex interactions described in Chapter 2 and how professional development plays an integral part in effective teaching. By the end of the winter break at the terminus of semester 1, I recognized that I needed to gather information about the preservice teachers concerning their beliefs about future professional development.

Design Development: Phase 2

I revisited the results reported above from Bandura (1986) and Pajares (1996). I became curious about whether self-efficacy might be a predictor of preservice teacher interest in further professional development in statistics. An important event during the winter break was a discussion that I had with Dr. Barbara Pence of San Jose State University. I explained to her the data I had gathered during phase 1 and I discussed with her my interest in professional development as an important part of the implementation of progressive teaching strategies. She suggested that I incorporate the professional development issue into the study during the second semester.

I completed a revision for my Institutional Review Board (IRB) approval. The new permissions maintained the original study plans and included two changes. The revision provided permission to add to the survey an instrument concerning participants' attitudes toward professional development. It also provided permission to use the ARTIST scales.

The additional instruments allowed for a research question with an expanded focus from the original research question. I integrated the original research question into the new design as a secondary question. I also added another secondary question: "How do the factors from the original study design correlate with preservice teachers' attitudes toward professional development?"

Timeline and Setting

In the summer prior to the study's commencement, I acquired IRB approval to approach students in specific courses in the Department of Mathematics at the University of Oklahoma. Each course served preservice teachers to some extent. The courses were: a fourth year mathematics content-for-teaching course for preservice secondary teachers (432), a third year mathematics content-for-teaching course for primary preservice teachers (3213), and a general introductory mathematics course that served as a prerequisite for the primary preservice teacher content-for-teaching course (1473). At this university, these courses provided the most efficient way to approach preservice teachers who are working on mathematical content.

The primary preservice teachers followed a track of mathematics courses that is worth clarification. As summarized in Table 3.10, preservice primary teachers take the general introductory mathematics course, a second general mathematics course, and then they take a second year course in mathematics content-for-teaching followed by a third year course in mathematics content-for-teaching. Students in the first and last of these four courses were invited to participate in the study.

Table 3.10 *Summary of Primary Preservice Teacher Minimum Mathematics Tract*

	Course 1 An Introductory Mathematics Course	Course 2 A Second General Mathematics Course	Course 3 A First Mathematics Content for Teaching Course	Course 4 A Second Mathematics Content for Teaching Course
Statistical Content	9 of the 43 course meeting hours are dedicated to statistical concepts	Usually no statistical content	A minimal amount of statistical concepts are addressed	8 of the 43 course meeting hours are dedicated to statistical concepts

The IRB approval for the study provided for collection of data during two semesters, referred to here as semester 1 and semester 2. During semester 1, participants were recruited from all three courses (1473, 3213, and 432). For semester 2, students were recruited only from the primary-level mathematics content-for-teaching courses. A summary of the timeline is provided in table 3.11.

Table 3.11 *Summary of Data Collected*

Introductory Mathematics Course	Preservice Primary Teachers Mathematics Content for Teaching Course	Preservice Secondary Teachers Mathematics Content for Teaching Course
Semester 1 • demographics questionnaire • current statistics self-efficacy instrument • self-efficacy to learn statistics instrument • attitudes toward statistics instrument • statistical content knowledge instrument (SCI)	Semester 1 • demographics questionnaire • current statistics self-efficacy instrument • self-efficacy to learn statistics instrument • attitudes toward statistics instrument • statistical content knowledge instrument (SCI)	Semester 1 • demographics questionnaire • current statistics self-efficacy instrument • self-efficacy to learn statistics instrument • attitudes toward statistics instrument • statistical content knowledge instrument (SCI) • grading project
Semester 2 no data collected	Semester 2 • demographics questionnaire • current statistics self-efficacy instrument • self-efficacy to learn statistics instrument • attitudes toward statistics instrument • statistical content knowledge instrument (ARTIST) • interest in professional development in statistics • grading project	Semester 2 • interest in professional development in statistics • statistical content knowledge instrument (ARTIST) [administered during semester 1 as part of the coursework – permissions to use the data in the study gained via follow-up contacts during semester 2]

Participants

Per IRB requirements, for each specific interaction that I wanted to execute with participants, I designed a specific consent form. The students who chose to participate signed each appropriate consent form. There were two basic types of consent forms. One was a form of consent to participate in the survey. The second type of consent form gave permission for work completed on grading projects to be included in the research analyses.

Students could choose to participate in the survey or the grading project, or both or neither. The protocol for the request to participate did include an explanation that those who participated in only non-survey activities might not be included in the final analysis. The goal for including participants from the general introductory course survey was to provide comparisons of growth in knowledge and/or attitudes toward statistics in preservice primary teachers as they progress through their mathematics content and mathematics content-for-teaching courses. Thus, participants from this first course only completed the survey; they were not offered opportunity to participate in the grading project. In semester 1, I approached a large section of the general introductory mathematics course. There were approximately 125 students in the section. Eight-one of the students participated in the survey (a 65% response rate).

From the primary preservice teacher courses, I invited students to participate in both the survey and the grading projects. Over the summer I had obtained the cooperation of the person originally assigned to teach the two sections of the course offered in semester 1, but that teaching assignment was changed at the last minute.

There were then two different instructors, on for each section of the course. During the early part of semester 1, I struggled to obtain cooperation from these last-minute instructors. One of the instructors was struggling as a first-time teacher of the course, so I was not permitted by the department administration to approach this instructor. The other instructor was cooperative once she learned that the study was minimally invasive to her course schedule. She was willing to allow one fifty-minute session during course time for the survey. It had previously been determined that participation in the qualitative portions of the data gathering would occur outside of class time.

Of the 35 students in the section, 22 participants volunteered for the survey portion of the study (a 63% response rate). Unfortunately, there was no interest from the students to participate in the grading project. So there were no grading project data for semester 1 for primary preservice teachers.

In semester 2, there were again two sections of the preservice primary teacher mathematics content-for-teaching course. Again, there was one instructor who was struggling, albeit a different instructor. I was able to gain permission to approach students in the section in which the instructor was not struggling. (Over the two semesters, and four sections of the course, there were four different instructors.) Coincidently, the same number of students participated. Of the 35 students in the cooperative section, 22 participated in the survey (63%). During this semester, two of the students also participated in the grading project.

The preservice secondary teacher mathematics content-for-teaching course is offered only in fall semesters. In semester 1, I was the instructor for the class, making the six enrolled students a vulnerable population according to IRB definitions. Thus, data collection in this class proceeded with all due care. For instructional reasons, I had decided to include the study instruments in the course design (most of the instruments are designed for course use in addition to research use). I believed that these instruments would provide meaningful growth for the preservice teachers. As required by the IRB plan, I explained at the beginning of the semester, and repeated each time an instrument was applied, that the work on the instruments was mandatory for the course, that the students could allow me permission to use these instruments in my research, and that they could chose not to give me permission if they wished. I repeatedly informed the students that I would turn in their grades for the course before I knew whether they had given me the permissions. Hence their decision to participate in the study had no relevance to their course grade even though completion of the instruments did.

I had a fellow graduate student approach the preservice teachers on the day that instructor evaluations were completed. I had to leave the classroom for the integrity of the evaluations and I took advantage of this situation for my colleague to present the protocol for participation in the research and invite the students to sign the informed consent forms. Afterwards, the graduate student delivered the signed informed consent forms to an advising faculty member who locked them in a filing cabinet. One week after the last day to submit final grades, I picked up the consent forms so that I could begin data analysis. All six students enrolled in the class agreed to allow me to use their surveys in my research.

Because the ARTIST scales were not acquired for the study until November of semester 1 (see subsequent section *Statistics content knowledge instrument*), they were not included as part of the initial IRB approval. I was able to implement the ARTIST scales into the coursework for the secondary teachers. During semester 1, the only intent of this action was for coursework. Once I gained IRB permission to include the ARTIST scales during semester 2, I had an IRB approved consent form that could be presented to 432 students. Because my contact with the 3213 section was only for purposes of the study, I could not ask them to perform the ARTIST scales.

During the break between semesters, I had gained IRB approval to add the ATPDS instrument to the survey. When I contacted the 432 participants for permission to use their ARTIST scales results for the study, I also asked if they were interested in completing the ATPDS instrument that was now IRB approved to be included as part of the survey (so long as new consent forms were completed). Four of the six were willing to take the time to complete the ATPDS instrument. All six granted permission to use the ARTIST scales in the study. A list of the numbers of particpants is provided in Table 3.12.

Table 3.12 *Number of Participants for the Survey*

	Introductory General Mathematics Course: Semester 1	Primary Mathematics Content-for-Teaching Course: Semester 1	Primary Mathematics Content-for-Teaching Course: Semester 2	Secondary Mathematics Content-for-Teaching Course: Semester 2
Potential Participants (approximate)	125	35	35	6
Actual Survey Participants	81	22	22	6
Actual ATPDS Participants	0	0	22	4

The Survey

The quantitative data were collected in one integrated survey. This survey consisted of a brief set of questions concerning student background and demographics followed by separate, previously developed instruments (Appendix A). The semester 1 survey had all three instruments that measure affective factors toward statistics (ATS, CSSE, and SELS); it also contained the SCI knowledge instrument. The semester 2 survey had all three affective-measure instruments; it also contained the ATPDS instrument. The knowledge instrument for the semester 2 survey – the ARTIST scales – was administered online.

The semester 1 survey participants completed a paper version of the survey, lasting about 40 to 50 minutes. The semester 2 survey participants completed the survey in two parts, one on paper and another online. The paper version contained the affective measures that the semester 1 participants completed and the new ATPDS instrument that the semester 1 participants did not complete. The ARTIST scales that I used for the knowledge instrument in semester 2 were completed online. The participants spent about 25 to 30 minutes on the semester 2 paper version. The protocol was for the group to move to a computer lab within 30 yards of the 3213 classroom. The participants spent about 20 to 30 minutes on the online portion of the survey.

Background and demographics instrument (Appendix A). The SCI was created for a comprehensive study led by an engineering faculty member to inventory undergraduate and graduate students' statistics concepts knowledge (Allen, 2006; Stone, 2006). The SCI development team typically included a background/demographics instrument in their data collection. With permission from the authors (A. Stone, personal communication, September, 2005), I modified that background/demographics instrument to fit the needs of this study (Appendix A).

The items on the background/demographics instrument asked participants for information such as their gender, university status (junior, senior, etc.), intended major, and statistical background. Most of the items allowed participants to circle or check an answer from among a short list of options. For these items, blank lines were left for participants to write any answer not included in the list.

For my own data collection in the general introductory mathematics course, I included the question about intended major to disaggregate participants from different majors. Specifically, I believed it would be beneficial to compare the students who were planning to major in elementary education to the other participants from the course. Such demographic information also enabled a comparison of this subgroup within the general introductory mathematics course and the preservice teacher participants in the primary preservice teacher mathematics content-for-teaching course. This comparison fed into an evaluation of the growth of preservice teachers' attitudes about, and content knowledge of, statistics between their first exposure to undergraduate statistics content in the general introductory mathematics course and their third course, that was their second exposure to statistics, in the mathematics content-for-teaching course.

The results from the instrument question about intended major were recorded into the data set as a qualitative variable. If the general introductory mathematics course participant indicated that they were not elementary education majors, then the variable value was recorded as a 0. If the participant indicated that they were majoring in elementary education, then the variable value was recorded as a 1. Participants from the mathematics content-for-teaching courses for primary teachers were automatically assigned a value of 1 for this variable. Correspondingly, participants from the mathematics content-for-teaching courses for secondary teachers were automatically assigned a value of 2 for this variable.

The statistics education background item provided specific types of answers from which participants could choose. I wanted to get more specific categories of answers for this item than I felt was likely if I allowed the participants to fill in the answers themselves. The participants were allowed to mark more than one of the answer options. The purpose of these options was to be able to check correlations between other measured factors and different levels of statistical experience.

The results from the instrument question about statistical experience were recorded into the data set as a qualitative variable. If the participant indicated that they had not completed a full semester course in statistics, then the variable value was recorded as a zero. If the participant indicated that they had completed at least one full semester course in statistics, then the variable value was recorded as a one.

Self-efficacy to use existing statistics reasoning instrument (Appendix A). I employed the *Current Statistics Self-Efficacy* (CSSE) instrument for two reasons. First, it helped to establish data concerning preservice teacher attitudes toward statistics. In addition, it gave a second attitudes measure that could be compared to the general *attitudes toward statistics* (ATS) measure from the instrument developed by Wise (1985). The ATS instrument is discussed shortly.

Finney and Schraw (2002) have suggested that past studies relying on more general attitudes are likely to be flawed and that more specific self-efficacy measures may improve the authenticity of the results. This claim is supported by Goldin (1998) in his discussion of global versus local affective factors. These claims motivated the initial design of the study. The incorporation of additional affective measures toward statistics allowed for an evaluation of the consistency of students' attitudes toward statistics where the attitudes were presented with different points of emphasis. It also increased the potential for one of the types of attitudes to correlate with content knowledge levels and/or interest in professional development in statistics. Since it was unclear whether either general or specific affective factors might be more likely to correlate with the non-affective factors, it seemed reasonable to consider both.

The goal of the CSSE instrument designed by Finney and Schraw (2002) was to measure the self efficacy of undergraduate students to use statistics. The instrument consisted of 14 items (Appendix A). For each item, the participant was asked to rate their confidence in their current ability to successfully complete a statistical task. The ratings were in Likert scale format with a range of 1 for "no confidence at all" to 6 for "complete confidence". Hence, participant scores for the CSSE ranged from 14 for marks of no confidence on every item to 84 for marks of complete confidence on every item.

Self-efficacy to learn statistics instrument (Appendix A). The *Self-Efficacy to Learn Statistics* (SELS) instrument was designed along with the self-efficacy to use existing statistics knowledge instrument (CSSE). Both instruments were shown to have sufficient validity and reliability (Finney & Shraw, 2003). The Self-Efficacy to Learn Statistics instrument contains the same 14 items as the self-efficacy to use statistics instrument. The same Likert scale format is used. In the CSSE instrument, the focus of the 14 items is in the context of participant confidence in their current ability to complete each task. In the SELS instrument, the focus of the 14 items is in the context of participant confidence in their ability to learn how to complete each task successfully. Again, the ratings are in Likert scale format with a range of 1 for "no confidence at all" to 6 for "complete confidence". Hence, participant scores for the SELS range from 14, for marks of no confidence on every item, to 84, for marks of complete confidence on every item. The numerical scores allow for quantitative analyses on the CSSE and SELS variables. High scores correspond to high levels of confidence and low scores correspond to low levels of confidence.

General attitudes toward statistics instrument (Appendix A). To measure general attitudes toward statistics, I chose the *Attitudes Toward Statistics* (ATS) instrument developed by Wise (1985) because of the number of studies that have verified its reliability and validity (Cashin & Elmore, 1997, 2005; Wise, 1985). This instrument consists of 29 items, each of which utilizes a Likert scale format that allows participants to choose between "strongly disagree", "disagree", "neutral", "agree", or "strongly agree". The 29 items cover topics such as the general benefit statistics provides for society, the benefit statistics will provide for the participant's career, and participant opinions of statistics courses at the undergraduate level.

The level of agreement a participant provides for each item has a numerical value associated with it. On 15 of the items, "strongly disagree" has a value of 1, "disagree" has a value of 2, "neutral" has a value of 3, "agree" has a value of 4, and "strongly agree" has a value of 5. The other 14 items use reverse order values. The reverse items format enables instrument administrators to check for participants who are randomly marking the same answer value (unless they randomly mark the value 3). It also provides the participants with an opportunity to think about their opinions from a different perspective.

To compute a quantitative score for this instrument, it is necessary to adjust the item response values. For example, a score of two on a reverse item is the equivalent of a score of four on the standard items. Items with answer values in reverse order were identified by Wise in a supplement to the instrument (Appendix A). An algorithm was designed to calculate the adjusted response, x, as a function of the original reverse response, y, as follows: $x = 6 - y$.

Once this adjustment was made, all 29 items could be summed to generate an ATS score. However, Wise separated the items into two mutually exclusive categories, attitudes toward the field of statistics and attitudes toward the current course in which statistics is being studied. These two categories are defined as the variables ATSfield (20 items) and ATScourse (9 items). The sum of these two variables is the ATS total score and is defined as ATSsum. High ATS scores correspond to strong positive attitudes toward statistics.

Statistics content knowledge instrument (Appendix A). To measure the preservice teacher participants' content knowledge in semester 1, I used the *Statistics Concept Inventory* (SCI), that had been developed and implemented for a separate study (Allen, 2006; Stone, 2006; Stone, A., personal communication, July 22, 2005). The SCI consisted of 38 multiple choice items. To make the instrument easier to score,

I generated an answer sheet on which participants wrote their answers. The 38 problems were given equal weight in the scoring of participant results.

Since this was a multiple choice instrument and not in Likert scale format, there were no weighted scores on the answers. If the answer was correct, then the participant was given one point for that item. If the answer was incorrect, then the participant was given zero points for that item. The score for each participant was the total number of correct responses out of the 38 items. The scores could also be converted to percentages for direct comparison purposes. For the majority of analyses reported in chapter 4, the raw scores were used.

Initially, there seemed to be certain advantages to using this instrument. First, there was evidence of reliability and validity (Allen, Stone, Rhoads, & Murphy, 2004). Second, I had access to the results of prior administrations of the SCI in different settings, this allowed comparisons between content knowledge levels of preservice primary teachers, preservice secondary teachers, and undergraduate students in mathematics, science, and engineering majors. Unfortunately, preliminary analysis of data from semester 1 revealed a danger threatening to compromise the results for the primary preservice teachers.

The percentage scores for both the 3213fa and the 1473 participants were near 21%; this value is below the expected scores of 23.6% for participants who guess on the multiple-choice SCI. (The number of response choices varies from three to six.) To rectify this concern, I searched for a new instrument to measure statistical content knowledge. Andrea Stone, who had provided access to the SCI content knowledge instrument, suggested a new set of instruments designed by Garfield, delMas, Chance, and Ooms as part of the ARTIST project (ARTIST group, 2005). These instruments were collectively called the *ARTIST scales*.

The ARTIST scales consist of eleven individual scales, each addressing a specific category of content within the broader context of statistics education at the undergraduate level. The labels for the eleven categories are: data representation; measures of center; measures of spread; probability; data collection; confidence intervals for one-sample; sampling variability; categorical, bivariate data; quantitative, normal distribution and measure of position; and tests of significance. The ARTIST scales were designed for the participants to access and complete the scales online. As the participants completed the ARTIST scales, their answers were entered into a database designed by the ARTIST scales authors. I was granted permission to use the scales on November 22, 2005 (B. delMas, personal communication, November 22, 2005). This late timing prevented use of this instrument with the primary preservice teachers in semester 1; thus, the statistics content knowledge data for this group from semester 1 came solely from the SCI. Content knowledge for the primary preservice teachers from semester 2 came from the ARTIST scales.

The online protocol designed by the ARTIST scale authors required that each participant be anonymous. Since I wanted to correlate the answers to the ARTIST scales with the attitude instruments for each participant, it was necessary to have the students write their name, and the answers to the questions on the scales, on a form I developed for this activity (Appendix B).

I chose only those scales corresponding to topics recommended in the NCTM standards and presented in most 7-12 curricula. The seven scales (of 11 total) that fit these criteria were data representation, probability, measures of center, measures of spread, data collection, confidence intervals for one-sample, and sampling variability. Analogously, I asked the primary preservice teacher participants to complete four of the scales. In this case, I chose only the scales that corresponded to topics recommended in the NCTM standards and presented in most K-6 curricula. The scales that fit these criteria were data representation, probability, measures of center, and measures of spread.

The ARTIST scales were scored in the same way as the SCI instrument: correct answers from the multiple choice instruments were awarded one point while incorrect answers were awarded zero points. The scores were recorded by item and then a sum was found for each participant for each scale. The sum of groups of ARTIST scales became variables as well. Since the primary teachers only completed four of the scales, a variable was defined for the sum of these four scales, SCsum4. This variable represents the score that would be accrued by the participants if the four scales were combined to generate a single instrument. Such an instrument would be similar to the SCI, that contains questions from a variety of statistical topics just as the four ARTIST scales contain questions about four different statistical topics. A variable similar to SCsum4, defining a sum of the seven designated scales, was defined for the secondary teachers only. This variable was labeled SCsum7. No group other than the preservice secondary teachers has participants with values for this seven scale sum.

Attitudes toward professional development in statistics instrument (Appendix B). As discussed in the introduction to this chapter, I wanted to incorporate a measure of the attitudes the preservice teachers demonstrated toward professional development. To maintain the focus of the study, these attitudes were measured only for statistics-related professional development. As there was not already an appropriate instrument available in the literature, I designed my own. This instrument is referred to as the *Attitudes Toward Professional Development in Statistics* (ATPDS) instrument.

The instrument first describes the hypothetical existence of short workshops of about three weeks, conducted in the summer months, designed to increase teacher ability to effectively teach statistics to their students. This context statement is followed by a set of six questions. Each question askes the participant to mark the level of interest they would have in such workshops if certain conditions existed.

The conditions change across the six items. The first item, "Do you believe that you would enjoy such a workshop?" is the benchmark question designed to get a general opinion of such professional development from the participants. The third question, "Do you believe that such a program would help you in your classroom teaching?" provides a contrast to item one. Observation of the results of items one and three allowes for comparison between the participants' personal feelings about the workshops and their view of what benefit the workshops might provide. Questions two, four, five, and six provide conditions upon the attendance of the workshop. For each item, the participant could choose from the following levels: "very much", "somewhat", "I am not sure", and "probably not".

For the analyses, item responses were quantified: the answer "very much" was given a value 4, the answer "somewhat" was given the value 3, the answer "I am not sure" was given the value 2, and the answer "probably not" was given the value 1. These values were summed to provide a total for the instrument. In the case of this particular instrument with four of the six items addressing variations of the same answer, individual item responses were very important.

Preservice teachers, by definition, have not had opportunities to participate in any form of continuing professional development beyond their collegiate coursework. Thus, this instrument measured *anticipated* beliefs. Regardless of the design of the data collection process for these beliefs, it would have been possible that the beliefs being measured would change for some of the teachers once they spent a few years teaching full time.

Organization of the Data for Quantitative Analyses

I generated Excel™ spreadsheets to organize the data from the surveys. A separate spreadsheet was generated for each instrument (ATS, CSSE, SELS, SCI, ARTIST, and ATPDS) for each participating group (1473, 3213, 432). Each participant was assigned a row on the appropriate spreadsheets. The columns of the spreadsheets designated items from the instrument each spreadsheet represented.

Each entry on each spreadsheet represented a numerical equivalent of the answer the participant provided on the survey. Since each of these instruments were designed to allow for a measure based on the sum of the item values, totals were calculated per participant via formulas generated within the Excel™ document. For the knowledge instruments, the values were 0 for an incorrect response and 1 for a correct response. These worksheets were too large to be included in the appendixes.

Once the total score for each participant for each instrument was calculated, a master spreadsheet was formed. This spreadsheet contained every participant for the entire study. Qualitative variables identified the group to which each participant belonged. These variables were defined in Table 3.7. Table 3.13 provides an example of the use of the qualitative variables in the master spreadsheet.

Each group (1473, 3213, 432) became a data set. Subgroups of these three groups were also defined into separate data sets. The 1473 data set was separated into 1473ed (primary education majors) and 1473not (all participants not primary education majors). The 432n4 data set (those who completed the ATPDS instrument) was defined as a subgroup of the 432n6 data set. The 3213 participants were separated into 3213fa (semester 1) and 3213sp (semester 2). The 3213sp21 (3213sp with a participant removed who did not complete the ATPDS instrument) and 3213sp19 (3213sp21 with two participants removed who did not complete all of the four ARTIST scales) data sets were defined as subgroups of the 3213sp data set.

Table 3.13

CODE	Course	Sem	Education Major	COLLSTAT
#1	1	0	0	0
#2	1	0	1	0
#3	2	0	0	0
#4	2	1	0	0
#5	3	0	0	1

In Table 3.13, participant #1 is in the data set 1473ed (course = 1 means 1473, sem = 0 means primary education). Participant #2 is in the data set 1473not. Participant #3 is in the data set 3213fa (course = 2 means 3213, sem = 0 means fall). Participant #4 is in the data set 3213sp. Participant #5 is in the data set 432n6. Participant #5 has had a semester course in statistics (COLLSTAT = 1).

The master spreadsheet contained every variable value defined. The individual item responses were not included on the master spreadsheet. Only the calculated totals from the individual spreadsheets were transferred to the master spreadsheet. The exception to this protocol was for the ATPDS scores. For this instrument, the individual item responses were included in the master spreadsheet as well as a variable representing the sum of the items.

Once the master spreadsheet was completed, every numerical value needed for the quantitative analyses was in this document. The master spreadsheet was entered into SPSS™ as a data file. During this transfer, all names were replaced by participant numbers. The participant numbers and the participant names were in the Excel™ document to allow for easy tracking of individual results when necessary. Only the participant numbers were included on the SPSS™ files. The master spreadsheet and subsequent SPSS™ data sheet were both too large to include in the appendixes.

Quantitative Analyses Procedures

Once the data were in an SPSS™ file, subsequent statistical procedures were performed using SPSS™. Among the categories of analyses performed on SPSS™ were ANOVA, stepwise linear regression, regression for predetermined models, scatterplots, box plots, Q-Q plots, generation of line graphs, and descriptive statistics (the mean and standard deviation of variables). These statistical procedures were used to complete broader statistical evaluations. These evaluations were:

- tests for normality and outliers,
- tests for differing means across data sets,
- Exploratory Data Analysis on the SCI versus ARTIST scales on the n = 6 data set 432n6,
- investigation of prediction models using regression analysis to identify

potential correlations between variables within data sets.

Subsequent paragraphs elaborate on each of these evaluations.

There were concerns that had to be addressed before performing the regression analyses. The data sets needed to be close to normally distributed to satisfy the assumptions made when performing regression analyses (Mendenhall & Sincich, 2003, p. 164). If outliers were affecting the data sets, then it was necessary to determine whether the outliers needed (or could) be removed. Multicollinearity was also a concern. If two or more of the independent variables in a regression model were moderately or highly correlated, then there could be errors in the estimates of the coefficients or the results of the regression analyses could be confusing or misleading (Mendenhall & Sincich, 2003, pp. 347, 359).

Checks for normality and outliers. Box plots were used to help identify potential outliers. SPSS™ was used since it indicates potential outliers on its box plot outputs. A lack of potential outliers in a box plot would indicate the possibility that the data are not normally distributed. Q-Q plots were also generated on SPSS™ to investigate normality in the data sets. If the Q-Q normal plots were not close to linear, then the variable was not close to normal in the data set analyzed. During the analyses, there were outliers that were identified to be removed.

Identification of differences in variable means across participant sets. I wanted to know if there were significant differences between participant groups in the means of any variables. While such differences would not have infered differences in more broadly defined populations, it would have provided suggestive evidence that a more focused study applied to a more broadly defined population would be warranted.

ANOVA were generated to investigate differences in means. Because the instruments included in the survey varied across different participant groups, the variable list was not consistent across all of these groups. Hence the ANOVA were generated one-at-a-time for each selected pair of participant groups. Only variables common to both participant sets were considered during each ANOVA. The pairs were selected based on the potential for different means to provide suggestive meanings in relation to the research question. For example, 1473ed and 1473not were considered for pairing because a significant difference in the means of (say) the SCI scores for these two groups would have suggested that the students planning to enter preservice primary education have significantly higher (or lower) statistics knowledge levels than does the general population of students who take the introductory general mathematics course at OU. Similar rationales applied to each of the pairings. The explanation for each appears as the data results are discussed in chapter 4. The participant group pairings and the variables considered are identified in Table 3.14.

Table 3.14

Participant Sets	Variables Considered
1473ed and 1473not	ATSfield, ATScourse, ATSsum, CSSE, SELS, SCI
1473 and 3213fa	ATSfield, ATScourse, ATSsum, CSSE, SELS, SCI
1473ed and 3213fa	ATSfield, ATScourse, ATSsum, CSSE, SELS, SCI
3213fa and 3213sp	ATSfield, ATScourse, ATSsum, CSSE, SELS
1473 and 3213sp	ATSfield, ATScourse, ATSsum, CSSE, SELS
1473ed and 3213	ATSfield, ATScourse, ATSsum, CSSE, SELS, SCI
1473 and 3213	ATSfield, ATScourse, ATSsum, CSSE, SELS
3213sp and 432	ATSfield, ATScourse, ATSsum, CSSE, SELS, SCgraphs, SCcenter, SCspread, SCprob, SCsum4

Once the ANOVA were generated, the reported p-values and the means were organized into tables that were easier to read than the original outputs. The ANOVA results were then considered. The significance level was fixed at $\alpha = .1$. The rationale for such a high alpha value was two-fold. First, I was searching for suggestive rather than conclusive results. Second, there were a few participant groups that had relatively small n-values (432n6, n = 6; 1473ed, n = 12).

The p-values from the ANOVA allowed a determination as to whether variable means were significantly different across any two groups. If the p-value was less than the fixed significance level, then the means of the variable were deemed significantly different between the two groups. (i.e. I rejected the null-hypothesis that the means were the same.)

Descriptive statistics were generated on SPSS™. These outputs were abridged into more easily read tables containing only the means for each variable for each participant group. Whenever a variable had significantly different means for two participant groups, I referred to the tables of the means to determine which of the participant groups had the higher mean for that variable.

Checks for Differences in the Two Statistics Knowledge Instruments Utilized. Due to the fact that the 432n6 participants completed both SCI and ARTIST scales instruments, there was opportunity to investigate any correlations that may exist between the two knowledge level instruments (for this population). A disadvantage to the 432n6 data was the small n of 6 for this data set. To compensate for this disadvantage, exploratory data analysis was relied upon more than inferential statistics (Tukey, 1977).

I used the small n of 6 to my advantage. A multiple line graph was generated that presented both SCI and ARTIST scores. This graph used each participant as a node on each line. This graph was easily read due to the small n of 6. The nodes could be compared one-by-one to look for consistency between the different knowledge instruments.

Comparisons of Statistics Knowledge Levels of Preservice Teachers and Other Undergraduate Student Groups. Through the project under which the SCI was developed, the instrument was administered to a broad range of science, mathematics, and engineering students at the University of Oklahoma (Allen, 2006; Stone, 2006). The sample set included upper division engineering students and first year graduate students among the participants. I wanted to know the answer to the question, "How did the preservice teachers compare to the body of participants for the SCI?" The 3213sp data set has no SCI scores, so that set was not included in the analysis. The 3213fa and 432n6 data sets were considered. A table was generated comparing the means of these three participant groups (3213fa, 432, and original SCI participants). The analysis for this data was comparison of means without ANOVA. The small n = 6 for the 432 data made direct comparison more meaningful.

Comparisons Between Preservice Primary and Preservice Secondary Teachers. The 3213sp and 432 participant groups were the only two groups to complete the ATPDS instrument. Since these two groups provided an interesting contrast (primary level and secondary level), I determined to compare these two groups across all variables.

It was expected that the 432 students would have higher statistics knowledge levels due to the fact that they have taken several more mathematics courses and have focused the majority of their undergraduate work to mathematics. It was not hypothesized that the attitude measures would necessarily have similar results. Preservice primary teachers may recognize the statistics concepts that they should know and believe that they do know those statistics topics. Hence a 3213sp participant could have outscored a 432 participant in the current self-efficacy to use statistics

(CSSE) measure. At the same time, such a 3213sp participant could have scored much lower than a 432 participant on any of the ARTIST scales.

Descriptive statistics were generated on SPSS™. These outputs were abridged into more easily read tables containing only the means for each variable for the 3213sp and 432 participant groups. ANOVA were generated for each of the common variables between these two groups. The 3213sp group did not have legitimate means for the ARTIST scales scores (the SCgraphs, SCcenter, SCspread, and SC prob variables). This was because of the two 3213sp participants who failed to complete the ARTIST scales. Hence all of the ARTIST scale variables were compared separately using the 3213n19 and the 432 participant sets. A table of means was generated. This table included the total number of possible correct items for each ARTIST scale, the mean number of correct responses per group, and the percentage of correct responses per group.

Investigation of Regression Models to Predict Variables of Interest. Regression models were generated to predict statistics knowledge levels and interest in CPD in statistics levels. Regression models were also generated to predict the SELS scores and the ATSfield scores for particular participant groups. The latter regression analyses were performed because of the ANOVA results.

The models for predicting statistics knowledge levels were generated using data from the 3213fa, 3213sp, and 432 participant groups. These were the participants from the study who represent the types of students PMTEs primarily influence. If there are preservice teacher characteristics that correlate with knowledge level, those relationships could be used by PMTEs to improve preservice teacher preparation in general. These analyses helped to satisfy some of the residual objectives of the study (see Chapter 1).

The models for predicting interest in CPD in statistics levels were generated using data from the 3213sp participant group. The 432 and 3213sp participant groups were the only groups to complete the ATPDS instrument. I did not use the 432 group since it had an n = 6 which was too small to generate meaningful regression models.

I was particularly interested in searching for potential correlations between variables based on particular results from the ANOVA results. The ATSfield means were significantly different between both the 3213sp and 1473 groups and the 3213sp and 1473ed groups. The SELS means were significantly different between the 3213fa and both the 1473 and 1473ed groups. The SELS means were also significantly different between 3213all and both the 1473 and 1473ed groups.

The ATSfield means were significantly higher for the 3213sp participants than for both the 1473 and 1473ed groups. The ATSfield means were not significantly higher for the 3213fa participants than for either the 1473 or 1473ed groups. I speculated that there was potential for instructor influence on such means. The 3213sp and 3213fa sections were taught by two different instructors. In the case that such influences exist, I determined to search for other variables that might correlate with ATSfield means. Knowledge of the variables that correspond to ATSfield could help PMTEs shape their course and program decision-making. This would be particularly valuable if future research indicated that instructor behavior does indeed affect ATSfield levels (see Chapter 4).

The SELS means were significantly higher for the 1473ed participants than for the 3213all group. This indicated that the measure of self-efficacy to learn statistics was much higher for preservice teachers in the introductory general mathematics course than for all of the participants (from both semesters) from the second mathematics content-for-teaching course. In the case that such results have a cause, I determined to investigate other variables that might correlate with SELS. As with the ATSfield variable, knowledge of the variables that correspond to SELS could help PMTEs shape their course and program decision-making. In fact, these investigations became more relevant by the end of the study. Regression analyses determined that SELS scores were moderately correlated with preservice primary teacher attitudes toward CPD in statistics (see Chapter 4).

I used a general procedure for developing each regression model. This procedure followed these nine steps:

1. I listed the variables that represented each characteristic I wished to investigate. Each of these variables became a dependent variable for a regression model.
2. For each dependent variable, I listed every potential variable that could be used as an independent variable in the model.
3. I restricted the list of variables considered for each independent variable list. Stepwise regression was to be performed for each dependent variable. A reduced list of independent variables helped to limit the number of t-tests performed in the analyses.

This was important because it reduced the probability of making one or more Type-I or Type-II errors during the stepwise regression (Mendenhall & Sincich, 2003, p. 325).

4. Stepwise linear regression was performed using SPSS™. The outputs for these SPSS™ analyses could then be observed to determine if the model was viable. The model was determined viable if it had statistical significance and if the independent variable coefficients each had standard errors that were statistically significant.

5. If the model was viable, then the coefficient of determination was investigated to see if the model provided any evidence of correlation between the independent variable(s) and the dependent variable.

6. Once the stepwise regression provided an indication of which, if any, of the independent variables might be predictors of the dependent variable, then those independent variables were put into scatterplots with the dependent variable. These scatterplots were observed to determine if a quadratic pattern appeared possible. Scatterplots of the unstandardized residual vs. unstandardized predicted were also observed for the same purpose. Once both of the scatterplots had been observed, a determination was made about whether to investigate a quadratic regression model.

7. If a potential quadratic pattern appeared in any scatterplot, then a quadratic model was developed for the variables used in the respective linear model. To develop such a model, it was necessary to create a "squared" variable by squaring each data point for the respective independent variable from the linear

model. The new variable was added to the initial linear model to make a quadratic model. So the original model was a nested model of the new quadratic model (Mendenhall & Sincich, 2003, p. 231). A regression analysis was used on the original independent variable and the new variable. SPSS™ was then used to generate the outputs for the quadratic model.

8. The values in the SPSS™ outputs were observed and steps 4 and 5 from this process were repeated for the quadratic model.

9. If the quadratic model satisfied the requirements from steps 4 and 5, then the coefficient of determination of the linear model was compared to that of the quadratic model. These values were considered, along with the number of variables involved in both the linear and the quadratic models, to determine which of the models was more parsimonious. That model was kept as the predictor model.

Once each model was generated, I evaluated the applicability of the model with respect to the variables involved and the preservice teacher characteristics these variables represented.

Details for the generation of the models to predict knowledge levels. To search for predictors of knowledge levels, a basic pattern of investigation was followed. The only variables used as independent variables were those representing affective factors (ATSfield, ATScourse, ATSsum, CSSE, and SELS) and the variable COLLSTAT. These variables were entered into the stepwise regression for step 4 of the general procedure. The dependent variable SCI was predicted using the data from the participant groups 3213fa and 3213fa with 432. The dependent variable SCsum4 was predicted using the data from the participant groups 3213sp19 and 3213sp19 with 432. The group 3213fa represented all preservice primary teachers in mathematics content-for-teaching courses who completed the SCI instrument. The 3213sp19 with 432 set represented all preservice teachers in mathematics content-for-teaching courses who completed the SCI instrument The group 3213sp19 represented all preservice primary teachers who completed the ARTIST scales. The 3213sp19 with 432 set represented all preservice teachers who completed the ARTIST scales.

Details for the generation of the models to predict interest in CPD in statistics. The independent variables used for the interest in CPD in statistics predictor models were the same variables used for the knowledge level predictor models. To make a prediction of participant interest in CPD in statistics, it was necessary to choose a particular measure for the dependent variable. There were six measured items on the ATPDS instrument. Each item provided a potential variable that measured attitudes toward CPD in statistics. A model could predict any one of these variables. To help with the determination of which measures to consider, a table was created to observe the means of each item in the ATPDS instrument.

The Pearson-Product correlation coefficient was used to determine if any of the six items were correlated. Several were, so no analyses were performed with more than one of the variables at a time. This reduced the possibility of multicollinearity in the models generated (Mendenhall & Sincich, 2003, p. 347). The WKen and WKhp variables were each predicted in separate models. The other variables were used only for qualitative analyses. They provided values that could be investigated per participant. The models were generated using the data from the 3213sp19 participant group.

Details for the generation of the models to predict SELS scores. Models to predict SELS scores were generated for the following groups of participants: 3213fa, 3213fa with 1473ed, 3213sp with 1473ed, and 3213fa with both 3213sp and 1473ed. To predict the SELS scores for the 3213fa and the 3213fa with 1473ed groups, the independent variables used were ATSfield, ATScourse, ATSsum, CSSE, SCI, and COLLSTAT. For the other groups of participants, the SCI variable was removed from the list since the 3213sp group data was involved, and that set had no data for the SCI variable. The groups used were chosen because they represent all of the preservice primary teacher participants in the study.

Details for the generation of the models to predict ATSfield scores. Models to predict ATSfield scores were generated for the 3213sp participant group. To predict the ATSfield scores, the independent variables used were, ATScourse, CSSE, SELS, and COLLSTAT. The 3213sp group was the only one used because that was the one participant group that had ATSfield scores significantly different from the mean scores from the introductory general mathematics course.

Multicollinearity issues. ATSfield and ATScourse have collinearity problems with ATSsum, since ATSsum is the sum of the other two variables. All three variables were left in the stepwise regression since it was unclear which, if any, of the three variables would provide the stronger predictability of the knowledge variable. If both

ATSfield and ATSsum or both ATScourse and ATSsum were chosen in a model, one would need to be eliminated due to multicollinearity concerns.

Pedagogical Content Knowledge for Statistics: Grading Student Work

There are many aspects of pedagogical content knowledge (PCK) that cannot be measured for preservice teachers. For one, it is difficult to find or create situations in which preservice teachers can be observed responding to authentic situations with age-appropriate K-12 students. There are, however, some aspects of PCK that can be measured in preservice teachers. I designed grading projects (Appendix C) to gather data, in qualitative form, about the preservice teachers' ability to grade statistics work from age-appropriate students. To this end, I needed to determine the age level to fix for the student work. For this grading project, intended for primary preservice teachers, I set the age level at fourth grade because (1) not all primary level teachers are expected to be able to teach mathematics content at the fifth or sixth grade levels, (2) it is preferable that primary teachers be proficient beyond first or second grade mathematics content, and (3) teachers are not always able to choose the grade level for which they will be offered a job.

Rationale for the grading project based on the literature. One measure of PCK is the ability of teachers to grade work completed by their students (Ball, 1999; Ball & Rowan, 2004; Hill & Ball, 2004; Hill, Schilling & Ball, 2004; Hill, Rowan & Ball, 2005). The ability to grade student work correctly combines both deep understanding of the content and an understanding of what the student was thinking based on what the student wrote in their solution. Carpenter and Lehrer (1999) asserted that teachers will only be able to respond successfully to a variety of alternative strategies from students when they possess an understanding both of the content and of student thinking.

Summary of the grading project. This project involved grade-level appropriate problems for 4th grade students with one or more invented student answers to each problem. Some of the answers were correct, others were incorrect. Some of the correct student answers utilized unorthodox strategies. The incorrect answers had varying degrees of inaccuracy and faulty mathematical and/or statistical reasoning. Each answer was designed to allow me to observe the preservice teacher response to the student answer. I was particularly interested in preservice teacher responses to problems that were worked correctly but utilized methods not involving more common algorithms (Appendix C, problem 3[b]). I was also interested in teacher response to problems that had subtle errors that might not be obvious (Appendix C, problem 2[b]). After each section of student work to be graded, there were follow-up questions, that I also designed. These follow-up questions were for the participants and would not have been part of any 4th grade assessment. The questions were designed to provide insight into the thinking of the preservice teachers during the grading process.

Protocols for the grading project. The primary preservice teachers wrote the answers to the follow-up questions by hand on paper copies that I distributed to them. The instructions clearly indicated that the participants should use a pen of some color other than dark blue or black so that I could easily identify their grading. When the hand-written forms were distributed to the participants, they were instructed that they could spend as many hours as they desired working on the project so long as they completed it within six weeks. However, I encouraged them to complete the tasks in two weeks since they might forget about the project altogether if they planned to wait six weeks to complete it.

Participants were also instructed that they were welcome to conduct the same activities while grading the project that they could use when grading student work if they were teaching full time in the classroom. That is, they could communicate with colleagues, they could verify procedures or content from texts, they could look up information from the internet, or they could take any other accuracy-verifying action. I reminded the participants that accuracy in grading was extremely important.

To help me qualitatively evaluate the participants' grading abilities, I added another instrument to the end of both the primary and secondary projects (Appendix C). Through this instrument, each participant was asked to rank the effort required for them to grade each separate problem. The participant could chose from the choices (1) I knew the answer and how to grade without any effort, (2) I had to think for a while, but did not need to reference any material, (3) I had to make a quick reference to a text or something like that, (4) I had to spend a long time with references to figure it out, but it was rewarding, and (5) I needed a lot of time to reference and had to give up on a satisfactory review of the material due to being short on time. Using this instrument, I was able to differentiate levels of knowledge and levels of motivation across the participants in the grading project.

Expected use of the grading project and rationale for non-statistical content. Data gathered for statistics content grading provided PCK measures that could be compared to other factors measured in the integrated survey. I realized that there was another set of comparisons that could be gained with a well constructed grading project. I mixed non-statistical content items into the surveys. This strategy allowed for a comparison between teacher grading of statistical content and their grading of other mathematical content. The goal was to see if relationships between statistics grading skills and other content grading skills correlated with any of the other factors measured.

Statistics content in the grading project. The grading project involved four items, two of which were two data representation problems. The first, which involved a frequency table, was titled "data representation" so that I could see if the participants would recognize the type of representation presented without a name provided. The other data representation problem was a stem-and-leaf plot. In addition to the data representation problems, there was one basic probability problem and one set of questions involving the mean, median, mode, and outliers of a single sample set.

The frequency table contained nothing unusual or unorthodox. It was not identified by type of table, but the participants could look in any 4^{th} grade text to find this type of table. The stem-and-leaf plot was designed to focus on a very specific type of error. I created the plot with stems of various weights. The 3-stem had only three leaves while the 5-stem had seven leaves. In the instrument, I provided a student response to this question that would have been correct if the stems had equal weight. I wanted to see how the participants graded this part of the question. I also specifically questioned the participants in the follow-up question regarding this issue.

The probability problem was also straightforward. However, for this problem, it was necessary to consider possible combinations of spins on a spinner to determine the correct answer. While the problem was straightforward in the sense of what it asked of the K-12 student, the answers I determined to write by hand were not straightforward. I designed one student's answer to have poor logic. The other student's answer started with the correct logic but that student failed to finish the problem with the necessary combinations. I wanted to see if the participants were able to notice the obvious error of the first student one and the more subtle error of the latter student.

The fourth problem asked the student to find the mean, median, and mode for a set of five small numbers. The problem then asked the student to add a sixth number, that was specified and was an outlier for the given set. As it was an outlier, it affected the mean more than it affected the median. I had the student respond that the new value affected the median more because it caused that student to work the problem in an entirely different way. I was particularly interested in observing the types of responses the participants would provide when they graded this last part. I was also interested in what types of explanations the participants would provide for the follow-up questions.

Non-statistical mathematics in the grading project. To provide contrast to the statistics grading, I included a division problem at fourth grade level, a prime and composite numbers problem, and a finding number patterns problem. The division problem involved dividing a three digit integer by a two digit integer, a typical fourth grade division problem. This problem was worked out by two different methods simulating two different 4^{th} grade level student's strategies. The goal of this problem was to see if the preservice teacher participants would recognize that the student who used the algorithm made an interpretational error while the student who used a more unorthodox method based on the concepts involved was able to correctly identify the solutions.

The prime and composite numbers problem was designed to probe for teacher ability/willingness to probe for unusual possibilities. Two of the numbers provided were a small prime number and a small composite number. A third number was a large composite number that may appear to be a prime as a first impression. This was the number 91. Since 91 equals 7 times 13, and since it has no other factors, it is easy to fail to identify any factors for 91. The key to this situation was that the preservice teacher was acting as a grader. Errors in grading can happen, but they can have serious consequences for K-12 student confidence and attitudes toward a subject. I was curious to see how seriously the participants considered their grading activities.

The number patterns problem provided three different patterns. Two of the patterns could be explained with somewhat obvious extensions, one being arithmetic and the other geometric. The K-12 student work provided explained the patterns using these somewhat obvious extensions. The third pattern was designed so that there was more than one possible way to extend the pattern. I intentionally provided handwritten

4th grade student work that identified an unusual but correct extension of the pattern. A more obvious pattern existed, but the instructions were intentionally vague about the type of pattern to be found. This problem would provide insight into the preservice teacher participant's ability to recognize that more than one pattern can exist. It would also indicate the depth on knowledge the participant possessed concerning number patterns.

Qualitative Analyses of the Grading Projects

I recorded the details of each participant's work grading each problem. I also recorded the answers the participants provided for the follow-up questions. For each problem, I compared the work provided of the two participants. I also considered the work of each participant with respect to mathematical accuracy. To provide context, I created summaries of the quantitative results for each of the participants.

The recorded details that provided interesting information are reported in the results section. I determined which recorded details should be reported by a set of rules. First, some problems were intentionally designed with the expectation that a participant might not be able to grade it correctly. For each case that a participant correctly graded such a problem, the results are reported. Second, all incorrectly graded problems are reported. Third, follow-up question answers are reported for each case that a participant made unexpected comments.

Summary

Are there indicators from preservice teacher attitudes toward, and knowledge of, statistics that might assist PMTE efforts to increase the possibility that these preservice teachers will pursue CPD in statistics? To gain a more complete answer to this question, I determined to look at potential solutions at two different levels. First, it is valuable to know those types of attitudes or knowledge that correspond to preservice teacher interest in CPD. Second, it is valuable to know if any of these types of attitudes and knowledge tend to be correlated in preservice teachers. If such correlations exist, then it might be possible to change behavior typically associated with one construct, such as self-efficacy to learn statistics, by improving another construct, such as teacher content knowledge.

There are many studies that provide insight into parts of this question. For example, it has been shown that general attitudes toward statistics and statistical performance levels have little or no correlation. But such studies have been suggested to be inconclusive due to the generality of the attitudes measures used. Few studies have measured preservice teacher attitudes toward learning statistics. By measuring preservice teacher attitudes in three different ways and preservice teacher statistical knowledge in two different ways, including a qualitative measure, this study provided much richer detail specifically concerning the research question than any previous individual study. Due to the unique questioning of preservice teacher attitudes toward CPD, this study reveals preservice teacher attitudes that have heretofore not been measured.

Chapter 4 Results

In this chapter I present nine separate sets of data analyses. Each of these sets of data analyses are presented in a section. Table 4.1 identifies each of these sections.

Table 4.1

Part	Title
1	Tests for Outliers and Normality; Reports on Data Set Observations
2	Tests for Differing Means Across the Groups 1473, 1473ed, 3213fa, and 3213sp
3	Tests for Different Results Between SCI and ARTIST Scales for the 432 Participants
4	Preservice Teacher Scores on the SCI Compared to the Original SCI Study Participants
5	Comparison of All the Variable Means Between the 3213sp and the 432 Participants
6	Searching for a Model to Predict Statistics Knowledge Levels
7	Searching for a Model to Predict SELS and ATSfield Scores
8	Searching for a Model to Predict Interest in CPD in Statistics
9	Qualitative Analysis of the Grading Projects

Analyses Part 1: Tests for Outliers and Normality; Reports on Data Set Observations

Before analyzing the data, I checked the data for outliers and influential points. It wanted to know how normal the data sets were. For non-normal data sets, it was important to know to what extent normality was not met and to report on this.

Some important features of the data could be identified by observation of the raw data. For example, ATSsum was calculated by adding the ATSfield and

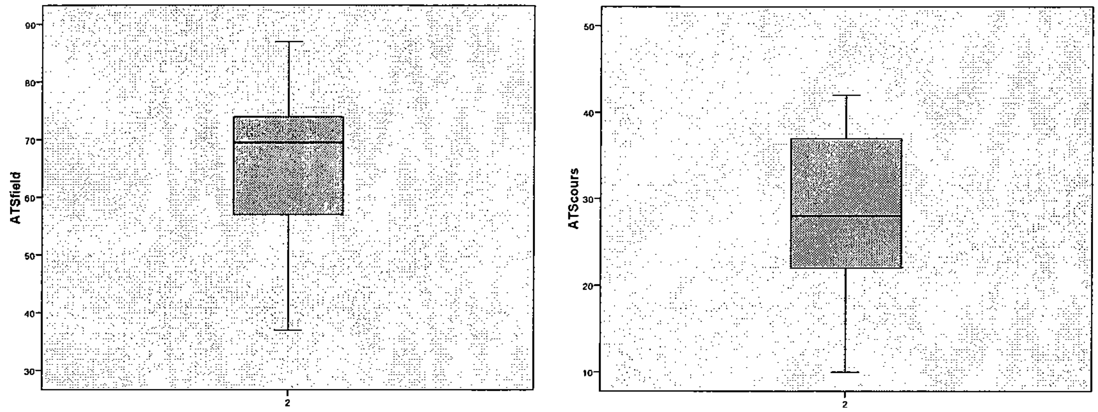
ATScourse values. So there is collinearity between ATSsum and both ATSfield and ATScourse during any regression analyses that is reported in subsequent sections. For the 3213sp data set, participants #17 and #18 did not answer any problems on any of the four ARTIST scales (Appendix D). Participant #1 did not answer any problems on two of the four ARTIST scales. Hence these participants were eliminated from any analyses involving the ARTIST scales (SCgraphs, SCcenter, SCspread, SCprob, and SCsum4). Also, participant #18 did not complete the Interest in CPD in Statistics (ATPDS) instrument. However, for all analyses that did not involve an instrument of concern, the results of the respective participants were included in the analyses. This generated the need for three different 3213sp data sets:

- the data from the original 3213sp (n=22) participants,
- the reduced 3213sp21 which does not contain participant #18, and that is used when Interest in CPD in Statistics results (any variable with WK involved) are considered, and
- the reduced 3213sp19 that does not contain participants #1, #17, and #18 and that is used when the knowledge levels (results from the ARTIST scales) are considered.

SPSS™ generated box plots for each original data set to inspect for outliers. Box plots were generated for each variable in each of the three original data sets. The primary purpose of these box plots was to investigate potential outliers. A secondary result of the box plots appears in the discussion of normality in the next subsection. The box plots for ATSfield and ATS course for the data set 3213fa are presented in figure 4.2. All other box plots are in appendix E.

Figure 4.2 SPSS™ *Generated Box Plots for the 3213fa Data Set*

ATSfield *ATScourse*



Details from the 3213fa box plots. From the six box plots for the 3213fa data set, it appeared that there were no outliers in the data sets for the 3213fa data. This claim was supported by investigation of the specific data entries for each participant. For example, there were no participants who skipped entire sections of any instruments. Also, it appeared that no participants randomly marked the same value for every answer for any of the instruments.

Details from the 1473 box plots. In the box plots for the 1473 data set (see Appendix E), four potential outliers were identified. The ATSfield score for participant #3 may have been an outlier. This caused the ATSsum score to also be skewed for this participant. The SELS scores for participants #6, #19, and #32 may have been outliers. Upon review of the item responses for each of these four cases, it appeared that the scores were not due to arbitrary marking of item responses (see Appendix D). Hence all four data values were kept in the data set.

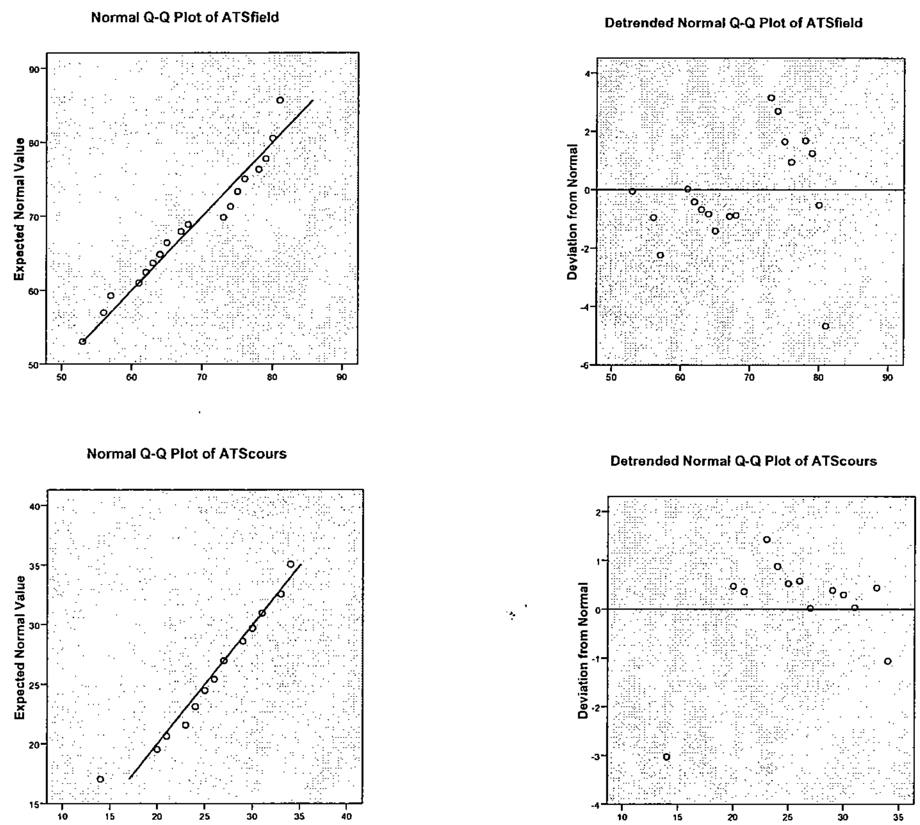
Details from the 3213sp box plots. The zero scores for participants #17 and #18 appeared in the box plots for SCgraphs and for SCsum. The zero scores for SCcenter, SCspread, and SCprob were graphed within the "whisker" (Tukey, 1977, p. 40) in the box plot, rather than as potential outliers. The zero scores for participant #1 on SCspread and SCprob were also graphed within the "whisker".

The tables indicate potential outliers for the ATScourse and SELS scores for participants #4 and #21. Based on inspection of the individual item responses for these two participants on each of the two instruments, the scores appeared to be not due to arbitrary marking of item responses (see Appendix D). There was variation in the answers across the items. All scores for participants #4 and #21 were kept in the data analyses.

The common potential outliers, from the same participants, for both ATScourse and SELS indicate a potential situation. It is possible that the two variables ATScourse and SELS are correlated. If this is the case, then low scores on one of these would predict low scores on the other instrument as well.

Overview of the Q-Q plots to investigate the normality of the sets. Q-Q plots were generated for each variable in each of the three original data sets. The primary purpose of these Q-Q plots was to investigate the normality of the data for each variable. The Q-Q plots for ATSfield and ATScourse for the data set 3213sp are presented in figure 4.3. The remainder of the box plots are in appendix E.

Figure Set 4.3 SPSS™ *Generated Q-Q Plots of Quantitative Variables for the 3213sp Data Set*



Details from the 3213sp Q-Q plots. From the Q-Q plots, ATScourse residuals appeared to be somewhat normal for the 3213sp data. The SELS Q-Q plot appeared to have good linearity except that the line is skewed toward the horizontal by the two points to the far left. The other plots did not indicate high levels of normality. This situation was not surprising since many of the box plots for these variables did not indicate outliers – a common trait of box plots for data sets that may not be normal.

Although high levels of normality were not indicated, none of the variables appeared to be so far off of normal as to indicate a need to transform the variable. Kolmogorov-Smirnov tests of normality could be performed, but this test is sensitive to the slightest departure from normality (Mendenhall & Sincich, 2003, p. 634). For the 3213sp data, no indication of normality was provided by Kolmogorov-Smirnov tests (see Appendix E).

Details from the 3213fa Q-Q plots .From the Q-Q plots, the SELS residuals appeared to be somewhat normal for the 3213fa data. The other plots did not indicate high levels of normality. This was not surprising since many of the box plots for these variables did not indicate outliers – a common trait of box plots for data sets that may not be normal. Although high levels of normality were not indicated, none of the variables appeared to be so far off of normal as to indicate a need to transform the variable.

Details from the 1473 Q-Q plots. From the Q-Q plot, all of the residuals except for SCI appeared to be somewhat normal for the 1473 data. The SCI plot seemed to fail the normality test at the extremes (the high and low values deviate from the linear normality approximation), but was otherwise relatively normal.

Analyses Part 2: Tests for Differing Means Across the Groups 1473, 1473ed, 3213fa, and 3213sp

Descriptive statistics were tabulated on SPSS™ for the variables ATSfield, ATScourse, ATSsum, CSSE, SELS, and SCI in the data sets 1473, 1473ed, 1473not, 3213fa, and 3213sp (see Appendix F). The means of the variables for each of these data sets are presented in table 4.4.

Table 4.4 *Means of the Variables Across Data set Groups*

	1473 all (n=81)	1473 elem. Ed. (n=12)	1473 not elem. Ed. (n=69)	3213fa (n=22)	3213sp (n=22)
ATSfield mean	64.30	62.33	64.64	65.23	69.36
ATScourse mean	25.42	25.75	25.36	28.32	26.05
ATSsum mean	89.72	88.08	90.00	93.55	95.41
CSSE mean	41.28	47.42	40.22	40.50	43.55
SELS mean	61.49	66.33	60.65	53.77	58.27
SCI mean	8.10	9.08	7.93	7.86	No data

The p-value significance results of the ANOVA are presented in table 4.5 (see Appendix G for original ANOVA tables from SPSS™). These results reflected p-values for the null hypothesis that the means were not the same. Thus any values on the table that were less than the determined significance level of $\alpha = .1$ indicated the means between the given variable were different for the two groups of comparison.

Table 4.5 *p-Values for the Attitudes Variables for the Participant Groups 1473, 1473ed, 1473not, 3213fa, and 3213sp*

	ATSfield	ATScours	ATSsum	CSSE	SELS	SCI
1473 Ed vs. 1473 not Ed	.521	.870	.710	.070*	.238	.238
1473 all vs. 3213fa	.747	.127	.366	.808	.048**	.758
1473 Ed vs. 3213fa	.568	.410	.478	.183	.057*	.320
3213fa vs. 3213sp	.242	.320	.729	.531	.396	No Data
1473 all vs. 3213sp	.055*	.715	.134	.489	.390	No Data
1473 Ed vs. 3213sp	.078*	.896	.209	.467	.179	No Data
1473 Ed vs. 3213 all	.215	.560	.285	.274	.071*	No Data
1473 all vs. 3213 all	.165	.211	.132	.777	.071*	No Data

*p<.1, **p<.05

At the $\alpha = .1$ significance level (in fact, at the $\alpha = .05$ level), the 22 elementary education majors who were in their second mathematics content course had lower self-efficacy to learn statistics than the 81 participants from the 1473 course (53.77 vs. 61.49, p=.048). These means were interesting since the mean for self-efficacy to learn statistics (SELS) for the 12 elementary education majors within the 1473 participant list was higher (though not significantly) than the other majors from the list of 1473 participants. From Table 4.15 we can see that the mean for the 1473ed SELS scores was 66.33 while the mean for the 1473not SELS scores was 60.65

At the $\alpha = .1$ significance level, the 22 elementary education majors who were in their second mathematics content course (the 3213fa data) had lower self-efficacy to learn statistics than the 12 preservice primary teacher participants (1473ed data) from the 1473 course. (53.77 vs. 66.33, p=.057). Even though the participants from the 3213sp data set did not have mean SELS scores significantly different from the 1473ed participants (p-value of .179), the 3213sp means were lower than the 1473ed means. (58.28 vs. 66.33). In fact, the SELS scores for the combined 3213fa and 3213sp data sets were significantly lower than the1473ed SELS mean (p=.071).

Note that these results are not longitudinal. None of the 3213fa or 3213sp participants were in the list of 1473 participants. But it is a fact that two different sections of 3213 students did have lower means than the single data set of preservice primary teachers enrolled in 1473.

The 12 preservice primary teachers in 1473 had significantly higher self-efficacy (at α =.1) to perform statistical procedures (CSSE) than did the other majors in the 1473 course (47.42 vs. 40.22, p=.070). Based only on the reported ANOVA results, it is unclear whether such confidence is warranted. A regression analysis indicated that CSSE is not a predictor of SCI scores for the 1473ed data set (Appendix H). Even though CSSE scores may not be useful for predicting SCI scores for the 1473ed data set, an initial investigation indicated that such confidence (high CSSE scores) may be warranted. There was no significant indication that the SCI means were different, but the 1473ed group (whose CSSE scores were significantly greater than the scores of the rest of the course participants) scored 14.5% higher on the SCI (knowledge of statistics) instrument than did the other 69 participants from the 1473 course.

The 3213sp participants had significantly higher ATSfield scores than did both the 1473 (n=81) participants and the 1473 preservice primary teachers (69.36 vs. 64.30, p=.055; 69.36 vs. 62.33, p=.078). This ATSfield result may appear contradictory to the SELS results. The attitudes toward statistics as a field (ATSfield) measure yielded results indicating more positive beliefs for participants who are in their second mathematics content course (3213sp). The self-efficacy to learn statistics (SELS) measure yielded results indicating more positive beliefs for participants who have not began their mathematics content courses (1473ed).

There are some important caveats in the SELS and ATSfield results. First, it is highly possible that SELS and ATSfield do indeed measure completely different affective constructs. Second, the SELS scores appear to be consistent across both groups of 3213 participants. The ATSfield scores are only significantly different between the 1473 and 3213sp participants. This could be due to the results occurring as an anomaly. It could also be due to potential influence on ATSfield scores from the instructor of the 3213 course. Recall that 3213fa and 3213sp were taught by different instructors.

It was interesting that no pairs of participant groups shared more than one variable that had significantly different means across the two participant groups. Several pairs of participant groups shared no variables that had significantly different means. The pairs of participant groups that had significantly different means for some variable are provided in Table 4.6.

Table 4.6 *Participant Sets with Significantly Different Means with the Variables Listed*

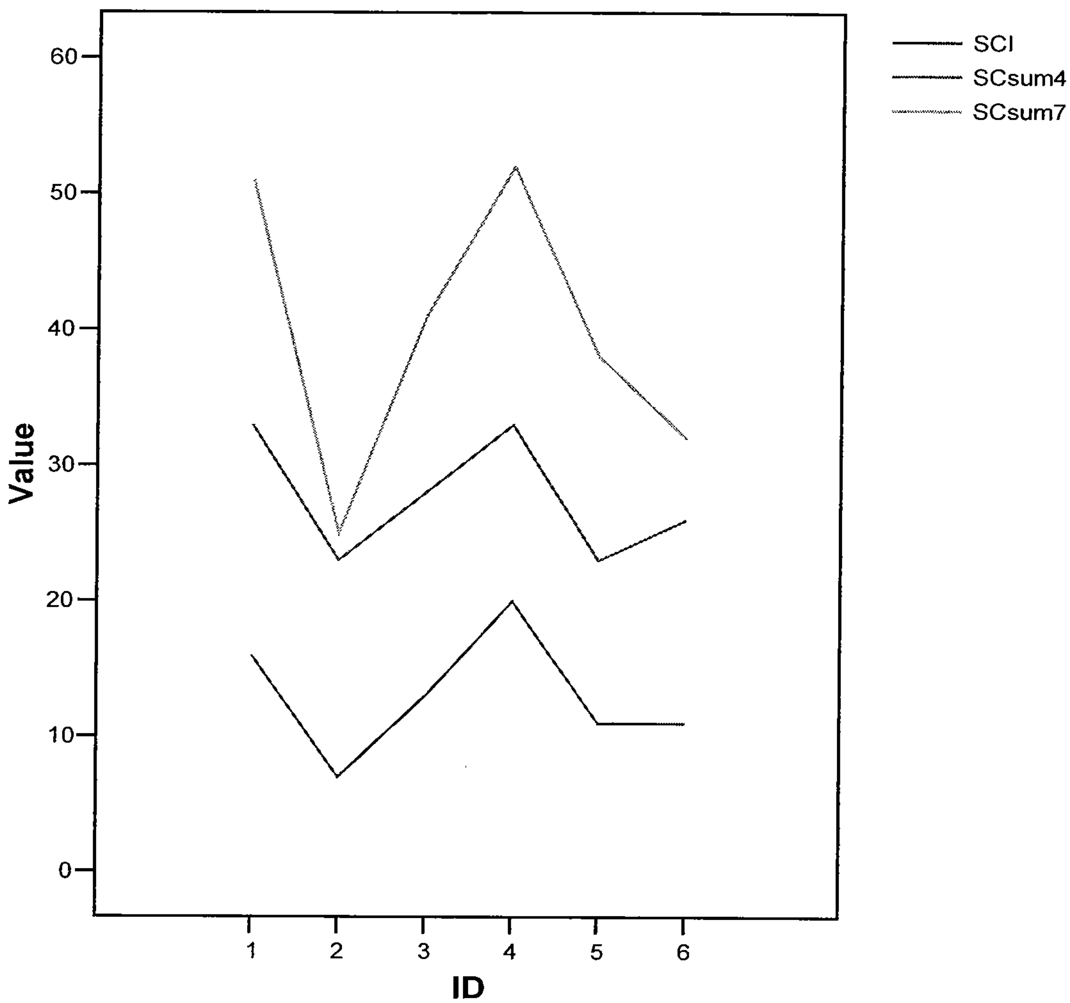
	1473ed	1473not	1473	3213fa	3213sp	3213all
1473ed	x	CSSE	None	SELS	ATSfield*	SELS
1473not	CSSE	x	None	None	None	None
1473	None	None	x	SELS	ATSfield	SELS
3213fa	SELS	None	SELS	x	None	None
3213sp	ATSfield*	None	ATSfield	None	x	None
3213all (3213fa and 3213sp combined)	SELS	None	SELS	None	None	x

Analyses Part 3: Tests for Different Results Between SCI and ARTIST Scales for the 432 Participants

Figures 4.7 and 4.8 contain the line graphs generated to compare the SCI and the ARTIST scales for the 432 participants. Figure 4.7 displays a pattern that indicates close correlation between the SCI instrument (SCI) and the sum of the first four ARTIST scales administered (SCsum4). These four scales are the same four scales used to measure the 3213sp participants' statistical knowledge. Other than a vertical shift due to different potential total raw scores on these two instruments, the graphs are close to identical. Only the results for participant #6 created even a moderate change of pattern between the SCI line and the SCsum4 line.

The SCsum7 score does not correlate with either of the other two as well as the other two measures correlate with each other. Although in general the SCsum7 line moves up when the other lines move up, and moves down when the other lines move down, the rate of change is more pronounced for the SCsum7 line. For participant #6, the SCsum7 line moves down while the SCsum4 line moves up.

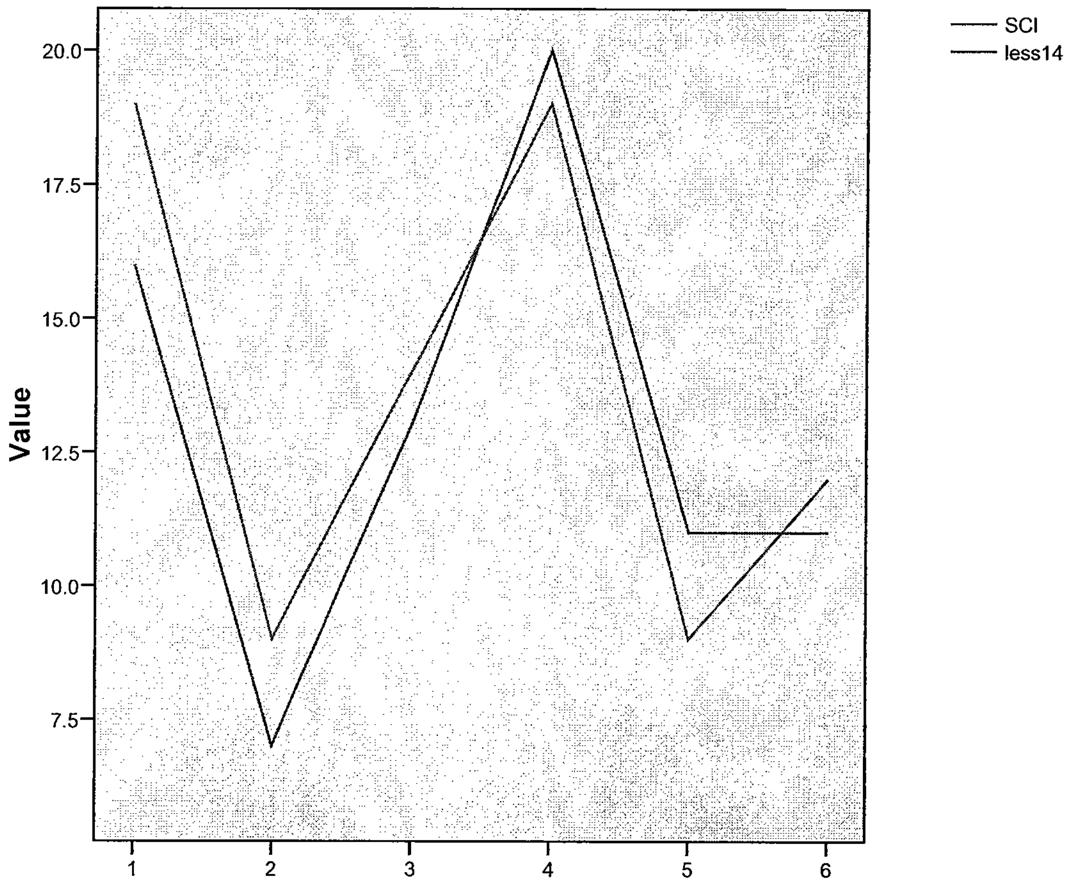
Figure 4.7 *Multiple Line Graphs for Individual 432 Scores on the SCI, SCsum4, and SCsum7 Instruments*



It appears that the SCsum7 measure separated the stronger from the weaker students at a more pronounced level than did the SCsum4 measure. The only deviation from this trend is for participant #6 who outscored participant #5 on the SCsum4 measure but scored lower than participant #5 on the SCsum7 measure. The SCsum4 measure appears to separate the stronger from the weaker students at the same level as does the SCI instrument.

To get a clearer view of the SCI vs. SCsum4 trend for the 432 participants, the vertical shift was decreased. The mean for SCI is 13.00. The mean for SCsum4 is 27.67. Adjusting for the differences in the means between the SCI and the SCsum4 scores yielded the multiple line graphs in Figure 4.8.

Figure 4.8 *Multiple Line Graph for Individual 432 Scores on the SCI, SCsum4, and SCsum7 Instruments*



The results described above concerning the SCI and SCsum4 scores remain unchanged in Figure 4.8. The results are easier to see and the issues surrounding participant #6 are more apparent. A benefit from figure 4.8 that is not apparent in figure 4.6 is this: compared to the means for the 432 participants on both knowledge measures, participants #1, #2, #3, and #6 performed better on the SCsum4 instruments than on the SCI while participants #4 and #5 performed better on the SCI instrument than on the SCsum4 instruments. It is unclear what ramifications these results provide. It is possible that particular participants are stronger in certain statistical topics. The two knowledge instruments may vary in the topics that tend to provide higher overall scores. Such detailed analyses of item responses were left for another study.

Analyses Part 4: Preservice Teacher Scores on the SCI Compared to the Original SCI Study Participants

The raw means for the data sets other than 432 are on Table 4.9. Since the raw means for the 432 data set are not posted in any tables, they are included in the table. The listed percentages in the table for the SCI study results are based on the table in appendix I.

Table 4.9 *SCI mean scores for 432, 3213fa, and SCI study*

Data Set	SCI Mean
432 raw scores	13.0
432 scores	34.2%
3213fa scores	20.7%
Original SCI Study highest scores	52.3%
Original SCI Study lowest scores	45.5%
Original SCI Study weighted scores	48.9%

There were few surprises in these results. The primary level preservice teachers from the fall 3213 course scored the lowest with 20.7% correct. This group had the lowest levels of total mathematics preparation. Few of the participants had full semester courses in statistics. The secondary level preservice teachers from the fall 3213 course scored above the primary teacher but below the original SCI study participants. Each of the six 432 participants had taken a semester length course in statistics. However, only one had taken more than one statistics course. The participants from the original SCI study scored the highest.

Analyses Part 5: Comparison of All the Variable Means Between the 3213sp and the 432 Participants

There were a large number of variables common to both the 3213sp and 432 data sets. Descriptive statistics tables were generated using SPSS™ for these two data sets. The complete tables are in Appendix F. In table 4.10, the means are presented by variable for each of these two data sets.

Table 4.10 *Means for the Common Variables for the 432 and 3213sp Data Sets*

Variable	432 Mean	3213sp Mean
ATSfield	78.50	69.57
ATScours	35.00	26.14
ATSsum	113.50	95.71
CSSE	43.50	42.71
SELS	70.33	57.95
SCgraphs	9.17	6.00
SCcenter	3.50	1.86
SCspread	7.67	4.38
SCprob	7.33	5.33
SCsum4	27.67	17.57
WKen	2.75	2.57
WKfr	3.25	3.05
WKhp	3.50	2.86
WKpd	3.25	3.10
WKin	4.00	3.71
WKls	2.75	1.57
WKtotal	19.50	16.86

These two data sets provide the information for all analyses involving participant interest in CPD in statistics. The secondary teachers indicated higher interest in CPD for statistics than did the primary teachers in all six questions on the Attitudes Toward CPD in Statistics instrument. ANOVA results (Table 4.11) indicate whether the 432 Attitudes Toward CPD in Statistics scores were significantly higher.

The means for current statistics self-efficacy (CSSE) appeared to be very close. Otherwise, the secondary teachers appeared to have scored much higher on every variable. To determine whether the apparent differences were significant, ANOVA were generated on SPSS™ comparing the two data sets across each variable.

The ANOVA table with all the original SPSS™ details is in Appendix G. An abridged version of this ANOVA table is in Table 4.11. For convenient reading, only the p-values indicating the significance are kept for Table 4.11.

Table 4.11 *The ANOVA p-values for Each Variable Common to 3213sp and 432*

ANOVA

	Sig.
ATSfield	.025
ATScours	.001
ATSsum	.004
CSSE	.913
SELS	.096
SCgraphs	.005
SCcenter	.028
SCspread	.026
SCprob	.070
SCsum4	.001

From table 4.11, at the $\alpha = .1$ level, the preservice secondary teachers scored significantly higher than did the preservice primary teachers on every instrument except for CSSE and SELS. At the $\alpha = .1$ (even the $\alpha = .05$) level, the preservice secondary teachers scored significantly higher than did the preservice primary teachers on every instrument except for CSSE, SELS, and SCprob. These significant results occured with only an n of 6 in one of the data sets.

Between the 3213sp and 432 data sets, the only variable that did not indicate at least some level of expectancy that the means are different was CSSE. It is reasonable to expect the 3213sp participants to have CSSE scores not different from the 432 participants. The 3213sp participants do not have reason to expect themselves to need to know the level of statistical material that the 432 participants should expect themselves to know. This could yield confidence scores between the two sets that are similar or even higher for the 3213sp participants.

ARTIST scale values. How do the secondary preservice teachers and the primary preservice teachers compare on the ARTIST scales? The maximum possible score on each of the ARTIST scales is as follows:

SCgraphs	13,
SCcenter	7,
SCspread	14,
SCprob	9,
SCdacoll	9,
SCcnfint	10,
SCsmpvar	15.

In table 4.12, I present the mean raw score and the percentage score for each ARTIST scale. These data are for the 3213sp and 432 data sets.

Table 4.12 *The 3213sp19 and 432 results for the ARTIST scales*

		3213sp19	3213sp19	432n6	432n6
	n	mean	Mean Percentage Correct	mean	Mean Percentage Correct
SCgraphs	13	6.26	48.15%	9.17	70.54%
SCcenter	7	1.74	24.86%	3.50	50.00%
SCspread	14	4.84	34.57%	7.67	54.79%
SCprob	9	5.89	65.44%	7.33	81.44%
SCdacoll	9	x	x	5.17	57.44%
SCcnfint	10	x	x	4.33	43.30%
SCsmpvar	15	x	x	2.67	17.80%
SCsum4	43	18.74	43.58%	27.67	64.35%
SCsum7	77	x	x	39.83	51.73%

The primary preservice teachers scored about 68% of what the secondary preservice teachers scored on data representation (SCgraphs). The ratio for measures of spread was similar, primary preservice teachers scored about 63% of what the secondary preservice teachers scored. Both groups scored the highest percentage on the probability scale. On this scale, the primary preservice teachers scored about 80% of what the secondary preservice teachers scored. The gap between primary and secondary teachers was lowest for probability. The gap between primary and secondary teachers was highest on the scale in which both groups had the lowest percentage correct (out of the first four scales), measures of center. The primary preservice teachers scored about 50% of what the secondary preservice teachers scored on this scale. The primary preservice teachers scored about 68% of what the secondary preservice teachers scored on the cumulative score for the first four scales. Both groups appeared to be much weaker at measures of center and spread than at data representation (tables and graphs) and probability. The secondary preservice teachers scored a higher percentage on data collection than they did on measures of center.

Their mean percentage score for the confidence intervals scale was slightly lower than it was for the measures of center scale. The secondary preservice teachers scored only 17.8% correct for sample variance. The sample variance scale was composed of 15 problems, divided into one problem with 5 choices for solution, five problems with 4 choices for solution, and nine problems with 3 choices for solution. Random guessing would provide for an expected mean percentage of about 30%. Hence the 17.8% mean percentage was below guessing rates.

Analyses Part 6: Search for a Model to Predict Statistics Knowledge Levels

A total of three models were generated to predict SCI scores. Four other models were generated to predict the sum score for the first four ARTIST scales. In table 4.12 are the data sets used, the knowledge measure used, and the table at which each regression analyses is presented. Each summary table (4.14, 4.17, 4.20, 4.21, 4.24, 4.26, and 4.29) is an abridgement of an SPSS™ output. Each SPSS™ output is provided in Appendix I.

Table 4.13 *Summary of the Output Tables for the Knowledge Level Predictor Models*

Location	Data Set	Knowledge Measure	Type of Regression
Table 4.14	3213fa	SCI	stepwise linear
Table 4.17	3213fa and 432	SCI	stepwise linear
Table 4.20	3213fa and 432	SCI	quadratic
Table 4.21	3213sp19	SCsum4	stepwise linear
Table 4.24	3213sp19	SCsum4	quadratic
Table 4.26	3213sp19 and 432	SCsum4	stepwise linear
Table 4.29	3213sp19 and 432	SCsum4	quadratic

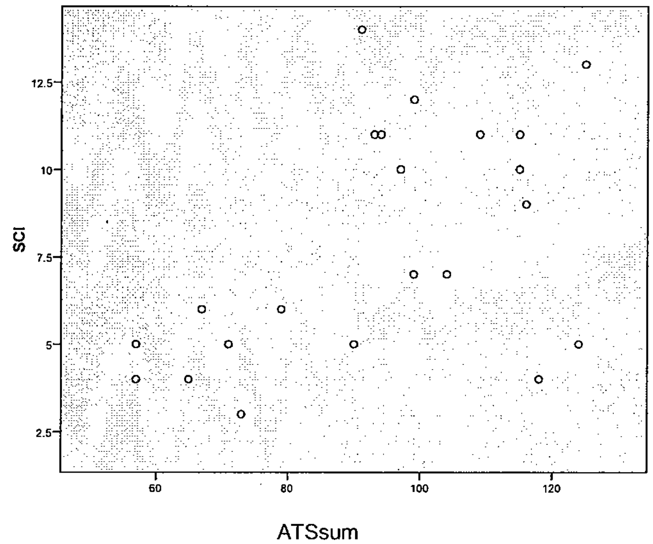
Table 4.14 *Coefficient of Determination, Statistical Significance of the Model, and Standard Error of the Coefficients of the Stepwise Regression to Predict SCI for the 3213fa Participants*

Coefficient of Determination	R^2	.257
Statistical Significance of the Model	p	.016
Standard Error of the Coefficients	Constant (β=.402) ATSsum (β=.080)	.891 .016

The resulting model is E(SCI) = .402 + .080(ATSsum) (see Appendix I). This model accounts for 25.7% of the variability in the 3213fa SCI scores ($R^2 = .257$). Hence ATSsum is a "weak" predictor of SCI scores for 3213fa participants (Mathbits, 2007). The model is statistically significant at the $\alpha = .05$ level ($p = .016$). The constant coefficient attempts to predict the SCI score when ATSsum is zero. ATSsum of zero is outside the range of the ATSsum data for this data set (all of the data sets). This is manifest in the standard error of the constant coefficient, which is .891. Thus the constant coefficient does not have a practical interpretation (Mendenhall & Sincich, 2003).

There is no apparent quadratic pattern in the scatterplot on figure 4.15. The residuals appear random in the scatterplot on figure 4.16. Hence there was no further regression analyses.

Graph 4.15 *Scatterplot of the variables ATSsum vs. SCI for the 3213fa Participants*



Graph 4.16 *Scatterplot of the Unstandardized Residual vs. Unstandardized Predicted for ATSsum Predicting SCI on the 3213fa Participants*

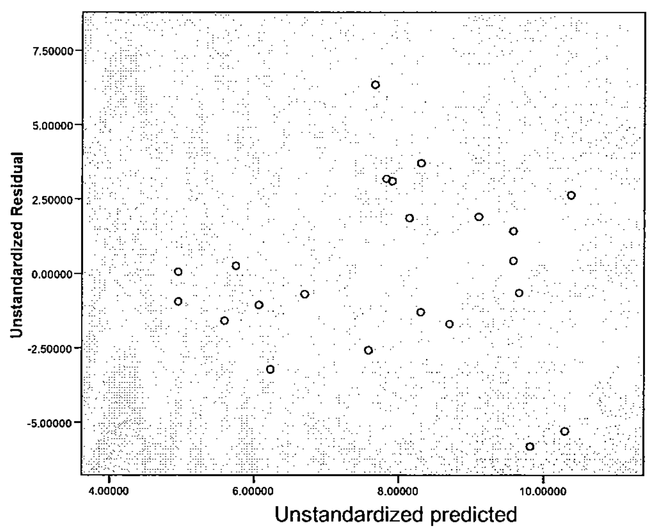


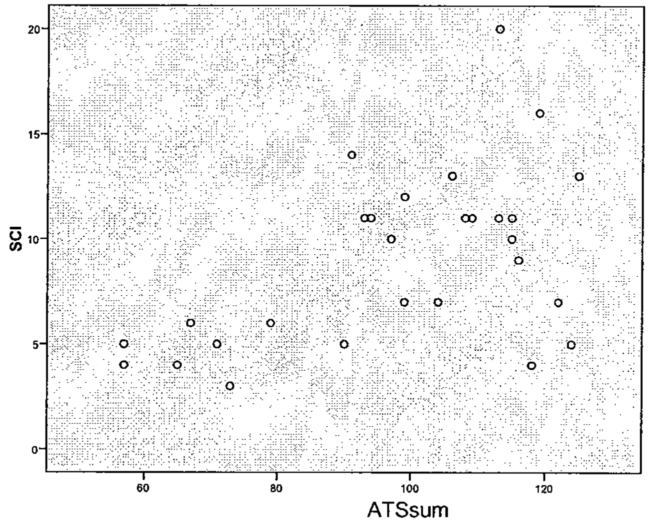
Table 4.17 *Coefficient of Determination, Statistical Significance of the Model, and Standard Error of the Coefficients of the Stepwise Regression to Predict SCI for the Combined 3213fa and 432 Participants*

Coefficient of Determination	R^2	.276
Statistical Significance of the Model	p	.004
Standard Error of the Coefficients	Constant (β= -1.266) ATSsum (β=.105)	.706 .004

The resulting model is E(SCI) = -1.266 + .105(ATSsum) (see Appendix I). This model accounts for 27.6% of the variability in the combined 3213fa and 432 SCI scores (R^2 = .276). When the 432 data is added to the 3213fa data, the model generated provides higher levels of explained variance than does the model developed from the 3213fa data alone. However, the increase is only slight. There is still not a good model for explaining the combined 3213fa and 432 SCI scores if the predictor variables of interest are attitude measures or number of college statistics courses taken. The model was statistically significant at the $\alpha = .01$ level ($p = .004$).

The scatterplot on figure 4.18 appears to have a potential concave up quadratic pattern. It was necessary to check for a potential quadratic regression model. The residuals appear random in the scatterplot on figure 4.19.

Graph 4.18 *Scatterplot of the variables ATSsum vs. SCI for the Combined 3213fa and 432 Participants*



Graph 4.19 *Scatterplot of the Unstandardized Residual vs. Unstandardized Predicted for ATSsum Predicting SCI on the Combined 3213fa and 432 Participants*

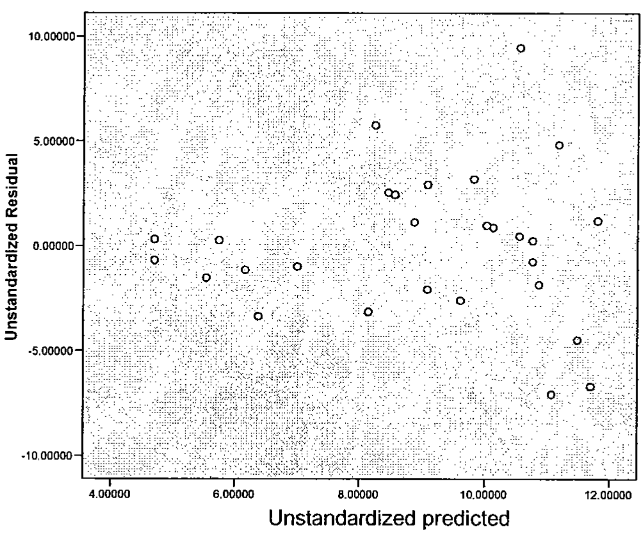


Table 4.20 *Coefficient of Determination, Statistical Significance of the Model, and Standard Error of the Coefficients of the Quadratic Regression to Predict SCI scores for the Combined 3213fa and 432 Participants Using the Variable ATSsum*

Coefficient of Determination	R^2	.333
Statistical Significance of the Model	*p*	.006
Standard Error of the Coefficients	Constant (β= -21.767) ATSsum (β=.578) ATSsumSq (β= -.003)	.142 .087 .155

The resulting model is E(SCI) = -21.767 + .578(ATSsum) - .003(ATSsumSq) (see Appendix I).. This model accounts for 33.3% of the variability (R^2 = .333). Adding the quadratic term increased the explained variance by 5.7%. However, the squared term coefficient is not significant at α = .1. Hence the best model for predicting SCI for the combined 3213fa and 432 participants is the first order model with ATSsum as a predictor variable.

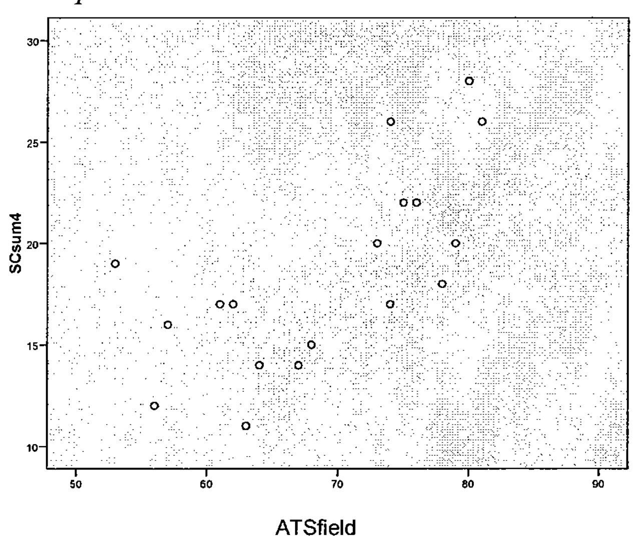
Table 4.21 *Coefficient of Determination, Statistical Significance of the Model, and Standard Error of the Coefficients of the Stepwise Regression to Predict SCsum4 scores for the 3213sp19 Participants*

Coefficient of Determination	R^2	.472
Statistical Significance of the Model	*p*	.001
Standard Error of the Coefficients	Constant (β= -.7.133) ATSfield (β=.374)	.301 .001

The resulting model is E(SCsum4) = -7.133 + .374(ATSfield) (see Appendix I). This model accounts for 47.2% of the variability in the 3213sp19 SCsum4 scores (R^2 = .472). The model is statistically significant at the α = .001 level (p = .001). This model provides evidence that the independent (ATSfield) and dependent (SCsum4) variables are "moderately" correlated for the 3213n19 data set (Mathbits, 2007). This model cannot be used to make meaningful predictions of SCsum4 scores. This is because of a lack of significance in the constant coefficient in the model.

The scatterplot on figure 4.22 appears to have a potential concave up quadratic pattern. It was necessary to check for a potential quadratic regression model. The scatterplot on figure 4.23 also indicates the potential value of a squared term. This conclusion was based on the concave up pattern in the scatterplot.

Graph 4.22 *Scatterplot of the variables ATSfield vs. SCsum4 for the 3213sp19 Participants*



Graph 4.23 *Scatterplot of the Unstandardized Residual vs. Unstandardized Predicted for ATSfield Predicting SCsum4 on the 3213sp19 Participants*

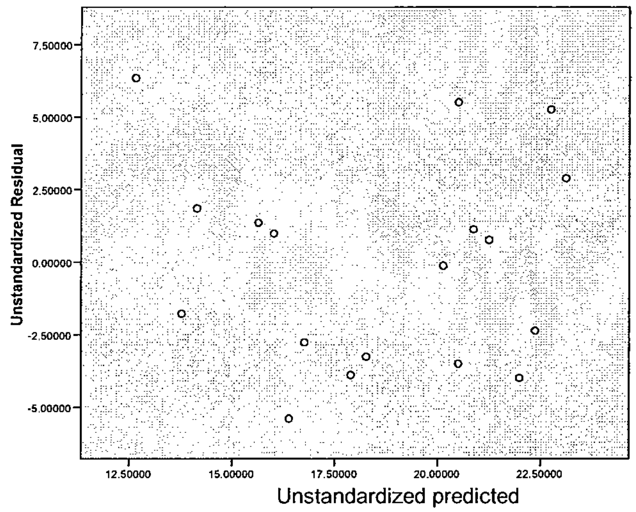


Table 4.24 *Coefficient of Determination, Statistical Significance of the Model, and Standard Error of the Coefficients of the Quadratic Regression to Predict SCsum4 scores for 3213sp19 Participants Using the Variable ATSfield*

Coefficient of Determination	R^2	.623
Statistical Significance of the Model	*p*	.000
Standard Error of the Coefficients	Constant (β=124.887) ATSfield (β= -3.594) ATSfldSq (β= -.029)	.030 .036 .022

The resulting model is E(SCsum4)=124.887 - 3.594(ATSfield) -.029(ATSfldSq) (see Appendix I). This model accounts for 62.3% of the variability in the Scsum4 results (R^2 = .623). Adding the quadratic term increased the explained variance by 15.1%. The model is statistically significant at the $\alpha = .001$ level ($p = .000$). All of the coefficients in this quadratic model are significant at the $\alpha = .05$ level. Hence the best model for predicting SCsum4 scores for the 3213sp19 participants is this quadratic model with ATSfield and ATSfldsq as the predictor variables. This model suggests that for the 3213sp participants, attitudes toward statistics as a field is a good predictor of statistical knowledge of the four topics data representation, probability, measures of center, and measures of spread.

To see the improvement of the new model over the original, we can revisit the scatterplot. This time the ATSfldSq variable was used with ATSfield in the scatterplot. The updated residual vs. predicted scatterplot appears more random. This scatterplot is presented below on table 4.25.

Graph 4.25 *Scatterplot of the Unstandardized Residual vs. Unstandardized Predicted for ATSfield and ATSfldsq Predicting SCsum4 on the 3213sp19 Participants*

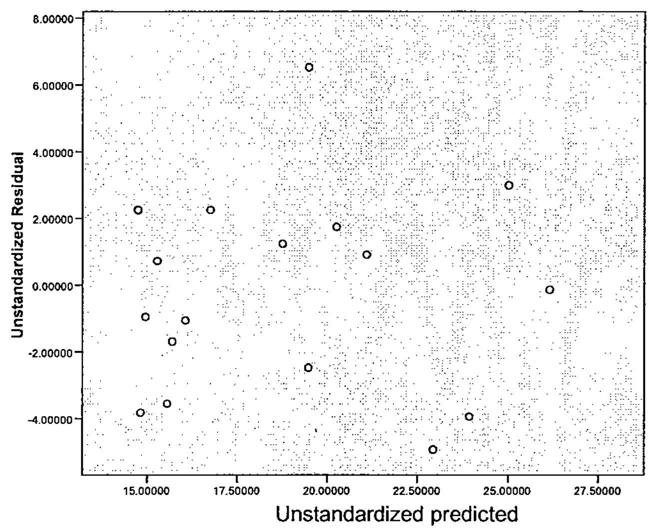


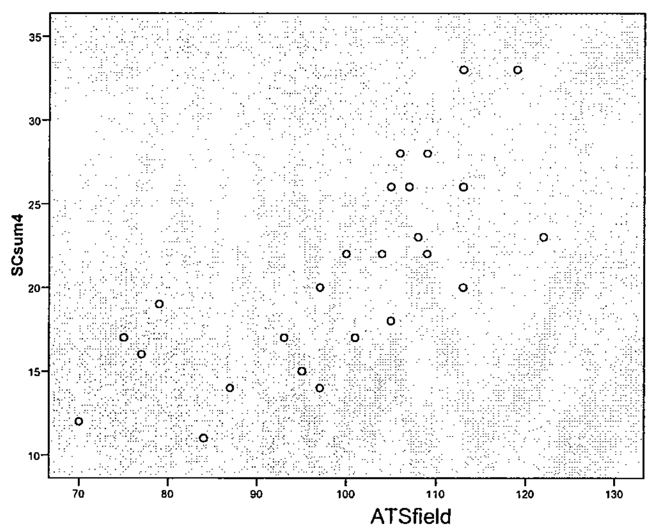
Table 4.26 *Coefficient of Determination, Statistical Significance of the Model, and Standard Error of the Coefficients of the Stepwise Regression to Predict SCsum4 scores for the Combined 3213sp19 and 432 Participants*

Coefficient of Determination	R^2	.553
Statistical Significance of the Model	*p*	.000
Standard Error of the Coefficients	Constant (β= -10.924) ATSsum (β=.320)	.083 .000

The resulting model is E(SCsum4) = -10.924 + .320(ATSsum) (see Appendix I). This model accounts for 55.3% of the variability in the SCsum4 scores for the combined 3213sp19 and 432 participants. This is an increase of 8.1% over the predictability of the model for the 3213sp19 alone. This model provides evidence of correlation between SCsum4 and ATSsum. This model utilizes the ATSsum score as the independent variable in place of the ATSfield score utilized in the 3213sp19 model.

The scatterplot on figure 4.27 appears to have a potential concave up quadratic pattern. It was necessary to check for a potential quadratic regression model. The residuals may not have been random in the scatterplot on figure 4.28.

Graph 4.27 *Scatterplot of the variables ATSsum vs. SCsum4 for the Combined 3213sp19 and 432 Participants*



Graph 4.28 *Scatterplot of the Unstandardized Residual vs. Unstandardized Predicted for ATSsum Predicting SCsum4 on the Combined 3213sp19 and 432 Participants*

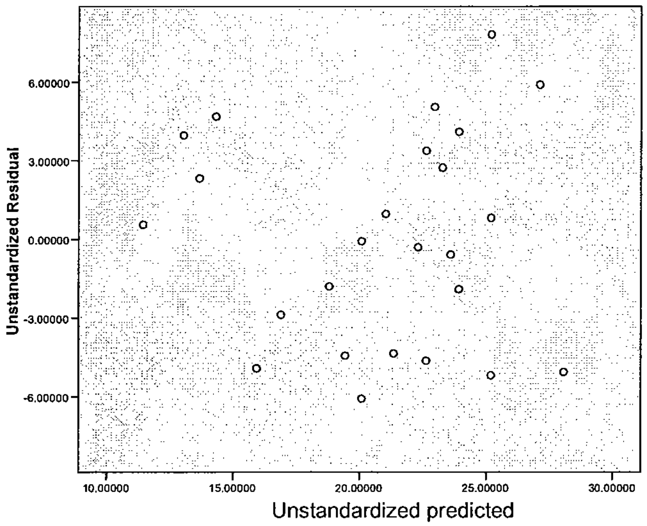


Table 4.29 *Coefficient of Determination, Statistical Significance of the Model, and Standard Error of the Coefficients of the Quadratic Regression to Predict SCsum4 scores for the Combined 3213sp19 and 432 Participants Using the Variable ATSsum*

Coefficient of Determination	R^2	.581
Statistical Significance of the Model	p	.000
Standard Error of the Coefficients	Constant (β=34.064) ATSsum (β= -.643) ATSsumSq (β=.005)	.371 .425 .235

On this model, the standard errors of the coefficients are not significant. So this model was not used. The best model for predicting SCsum4 scores for the combined 3213sp19 and 432 participants is the linear model using ATSsum as the independent variable. Thus the model E(SCsum4) = -10.924 + .320(ATSsum) provides suggestion that for all participants who completed the ARTIST scales, participant attitudes toward statistics (both as a field and as coursework) is an indicator of participant statistical knowledge of the four topics data representation, probability, measures of center, and measures of spread.

Analyses Part 7: Search for a Model to Predict SELS and ATSfield Scores

In this section I present the regression models generated to predict the SELS levels of the three preservice primary teacher participant sets (1473ed, 3213fa, 3213sp). I also present one regression model generated to predict ATSfield scores for the 3213sp participant set. These analyses were founded on the ANOVA results previously presented.

Table 4.30 provides the data sets used, the affective measure used, and the table at which the summary of each regression analyses is presented. Each summary table provided is an abridgment of an SPSS™ output which is provided in Appendix I. The significance level continued to be $\alpha = .1$ for these analyses.

Table 4.30 *Summary of the Output Tables for the SELS and ATSfield Predictor Models*

Location	Data Set	Affective Measure	Type of Regression
Table 4.31	3213sp and 1473ed	SELS	stepwise linear
Table 4.34	3213fa and 1473ed	SELS	stepwise linear
Table 4.37	3213fa, 3213sp, and 1473ed	SELS	stepwise linear
Table 4.38	3213fa, 3213sp, and 1473ed	SELS	stepwise linear
Table 4.43	3213fa, 3213sp, and 1473ed	SELS	quadratic
Table 4.44	3213fa	SELS	stepwise linear
Table 4.45	3213fa	SELS	stepwise linear
Table 4.50	3213sp	ATSfield	stepwise linear

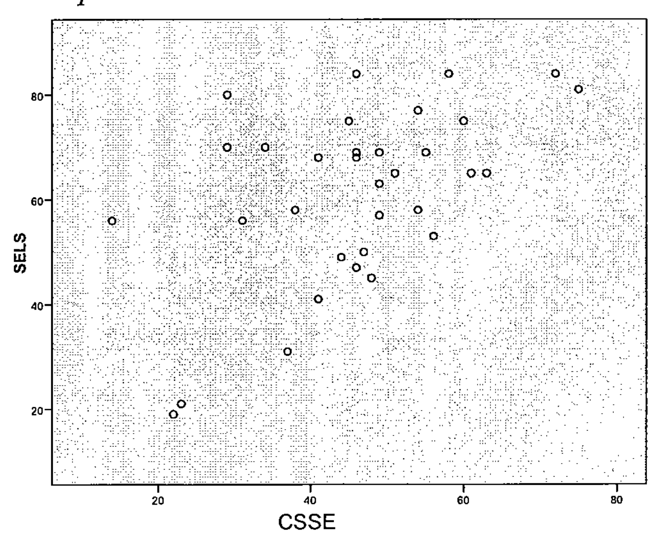
Table 4.31 *Coefficient of Determination, Statistical Significance of the Model, and Standard Error of the Coefficients of the Stepwise Regression to Predict SELS Results for the 3213sp and 1473ed Participants Combined*

Coefficient of Determination	R^2	.250
Statistical Significance of the Model	p	.003
Standard Error of the Coefficients	Constant (β=35.562) CSSE (β=.569)	.000 .003

The resulting model is E(SELS) = 35.562 + .569(CSSE) (see Appendix I). This model accounts for 25% of the variability in the combined 3213sp and 1473ed SELS scores. In this model, the constant and CSSE coefficients are both statistically significant.

There is no apparent quadratic pattern in the scatterplot on figure 4.32. The residuals appear random in the scatterplot on figure 4.33. Hence there is no further regression analysis.

Graph 4.32 *Scatterplot of the variables CSSE vs. SELS for the 3213sp and 1473ed Participants Combined*



Graph 4.33 *Scatterplot of the Unstandardized Residual vs. Unstandardized Predicted for CSSE Predicting SELS on the Combined 3213sp and 1473ed Participants*

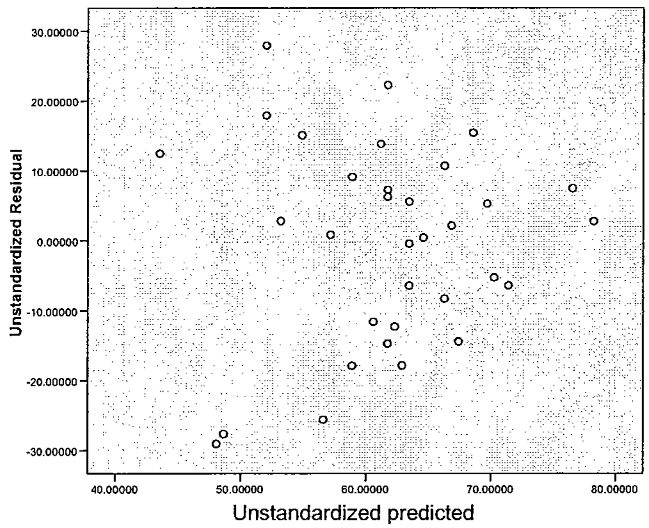


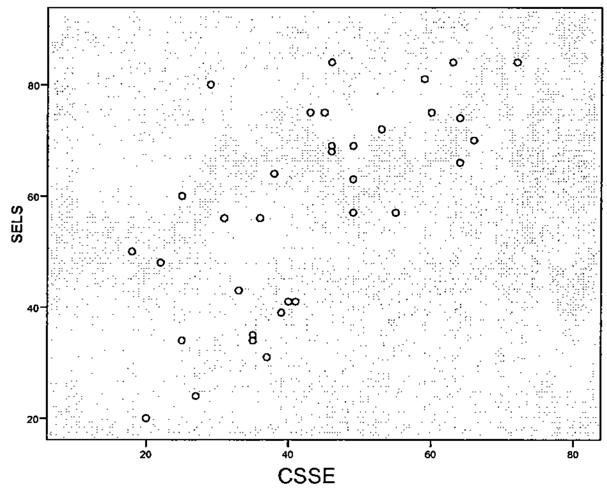
Table 4.34 *Coefficient of Determination, Statistical Significance of the Model, and Standard Error of the Coefficients of the Stepwise Regression to Predict SELS Results for the Combined 3213fa and 1473ed Participants*

Coefficient of Determination	R^2	.450
Statistical Significance of the Model	p	.000
Standard Error of the Coefficients	Constant (β=21.014) CSSE (β=.866)	.010 .000

The resulting model is E(SELS) = 21.014 + .866(CSSE) (see Appendix I). This model accounts for 45.0% of the variability in the 3213fa and 1473ed SELS scores. In this model, the constant and CSSE coefficients are both statistically significant. Hence for all the preservice primary teacher participants from semester 1, current self-efficacy to perform statistical procedures is a weak to moderate predictor of self-efficacy to learn statistics.

There is no apparent quadratic pattern in the scatterplot on figure 4.35. The residuals appear random in the scatterplot on figure 4.36. Hence there is no further regression analysis.

Graph 4.35 *Scatterplot of the variables CSSE vs. SELS for the Combined 3213fa and 1473ed Participants*



Graph 4.36 *Scatterplot of the Unstandardized Residual vs. Unstandardized Predicted for CSSE Predicting SELS on the Combined 3213fa and 1473ed Participants*

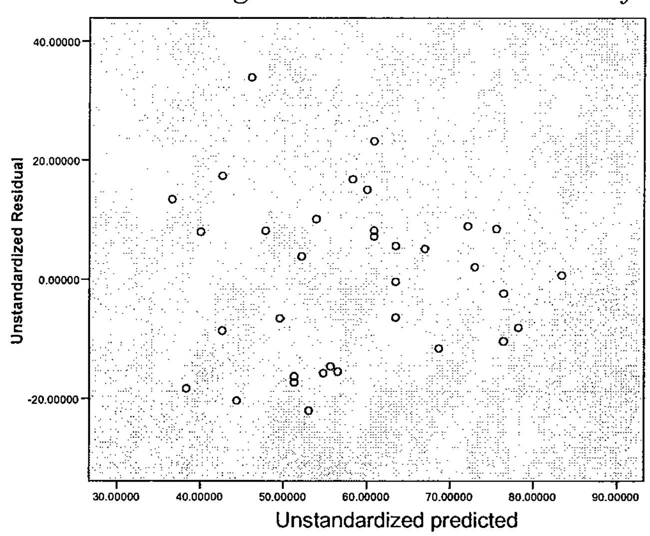


Table 4.37 *Coefficient of Determination, Statistical Significance of the Model, and Standard Error of the Coefficients of the Stepwise Regression to Predict SELS Results Using CSSE for the Combined 3213sp, 3213fa, and 1473ed Participants*

Coefficient of Determination	R^2	.375
Statistical Significance of the Model	p	.000
Standard Error of the Coefficients	Constant (β=27.372) CSSE (β=.715)	.000 .000

Table 4.38 *Coefficient of Determination, Statistical Significance of the Model, and Standard Error of the Coefficients of the Stepwise Regression to Predict SELS Results Using CSSE and ATScourse for the Combined 3213sp, 3213fa, and 1473ed Participants*

Coefficient of Determination	R^2	.440
Statistical Significance of the Model	p	.000
Standard Error of the Coefficients	Constant (β=15.799) CSSE (β=.561) ATScourse (β=.677)	.032 .000 .016

The resulting model #1 is E(SELS) = 27.372 + .715(CSSE) (see Appendix I). This model accounts for 37.5% of the variability in the combined 3213sp, 3213fa, and 1473ed SELS scores. The resulting model #2 is

E(SELS) = 15.799 + .561(CSSE) + .677(ATScourse)

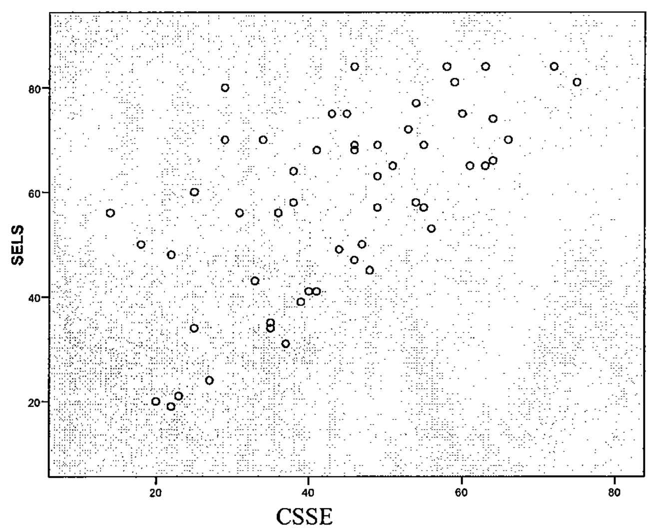
(see Appendix I). This model accounts for 44.0% of the variability in the combined 3213sp, 3213fa, and 1473ed SELS scores. Both models have statistically significant coefficients for each variable and for the constant coefficient.

The first model is more parsimonious with only one independent variable (CSSE). In this case, the first model is preferred over the second model. The 6.5%

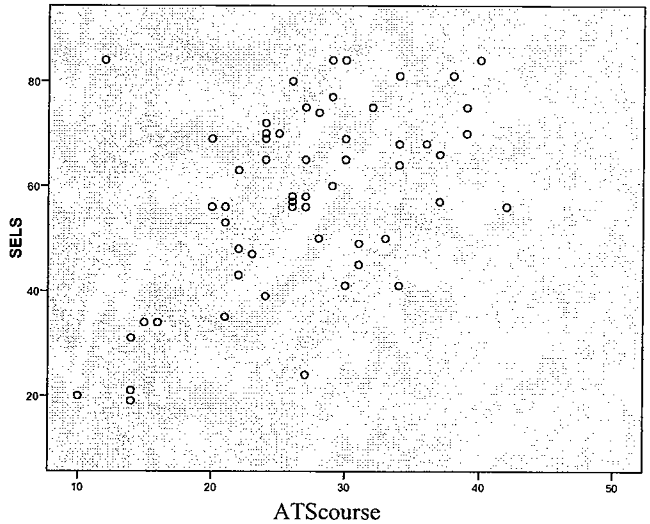
additional explained variance in the SELS scores is not worth including the second independent variable (ATScourse) in the prediction model. Hence for all preservice teacher participants in mathematics content-for-teaching courses, current self-efficacy to perform statistical procedures is a weak to moderate predictor of self-efficacy to learn statistics.

There is potential for an underlying concave down quadratic pattern in the CSSE scatterplot on figure 4.39. There is no apparent quadratic pattern in the ATScourse scatterplot on figure 4.40. A quadratic model involving CSSE was considered. The scatterplot of residuals from figure 4.40 and figure 4.41 provide no indications of residual patterns.

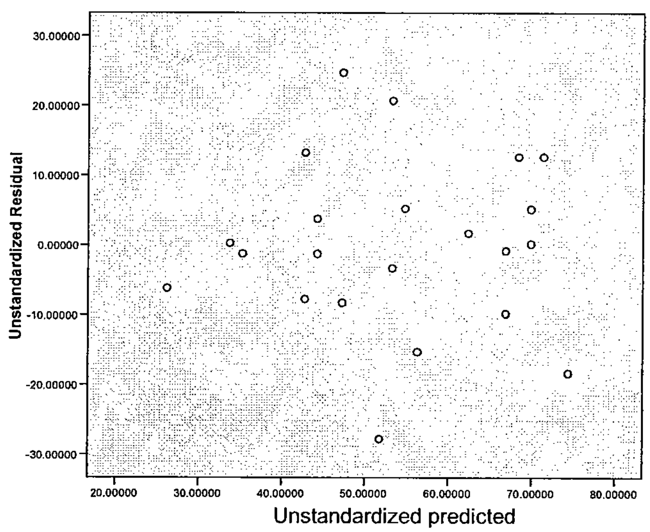
Graph 4.39 *Scatterplot of the CSSE vs. SELS variables on the Combined 3213fa and 1473ed Participants*



Graph 4.40 *Scatterplot of the ATScourse vs. SELS variables on the Combined 3213fa and 1473ed Participants*



Graph 4.41 *Scatterplot of the Unstandardized Residual vs. Unstandardized Predicted for CSSE Predicting SELS on the Combined 3213fa and 1473ed Participants*



Graph 4.42 *Scatterplot of the Unstandardized Residual vs. Unstandardized Predicted for CSSE Predicting SELS on the Combined 3213fa and 1473ed Participants*

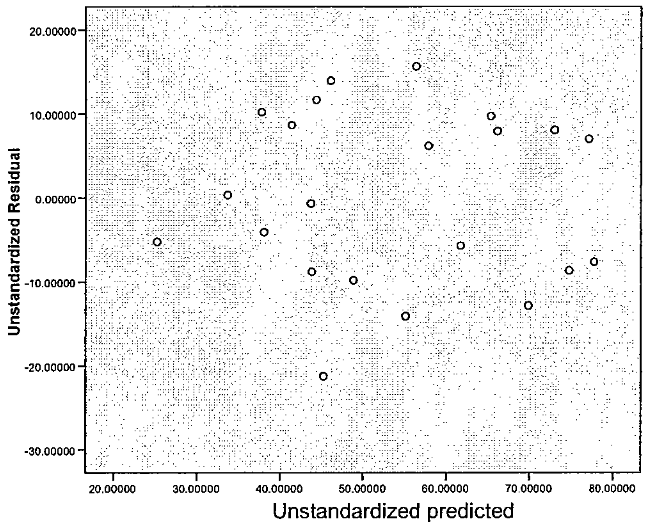


Table 4.43 *Coefficient of Determination, Statistical Significance of the Model, and Standard Error of the Coefficients of the Quadratic Regression Model to Predict SELS scores for the Combined 3213sp, 3213fa, and 1473ed Participants Using the Variables ATSfield and ATSfldSq*

Coefficient of Determination	R^2	.380
Statistical Significance of the Model	*p*	.000
Standard Error of the Coefficients	Constant (β=35.702) ATSfield (β=.274) ATSfldSq (β=.005)	.011 .680 .499

In both of the quadratic models generated to predict SCsum4 scores for the combined 3213sp, 3213fa, and 1473ed participants, the coefficients are not significant. These models were not used. This fits with the apparently random scatterplots for the errors in the initial linear model.

Table 4.44 *Coefficient of Determination, Statistical Significance of the Model, and Standard Error of the Coefficients of the Stepwise Regression to Predict SELS Results for the 3213fa Participants*

Coefficient of Determination	R^2	.547
Statistical Significance of the Model	p	.000
Standard Error of the Coefficients	Constant ATScourse (β=1.505)	.234 .000

Table 4.45 *Coefficient of Determination, Statistical Significance of the Model, and Standard Error of the Coefficients of the Stepwise Regression to Predict SELS Results for the 3213fa Participants*

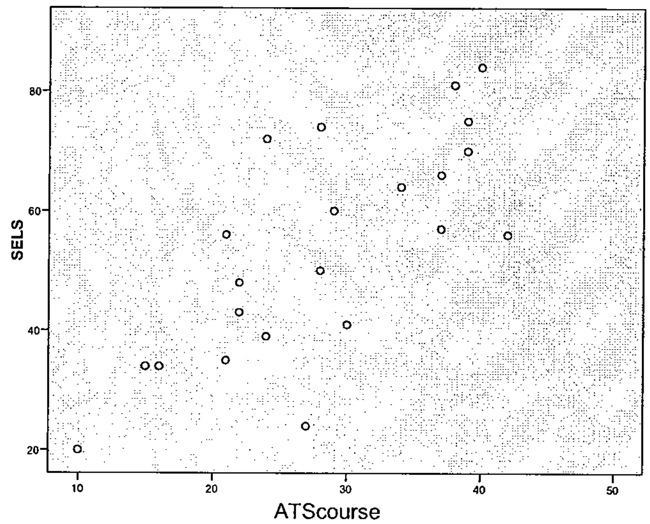
Coefficient of Determination	R^2	.683
Statistical Significance of the Model	p	.000
Standard Error of the Coefficients	Constant (β=4.958) ATScourse (β=.538) CSSE (β=.955)	.548 .000 .010

The resulting model 1 is E(SELS) = 11.157 + 1.505(ATScourse) (see Appendix I). This model accounts for 54.7% of the variability in the 3213fa SELS scores. The resulting model 2 is E(SELS) = 4.958 + .538(CSSE) + .955(ATScourse). This model accounts for 68.3% of the variability in the 3213fa SELS scores.

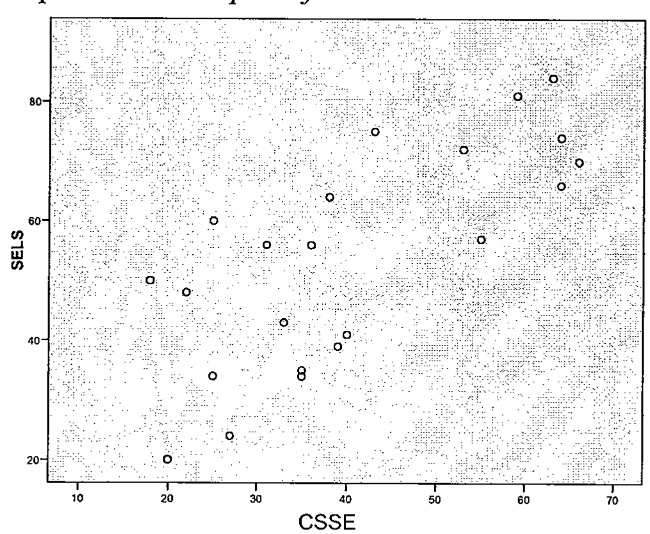
The first model is more parsimonious with only one independent variable (ATScourse). However, in this case, the second model is preferred over the first model. The 13.6% additional explained variance in the SELS scores is worth the second independent variable (CSSE) in the model. Thus a combination of participant attitudes toward the course currently providing statistical content and participant self-efficacy to currently perform statistical procedures was a good indicator of participant self-efficacy to learn statistics.

There are no apparent quadratic patterns in the scatterplots on figure 4.46 and figure 4.47. It is not obvious that a quadratic model is needed. Thus graphs of the residuals were generated to help clarify the decision to generate a quadratic model. From graphs 4.48 and 4.49 below, the residuals appear random.

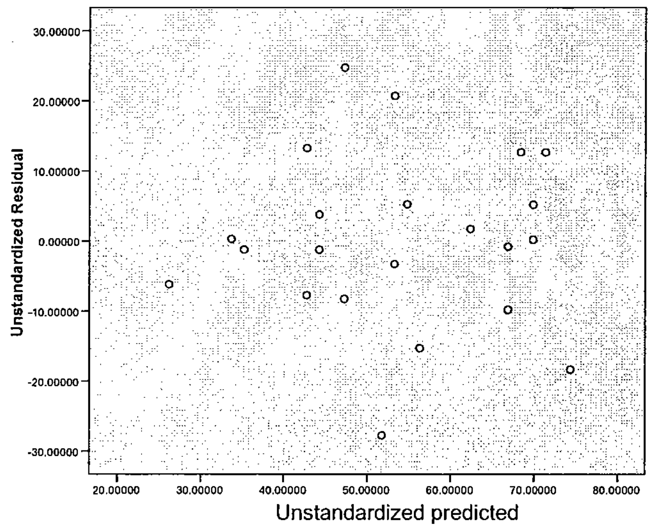
Graph 4.46 *Scatterplot of ATScourse vs. SELS on the 3213fa Participants*



Graph 4.47 *Scatterplot of CSSE vs. SELS on the 3213fa Participants*



Graph 4.48 *Scatterplot of the Unstandardized Residual vs. Unstandardized Predicted for ATScourse Predicting SELS on the 3213fa Participants*



Graph 4.49 *Scatterplot of the Unstandardized Residual vs. Unstandardized Predicted for CSSE Predicting SELS on the 3213fa Participants*

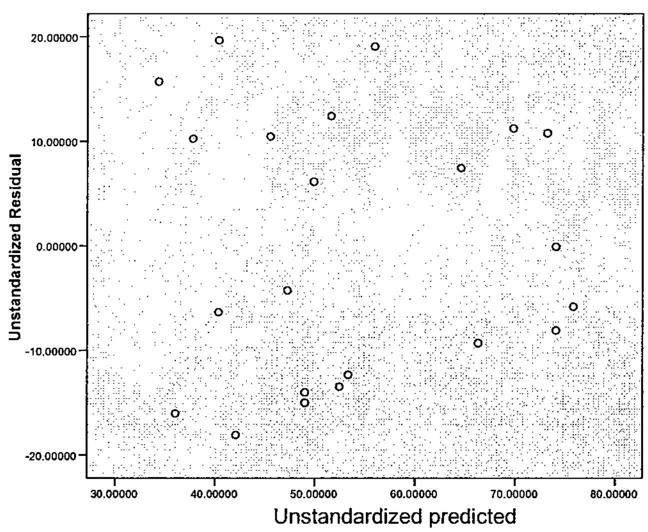


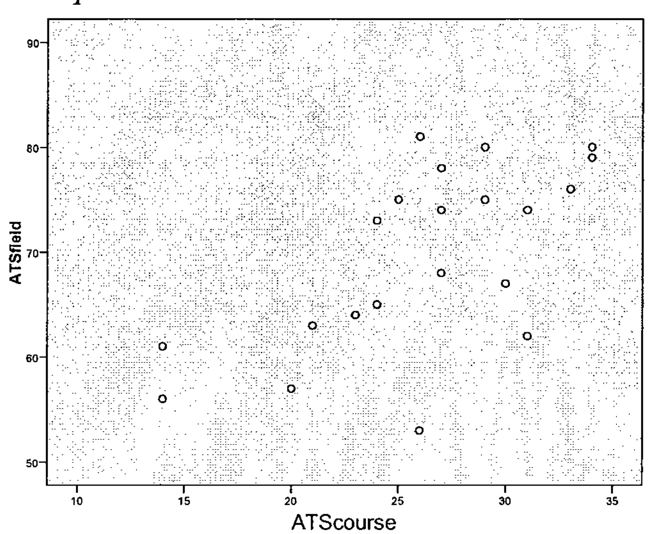
Table 4.50 *Coefficient of Determination, Statistical Significance of the Model, and Standard Error of the Coefficients of the Stepwise Regression to Predict ATSfield Results for the 3213sp Participants*

Coefficient of Determination	R^2	.413
Statistical Significance of the Model	p	.001
Standard Error of the Coefficients	Constant (β=43.372) ATScourse (β=.998)	.000 .001

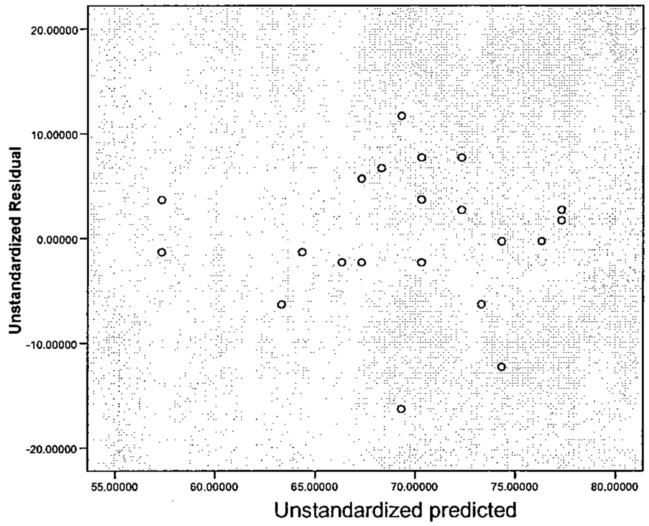
The resulting model is E(ATSfield) = 43.372 + .998(ATScourse) (see Appendix I). This model accounts for 41.3% of the variability in the 3213sp ATSfield scores. All coefficients are statistically significant for this model. Hence for the 3213sp participants, attitudes toward the current course providing statistical content are a moderate indicator of participant attitudes toward statistics as a field.

There is no apparent quadratic pattern in the scatterplot on figure 4.51. The residuals appear random in the scatterplot on figure 4.52 below. So no quadratic model has been generated.

Graph 4.51 *Scatterplot of the variables ATScourse vs. ATSfield for the 3213sp Participants*



Graph 4.52 *Scatterplot of the Unstandardized Residual vs. Unstandardized Predicted for ATScourse Predicting ATSfield on the 3213sp Participants*



Analyses Part 8: Search for a Model to Predict Interest in CPD in Statistics

To predict interest in CPD in statistics, it was necessary to identify a measure for this construct. The ATPDS instrument provided six variables. Any one of these variable could have been used to measure the desired construct. The choice of which variable(s) to use was based on applicability of the variable and quantitative evaluations of the variables. First, the variable means were determined for both the 3213sp and 432 data sets. The original descriptive statistics outputs for table 4.53 are provided in Appendix F.

Table 4.53 *The Mean for Each Item and for the Totals on the ATPDS Instrument – for Each of the 3213sp21 and 432 Participant Sets*

	3213sp	432
Enjoy a CPD stat workshop	2.57	2.75
Attend with colleagues	3.05	3.25
Help with classroom teaching	2.86	3.50
Attend if cost is covered	3.10	3.25
Attend if a stipend provided	3.71	4.00
Attend if participant had to pay	1.57	2.75
TOTAL of the six items	16.86	19.50
Average of each item	2.81	3.25

It is useful to know if any of the items on the ATPDS instrument are correlated. To check for such correlation, the Pearson-Product correlation coefficient was generated for each bivariate pairing of the items. The SPSS™ results of the Pearson-Product correlations are in table 4.54.

Table 4.54 *Correlations Between the Items on the Interest in CPS in Statistics Instrument (ATPDS)for the 3213sp Data Set*

Correlations

		WKen	WKfr	WKhp	WKpd	WKin	WKls
WKen	Pearson Correlation	1	.602(**)	.531(*)	.751(**)	.595(**)	.396
	Sig. (2-tailed)		.004	.013	.000	.004	.076
	N	21	21	21	21	21	21
WKfr	Pearson Correlation	.602(**)	1	.525(*)	.741(**)	.718(**)	.285
	Sig. (2-tailed)	.004		.015	.000	.000	.210
	N	21	21	21	21	21	21
WKhp	Pearson Correlation	.531(*)	.525(*)	1	.603(**)	.397	.342
	Sig. (2-tailed)	.013	.015		.004	.075	.129
	N	21	21	21	21	21	21
WKpd	Pearson Correlation	.751(**)	.741(**)	.603(**)	1	.719(**)	.633(**)
	Sig. (2-tailed)	.000	.000	.004		.000	.002
	N	21	21	21	21	21	21
WKin	Pearson Correlation	.595(**)	.718(**)	.397	.719(**)	1	.320
	Sig. (2-tailed)	.004	.000	.075	.000		.157
	N	21	21	21	21	21	21
WKls	Pearson Correlation	.396	.285	.342	.633(**)	.320	1
	Sig. (2-tailed)	.076	.210	.129	.002	.157	
	N	21	21	21	21	21	21

** Correlation is significant at the 0.01 level (2-tailed).
* Correlation is significant at the 0.05 level (2-tailed).

From table 4.54 above, WKen, the participant belief that they would enjoy a CPD in statistics workshop, correlates with WKfr, WKhp, WKpd, and WKin. WKhp, the participant belief that a CPD in statistics workshop would help with their classroom teaching, correlates with WKen, WKfr, and WKpd. With such strong correlations, it may not hold that different items will generate much different prediction models for attitudes toward CPD in statistics. At least a few different dependent variables will be predicted. Such modeling diversity is important for an investigation into any affective variables that may provide PMTE's some insight into ways to increase preservice teacher interests in CPD in statistics.

Table 4.55 *Summary of the Output Tables for the SELS and ATSfield Predictor Models*

Location	Data Set	ATPDS measure	Type of Regression
Table 4.56	3213sp	WKen	stepwise linear
Table 4.59	3213sp	WKhp	stepwise linear
Table 4.62	3213sp	WKhp	quadratic

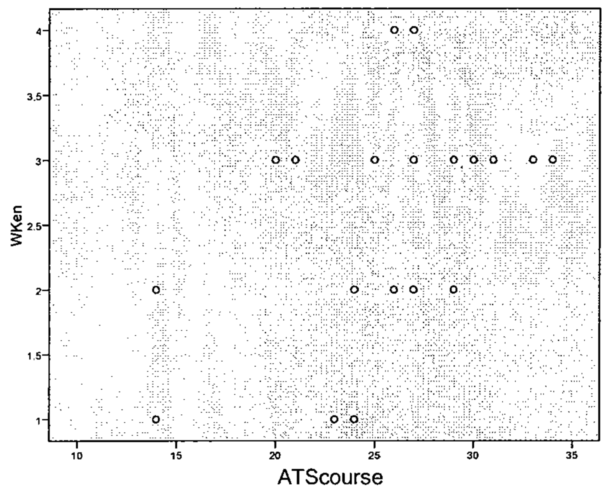
Table 4.56 *Coefficient of Determination, Statistical Significance of the Model, and Standard Error of the Coefficients of the Stepwise Regression to Predict WKen Results for the 3213sp19 Participants*

Coefficient of Determination	R^2	.228
Statistical Significance of the Model	*p*	.039
Standard Error of the Coefficients	Constant (β=.799) ATScourse (β=.071)	.353 .039

The resulting model is E(WKen) = .799 + .071(ATScourse) (see Appendix I). This model accounts for 22.8% of the variability for the 3213sp19 participants in the answers to the question "Do you believe that you would enjoy (a CPD program in statistics)?" While this model provides evidence that the independent and dependent variables are weakly correlated, it cannot be used to make meaningful predictions of SCsum4 scores. This is because of a lack of significance in the constant coefficient in the model.

There is no apparent quadratic pattern in the scatterplot on graph 4.57. The residuals appear random in the scatterplot on graph 4.58. So no quadratic model has been generated.

Graph 4.57 *Scatterplot of the variables ATScourse vs. WKen for the 3213sp19 Participants*



Graph 4.58 *Scatterplot of the Unstandardized Residual vs. Unstandardized Predicted for ATScourse Predicting WKen on the 3213fa Participants*

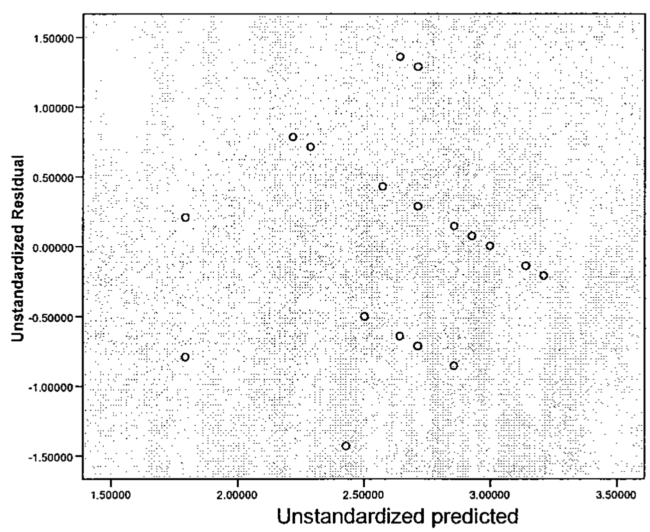


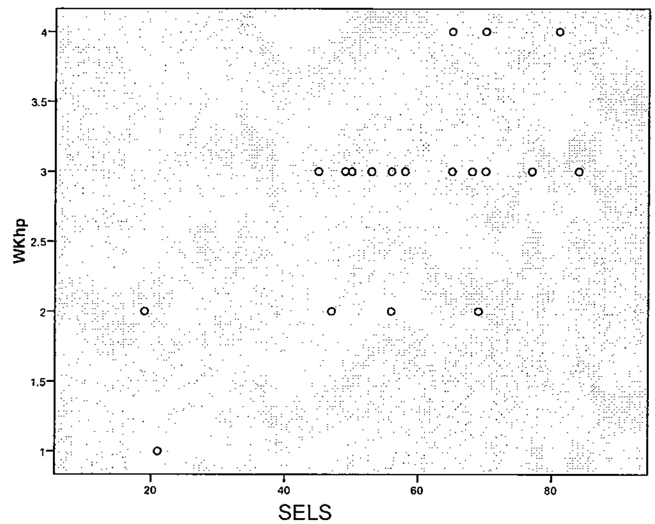
Table 4.59 *Coefficient of Determination, Statistical Significance of the Model, and Standard Error of the Coefficients of the Stepwise Regression to Predict WKhp Results for the 3213sp19 Participants*

Coefficient of Determination	R^2	.473
Statistical Significance of the Model	p	.001
Standard Error of the Coefficients	Constant (β=1.217) SELS (β=.029)	.012 .001

The resulting model is E(WKhp) = 1.217 + .029(SELS) (see Appendix I). This model accounts for 47.3% of the variability for the 3213sp19 participants in the answers to the question "Do you believe that (a CPD program in statistics) would help you in your classroom teaching?" The coefficients for this model are all statistically significant. Hence for the 3213n19 participants, the data suggest that participant self-efficacy to learn statistics provides some indication of participant belief that CPD in statistics helps classroom teaching.

This scatterplot on figure 4.60 appears to have a potential concave up quadratic pattern. It is necessary to check for a potential quadratic regression model. The residuals do not appear random in the scatterplot on figure 4.61.

Graph 4.60 *Scatterplot of the variables SELS vs. WKhp for the 3213sp19 Participants*



Graph 4.61 *Scatterplot of the Unstandardized Residual vs. Unstandardized Predicted for SELS Predicting WKhp on the 3213sp19 Participants*

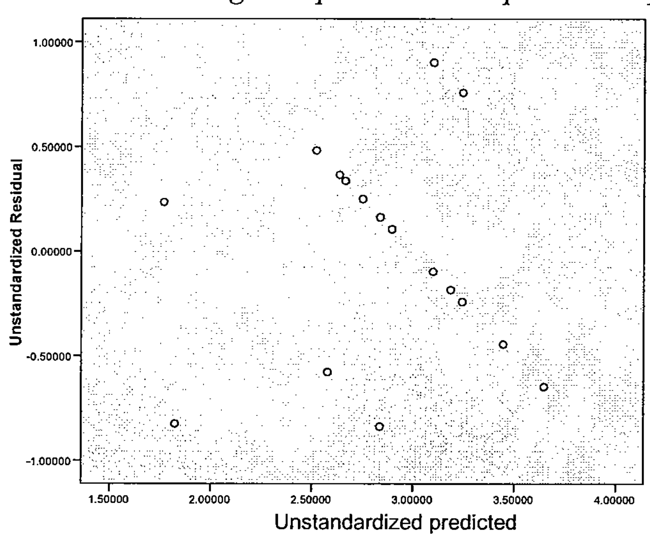


Table 4.62 *Coefficient of Determination, Statistical Significance of the Model, and Standard Error of the Coefficients of the Quadratic Regression Model to Predict WKhp scores for the 3213sp19 Participants Using the Variables SELS and SELSsqrd*

Coefficient of Determination	R^2	.544
Statistical Significance of the Model	p	.002
Standard Error of the Coefficients	Constant (β=.108) SELS (β=.079) SELSsqrd (β= -.001)	.897 .027 .134

The resulting model is E(WKhp) = .108 + .079(SELS) - .001(SELSsqrd) (see Appendix I). This model accounts for 54.4% of the variability in the WKhp variable. Adding the quadratic term increased the explained variance by over 7%. However, the squared term coefficient is not significant. This quadratic model is also not as parsimonious as the E(WKhp) = 1.217+.029(SELS) model since the quadratic model adds an extra variable while only improving the explained variability of the dependent variable by 7.1%. Hence the best model is the first order model with SELS as a predictor variable.

Other WK models. It is important to provide as complete an understanding of the predictability of preservice teacher interest in CPD in statistics as possible. To that extent, I attempted to model WKpd, the interest participants indicated they would have in CPD workshops in statistics if they were paid for the effort. However, there were no independent variables that provided a WKpd prediction model with statistically significant coefficients.

Analyses Part 9: Qualitative Analysis of the Grading Projects

There were two participants in the primary level grading projects. These participants are referred to as El and Cy. El and Cy were in the same section of 3213. The analysis is approached in two ways. Details of each participant's grading project results are discussed. But first, quantitative results for the two participants are presented to help provide context for the qualitative discussion.

The grading projects are provided in their entirety in Appendix C. The projects contained 7 questions appropriate for 4^{th}-grade level students. I provided answers to the questions. Some answers were correct, some incorrect. Some were solved using unorthodox methods. The teachers were to grade the answers to the questions. There were follow-up questions designed to probe into the participant's approach to the grading.

Quantitative results for both primary level participants. To develop a context for comparing the qualitative analysis for each primary level grading projects participant, knowledge and attitude measures from the quantitative analyses are presented. The context includes direct comparison of each measure between the two participants. The context also includes measure-by-measure comparisons between each participant and 3213sp19 group means from the quantitative analyses.

By coincidence, El and Cy represent statistical extremes for their section of the mathematics content course (3213). As evident in Table 4.79, El is relatively strong in statistical content knowledge of graphs and measures of center based on the SCgraphs and SCcenter scores. El did not participate in the measures of spread and probability scales. El is also strong in all affective measures related to statistics, with respect to the 3213sp19 means (see chart 4.79). Cy is relatively weak in statistical content knowledge based on the SCgraph, SCcenter, SCspread, and SCprob scores. Cy is also weak in all affective measures related to statistics, with respect to the 3213sp19 means (see chart 4.79).

The original data values for chart 4.79 are in appendix D. In the SPSS™ charts, El is participant #1 and Cy is participant #21. These participants also appear in previous sections of this chapter as participants #1 and #21 for data set 3213sp.

Chart 4.79 *Table of Quantitative Measures for the Primary Level Grading Project Participants and the Mean for Their Mathematics Content Course Section*

	3213sp	3213sp	3213sp.n19
	El	Cy	Section Mean
collstat	1	0	x
ATSfield	80	61	69.26
ATScourse	34	14	25.84
ATSsum	114	75	95.11
CSSE	75	23	40.37
SELS	81	21	56.16
SCgraphs	7	8	6.26
SCcenter	6	3	1.74
SCspread	0	4	4.84
SCprob	0	2	5.89
SCsum4	13	17	18.74
WKen	3	2	2.63
WKfr	3	3	3.16
WKhp	4	1	2.84
WKpd	4	3	3.16
WKin	4	4	3.84
WKls	2	1	1.58
WKtotal	20	14	17.21

A caveat exists for the knowledge levels for El. El scored above average for data representation (graphs) and scored the highest 3213sp score for measures of center. However, El did not participate in the measures of spread and probability portions of the ARTIST scales. Direct observation of the original documents completed by El revealed that the first two scales were answered in full while the last two scales had no answers provided. This forced the removal of El from the adjusted data set 3213sp19. Without this removal (and two other similar removals), the SCsum4 scores would have been skewed for the 3213sp data set.

Although we do not know exactly why El did not complete the last two ARTIST scales, we need to acknowledge that it is always possible that a participant will opt out of some questions from such affective factors as insecurity concerning the topics. However, there were circumstances surrounding the data collection event that could give a different explanation for El's decision. The participants completed the ARTIST scales as part of the broader study survey. They completed the affective portions of the survey in the 3213 classroom over an approximately 20 minutes. The participants then went to a computer lab to complete the ARTIST scales online. While the participants were in the lab, non- participants were getting individual attention from the 3213 section instructor in preparation for an upcoming examination. The instructor advised participants to return to class to participate in the review as soon as they were finished with the ARTIST scales. It is possible that El chose to skip the last two of the four ARTIST scales to utilize the opportunity to review for the upcoming examination.

Quantitative results for El. El had participated in a full semester undergraduate statistics course. El's affective measures for statistics are all above the 3213sp average. While the caveat in El's knowledge measures causes difficulty for analyzing El's statistical knowledge, those measures that do exist indicate that El has an above average understanding of those statistical concepts measured, with respect to the 3213sp19 data set. A particular measure of interest is El's WKhp score of 4, which indicates that El has the highest possible belief (from the Likert format) that a CPD workshop in statistics would help El's classroom teaching.

Quantitative results for Cy. In contrast to El's score of 4 for the WKhp variable, Cy scored a 1. Thus Cy has the lowest possible belief that a CPD workshop in statistics would help Cy's classroom teaching. There could be a variety of reasons why Cy might lack confidence in CPD workshops in statistics. Although this current study did not collect detailed explanatory data, I offer some conjectures. For example, Cy may believe that such workshops are generally ineffective. Alternatively, Cy may believe that Cy's personal abilities to learn statistics will prevent such a workshop from being successful. This latter possibility matches with the low score (relative to the 3213sp19 mean of 40.37) that Cy received – far below the 3213sp19 mean on the SELS. In fact, the score of 21 was the second lowest of the 22 participants from the 3213sp group (Appendix D).

All of Cy's affective measures were low. The 23 for CSSE was much lower than the group mean of 40.37. There were only a few other participants with lower SELS scores (22, 14, 14). Although Cy's knowledge scores were slightly below average overall with an SCsum4 score of 17 compared to the mean of 18.74, Cy scored above average in both data representation and measures of center. Cy's measure of spread score was slightly below average and Cy's probability score was the lowest score of all participants who complete that scale, with a score of 2.

Comments about El's grading project. El erred on only one graded problem, the composite number 91. Not only were problems graded correctly, but also explanations for partial credit decisions were provided by El throughout the completed grading project. El also indicated that El did not need to use nay resource materials to help with the grading project.

Comments about Cy's grading project. Even though the instructions called for the participants to grade each problem on a ten point scale, to allow for partial credit, Cy did not use a partial credit grading system. Instead, Cy marked each problem as either totally correct (with a plus sign) or as completely wrong if any error occurred (with a minus sign). Cy was unable to correctly grade 7 of the 15 student responses during the grading project. Cy's use of resources varied across the questions, the level of resource use for each question will be detailed within the report for each question.

Grading project results for the topic: Dividing by two digit divisors. The first problem to be graded was a division problem, which was presented as a word problem. It was designed to test student understanding of quotients and remainders. There were two hypothetical student answers provided for the preservice teachers to evaluate. Student 1 used a traditional long division method to derive the desired values. This student correctly found the two numbers needed; however, the student confused the two numbers and interpreted the quotient as the remainder and vice versa. Student 2 used a less conventional method, employing a guess-and-check method with multiplication. The student tested multiples of the divisor, 14, and stopped when the multiples exceeded the dividend value of 128. The student showed all work and provided the correct answers.

Both El and Cy correctly marked the work for Student 1 as incorrect. Both graders also correctly marked the work for Student 2 as correct. Thus both graders recognized correct work that did not utilize the conventional algorithm but was based on correct understanding of the mathematical concepts. Both graders indicated that they knew the answer and how to grade the problem without needing to use resources.

The grader responses for the follow-up questions to the division problem explained the errors Student 1 made. El provided a more complete explanation than did Cy for correcting Student 1. However, when asked if the work of Student 2 could be used to help Student 1 understand the mathematical concepts involved, Cy provided a more detailed explanation than did El. Cy's answer included a pedagogical insight, that if Student 2 were the one to help Student 1, then both students might benefit.

Grading project results for the topic: Prime and composite numbers. The second problem tested understanding of factors. The problem presented three numbers. The student was to identify each number as prime or composite.

The solutions of a single hypothetical student were provided for the grader. Two of the numbers, 21 and 17, were relatively small (within the multiplication tables up to nine times nine). The student answered both correctly: 21 was labeled as a composite number. Seventeen was labeled as a prime number. The third number, 91, presented in the problem was somewhat larger and could appear to be a prime number since it does not appear on multiplication tables up to twelve times twelve. The student stated that 91 was a prime number. However, 91, is equal to seven times thirteen and hence is a composite number.

Both graders correctly graded the student solutions for the two smaller numbers, indicating that the graders probably understood the basic concept of prime and composite numbers. However, both graders incorrectly graded the student solution for the number 91. The graders agreed with the student that 91 is prime. El stated that El knew the answer and how to grade the problem without needing to use resources. Cy stated that Cy had to spend a long time with references to figure it out, but that the activity was rewarding. Hence on the one hand, El did not feel compelled to investigate the topic even though El did not grade the problem correctly. On the other hand, Cy felt compelled to investigate the topic further and yet was still unable to recognize the error in grading.

Grading project results for the topic: Number patterns. The third problem tested ability to recognize number patterns. The question was divided into three parts. Each part presented the student with a sequence of three numbers: each sequence had either an arithmetic or geometric pattern. The student was asked to identify the next (fourth) number in the pattern.

The solutions of a single hypothetical student were provided for the grader. The student's solutions for the first and third sequences were based on simple patterns:

the first pattern was arithmetic with 2 as the addend, and the third pattern was geometric with 3 as the multiplier. The second sequence, {1, 3, 7, …} was intentionally designed to match two different possible patterns. This sequence could lead a student or grader to use a pattern in which 2 is added to the initial value of 1 to get 3 then 4 was added to the value of 3 to get 7.With this pattern, the student or grader might add 6 to the third term, 7, and get 13. The student's answer, 15, made use of an alternate possible pattern based on the formula $x_n = 2n-1$ where x_n is the nth term in the sequence and n is an integer starting at one.

The instructions intentionally specified that the students did not need to show their work. Although these instructions created some concern from the graders during the follow-up questions, I made this decision to strengthen my ability to evaluate the participant's grading. If the student showed work on the second sequence then the grader would be provided with the pattern even if the grader were not capable of identifying such a pattern. If the instructions did not specify that showing work was not required, then the grader might mark the problem incorrect based on not showing work. In either of those two cased, I would have been unable to detect whether the grader was able to identify the pattern that led to the student's answer.

El correctly recognized the student's answer as correct based on a formula the student might have used. In fact, El wrote out a pattern that allowed the answer of 15 to be correct. Cy, on the other hand, did not correctly grade the student solution. Since the student did not provide the answer that Cy expected based on Cy's solution to the sequence, Cy marked the problem as incorrect. Both graders stated that they knew the answer and how to grade the problem without needing to use resources.

Grading project results for the topic: Data representation. To investigate understanding of data representation, the fourth question used a frequency table. This question was a word problem, giving a table, pre-designed with the categories provided for the student. The student made tabulations in the second column and cumulative frequency totals in the third column.

The solutions of a single hypothetical student were provided for the grader. The student correctly tabulated the frequencies in column two, but made an error in column three. The correct sequence of answers, from top to bottom, of the third column would be 11, 15, 21, 24; the student answered 24, 24, 24, 24, indicating that the student gave the total cumulative total for all four categories in every row.

El correctly recognized column two as correct for the student and column three as incorrect for the student. El not only recognized the error, El explained in the follow-up questions that the student was likely confused by the title of "Cumulative Frequency" at the top of column three. In contrast, Cy gave the student full credit for the frequency table, indicating that Cy was unable to identify the error. Cy explained in the follow-up question that Cy did not check for the error. When asked how to provide an explanation to the student for what the student did wrong, Cy left the item blank. El stated that El knew the answer and how to grade the problem without needing to use resources. Cy stated that Cy made a quick reference (not a thorough one) yet was still unable to recognize the error in grading.

Grading project results for the topic: Probability. The fifth problem tested understanding of probability. A spinner with three equally divided regions was the random generator for the problem. The students were asked to identify the more likely sum from among the solutions 2, 3, or 4. The values in the regions on the spinner were 1, 2, and 3. Hence the correct solution would be the number out of the choices 2, 3, or 4 that is the sum for the most combinations of two spins of the spinner.

There were two hypothetical student answers provided for the preservice teachers to evaluate. Student 1 gave an answer that did not consider two separate spins. The rationale given by Student 1 did not have any support from probability theory. Student 2 recognized the correct way to solve the problem, provided the correct answer, and showed the work to support that answer, listing all combinations of spins and tabulating which sum occurs most often. However, Student 2 did not find all of the possible combinations: instead of the three ways to achieve a sum of 4, the student listed only two ways. Thus the "correct" solution for Student 2 was found partly by chance.

Both graders correctly marked the Student 2 work as mostly correct. El correctly recognized that Student 2 understood the underlying concepts by rewarding student 2 with 8 out of 10 points. El also noticed the minor error of failing to find all combinations and explained that the last two points were not given because of this omission. Cy gave Student 2 full credit for the problem. However, when Cy was prompted in the follow-up question, Cy was able to recognize that Student 2 should have found three ways the spins could add up to four. Hence it is unclear whether Cy failed to identify the error until the follow-up question or whether Cy determined that the student had enough understanding of the concept to award the student full credit. Both graders stated that they knew the answer and how to grade the problem without needing to use resources.

Grading project results for the topic: Mean, median, mode, and outliers. The mean, median, and mode problem was presented in three parts. First the student was asked to find the mean, median, and mode of the following list: 6, 7, 7, 9, and 11. This is a five-number list and so the median is easier to find than for an even-numbered list. Second, the student was asked to find the median of the list when the number 17 is added to the list above. Third, the student was asked to determine whether the mean or the median changed more when the 17 was added to the list.

The solutions of a single hypothetical student were provided for the grader. The student found the correct mean, median, and mode for the first part of the question and showed correct work. On the second part, the student correctly calculated the new median using the midpoint of the middle two numbers procedure. On the third part, the student answered the question by saying, "It affected the median the most since I had to change the way that I found the answer." The question was meant to refer to the magnitude of the statistic, not the process for calculating the statistics. Thus, a grader has a choice to make. They could base grading of the student's answer on interpretation of the wording and not deduct points for mathematics content understanding. Alternatively, they could also determine that the answer the student provided was incorrect and deduct points accordingly. One goal for this problem was to see if the grader recognizes that the mean will change more than the median, that cannot occur unless the grader first recognizes the correct meaning of the question. If the grader allows the student's incorrect interpretation to affect the grading, then the grader may never identify the correct answer.

El correctly graded all three parts of the problem. On the third part, El asked the student to consider the mean in the comments provided on the graded project. Cy correctly gave the student credit for the answers on the first two parts. Cy chose to give the student credit for the answer on the third part. Both graders indicated that they knew the answer and how to grade the problem without needing to use resources.

Grading project results for the topic: Stem-and-leaf plots. The seventh question asked the student to identify the mode and the median for a stem-and-leaf plot. The plot has three stems, each of which has a different number of leaves: as the stem value grows larger, the number of leaves grows larger.

The solutions of a single hypothetical student were provided for the grader. The student recognized the mode correctly. The student identified the median by an incorrect, if somewhat inventive, method. The student believed that, since the median is the "middle" term, it can be identified by finding the middle stem value and then finding the middle leaf value for that stem. This procedure would in fact be correct if there were an appropriate symmetry in the numbers of leaves for the stems. However, for this problem, the student is incorrect in the solution.

El correctly gave the student credit for the mode and for identifying the median as the "middle term". El also correctly did not award any points to the student for the solution to the median. Cy incorrectly gave the student full credit for the problem. In the follow-up questions, Cy explained, "Honestly, [I] did not know how to do [the] stem-and-leaf problem." Cy indicated that Cy "needed a lot of time to reference and had to give up on a satisfactory review of the material due to being short on time." El indicated that El knew the answer and how to grade the problem without needing to use resources.

Summary

All of the participants in the fourth mathematics course for preservice primary teachers (3213all) scored significantly lower self-efficacy to learn statistics in the future (SELS) than did both the preservice primary education majors in the 1473 course (1473ed) and all of the 1473 course participants (1473). This indicates a possibility that preservice primary teachers at the University of Oklahoma lose confidence in their ability to learn statistics as they progress through their preservice teacher program. In fact, the average SELS scores of the primary preservice teachers was higher than the average of the entire 1473 participant group hence the following inequality holds: 3213all SELS scores < 1473all SELS scores < 1473ed SELS scores.

The current self-efficacy to use statistics for the preservice primary teachers in the 1473 course (1473ed) was significantly higher than that of the other 1473 participants (1473not). This might be warranted based on statistics knowledge levels. Although the SCI means are not significantly different, the 1473ed group scored 14.5% higher (9.08 to 8.10) on the SCI (knowledge of statistics) instrument than did the other 69 participants from the 1473 course. Thus it appears that the preservice primary teachers have neither lower confidence to perform statistics nor lower statistics knowledge levels than other undergraduate majors who take 1473 at the University of Oklahoma.

The ATSfield mean score was significantly higher for the participants from the 3213 section in semester 2 (3213sp) than they were for both the 1473 and 1473ed groups. In this case, the participants from the fourth mathematics course for preservice primary teachers scored higher than the participants from the first mathematics course for preservice primary teachers. In fact, the mean for the preservice primary teachers was lower than that of the 1473 participant set as a whole. Hence the following inequality holds: 1473ed ATSfield scores < 1473all ATSfield scores < 3213sp ATSfield scores. The order of the inequality is contrary to the inequality for the self-efficacy to learn statistics results. The inequality holds with the 3213fa ATSfield scores in place of the 3213sp ATSfield scores. However, the results were not significant for the 3213fa group. Among potential reasons for this discrepancy is that (1) the instructor plays a role in such attitudes and (2) the results are an anomaly due to relatively small numbers of participants. Part (1) above is possible since the two sections were taught by two different instructors.

The SCI instrument and the sum (SCsum4) of the four ARTIST scales (1) data representation, (2) measures of center, (3) measures of spread, and (4) probability, both appear to measure statistics knowledge for preservice secondary teachers at the University of Oklahoma at the same levels. This consistency fails to hold between the SCI instrument and the sum (SCsum7) of the seven ARTIST scales (1) data representation, (2) measures of center, (3) measures of spread, (4) probability, (5) sample variability, (6) confidence intervals, and (7) data collection. It is unclear why the SCsum4 provides such similar scores to the SCI while the SCsum7 does not.

Two measures of interest in CPD used in the study were WKen and WKhp, interest in participating in CPD workshops in statistics and belief that such workshops are beneficial to classroom teaching. These data were analyzed for the 3213sp19 data set. There were no highly correlated predictor variables for WKen. The best predictor of WKhp was SELS. The regression model generated to predict beliefs that CPD workshops can be beneficial to statistical classroom teaching (WKhp) used self-efficacy to learn statistics (SELS) as a predictor variable. This model accounted for 47.3% of the variability in the WKhp values. Hence self-efficacy to learn statistics is a mild predictor of participant belief that such workshops are beneficial to classroom teaching.

The difference in the means of the SELS scores across participant groups made SELS a variable of particular interest. A regression model was found that predicted 68.3% of the variability of the SELS values. The model used ATScourse and CSSE, attitudes toward the current course in which the participants are learning statistics and participant self-efficacy to use statistics now, to predict SELS.

Due to the literature that indicates the importance of statistics knowledge in effective teaching, SCsum4 and SCI scores were of particular interest. There were no models generated to predict SCI levels that explained much of the variance of the SCI values. The best model used ATSsum to predict SCI and only explained 27.6% of the variability in the SCI values.

When considering preservice primary teachers alone (3213sp19), general attitudes toward statistics as a field (ATSfield) was found to be a moderate predictor of statistics knowledge (SCsum4). A model was generated that used ATSfield to predict SCsum4 with 62.3% of the variability in the SCsum4 values explained. When considering both preservice primary teachers (3213sp19) and preservice secondary teachers (4232) as one group of preservice teachers, the combined general attitudes toward statistics measure (ATSsum) was found to be a moderate predictor of statistics knowledge (SCsum4). ATSsum is the combination of attitudes toward statistics as a field (ATSfield) and attitudes toward the current statistics course (ATScourse). A model was generated that used ATSsum to predict SCsum4 with 58.1% of the variability in the SCsum4 values explained.

The qualitative results indicated that there is a discrepancy between the levels of grading accuracy that can exist between two preservice primary teachers. Between the two grading project participants, the participant who graded many more of the problems correctly was also the participant who scored much higher in every affective measure. This same participant also scored higher in the two knowledge scales that both participants complete. The participant who graded the project problems more accurately did not complete the other two ARTIST scales. This may have been due to the classroom activities that were taking place during the administration of the survey to that particular section of 3213.

Chapter 5 Discussion and Conclusions

Preservice mathematics teacher educators, PMTEs, work to improve K-12 classroom learning by improving teacher preparation. Research shows that improved classroom learning occurs when progressive teaching methods such as discovery learning and other constructivist philosophies are utilized in place of traditional mathematics instruction (Carpenter, et al., 1989; Fennema, et al., 1993; Goldin, 1990; Steffe, 1990). Other research indicates that it is difficult for teachers to master, or even embrace, such philosophies until after the teacher has experienced years of classroom practice (Kagan,1992). This leads to a need that can only be fulfilled by CPD, continuing professional development.

Although CPD is not necessarily a service that all PMTEs provide, attitudes that teachers take with them as they leave their preservice training and enter the teaching field are very much related to the service that PMTEs provide (Wenger, 1998; Lave & Wenger, 1991). Research indicates that attitudes a person has toward a construct such as mathematics or statistics affect that person's willingness to participate in activities, like CPD, related to that construct (Bandura, 1986, 1997; Pajares, 1996). Other research indicates that the attitudes a teacher has toward specific mathematics topics affect that teacher's willingness to explore new methods of delivery (Lubinsky & Jaberg, 1997; Philippou & Christou, 2002).

Based on the results above, the attitudes preservice teachers develop and maintain are important. Since undergraduate students have been shown to often have negative attitudes toward statistics, PMTEs should take particular interest in preservice teacher attitudes toward statistics (Ball, 1990; Gal & Ginsberg, 1994; Rhoades, 2000). To improve preservice teacher attitudes toward statistics, it is useful to know if particular preservice teacher characteristics, such as statistical content knowledge and number of statistics courses taken, correlate with teacher attitudes toward statistics. When investigating such potential correlations, it is useful to avoid addressing attitudes toward statistics in too general of a manner. Research indicates that the results of attitudes measures are more useful if those measures are of specific types of attitudes rather than of general attitude measures (Finney & Schraw, 2002).

Due to a lack of studies that have addressed attitudes toward statistics in multiple ways, I determined to implement the study presented in this document. The unique service that this study provides is the multiple approaches toward measuring preservice teacher attitudes toward statistics. This study measured preservice teacher attitudes toward statistics using:

- an instrument that measures general attitudes toward statistics
- an instrument that measures self-efficacy to use current statistical knowledge
- an instrument that measures self-efficacy to learn statistical concepts
- an instrument that measures attitudes toward CPD in statistics.

To complement the measures of attitudes toward statistics, other measures were taken to check for potential correlations with any of the attitudes measured. These measures included participant knowledge of statistical content and the number of statistics courses the participant has taken. To provide comparative results, many of the instruments were administered to multiple participant groups. These groups included preservice primary teachers who are taking their first undergraduate mathematics course, preservice primary teachers who are taking their second mathematics content-for-teaching course, and preservice secondary teachers.

To search for details that may have provided insight into why any potential correlations exist, especially any potential correlations between preservice teacher content knowledge and preservice teacher attitudes, a qualitative instrument (the grading project) was included in the study. This instrument allowed me to inspect preservice teacher ability to grade student work. To provide a comparison for the results of the statistics grading skills, non-statistical mathematics content was included in the grading project.

Discussion of the Results

There were three primary investigations that occurred during the analyses of the data. I compared the means of each variable across the participant groups involved in the study to see if attitudes or knowledge levels might be different between two groups. I sought correlations between variables within particular groups, especially the 3213sp data set. The 3213sp data set was uniquely important because it was the one group that participated in the Attitudes Toward CPD in Statistics instrument and also had a large enough group (n=22) to provide potential statistical significance. This was also the group from which the two grading project participants emerged. Finally, I compared the qualitative results of the grading projects of the two participants. This qualitative analysis included a report of the similarities and differences between the two participant's works from the grading project. This analysis also included a comparison of each participant's work from the grading project and of each participant's results from the quantitative instruments.

Differences in the means of variables across participant groups. There were three variable for which the means were significantly different across participant groups. As a group, the participants in the fourth mathematics course for preservice primary teachers (3213all) scored significantly lower self-efficacy to learn statistics in the future (SELS) than did both the preservice primary education majors in the 1473 course (1473ed) and all of the 1473 course participants (1473). The current self-efficacy to use statistics for the preservice primary teachers in the 1473 course (1473ed) was significantly higher than that of the other 1473 participants (1473not). The ATSfield mean score was significantly higher for the participants from the 3213 section in semester 2 (3213sp) than they were for both the 1473 and 1473ed groups.

The lower SELS scores of the 3213 participants compared to the 1473ed participants indicates that preservice primary teachers at the University of Oklahoma lose confidence in their ability to learn statistics as they progress through their preservice teacher program. This could be due to participant exposure to more statistics. If more difficult statistics is studied in the later coursework, then the confidence to perform statistics could diminish as a result. Another explanation is that instructor methods were creating greater negative feeling toward the material. I observed more negative attitudes toward the study itself from the 3213 instructors than I did from the 1473 instructor. In fact, the 1473 instructor was very enthusiastic about the study. The 3213 instructors were less than enthusiastic.

The higher CSSE scores for the 1473ed participants compared to the 1473not participants might have been warranted based on statistics knowledge levels. The 14.5% higher (9.08 to 8.10) mean scores on the SCI (knowledge of statistics) for the 1473ed participants compared to the 1473not participants indicates that the preservice primary teachers have neither lower confidence to perform statistics nor lower statistics knowledge levels than other undergraduate majors who take 1473 at the University of Oklahoma. Results such as those from Ball (1990) that indicate weak mathematics skills from beginning preservice primary teachers might lead to assumptions that this group is weaker in mathematics, and possibly in statistics, than other undergraduate groups. These attitudes, CSSE, and knowledge, SCI, results were contrary to such possible assumptions.

The ATSfield means were highest for the 3213sp participants. The other ATSfield means, in descending order, were from the 3213fa, 1473not, and 1473ed participants. It was unclear why the 3213sp means were significantly higher than those of the 1473ed means while the 3213fa means were not significantly higher. As mentioned in chapter 4, potential reasons for this discrepancy are: (1) the instructor plays a role in such attitudes and (2) the results are an anomaly due to relatively small numbers of participants. The order of the ATSfield means from highest to lowest was 3213sp, 3213fa, 1473not, 1473ed; while the order of the SELS means from highest to lowest was 1473ed, 1473not, 3213sp, and 3213fa. Why these lists are in almost completely opposite order of each other is unclear. These lists appear to indicate that the two instruments measure different affective traits. The lists also indicate that the participants do not answer affective instruments about statistics in a general manner. The participants appear to carefully consider the questions and answer accordingly. Otherwise we would expect similar scores between the two instruments across participant sets.

Comparisons between the preservice secondary and preservice primary teachers. A secondary objective of the study was to compare the preservice secondary and preservice primary teacher results. Since the 3213sp19 participant shared more variables with the preservice secondary teachers than any other group, that participant set was used to represent the preservice primary teachers. Due to the higher level of mathematics preparation, I expected the preservice secondary teachers (432) to likely outscore the preservice primary teachers (3213sp19) in statistical knowledge. I also expected at least some of the attitudes to be better for the 432 group. This was based on the rationale that stronger knowledge levels might transfer to stronger attitudes.

With only an n of 6, the preservice secondary teachers scored significantly higher (at the α=.1 level) than the preservice primary teachers for the variables ATSfield, ATScourse, ATSsum, SELS, SCgraphs, SCcenter, SCspread, SCprob, and SCsum4. These included every variable tested except for CSSE. The SCsum4 mean difference indicated that the preservice secondary teachers did indeed have higher statistics knowledge levels than did the preservice primary teachers. Not only was CSSE not significantly higher for the 432 participants, the p-value was near 1 (p = .913). The 432 participants did have a higher measured average.

Why might the CSSE scores be near the same while all of the other scores are significantly higher for the 432 participants? There is no precedent in the literature for why both secondary and primary level preservice teachers would score similarly on one attitudes measure while the secondary teachers score significantly higher on the other attitudes measures. Perhaps the amount of mathematics the secondary teachers had learned gave them a confidence to learn statistics (SELS) that the primary teachers did not have. On the other hand, the lack of mathematics background may not have hindered the preservice primary teachers' confidence to perform statistical tasks (CSSE). This could be caused in part by the proverbial "the more you know, the more you realize how much you do not know." Thus the secondary teachers, while knowing more statistics, may not have any more confidence in their ability to do statistics.

Lack of evidence that full semester courses in statistics affect preservice teacher knowledge or attitudes. There was no evidence that having completed at least one full semester course in statistics improved preservice teacher knowledge of statistics. This conclusion is based on the absence of regression models using the COLLSTAT variable to predict any other variable. Such regression models were attempted using stepwise regression. The same result occurred with respect to preservice teacher attitudes toward statistics, including attitudes toward CPD in statistics. These results were consistent across all data sets.

Correlations between measured variables and statistical content knowledge within participant groups. The literature indicated that there was little correlation between preservice teacher attitudes toward statistics and their knowledge of statistics.

In large part, this study supports such results. Most of the affective measures could not predict knowledge levels in regression models. However, unlike previous studies, there was one predictor model found that linked attitudes to knowledge levels.

The 3213sp19 data set was the only preservice primary teacher group to complete the ARTIST scales. Recall that I chose only four of the eleven available scales for the 3213sp participants to complete. This was because the other seven scales contained statistical material beyond the expected statistical knowledge of this participant group. There were no reasonable predictor models generated for the dependent variable SCI (for the other preservice primary teacher groups). So there was no indication that any of the measured affective variables predicts the knowledge level as measured using the SCI instrument. I did, however, find a model to predict the SCsum4 score using the independent variable ATSfield. This model was generated using the 3213sp19 data. Recall that 3213sp19 was the set of 3213sp participants who completed all four ARTIST scales.

The model generated was a linear model using only ATSfield as an independent variable. The model predicted 62.3% of the variability in the SCsum4 scores. This indicates that there could be a correlation between the ATSfield variable and the SCsum4 variable for the 3213sp19 data set. If this correlation exists, then participants from the semester 2 preservice primary teacher course are more likely to understand statistical content when they have positive attitudes toward statistics as a field.

Although the ATSfield variable is from the Wise ATS instrument, which is a general measure of statistics attitudes, the ATSsum score – the total score on the instrument – was not a predictor of the SCsum4 scores. Only the more specific attitude measure, ATSfield, provided indication of knowledge levels. The other part of the instrument, attitudes toward the current statistics content course, was not correlated with SCsum4 knowledge levels. This supports the need to use more specific attitudes measures rather than general measures as has been suggested by Finney & Schraw (2002).

Correlations between measured variables and self-efficacy to learn statistics (SELS) within participant groups. The significant difference in the SELS scores between the participants from the first mathematics course for preservice primary teachers and the last mathematics course for preservice primary teachers (the second mathematics content-for-teaching course) made further investigation into SELS scores warranted. I sought to construct regression models using the measured variables to predict SELS scores. Because the primary focus of PMTEs would be the first and second mathematics content-for-teaching courses, I did not use the 1473 data for the model. The 3213fa data was used since this group had the lowest SELS scores. A model was found that used ATScourse and CSSE to predict SELS. This model explained 68.3% of the variance in the SELS scores for the 3213fa participants.
It is unclear why these variables correlate somewhat with the SELS scores while other variables do not. Future studies could have follow-up interviews that might add insight into such attitudes scores.

Correlations between measured variables and attitudes toward continuing professional development within participant groups. A primary focus of the study in phase 2 was the investigation into preservice teacher interest in CPD in statistics. Of particular interest were any potential correlations between preservice teacher interest in CPD in statistics and other preservice teacher characteristics such as statistical content knowledge, number of statistics courses taken, and various attitudes toward statistics. To search for such potential correlations, regression analyses were conducted using stepwise regression techniques to predict preservice teacher interest in CPD in statistics. It was necessary to determine a particular measure of interest in CPD in statistics. This measure was used as the dependent variable in the analyses. I determined to predict two variables in separate analyses, WKen and WKhp. The variable WKen represented participant beliefs that they would enjoy workshops that provide CPD in statistics. The variable WKhp represented participant beliefs that workshops providing CPD in statistics would help to improve the participant's classroom teaching.

Only moderate to weak predictor models could be generated to predict interest in CPD in statistics. The best model found for predicting WKen used ATScourse as the independent variable and only accounted for 22.8% of the variability in the WKen scores. This model was not considered practical for use to predict participant beliefs that they would enjoy workshops for CPD in statistics. The best model found for predicting WKhp used SELS scores and accounted for 47.3% of the variability in the WKhp scores. This model was considered to be of weak to moderate use in predicting participant beliefs that workshops for CPD in statistics will help their classroom teaching.

There were no correlations between ATScourse, CSSE, and WKhp. However, WKhp was only measured for the 3213sp participants (there was no WKhp data for the 3213fa participants). For the 3213fa participants, a model was found that used ATScourse and CSSE to predict SELS. This model explained 68.3% of the variation in the SELS scores. The results from the study indicate that, under certain circumstances, attitudes toward (1) the current statistics course (ATScourse), (2) self-efficacy to use statistics (CSSE), (3) self-efficacy to learn statistics (SELS), and (4) beliefs that CPD in statistics can help classroom teaching (WKhp), might be weakly to moderately correlated. Caution must be used when extrapolating results that were not consistent across all of the preservice primary teacher groups (3213fa and 3213sp). Table 5.1 illustrates the different results drawn from the 3213fa and 3213sp data sets.

Table 5.1 *Regression Models from the 3213fa and 3213sp Data Sets*

participant data set	3213fa	3213sp
results	ATScourse and CSSE together can predict SELS with 68.3% explained variance	SELS can predict WKen with 47.3 % explained variance

Results from the grading project. Pedagogical content knowledge is an important aspect of effective teaching (Shulman, 1986). I struggled to find a way to incorporate pedagogical content knowledge into the quantitative portion of the study. It is difficult to quantify pedagogical content knowledge. By definition, pedagogical content knowledge measures require not only a measure of the teacher, a human subject, it also requires a measure that involves students – or at least student work. Hence I determined that the pedagogical content knowledge analyses would be qualitative based on a grading project. Unfortunately, there was a low level of response for the grading project from potential participants.

The results of analyses for the two participants indicated that the preservice teacher (#21) who was less able to correctly grade 4th grade statistics problems was also the participant who indicated that they were more likely to need to look up the correct solutions. This same participant was often not inclined to look up solutions – even when the participant clearly understood that they did not know how to grade a particular problem. Yet this participant showed insight into some problems, such as the probability problem, requiring more than superficial procedural understanding of the material.

It is difficult to compare the motivation levels of the two participants. The participant (#21) who was weaker at grading correctly was more likely to ignore the need to look up the material required to help them correctly grade a problem. However, since the participant who was stronger at grading (#1) had much less need to look up material, it is unclear how motivated that participant would be to look up material if it became necessary.

Both participants were unable to grade every problem correctly, even though they had open access to any supplementary materials they desired. This is consistent with previous results that indicate preservice teachers often have weak understanding of mathematics and statistics for teaching (Ball, 1990; Mickelson & Heaton, 2004).

Compared to participant #1, participant #21, who was less capable of correctly grading 4^{th} grade student problems, indicated less interest in investigating the content to account for the lack of understanding. Participant #21 also had low attitudes (including self-efficacy) scores. This is consistent with previous results that indicate the confidence a person has in their ability to perform a task affect their willingness to participate in activities that involve the task.

Comparisons between the quantitative results and the grading project (qualitative results) revealed a consistency between several factors. Participant #21, who was unable to correctly grade as many of the problems on the grading project, was also the participant who had lower attitudes scores on every affective variable. In fact, not only did participant #21 have lower attitudes scores than did participant #1, participant #21 had attitudes variable scores much lower than the 3213sp mean while participant #1 had attitudes scores higher than the 3213sp mean. Although participant #1 did not complete all four ARTIST scales, as discussed in chapter 4, the participant #1 had higher scores than did participant #21 on the two scales that were completed by both. Compared to participant #21, participant #1 also had equal or higher attitudes toward CPD in statistics scores on every item from the ATPDS instrument.

Significance and Implications

The research question was, "Are there indicators from preservice teacher attitudes toward, and knowledge of, statistics that might assist PMTE efforts to increase the possibility that these preservice teachers will pursue CPD in statistics?"

The answer is yes. These indicators involve preservice teacher characteristics that may be correlated and preservice teacher characteristics that vary across participant groups.

The difference in the means of the SELS scores across all three preservice primary teacher participant groups has potential implications. It is not clear that instructor differences were the cause of the difference in the means. However, instructor differences may be the cause. There is evidence from the literature that instructor characteristics can play a role in teacher attitudes (Borasi et al., 1999; Kagan, 1992). Thus PMTEs can use the suggestive results of this study as reason to think about how their methods are affecting teacher confidence to learn statistics.

The SELS and WKhp correlation, although not a strong correlation, could be useful to PMTEs. If self-efficacy to learn statistics is an indicator of beliefs that CPD in statistics can help classroom teaching, then PMTEs could seek methods both to improve self-efficacy to learn statistics levels as well as to improve the beliefs that CPD in statistics will help classroom teaching. The levels of participant beliefs that they would enjoy CPD in statistics (WKen scores) are not included in the discussions of these correlations. There is a possibility that WKen scores are much less important than the WKhp scores. Consider this. If teachers learn to own the responsibility of delivering effective classroom teaching, then whether they will enjoy CPD in statistics should not be as important to them as whether they will be better classroom teachers due to CPD in statistics. In such a scenario, WKhp scores would be a better predictor of participation in CPD in statistics than would WKen scores.

There is evidence that teachers can be guided to own the delivery of effective classroom teaching. One type of teacher preparation that can help create such ownership is *reflective practices* (Borasi, et al., 1999; Krainer, 1999; Llinares, 2002). Teachers who use reflective practices constantly re-evaluate their teaching and consider ways that they can improve their teaching. PMTEs who teach preservice teachers to use reflective practices may directly improve preservice teacher interest levels in CPD. Also, teachers who leave preservice teacher preparation programs with reflective practices may be more likely to develop recognition of the importance of CPD in statistics. There are classroom needs that CPD in statistics are designed to address. As these needs develop in a teacher's classroom, the teacher will either recognize or fail to recognize these needs. A teacher who uses reflective practices is more likely to recognize such needs (Borasi, et al., 1999; Krainer, 1999; Llinares, 2002). A teacher who recognizes specific classroom needs may be more likely to pursue CPD to address those needs.

The majority of the results related to statistical knowledge and statistical pedagogical knowledge were consistent with previous results. One exception is that the attitudes toward statistics as a field, ASTfield, of preservice primary teachers in their last required mathematics course may be correlated to their statistics knowledge levels as measured by the sum of the four ARTIST scales data representation, measures of center, measures of spread, and probability. Thus it may be worth the effort for PMTEs of primary preservice teachers to attempt to nurture attitudes toward statistics as a field as well as to impart statistical knowledge. It should be noted that the study provides no evidence for explaining why the two variables are correlated. Maybe students who are naturally strong in statistics tend to look more favorably on the subject. Maybe students who appreciate the subject find added motivation to learn the subject. It is possible that a combination of both the previous cause-and-effect statements holds. Further investigation into the cause of this correlation is warranted.

Weaknesses of the Study

The small n-values for many of the data sets restricts the extent to which the results can be generalized. This situation was exacerbated by the fact that the 3213fa and 3213sp data sets did not contain the same sets of variables. This limited the number of participants to n=22 for certain variables, such as SCI scores, when the potential existed for an n of 44.

The study was investigatory. It was designed to search for evidence of potential correlations. A combination of low n-values and a lack of controlled structure limited the results to suggestive rather than conclusive. Although the study revealed actual differences in means across data sets, the conclusions that can be drawn from these differences are limited. This is in part because the study was not experimental. The data sets were not carefully controlled. Although there is potential for explaining results such as differences in SELS means by using important variables such as instructor characteristics, there are too many uncontrolled variables to make any assumption of the certainty of claimed cause and effect.

Ideas for Improving the Study

I chose to collect data that reflected a snapshot of a moment in time. Although the data was collected over two semesters, the data collected in semester 2 (3213sp) was of participants who were at the same level of progression in their teacher training program as other participants (3213fa). In this sense, the semester 2 data was designed to increase the n-value of preservice primary teachers. It did not provide longitudinal data. There were no participants other than those from 432 who provided data across both semesters. The 432 data for 4 of the participants was collected in both semesters, but the two different data collection dates were within two months of each other. For data which measure characteristics that are likely to change at semester rates rather than weekly rates, the 432 data remains a snapshot in time.

To create a more complete representation of the issues which this study addresses, it would be beneficial to administer a longitudinal study. Within such a study, preservice teachers could participate as they did for this study. However, in a longitudinal study, those preservice teachers could continue to participate in the study by reporting their participation in professional development after they enter the teaching field. Questionnaires could be distributed to previous participants requesting information related to the number of professional development opportunities they have been exposed to, how many of those opportunities involved mathematics, how many of those opportunities involved statistics, and how many of those opportunities they participated in within each of the above categories.

A longitudinal study could also have followed the preservice teachers through a full preservice teacher program. Such longitudinal data involving preservice primary teachers would also have provided valuable results concerning changing SELS scores as these students progressed through their preservice teacher program. Such a study might also allow for comparison of the effects of particular instructors, and those instructors' methods, on the preservice teacher attitudes.

It would have been beneficial to the study to have asked for participants from among the preservice teachers to keep a log during a semester of one of the mathematics content courses in which they participated. Within the log, they could be instructed to enter the mathematics topic(s) of focus for each class session, to enter their feelings related to how well they understood each topic before and after that session, to enter their feelings related to how intimidating the lesson felt, and to discuss alternative methods of learning the topic which they would find interesting. Although students may not be likely to keep diligent log entries, even sporadic entries would provide some extra insight as to how preservice teachers perceive their experience of learning mathematics concepts for K12 teaching.

The attitudes toward CPD in statistics measure would have been more meaningful if the results came from interviews as well as an instrument. The lack of interest from the 3213 students across both semesters limited the practicality of this research method. The inclusion of interview data is discussion in the Ideas for Improving the Study section below.

The attitudes toward CPD in statistics instrument, while short and simple, may have been more complicated than it should have been. There were four questions each related to participant willingness to participate in CPD in statistics. Each of these was based on various levels of support – either monetary or social. Such variation did not provide insight into participant attitudes. First, it is a risky business to rely on answers dependent on participant projections of whether they will participate in an activity in the future. The participants approached for the study were not in a position to participate in such CPD when they completed the survey. The two variables used to measure participant expectations (1) WKen – that they would enjoy CPD workshops, and (2) WKhp – that CPD workshops would improve classroom teaching, were more meaningful than the other four variables.

It would have been better to use 4^{th} grade appropriate language levels in the hand-written student answers for the grading project. Some of my answers were in a sentence structure more advanced than would be expected. A solution would be to seek permission to acquire anonymous solutions written by actual 4^{th} graders.

The instructions on the background/demographics questionnaire were more complicated than would be desired. It does not appear that any participants declined to complete the survey due to this condition. However, in future applications of this and similar questionnaires, such conditions should be avoided.

A significant improvement to this study would have been made if I could have found a way to increase the n-values of the 3213 and 432 participants. I was restricted by low enrollment levels and the desire to create a snapshot in time. To accumulate the data over a few years would introduce more unexplained variables. Theoretically, there were enough 3213 students (140) across the two semesters (35 per class x 2 sections x 2 semesters) to provide adequate n-values. At the observed participation rate of 63%, there would have been about 88 participants. The difficulties related to the troubled instructors affected this opportunity. The 432 participation rates were excellent, but there are few students who participate in that program.

Recommendations for Future Research

The different SELS scores between groups of preservice primary teachers who are beginning versus ending their mathematics training is a concern. Further studies must investigate the possibility that instructor practices are playing a role in such results. This investigation could be incorporated into a longitudinal study.

The longitudinal study designs discussed in the "Ideas for Improving the Study" section could be implemented in a future study. Those combinations of variables and participant groups that did not indicate potential for either correlations between variables or different means across participant groups could be diminished in future studies. Those combinations of variables and participant groups that did indicate potential for either correlations between variables or different means across participant groups could be expanded in future studies.

A longitudinal study could focus on the progress of preservice primary teacher SELS scores across the length of a teacher preparation program. This same study could investigate differences in SELS scores based on instructor characteristics and methods. A similar investigation could be made into ATSfield scores. For the latter variable, the longitudinal aspects would be diminished while the influence of the instructor characteristics and methods could be emphasized. The longitudinal study would reinforce or discredit the initial results from this study. Those initial results indicated that attitudes toward statistics as a field (ATSfield) can significantly increase as students progress through the mathematics courses in a preservice primary teacher program, but that such gains were not universal across all sections. Interviews with both instructors and their participant students would help to clarify whether the instructor was one of the causes behind the significantly higher ATSfield scores for the 3213sp participants over the 3213fa participants.

This study was broadly investigatory in nature. There were many indications from the study of potential cause-and-effect scenarios. The study indicated potential correlations between certain factors in preservice teacher statistics education. All of these indicated results need to be investigated in detail using studies that are not broadly investigatory, but rather highly specific with respect to the research questions asked, the participant groups approached, and the constructs investigated. These studies should seek to approach each specific question using multiple data collecting techniques. Some of these collection techniques should include repeated use of the instruments used in this study, new instruments that may prove useful, interview data, and open-ended questionnaires as appropriate.

This study revealed conditions within a preservice teacher preparation program. Each of these conditions, such as differences in SELS scores, were revealed due to the use of more specific attitudes measures than had been previously used. Such results can benefit preservice mathematics teacher educators as they seek ways to improve teacher attitudes toward CPD in statistics.

REFERENCES

Allen, K., (2006). The Statistics Concept Inventory: Development and Analysis of a Cognitive Assessment Instrument in Statistics. Unpublished doctoral dissertation University of Oklahoma.

Allen, K., A. Stone, T.R. Rhoads, and T.J. Murphy. (2004). The Statistics Concept Inventory: Developing a Valid and Reliable Instrument. Proceedings of the 2004 American Society for Engineering Education Annual Conference and Exposition. Session 3230.

American Mathematical Society: Conference Board of Mathematical Sciences (2001). The mathematical education of teachers. Washington, D.C.: Author.

American Statistical Association, (2005). Letter of endorsement for the GAISE reports for prek-12 and undergraduate statistics education. <http://it.stlawu.edu/~rlock/gaise/ASAEndorse.htm> (2006, January 25).

Ball, D. L. (1990). The mathematical understandings that prospective teachers bring to teacher education. *The Elementary School Journal*, 90, 449-465.

Ball, D.L. (1999) Crossing boundaries to examine the mathematical entailed in elementary teaching. *Contemporary Mathematics*, 243, 15-36.

Ball, D. L. (2000). Bridging practices: Intertwining content and pedagogy in teaching and learning to teach. *Journal of Teacher Education*, 51(3), 241-247.

Ball, D. L., & Cohen, D. K. (1999). Developing practice, developing practitioners: Toward a practice-based theory of professional education. In L. Darling-Hammond & G. Sykes (Eds.), Teaching as the learning profession: Handbook of policy and practice (pp. 3-32). San Francisco, CA: Jossey-Bass.

Ball, D. L., Ferrini-Mundy, J., Kilpatrick, J., Milgram, R. J., Schmid, M., & Schaar, R., (2005). Reaching for common ground in K–12 mathematics education. *Notices of the AMS*, 52(9), 1055-1058.

Ball, Lubienski, & Mewborn (2001). Research on teaching mathematics: The unsolved problem of teachers' mathematical knowledge. Handbook of Research on Teaching 4th Ed. 433-456, New York: Macmillan.

Ball, D. L. & Rowan, B. (2004). Introduction: Measuring instruction. *The Elementary School Journal*, 105 (1), 3-10.

Bandura, A. (1986). Social foundations of thought and actions: A social cognitive theory. Englewood Cliffs, NJ: Prentice Hall.

Bandura, A. (1997). Self-efficacy: The exercise of control. New York: Freeman.

Batanero, C., Godino, J. D., & Roa, R. (2004). Training teachers to teach probability. *Journal of Statistics Education*, 12(1).

Batanero, C., Merino, B., & Diaz, C., (2003). Assessing secondary student's understanding of averages. *European Research in Mathematics Education III*. <http://www.dm.unipi.it/~didattica/CERME3/proceedings/Groups/TG5/TG5_batanero_cerme3.pdf> (2006, January 25).

Batanero, C., & Serrano, L. (1999). The meaning of randomness for secondary school students. *Journal for Research in Mathematics Education, 30*(5), 558-567.

Begg, A. & Edwards, R. (1999). Teachers' ideas about teaching statistics. Paper presented at the joint conference of the AARE & NZARE, Melbourne.

Ben-Zvi, D. & Arcavi, A. (2001). Junior high school students' construction of global views of data and data representations. *Educational Studies in Mathematics, 45*,35-65.

Borasi, R., Fonzi, J., Smith, C. F. & Rose, B. J. (1999). Beginning the process of rethinking mathematics instruction: A professional development program. *Journal of Mathematics Teacher Education*, 2(1), 49–78.

Borko, H., Eisenhart, M., Brown, C. A., Underhill, R. G., Jones, D. & Agard, P. C. (1992). Learning to teach hard mathematics: Do novice teachers and their instructors give up too easily? *Journal for Research in Mathematics Education*, 23(2), 194–222.

Bright, G. W., Friel, S. (1993). Elementary teachers' representation of relationships among statistics concepts. Paper presented at the Annual Meeting of the American Educational Research Association, Atlanta GA.

Callingham, R.A., Watson, J.M., Collis, K.F. & Moritz, J.B. (1995). Teacher attitudes towards chance and data. In B. Atweh & S. Flavel (Eds.) Galtha. Proc. 18th Annual Conference of Mathematics Education Research Group of Australasia held at Northern Territory University, pp 143-150. MERGA: Darwin.

Canada, D. (2004). Preservice teachers' understanding of variation. Unpublished doctoral dissertation, Portland State University, Portland.

Carpenter, T. P. (1985). Learning to add and subtract: An exercise in problem solving.

In E. A. Silver (Ed.), Teaching and learning mathematical problem solving: Multiple research perspectives (pp. 17-40). Hillsdale, NJ: Erlbaum.

Carpenter, T. P., & Fennema, E. (1992) Cognitively Guided Instruction: Building on the knowledge of students and teachers. *International Journal of Education Research*, 17, 457-470.

Carpenter, T.P., Fennema, E., & Franke, M.L. (1996) Cognitively Guided Instruction: a knowledge base for reform in primary mathematics instruction. *The Elementary School Journal*,3-20.

Carpenter, T. P., Fennema, E., Peterson, P. L., & Carey, D. A. (1988). Teachers' pedagogical content knowledge of students' problem solving in elementary arithmetic. *Journal for Research in Mathematics Education*, 19(5), 385–401.

Carpenter, T. P., Fennema, E., Peterson, P. L., Chiang, C.-P., & Loef, M. (1989). Using knowledge of children's mathematical thinking in classroom teaching: An experimental study. *American Educational Research Journal*, 26, 499–531.

Carpenter, T. P. & Lehrer, R. (1999). Teaching and learning with understanding. In E. Fennema and T. Romberg (Eds.) Mathematics Classrooms that Promote Understanding, (pp.30-32). Mahwah, NJ: Erlbaum.

Carvalho, C. & Cesar, M., (2002). Sharing ideas and statistics learning: The role of peer interactions in school context. ICOTS6, 2002.

Cashin, S. E., & Elmore, P. B. (1997). Instruments used to assess attitudes toward statistics: A psychometric evaluation. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.

Cashin, S. E., & Elmore, P. B. (2005). The survey of attitudes toward statistics scale: A construct validity study. *Educational and Psychological Measurement,* Vol. 65 No. 3, June 2005 509-524.

Chadjipadelis, T., (1999). Teaching teachers to teach statistics. Presented at the 52nd session of the International Statistical Institute, Finland, 1999.

Cobb, P. (2000). Conducting teaching experiments in collaboration with teachers. In R. Lesh & A. Kelly (Eds.), Handbook of Research Design in Mathematics and Science Education (pp. 307-333). Mahwah, NJ: Lawrence Erlbaum Associates.

Cobb, P., & McClain, K. (2004). Principles of instructional design for supporting the development of students' statistical reasoning. In D. Ben-Zvi & J. Garfield (Eds.), The challenge of developing statistical literacy, reasoning, and thinking. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Cobb, G. W., & Moore, D. S. (1997). Mathematics, statistics, and teaching. *American Mathematical Monthly,* 104, 801-823.

Cobb, P. W., Yackel, E., & Wood, T., (1997). A Constructivist Alternative to the Representational View of Mind in Mathematics education. *Journal for Research in Mathematics Education*, 23(1), 2-33.

Cooney, T. J. (1994). Teacher education as an exercise in adaptation. In D. B. Aichele and A. Coxford (Eds.), Professional development for teachers of mathematics. Yearbook of the National Councel of Teachers of Mathematics (pp. 9-22). Reston, VA: NCTM

Cooney T.J., (1998). Extending our horizons. *Journal of Mathematics Teacher Education*, 1(2), 125-126.

Cooney, T. J. & Shealy, B. (1997). On Understanding the structure of teachers' beliefs and their relationship to change, in E. Fennema and B. S. Nelson (Eds.) Mathematics Teachers in Transition (pp 87-109). Mahwah, NJ: Lawrence Erlbaum Associates.

Eisenhart, M., Borko, H., Underhill, R., Brown, C., Jones, D., & Agard, P. (1993). Conceptual knowledge falls through the cracks: Complexities of learning to teach mathematics for understanding. *Journal for Research in Mathematics Education*, 24, 8-40.

Fennema, E., & Franke, M.L. (1992). Teachers' knowledge and its impact. In D. A. Grouws (Ed.), Handbook of research on mathematics teaching and learning (pp. 147-164). New York: Macmillian.

Fennema, E., Franke, M. L., Carpenter, T. P. & Carey, D. A. (1993). Using children's mathematics knowledge in instruction. *American Educational Research Journal*, 30, 555–583.

Finney, S.J., & Schraw, G. (2003). Self-efficacy beliefs in college statistics courses. *Contemporary Educational Psychology*, 28, 161 – 186.

Franke, M. L., Fennema, E., & Carpenter, T. (1997). Teachers Creating Change: Examining Evolving Beliefs and Classroom Practice. In E. Fennama & B. Scott Nelson (Eds.), *Mathematics Teachers in Transition* (pp. 255-282). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Friel, S. N. & Bright, G. W. (1998). Teach-Stat: A model for professional development in data analysis and statistics for teachers K-6. In S. P. Lajoie

(Ed.), Reflections on statistics: Learning, teaching, and assessment in grades K-12 (pp. 63-117). Mahwah, NJ: Lawrence Erlbaum Associates.

GAISE project. (2005a). A curriculum framework for prek-12 statistics education. A report presented to the american statistical association, March, 2005. <http://it.stlawu.edu/~rlock/gaise/GAISEPreK-12.pdf> (2006, January, 25)

GAISE project. (2005b). A curriculum framework for prek-12 statistics education: Executive summary. A report presented to the american statistical association, March, 2005. <http://it.stlawu.edu/~rlock/gaise/PreK-12Summary.pdf> (2006, January, 25)

GAISE project. (2005c). GAISE college report appendix. A report presented to the American Statistical Association in 2005. <http://it.stlawu.edu/~rlock/gaise/GAISECollegeAppendix.pdf> (2006, January, 25)

GAISE project. (2005d).GAISE college report. A report presented to the American Statistical Association in 2005. <http://it.stlawu.edu/~rlock/gaise>, <http://it.stlawu.edu/~rlock/gaise/GAISECollege.pdf> (2006, January, 25)

Gal, I. (2002). Adult's statistical literacy: Meanings, components, responsibilities, International Statistical Review, 70(1), 1-25.

Gal, I. (2004). Statistical literacy: Meanings, components, responsibilities. In D. Ben-Zvi & J. Garfield (Eds.), The challenge of developing statistical literacy, reasoning, and thinking. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Gal I. & Ginsberg, L., (1994). The role of beliefs and attitudes on learning statistics: Towards an assessment framework. *Journal of Statistics Education*, 2(2).

Garfield, J. (1995) How students learn statistics. *International Statistical Review*, 63, 25-34.

Garfield, J. (1998) Challenges in assessing statistical reasoning. AERA Annual Meeting presentation, San Diego.

Garfield, J. & Chance, B. L. (2000). Assessment in statistics education: Issues and challenges. *Mathematical Thinking and Learning, 2*(1&2), 99-126.

Goldin, G. (1990). Epistemology, constructivism, and discovery learning in mathematics. In the JRME Monograph number 4: Constructivist Views on the Teaching and Learning of Mathematics. Reston, VA: NCTM.

Goldin, G. (1998). Representational systems, learning, and problem solving in mathematics. *Journal of Mathematical Behavior*, 17(2), 137-165.

Goldin, G. (2002). Affect, meta-affect and mathematical belief structures, in G. C. Leder, E. Pehkonen & G. Torner (eds.), Beliefs: A hidden variable in mathematics education?, Vol. 31, Kluwer Academic Publishers, Dordrecht/Boston/London, pp. 59-72.

Gould, R. & Peck, R., (2005). Preparing Secondary Teachers to Teach Statistics. ICME 10, Copenhagen 2004

Greer, B. & Ritson, R. (1993). Teaching data handling with the Northern Ireland Mathematics Curriculum: Report on survey in schools. Belfast: Queen's University.

Guskey, T. R. (1986). Staff Development and the Process of Teacher Change. *Educational Researcher 15*(5), 5-12.

Heaton, R. (2000). Teaching mathematics to the new standards: Relearning the dance. New York: Teachers College Press.

Heaton, R., & Mickelson, W. (2002). The learning and teaching of statistical investigation in teaching and teacher education. *Journal of Mathematics Teacher Education*, *5*(1), 35-59.

Hiebert, J., (1999). Relationships between research and the NCTM standards. *Journal for Research in Mathematics Education*, 30(1), 3-19.

Hill, H., & Ball, D. L. (2004). Learning mathematics for teaching: Results from California's mathematics professional development institutes. *Journal for Research in Mathematics Education*, 35 (5), 330-351.

Hill, H., Rowan, B., & Ball, D. (2005). Effects of teachers' mathematical knowledge for teaching on student achievement. *American Educational Research Journal*.

Hill, H., Schilling, S., & Ball, D. (2004). Developing measures of teachers' mathematics knowledge for teaching. *The Elementary School Journal*, 105(1).

Hirabayashi, H. (1999). How do we arouse student's interest in statistics? Presented at ISI-52, August, 1999, Helsinki, Finland

Jones, G., Langrall, C., Thornton, C., & Mogill, A. T. (1999). Students' probabilistic thinking in instruction. *Journal for Research in Mathematics Education, 30*(5), 487-519.

Kagan, D. M. (1992). Professional Growth Among Preservice and Beginning Teachers. *Review of Educational Research 62*(2), 129-169.

Krainer, C. (1999). Promoting reflection and networking as an intervention strategy in professional development programs for mathematics teachers and mathematics teachers' educators. In O. Zaslavshy (Ed.), Proceedings of the Conference PME23 Vol. 1. (pp. 59-168), Haife: Technion.

Lajoie, S. P. (Ed.). (1998). Reflections on statistics: Learning, teaching, and assessment in grades K-12 (pp. 1-336). Mahwah, NJ: Erlbaum.

Comment [c1]: 97

Lajoie, S. P., (1999). Understanding of Statistics. In E. Fennema & T. A. Romberg (Eds.), Mathematics classrooms that promote understanding (pp.109-132). Mahwah, NJ: Erlbaum.

Lakoma, E., (2000). Stochastics Teaching and cognitive development. In John Fauvel and Jan van Maanen (Eds.), History in Mathematics Education (pp 74-77). Dordrecht, The Netherlands: Kluwer.

Landwehr, J., (2004). GAISE report discussion. Presented at the Joint Statistical Meetings, Toronto, Canada, August 10, 2004. <http://it.stlawu.edu/~rlock/gaise/LandwehrJSM04.ppt#293,5,Dissemination> (2006, January 25).

Lave, J. and Wenger, E. (1991). Situated Learning: Legitimate Peripheral Participation. New York: Cambridge University Press

Lehrer, R., & Schauble, L. (2000). Inventing data structures for representational purposes: Elementary grade students' classification models. *Mathematical Thinking and Learning, 2*(1&2), 51-74.

Li, J. (2004). Teaching approach: Theoretical or experimental? In Fan, L., Wong, N., Cai, J., Li, S. (Eds.), How Chinese Learn Mathematics: Perspectives from insiders. Hong Kong: World Scientific.

Llinares, S., (2002). Participation and reification in learning to teach: the role of knowledge and beliefs, in G. C. Leder, E. Pehkonen & G. Torner (eds.), Beliefs: A hidden variable in mathematics education?, Vol. 31, Kluwer Academic Publishers, Dordrecht/Boston/London, pp. 195-209.

Lubinski, C.A. & Jaberg, P.A. (1997). Teacher change and mathematics K-4: Developing a theoretical perspective, in E. Fennema and B. S. Nelson (Eds.)

Mathematics Teachers in Transition (pp 223-254). Mahwah, NJ: Lawrence Erlbaum Associates.

Maher, C., Alston, A. (1990). Teacher development in mathematics in a constructivist framework. In the JRME Monograph number 4: Constructivist Views on the Teaching and Learning of Mathematics. Reston, VA: NCTM.

Makar, K. (2004). Developing statistical inquiry: Prospective secondary mathematics and science teachers'I nvestigations of equity and fairness through analysis of accountability data. Doctoral Dissertation, University of Texas at Austin, Austin, Texas.

Makar, K., & Confrey, J. (2004). Secondary teachers' reasoning about comparing two groups. In D. Ben-Zvi & J. Garfield (Eds.), The challenge of developing statistical literacy, reasoning, and thinking. Dordrecht, the Netherlands: Kluwer Academic Publisher.

Mathbits. Retrieved March15, 2007, from http://mathbits.com/MathBits/TISection/Statistics2/correlation.htm

Mathematical Sciences Education Board. (1996). The preparation of teachers of mathematics: Considerations and challenges. National Research Council, Washington, D.C.: National Academy Press. <http://www.nap.edu/catalog/10055.html> (2005, November 14).

Mathematical Sciences Education Board (1997). Toward excellence in K-8 mathematics. <http://www.nap.edu/catalog/10054.html> (2005, November 14).

McClain, K., McGatha, M., Hodge, L., (2000). Improving data analysis. Mathematics *Teaching in the Middle School*. 5(8), 548-553.

McKnight, C., Crosswhite, F., Dossey, J., Kifer, E., Swafford, J., Trowers, K., & Cooney, T. (1987). The underachieving curriculum: Assessing U.S. school mathematics from an international perspective. Champaign, IL: Stipes Publishing Co.

Mendenhall, W. & Sincich T. (2003). Regression Analysis (6^{th} edition). Pearson Education: NJ.

Mickelson, W., & Heaton, R. (2004). Primary teachers' statistical reasoning about data. In D. Ben-Zvi & J. Garfield (Eds.), The challenge of developing statistical literacy, reasoning, and thinking. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Mvududu, Nyaradzo, (2003). A cross-cultural study of the connection between students' attitudes toward statistics and the use of constructivist strategies in the course. *Journal of Statistics Education*, 11(3).

National Center for Improving Student Learning and Achievement in Mathematics and Science, (2004). Designing statistics instruction for middle school students. *In Brief*, 4(1). Winter 2004.

National Center for Research in Mathematical Sciences Education, (1994a). Statistics and school mathematics. *NCRMSE Research Review*, 3(2), 1-5.

National Center for Research in Mathematical Sciences Education, (1994b). Integrating statistics into the school curriculum. *NCRMSE Research Review*, 3(2), 6-12.

National Council for Accreditation of Teacher Education. (2003). Programs for initial preparation of mathematics teachers. Reston, VA, Author.

National Council of Teachers of Mathematics. (1989). Curriculum and evaluation standards for school mathematics. Reston, VA: The National Council of Teachers of Mathematics.

National Council of Teachers of Mathematics. (1992). Data Analysis and Statistics across the Curriculum: Addenda Series, Grades 9–12. Reston, VA: Author.

National Council of Teachers of Mathematics. (2000*)*. Principles and Standards for school mathematics. Reston, VA: National Council of Teachers of Mathematics.

Pajares, F. (1996). Self-efficacy beliefs in academic settings. *Review of Educational Research*, 66(4), 543-578.

Peck, R. & Gould, R., (2005). Preparing Secondary Teachers to Teach Statistics: A Distance Education Model. Proceedings of the 55th Session of the International Statistical Institute, April 5-12, 2005. Sydney, Australia.

Philippou, G., & Christou, C., (1998). The effects of a preparatory mathematics program in changing prospective teachers' attutudes toward mathematics. *Educational Studies in Mathematics*, 35, 189-206.

Philippou, G., & Christou, C., (2002). A study of the mathematics teaching efficacy beliefs of primary teachers, in G. C. Leder, E. Pehkonen & G. Torner (eds.), Beliefs: A hidden variable in mathematics education?, Vol. 31, Kluwer

Academic Publishers, Dordrecht/Boston/London, pp. 211-231.

Rhoads, T. R., (2000).Student attitudes toward statistics before and after a computer-integrated introductory statistics course. *IEEE Transactions on Education*, 43(2), 182-187.

Rhoads, T. R., (2006).The chart in the appendixes: INTERNET DOWNLOAD from https://engineering.purdue.edu/SCI/results.htm

Richardson, V., (1996). The role of attitudes and beliefs in learning to teach. In J. Sikula, T Buttery, & E. Guyton (Eds.), Handbook of research on teacher education (pp. 102-119). New York: Macmillan.

Richardson, V., Anders, P., Tidwell, D., & Lloyd, C. (1991). The relationship between teachers' beliefs and practices in reading comprehension instruction. American *Educational Research Jounal*, 28, 559-586.

Schau, C., Stevens, J. J., Dauphinee, T. L., & DelVecchio, A. (1995) The development and validation of the Survey of Attitudes Towards Statistics. Educational and Psychological Measurement, 55, 868-875.

Schmidt, W. H., McKnight, C. C., & Raizon, S. A. (1996). A splintered vision: An investigation of U.S. science and mathematics instruction. Boston: Kluwar.

Schoenfeld, A. H. (1985). Mathematical Problem Solving. Orlando, FL: Academic Press.

Schoenfeld, A. H. (1994). What do we know about mathematics curricula? *Journal of Mathematical Behavior*, 13(1), pp. 55-80.

Shulman, L. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, 15(2), 4–14.

Shulman, L.(1987). Knowledge and Teaching: Foundations of the New Reform. Harvard Educational Review, 57 (1): 1-22.

Simon, M. (1993). Prospective elementary teachers’ knowledge of division. *Journal for Research in Mathematics Education*, 24, 233-254.

Simon, M. (1995). Reconstructing mathematics pedagogy from a constructivist perspective. *Journal for Research in Mathematics Education, 26*(2), 114-145.

Smith, G. (1998). Learning Statistics by Doing Statistics, *Journal of Statistics Education*, 6(3).

Sorto, M. A., (2004). Prospective middle school teachers' knowledge about data analysis and it's application to teaching. Doctoral Dissertation, Michigan State University, Lansing, MI.

Sorto M. A., & White, (2004). Statistical Knowledge for teaching. Presented at the ICME 10, Copenhagen 2004

Steffe, L. (1990). On the knowledge of mathematics teachers. In the JRME Monograph number 4: Constructivist Views on the Teaching and Learning of Mathematics. Reston, VA: NCTM.

Stigler, J.W., Heibert, J. (1997). Understanding and improving classroom mathematics instruction: An overview of the TMSS video study. *Phi Delta Kappan* 79(1), 14-21.

Stone, A. 2006. A Psychometric Analysis of the Statistics Concept Inventory. Unpublished doctoral dissertation. University of Oklahoma.

Tempelaar, D. (2003). Statistical reasoning and its relationship to attitudes towards statistics and achievement. Presented at the ISI 54th session, Berlin, Germany, 2003.

Thompson, P. W. (1984). Content versus method. *College Mathematics Journal, 15*(5), 394-395.

Thompson, A. G. (1992). Teachers' beliefs and conceptions: A synthesis of the research. In D. A. Grouws (Ed.), Handbook of research on mathematics teaching and learning

Tukey, J. W. (1977). Exploratory Data Analysis. Reading, MA: Addison-Wesley.

Vacc, N., & Bright, G. (1999). Elementary Preservice Teachers’ Changing Beliefs and Instructional Use of Children’s Mathematical Thinking. *Journal for Research in Mathematics Education*, 30(1), 89 – 110.

Vygotsky, L. S. (1962). Thought and language. Massachusetts: MIT

Vygotsky, L. S. (1978). Mind and society: The development of higher psychological process. Cambridge MA: Harvard University Press.

Watson, J. (2002). Can grade 3 students learn about variation?. Paper presented at the Sixth International Conference on Teaching Statistics, Cape Town.

Watson, J., Kelly, B., Callingham, R., Shaughnessy, J. M. (2003). The measurement of school students' understanding of statistical variation. International Journal of Mathematics Education Science and Technology, 34(1), 1-29.

Wenger, E. (1998). Communities of practice: Learning, meaning, and identity. New York: Cambridge University Press.

Wild, C. J., & Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. International Statistical Review, 67(3), 223-265.

Wilson, M. S., & Cooney, T. J. (2002). 'Mathematics teacher change and development. The role of beliefs', in G. C. Leder, E. Pehkonen & G. Torner (eds.), Beliefs: A hidden variable in mathematics education?, Vol. 31, Kluwer Academic Publishers, Dordrecht/Boston/London, pp. 127-148.

Wilson, P., (2004). GAISE report discussion. Presented at the Joint Statistical Meetings, Toronto, Canada, August 10, 2004. <http://it.stlawu.edu/~rlock/gaise/WilsonJSM04.ppt#265,1,GAISE Report Discussion> (2006, January 25).

Wilson, S. M., & Berne, J. (1999). Teacher learning and the acquisition of professional knowledge: An examination of research on contemporary professional development. *Review of Research in Education*, *24*, 173-209.

Wise, S. L. (1985). The development and validation of a scale measuring attitudes toward statistics. *Educational and Psychological Measurement*, 45, 401 – 405.

Zaslavsky, O., Leikin, R. (1999). Interweaving the training of mathematics teacher educators and professional development of mathmatics teachers. In O. Zaslavsky (Ed.) Proceedings of the Conference PME23 Vol 1. (pp. 143-158), Haifa: Technion.

Appendixes: Table of Contents

Appendix A

All Instruments in the Survey

Survey part 1

Background Information:

Demographic Questionnaire

and

Statistics and mathematics background information

adapted from
<Rhoads, T. R., et al, Statistics Concept Inventory, Summer 2005>

http://coecs.ou.edu/sci/

Demographic Questionnaire

Please indicate your gender (circle one): F M

Please indicate the student status that best describes you (circle one):
Freshman Sophomore Junior Senior Graduate Other

Please indicate your intended major (circle all that apply):
Mathematics / Math Education
Elementary Education
Fine or Applied Arts
Commerce or Business-Related Majors
Humanities, Liberal Arts, or Social Science (English, History, Psychology, Sociology, etc.)
Specify
Engineering
Physical Science - Specify topic:
Life Science – Specify topic:
Undecided
Other Specify:

Statistics and math information

What is your experience with statistics? (check all that apply)
________I studied some statistics in high school as part of another class.
________I took a statistics course in high school.
________I have studied some statistics in college as part of another class.
________I have taken a statistics course in college before this one.
________I have taken more than one statistics course in college before this one.
________I have had no statistics experience.
________I studied some statistics in one or more math classes before 9^{th} grade.

Mark only one of the following three choices.
________I studied some statistics in every mathematics class I had from first to sixth grades.
________I do not remember studying any statistics in the first to the sixth grades.
________I do not remember the first to the sixth grades very well.

What statistics course are you enrolled in currently? ________________________________
Is this the first time you are taking this statistics course? (circle one) Yes No
If you are repeating the course this semester, why? (circle one)
Failed the course the first time.
Dropped the course due to a failing grade.
Dropped the course for other reasons.
Did not fail the course, but am repeating it for other reasons.

Name two specific mathematics topics which you are likely to teach (appropriate age level) which you think
that you will enjoy teaching very much.

________________________________ ________________________________

Name two specific mathematics topics which you are likely to teach (appropriate age level) which you think
will be very hard for you to teach.

________________________________ ________________________________

Survey part 2

Current Statistics Self-Efficacy

adapted from
< Finney, S. J. & Schraw, G., Current Statistics Self-Efficacy, 2003>

https://ore.gen.umn.edu/artist//cse.html

Current Statistics Self-efficacy

Please rate your confidence in your **current ability** to successfully complete the following tasks. The item scale has 6 possible responses: (1) no confidence at all, (2) a little confidence, (3) a fair amount of confidence, (4) much confidence, (5) very much confidence, (6) complete confidence. For each task, please mark the one response that represents your confidence in your **current ability** to successfully complete the task.

	No confidence at all					Complete Confidence
1. Identify the scale of measurement for a variable.	1	2	3	4	5	6
2. Interpret the probability value (p-value) from a statistical procedure.	1	2	3	4	5	6
3. Identify if a distribution is skewed when given the values of three measures of central tendency.	1	2	3	4	5	6
4. Select the correct statistical procedure to be used to answer a research question.	1	2	3	4	5	6
5. Interpret the results of a statistical procedure in terms of the research question.	1	2	3	4	5	6
6. Identify the factors that influence power.	1	2	3	4	5	6
7. Explain what the value of the standard deviation means in terms of the variable being measured.	1	2	3	4	5	6
8. Distinguish between a Type I error and a Type II error in hypothesis testing.	1	2	3	4	5	6
9. Explain what the numeric value of the standard error is measuring.	1	2	3	4	5	6
10. Distinguish between the objectives of descriptive versus inferential statistical procedures.	1	2	3	4	5	6
11. Distinguish between the information given by the three measures of central tendency.	1	2	3	4	5	6
12. Distinguish between a population parameter and a sample statistic.	1	2	3	4	5	6
13. Identify when the mean, median and mode should be used as a measure of central tendency.	1	2	3	4	5	6
14. Explain the difference between a sampling distribution and a population distribution.	1	2	3	4	5	6

Survey part 3

Self-Efficacy to Learn Statistics

adapted from
< Finney, S. J. & Schraw, G., Self-Efficacy to Learn Statistics, 2003>

https://ore.gen.umn.edu/artist//cse.html

Self-efficacy to Learn Statistics

Please rate your confidence in **learning** the skills necessary while you're in this class to successfully complete the following tasks. The item scale has 6 possible responses: (1) no confidence at all, (2) a little confidence, (3) a fair amount of confidence, (4) much confidence, (5) very much confidence, (6) complete confidence. For each task, please mark the one response that represents your confidence in **learning** the skills necessary in this course to successfully complete the task.

	No confidence at all					Complete Confidence
1. Identify the scale of measurement for a variable.	1	2	3	4	5	6
2. Interpret the probability value (p-value) from a statistical procedure.	1	2	3	4	5	6
3. Identify if a distribution is skewed when given the values of three measures of central tendency.	1	2	3	4	5	6
4. Select the correct statistical procedure to be used to answer a research question.	1	2	3	4	5	6
5. Interpret the results of a statistical procedure in terms of the research question.	1	2	3	4	5	6
6. Identify the factors that influence power.	1	2	3	4	5	6
7. Explain what the value of the standard deviation means in terms of the variable being measured.	1	2	3	4	5	6
8. Distinguish between a Type I error and a Type II error in hypothesis testing.	1	2	3	4	5	6
9. Explain what the numeric value of the standard error is measuring.	1	2	3	4	5	6
10. Distinguish between the objectives of descriptive versus inferential statistical procedures.	1	2	3	4	5	6
11. Distinguish between the information given by the three measures of central tendency.	1	2	3	4	5	6
12. Distinguish between a population parameter and a sample statistic.	1	2	3	4	5	6
13. Identify when the mean, median and mode should be used as a measure of central tendency.	1	2	3	4	5	6
14. Explain the difference between a sampling distribution and a population distribution.	1	2	3	4	5	6

Survey part 4

Attitudes Toward Statistics

adapted from
<Wise, S. L., Attitudes Toward Statistics, 1985>

https://ore.gen.umn.edu/artist//ats.html

ATTITUDES TOWARD STATISTICS

Directions: For each of the following statements mark the rating category that most indicates how you currently feel about the statement. Please respond to all of the items.

	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
1. I feel that statistics will be useful to me in my profession.	______	______	______	______	______
2. The thought of being enrolled in a statistics course makes me nervous.	______	______	______	______	______
3. A good researcher must have training in statistics.	______	______	______	______	______
4. Statistics seems very mysterious to me.	______	______	______	______	______
5. Most people would benefit from taking a statistics course.	______	______	______	______	______
6. I have difficulty seeing how statistics relates to my field of study.	______	______	______	______	______
7. I see being enrolled in a statistics course as a very unpleasant experience.	______	______	______	______	______
8. I would like to continue my statistical training in an advanced course.	______	______	______	______	______
9. Statistics will be useful to me in comparing the relative merits of different objects, methods, programs, etc.	______	______	______	______	______
10. Statistics is not really very useful because it tells us what we already know anyway.	______	______	______	______	______
11. Statistical training is relevant to my performance in my field of study.	______	______	______	______	______
12. I wish that I could have avoided taking my statistics course.	______	______	______	______	______
13. Statistics is a worthwhile part of my professional training.	______	______	______	______	______
14. Statistics is too math oriented to be of much use to me in the future.	______	______	______	______	______
	Strongly				Strongly

	Disagree	Disagree	Neutral	Agree	Agree
15. I get upset at the thought of enrolling in another statistics course.	______	______	______	______	______
16. Statistical analysis is best left to the "experts" and should not be part of a lay professional's job.	______	______	______	______	______
17. Statistics is an inseparable aspect of scientific research.	______	______	______	______	______
18. I feel intimidated when I have to deal with mathematical formulas.	______	______	______	______	______
19. I am excited at the prospect of actually using statistics in my job.	______	______	______	______	______
20. Studying statistics is a waste of time.	______	______	______	______	______
21. My statistical training will help me better understand the research being done in my field of study.	______	______	______	______	______
22. One becomes a more effective "consumer" of research findings if one has some training in statistics.	______	______	______	______	______
23. Training in statistics makes for a more well-rounded professional experience.	______	______	______	______	______
24. Statistical thinking can play a useful role in everyday life.	______	______	______	______	______
25. Dealing with numbers makes me uneasy.	______	______	______	______	______
26. I feel that statistics should be required early in one's professional training.	______	______	______	______	______
27. Statistics is too complicated for me to use effectively.	______	______	______	______	______
28. Statistical training is not really useful for most professionals.	______	______	______	______	______
29. Statistical thinking will one day be as necessary for efficient citizenship as the ability to read and write.	______	______	______	______	______

Survey part 5

Knowledge of Statistics

adapted from
<Rhoads, T. R., et al, Statistics Concept Inventory, Summer 2005>

http://coecs.ou.edu/sci/

PLEASE DO NOT WRITE IN THE FOLLOWING TEST BOOKLET
(pages 10 to 20)

USE THE ANSWER SHEET ON PAGE 9

Please write your answers on the lines below. Do not write in the test booklet.

Answers

1. _________
2. _________
3. _________
4. _________
5. _________
6. _________
7. _________
8. _________
9. _________
10. ________
11. ________
12. ________
13. ________
14. ________
15. ________
16. ________
17. ________
18. ________
19. ________
20. ________
21. ________
22. ________
23. ________
24. ________
25. ________
26. ________
27. ________
28. ________
29. ________
30. ________
31. ________
32. ________
33. ________
34. ________
35. ________
36. ________
37. ________
38. ________

1. You are a doctor testing a blood-born disease. You know that in the overall population, 2 out of 100 people have the disease. All positives are accurately detected. You also know that the test returns a positive result for 5 out of 100 people tested who do not have the disease. Portions of the related contingency table are given below. What is the probability that a patient will test positive?

	Has the disease (+)	Does not have the disease (-)
Tests positive (+)		
Tests negative (-)		0.95*0.98
	0.02	

   a) 0.02
   b) 0.05*0.98
   c) 0.02 + 0.05*0.98
   d) 0.95*0.98
   e) 0.02 + 0.05

2. A certain diet plan claims that subjects lose an average of 20 pounds in 6 months on their plan. A dietitian wishes to test this claim and recruits 15 people to participate in an experiment. Their weight is measured before and after the 6-month period. Which is the appropriate test statistic to test the diet company's claim?

   a) two-sample Z test
   b) paired comparison t test
   c) two-sample t test

3. In practice, which data collection strategy would be the best way to estimate the *mean household income* in the United States? One should measure the income level of

   a) every individual within the United States
   b) every household within the United States
   c) 1500 randomly selected individuals in the United States
   d) 1500 randomly selected households in the United States
   e) 10 random individuals within each of 150 random US counties
   f) 10 random households within each of 150 random US counties

4. Which would be *more likely* to have 70% boys born on a given day: A small rural hospital or a large urban hospital?

   a) Rural
   b) Urban
   c) Equally likely
   d) Both are extremely unlikely

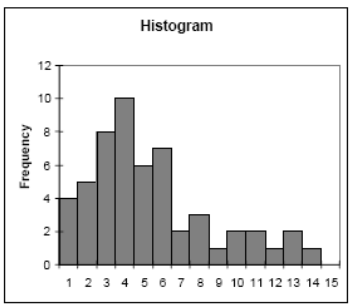
5. A coin of unknown origin is flipped twelve times in a row, each time landing with heads up. What is the most likely outcome if the coin is flipped a thirteenth time?

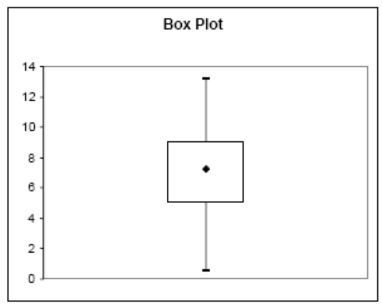
    a) Tails, because even though for each flip heads and tails are equally likely, since there have been twelve heads, tails is slightly more likely
    b) Heads, because this coin has a pattern of landing heads up
    c) Tails, because in any sequence of tosses, there should be about the same number of heads and tails
    d) Heads and tails are equally likely

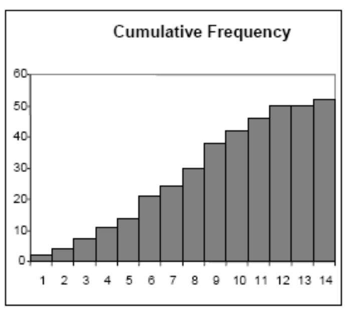
6. An Olympic track team consists of 6 sprinters (2 compete in the 100 meter event, 2 compete in the 200 meter event, and the remaining 2 compete in the 400 meter event). For which of the following samples would you expect to calculate the largest variance?

    a) A randomly selected sprinter's running times for 15 trials of the 200 meter event
    b) The track team's (all six members) running times for the 200 meter event
    c) A randomly selected sprinter's running times for 5 trials each of the 100 meter, 200 meter and 400 meter events
    d) The track team's running times for the 100 meter, 200 meter, and 400 meter events, each person running all three events

7. Three of the following are graphical presentations of the same set of data. Which of the graphs is of a different data set?







Stem	Leaf
0	55
1	79
2	124
3	1355
4	6
5	00235679
6	179
7	033348
8	0136799
9	00358
10	2679
11	1455
12	
13	22

    a) Histogram
    b) Box Plot
    c) Cumulative Frequency
    d) Stem and Leaf

8. A student scored in the 90^{th} percentile in his Chemistry class. Which is always true?

   a) His grade will be an A
   b) He earned at least 90% of the total possible points
   c) His grade is at least as high as 90% of his classmates
   d) None of these are always true

9. The following are temperatures for a week in August: 94, 93, 98, 101, 98, 96, and 93. By how much could the highest temperature increase without changing the median?

   a) Increase by 8°
   b) Increase by 2°
   c) It can increase by any amount.
   d) It cannot increase without changing the median.

10. A bottling company believes a machine is under-filling 20-ounce bottles. What will be the alternate hypothesis to test this belief?

    a) On average, the bottles are being filled to 20 ounces.
    b) On average, the bottles are not being filled to 20 ounces.
    c) On average, the bottles are being filled with more than 20 ounces.
    d) On average, the bottles are being filled with less than 20 ounces.

11. Which of the following statistics is least impacted by extreme outliers?

    a) range
    b) 3rd quartile
    c) mean
    d) variance

12. A student attended college A for two semesters and earned a 3.24 GPA (grade point average). The same student then attended college B for four semesters and earned a 3.80 GPA for his work there. How would you calculate the student's GPA for all of his college work? Assume that the student took the same number of hours each semester.

    a) $\dfrac{3.24+3.80}{2}$

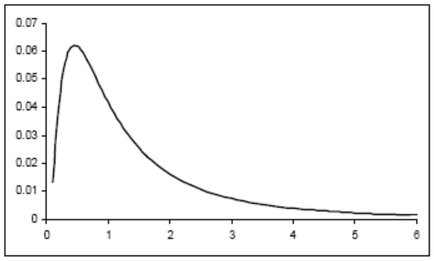
    b) $\dfrac{3.24(2)+3.80(4)}{2}$

    c) $\dfrac{3.24(2)+3.80(4)}{6}$

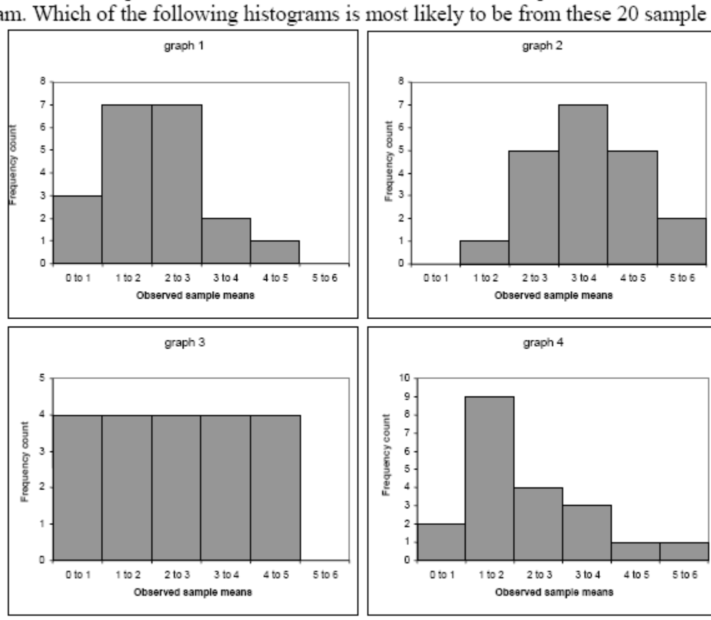
    d) It is not possible to calculate the students overall GPA without knowing his GPA for each individual semester.

13. You have called your cell phone provider to discuss a discrepancy on your billing statement. Your call was received and placed on hold to "await the next available service representative". You are told that the average waiting time is 6 minutes. You have been waiting on hold for 4 minutes. How many more minutes do you anticipate you will have to wait before you speak to a service representative?

   a) 2
   b) 4
   c) 6
   d) there is no way to estimate



14. From the above probability density function, 10 random data points are drawn and the mean is computed. This is repeated 20 times. The observed means were placed into six bins to construct a histogram. Which of the following histograms is most likely to be from these 20 sample means?



   a) graph 1
   b) graph 2
   c) graph 3
   d) graph 4

15. For the following set of numbers, which measure will most accurately describe the central tendency?
3, 4, 5, 6, 6, 8, 10, 12, 19, 36, 83

a) Mean
b) Median
c) Mode
d) Standard deviation

16. A standard deck of 52 cards consists of 13 cards in each of 4 suits: hearts (♥), diamonds (♦), clubs (♣), and spades (♠). Five separate, standard decks of cards are shuffled and the top card is drawn from each deck. Which of the following sequences is least likely?

a) ♥♥♥♥♥
b) ♣♦♥♠♣
c) ♠♥♠♥♠
d) All three are equally likely.

17. A researcher conducts an experiment and reports a 95% confidence interval for the mean. Which of the following must be true?

a) 95% of the measurements can be considered valid
b) 95% of the measurements will be between the upper and lower limits of the confidence interval
c) 95% of the time, the experiment will produce an interval that contains the population mean
d) 5% of the measurements should be considered outliers

18. A researcher performs a t-test to test the following hypotheses:

$$H_0 : \mu \leq \mu_0$$
$$H_1 : \mu > \mu_0$$

He rejects the null hypothesis and reports a p-value of 0.10. Which of the following must be correct?

a) The test statistic fell within the rejection region at the $\alpha = 0.05$ significance level
b) The power of the test statistic used was 90%
c) Assuming H_0 is true, there is a 10% possibility that the observed value is due to chance
d) The probability that the null hypothesis is not true is 0.10
e) The probability that the null hypothesis is actually true is 0.9

19. Which is true of a t-distribution?

a) It is used for small samples
b) It is used when the population standard deviation is not known
c) It has the same basic shape as a normal distribution but has less area in the tails
d) a & b are both true
e) a, b & c are all true

20. The mean height of American college men is 70 inches, with standard deviation 3 inches. The mean height of American college women is 65 inches, with standard deviation 4 inches. You conduct an experiment at your university measuring the height of 100 American men and 100 American women. Which result would *most* surprise you?

a) One man with height 79 inches
b) One woman with height 74 inches
c) The average height of women at your university is 68 inches
d) The average height of men at your university is 73 inches

21. A meteorologist predicts a 40% chance of rain in London and a 70% chance in Chicago. What is the most likely outcome?

a) It rains only in London
b) It rains only in Chicago
c) It rains in London and Chicago
d) It rains in at least one city

22. You perform the same two significance tests on large samples from the same population. The two samples have the same mean and the same standard deviation. The first test results in a p-value of 0.01; the second, a p-value of 0.02. The sample mean is equal for the 2 tests. Which test has a larger sample size?

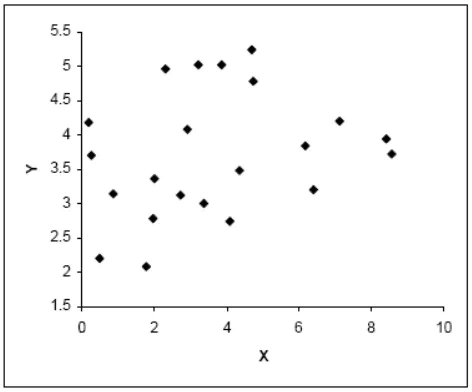
a) First test
b) Second test
c) Sample sizes equal
d) Sample sizes are not equal but there is not enough information to determine which sample is larger

23. Which statistic would you expect to have a normal distribution?

I) Height of women
II) Shoe size of men
III) Age in years of college freshmen

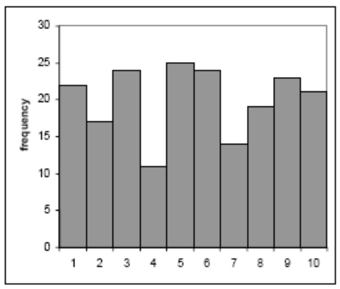
a) I & II
b) II & III
c) I & III
d) All 3

24. Estimate the correlation coefficient for the two variables X and Y from the scatter plot below.



a) -0.3
b) 0
c) 0.3
d) 0.9
e) 1.6

25. Consider the sample distribution below. This sample was *most likely* taken from what kind of population distribution?



a) Normal
b) Uniform
c) Skewed
d) Bimodal

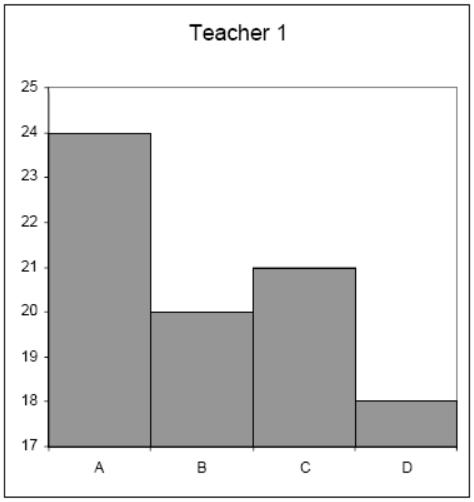
26. You have a set of 30 numbers. The standard deviation from these numbers is reported as zero. You can be certain that:

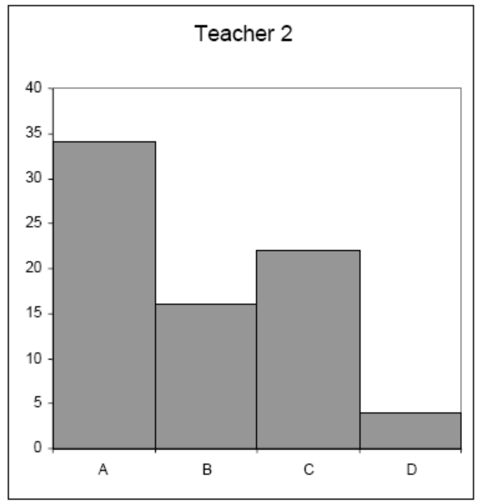
a) Half of the numbers are above the mean
b) All of the numbers in the set are zero
c) All of the numbers in the set are equal
d) The numbers are evenly spaced on both sides of the mean

27. In order to determine the mean height of American college students, which sampling method would *not* introduce bias?

   a) You randomly select from the university basketball team
   b) You use a random number table to select students based on their student ID
   c) You roll a pair of dice to select from among your friends
   d) None of the methods will have bias

28. The following histograms show the number of students receiving each letter grade for two separate physics classes. Which conclusion about the grades is valid?



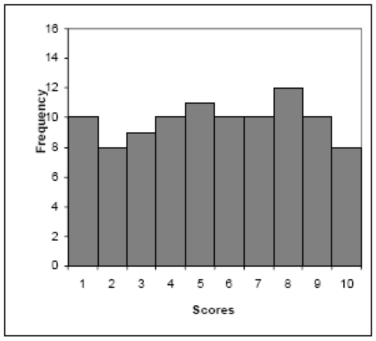


   a) Teacher 1 gave more B's and C's but approximately the same number of A's and D's as Teacher 2
   b) Teacher 2 gave more A's and fewer D's than Teacher 1
   c) Teacher 2 gave more B's and C's than Teacher 1
   d) The overall grade distribution for the two Teachers is approximately equal

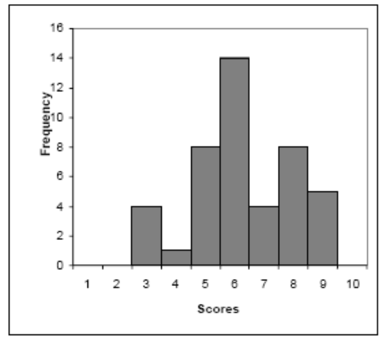
29. A scientist takes a set of 50 measurements. The standard deviation is reported as -2.30. Which of the following must be true?

   a) Most of the measurements were negative
   b) All of the measurements less than the mean
   c) All of the measurements were negative
   d) The standard deviation was calculated incorrectly

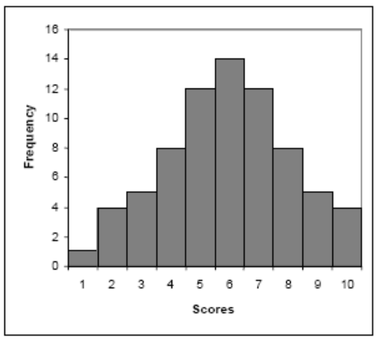
30. The following are histograms of quiz scores for four different classes.
Which distribution shows the most variability?



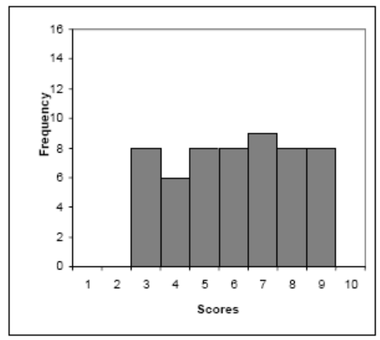
I



II



III



IV

a) I
b) II
c) III
d) IV

31. In a manufacturing process, the error rate is 1 in 1000. However, errors often occur in groups, that is, they are not independent. Given that the previous output contained an error, what is the probability that the next unit will also contain an error?

a) Less than 1 in 1000
b) Greater than 1 in 1000
c) Equal to 1 in 1000
d) Insufficient information

32. An engineer performs a hypothesis test and reports a p-value of 0.03. Based on a significance level of 0.05, what is the correct conclusion?

a) The null hypothesis is true.
b) The alternate hypothesis is true.
c) Do not reject the null hypothesis.
d) Reject the null hypothesis.

33. For the past 100 years, the average high temperature on October 1 is 78° with a standard deviation of 5°. What is the probability that the high temperature on October 1 of next year will be between 73° and 83°?

a) 0.68
b) 0.95
c) 0.997
d) 1.00

34. You are rolling dice. You roll 2 dice and compute the mean of the number rolled, then 6 dice and compute the mean, then 10 dice and compute the mean. One of the rolls has an average of 1.5. Which trial would you be *most surprised* to find this result?

a) Rolling 2 dice
b) Rolling 6 dice
c) Rolling 10 dice
d) There is no way this can happen

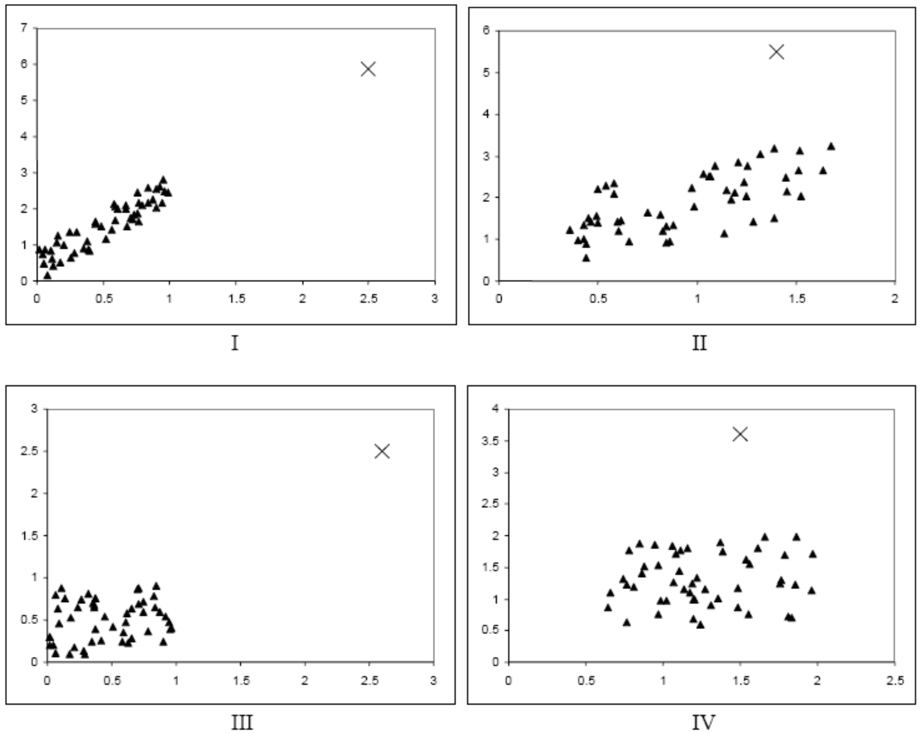
35. Two confidence intervals are calculated for two samples from a given population. Assume the two samples have the same standard deviation and that the confidence level is fixed. Compared to the smaller sample, the confidence interval for the larger sample will be:

a) Narrower
b) Wider
c) The same width
d) It depends on the confidence level

36. A company has decided to begin producing a new product. They want to use existing equipment. An engineer is assigned to determine which of two equipment settings will yield the highest quality product. He performs ten runs at each of the settings and measures the quality. Which test is most appropriate for this analysis?

a) two-sample Z test
b) paired comparison t test
c) two-sample t test
d) one-sample t test

37. Consider the correlation coefficients of the scatter plots below. If the data point that is marked by an × is *removed*, which of the following statements would be true?



a) correlation of (I) decreases, correlation of (II) stays the same
b) correlation of (III) increases, correlation of (IV) increases
c) correlation of (I) stays the same, correlation of (III) decreases
d) correlation of (II) increases, correlation of (III) increases

38. Information about different car models is routinely printed in public sources such as *Consumer Reports* and new car buying guides. Data was obtained from these sources on 1993 models of cars. For each car, engine size in liters was compared to the number engine revolutions per mile. The correlation between the two was found to be -0.824. Which of the following statements would you most agree with?

a) A car with a large engine size would be predicted to have a high number of engine revolutions per mile
b) A car with a large engine size would be predicted to have a low number of engine revolutions per mile
c) Engine size is a poor predictor of engine revolutions per mile
d) Engine size is independent of engine revolutions per mile

Answer Sheet for Correlating Participant Answers to the ARTIST Scales

<Lancaster, S. M., 2006>

ARTIST scales

Knowledge of statistics measures

Write the answer you give for each online question on the correct space below.

Data representation

1	
2	
3	
4	
5	
6	
7	
8	
9	
10	
11	
12	
13	

Measures of center

1	
2	
3	
4	
5	
6	
7	

Measures of spread

1	
2	
3	
4	
5	
6	
7	
8	
9	
10	
11	
12	
13	
14	

Probability

1	
2	
3	
4	
5	
6	
7	
8	
9	

Sampling variability

1	
2	
3	
4	
5	
6	
7	
8	
9	
10	
11	
12	
13	
14	
15	

Data collection

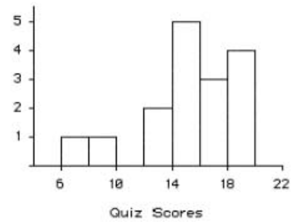
1	
2	
3	
4	
5	
6	
7	
8	
9	

Confidence intervals, one-sample

1	
2	
3	
4	
5	
6	
7	
8	
9	
10	

ARTIST SCALE: **DATA REPRESENTATION**

1. Scores for a quiz were calculated as the number of correct responses. Below is a graphical display of the quiz scores. How many of the scores are above 15? (Note: all scores are integers and bars begin at left endpoints)



a. 6

B. 7

c. 12

d. 13

e. Can't be determined.

2. In order to determine which kind of data display (e.g., histogram versus bar graph) is appropriate for a given variable, one should consider which of the following:

A. whether the relevant variable is quantitative or categorical

b. whether the study is observational or experimental

c. the range of the data

3. A class survey asked students to indicate if they are MAC or PC users. Of the following graphs, which is most appropriate to display their results?

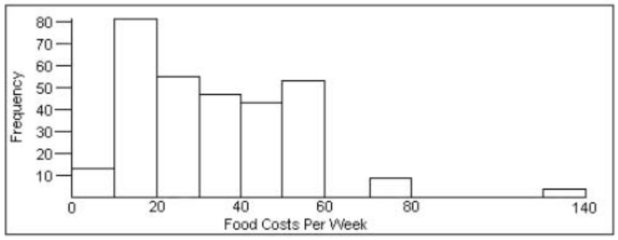
A. Pie chart

b. Histogram

c. Either a pie chart or a histogram

d. None of the above

4. A college statistics class conducted a survey. They gathered data from a large random sample of students who estimated how much money they typically spent each week in different categories (e.g., food, entertainment, etc.). A distribution of the survey results is presented below. One student claims the distribution of food costs basically looks bell-shaped, with one outlier. How would you respond?



a. Agree, it looks pretty symmetric if you ignore the outlier.

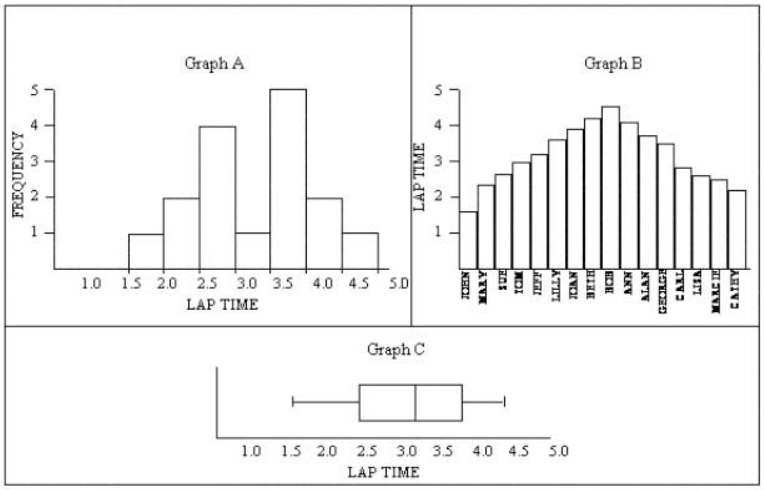
b. Agree, most distributions are bell-shaped.

c. Disagree, it looks more skewed to the left.

D. Disagree, it looks more skewed to the right.

e. Disagree, it looks more bimodal.

Items 5 and 6 refer to the following situation:
A local running club has its own track and keeps accurate records of each member's individual best lap time around the track, so members can make comparisons with their peers. Here are graphs of these data.



5. Which of the above graphs allows you to most easily estimate the median running time.

   a. Graph A.

   b. Graph B.

   C. Graph C.

   d. All of the above.

6. Which of the above graphs allows you to most easily see the shape of the distribution of running times?

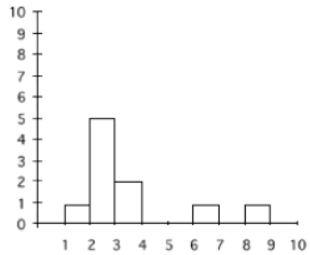
   A. Graph A.

   b. Graph B.

   c. Graph C.

   d. All of the above.

Items 7 to 9 refer to the following situation:
Here is a histogram for a set of test scores from a 10-item makeup quiz given to a group of students who were absent on the day the quiz was given.

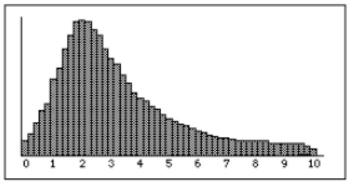


7. What do the numbers on the vertical axis represent?
   a. Independent variable
   b. Scores on the test
   c. Dependent variable
   D. Number of Students

8. How many people received scores higher than 4?
   a. 1
   B. 2
   c. 3
   d. 4

9. How many people took the test and have scores represented in the graph?
   a. 5
   B. 10
   c. 20
   d. 30

10. Select the description that best represents the shape of the following distribution.



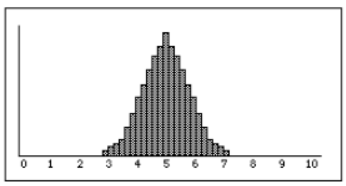
a. Left (negatively) skewed

B. Right (positively) skewed

c. Normal leaning right

d. Normal leaning left

11. Select the description that best represents the shape of the following distribution.



A. Normal

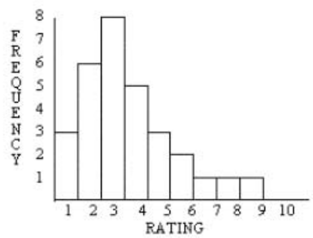
b. Skewed

c. Bimodal

d. Uniform

Items 12 and 13 refer to the following situation:
One of the items on the student survey for an introductory statistics course was "Rate your aptitude to succeed in this class on a scale of 1 to 10" where 1 = Lowest Aptitude and 10 = Highest Aptitude. The instructor examined the data for men and women separately. Below is the distribution of this variable for the 30 women in the class.

12. How should the instructor interpret the women's perceptions regarding their success in the class?

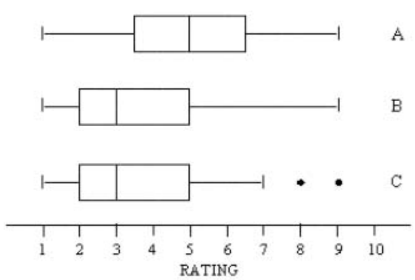


A. A majority of women in the class do not feel that they will succeed in statistics although a few feel confident about succeeding.

b. The women in the class see themselves as having lower aptitude for statistics than the men in the class.

c. If you remove the three women with the highest ratings, then the result will show an approximately normal distribution.

13. Which of the following boxplots represents the same data set as the histogram shown above?



a. Graph A.

B. Graph B.

c. Graph C.

ARTIST SCALE: **MEASURES OF CENTER**

1. The school committee of a small town wanted to determine the average number of children per household in their town. They divided the total number of children in the town by 50, the total number of households. Which of the following statements must be true if the average children per household is 2.2 children?

   a. Half the households in the town have more than 2 children.

   B. There are a total of 110 children in the town.

   c. The most common number of children in a household is 2.2.

   d. None of the above.

2. The distribution of the top 1% of individual incomes in the US is strongly skewed to the right. In 1997, the two measures of center for the top 1% of individual incomes were $330,000 and $675,000. Which number represents the mean income of the top 1% and which number represents the median income of the top 1%? Choose the best answer.

   a. $330,000 is the mean and $675,000 is the median.

   B. $330,000 is the median and $675,000 is the mean.

   c. Not enough information to tell which is which.

3. For this graphical display of Quiz Scores, which estimates of the mean and median are most plausible?

   a. median = 13 and mean = 12

   b. median = 14 and mean = 15

   C. median = 15.5 and mean = 14.1

   d. median = 16.5 and mean = 16.2

4. You give a test to 100 students and determine the median score. After grading the test, you realize that the 10 students with the highest scores did exceptionally well. You decide to award these 10 students a bonus of 5 more points. The median of the new score distribution will be ___________ that of the original score distribution.

   a. lower than

   B. equal to

   c. higher than

   d. depending on skewness, higher or lower than

Items 5 and 6 refer to the following situation:
A college statistics class conducted a survey of how students spend their money. They gathered data from a large random sample of college students who estimated how much money they typically spent each week in different categories (e.g., food, entertainment, etc.). The following statistics were calculated for money spent weekly on food: mean = \$31.52; median = \$30.00; interquartile range = \$34.00; standard deviation = \$21.60; range = \$132.50.

5. A student states that the median food cost tells you that a majority of students in this sample spend about \$30 each week on food. How do you respond?

   a. Agree, the median is an average and that is what an average tells you.

   b. Agree, \$30 is representative of the data.

   c. Disagree, a majority of students spend more than \$30.

   D. Disagree, the median tells you only that 50% of the sample spent less than \$30 and 50% of the sample spent more.

6. The class determined that a mistake had been made and a value entered as 138 should have been entered as 38. They recalculate all of the statistics. Which of the following would be true?

   a. The value of the median decreases, the value of the mean stays the same.

   b. The values of the median and mean both decrease.

   C. The value of the median stays the same, the value of the mean decreases.

7. The Sydney Morning Herald - February 8, 2004 reported that in 1961, the average number of children born to Australian women (3.55) was at its highest level since reliable records began in the 1920s. With only the fact of the mean being 3.55 children, can one make the claim that women were statistically more likely to have four children than any other number of children? Choose the best response.

   a. Agree, 4 children is the most typical number of children.

   b. Agree, 3.55 is closer to 4 than any other whole number.

   c. Disagree, the most typical number of children is 3.

   D. Disagree, it is impossible to tell which number of children is most likely.

ARTIST SCALE: **MEASURES OF SPREAD**

1. A class of 30 introductory statistics students took a 15 item quiz, with each item worth 1 point. The standard deviation for the resulting score distribution is 0. You know that:

   a. about half of the scores were above the mean.

   b. an arithmetic error must have been made.

   C. everyone correctly answered the same number of items.

   d. the mean, median, and mode must all be 0.

2. The 30 introductory statistics students took another quiz worth 30 points. On this quiz, the standard deviation of the scores of that quiz was 1 point. Which of the following gives the most suitable interpretation?

   a. all of the individual scores are one point apart

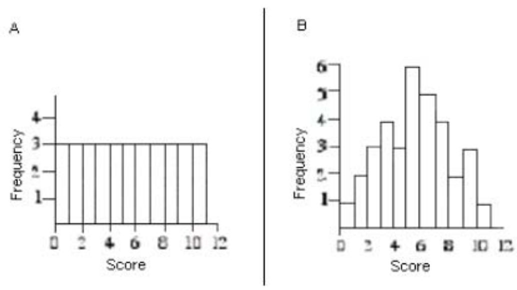
   b. the difference between the highest and lowest score is 1

   c. the difference between the upper and lower quartile is 1

   D. a typical score is within 1 point of the mean

Items 3 and 4 refer to the following situation:
For each pair of graphs, determine which graph has the higher standard deviation (it is not necessary to do any calculations to answer these questions).

3. Which distribution has the higher standard deviation.

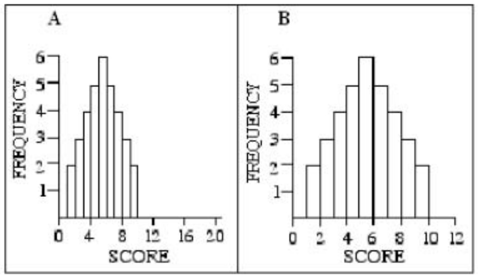


   A. A has a larger standard deviation than B.

   b. B has a larger standard deviation than A.

   c. Both graphs have the same standard deviation.

4. Which distribution has the higher standard deviation.



a. A has a larger standard deviation than B

b. B has a larger standard deviation than A

C. Both graphs have the same standard deviation

5. A teacher gives a 15 item science test. For each item, a student receives one point for a correct answer; 0 points for no answer; and loses one point for an incorrect answer. Total test scores could range from +15 points to -15 points. The teacher computes the standard deviation of the test scores for the class to be -2.30. What do we know?

A. The standard deviation was calculated incorrectly.

b. Most students received negative scores.

c. Most students scored below the mean.

d. None of the above.

6. Consider two populations in the same state. Both populations are the same size (22,000). Population 1 consists of all students at the State university. Population 2 consists of all residents in a small town. Consider the variable Age. Which population would most likely have the largest standard deviation?

a. Population 1 would more likely have a higher standard deviation(SD) than Population 2.

B. Population 2 would more likely have a higher standard deviation(SD) than Population 1.

c. They would likely have the same standard deviation(SD) for age because they have the same population size.

d. There is not enough information to tell.

Items 7 and 8 refer to the following situation:
For each list of test scores presented below (List A and List B), select the best estimate for the standard deviation. The mean for each list is 50. No calculations are required to answer these questions.

7. LIST A: 49, 51, 49, 51, 49, 51, 49, 51, 49, 51

   A. 1

   b. 2

   c. 5

   d. 10

8. LIST B: 31, 36, 48, 50, 50, 53, 54, 56, 60, 62

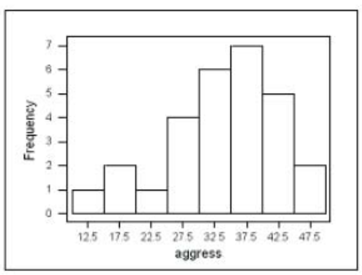
   a. 1

   b. 3

   C. 8

   d. 20

9. A test to measure aggressive tendencies was given to a group of teenage boys who were members of a street gang. The test is scored from 10 to 60, with a high score indicating more aggression. The histogram represents the results for these 28 boys. Which two measures would be most appropriate to describe center and spread for this distribution?



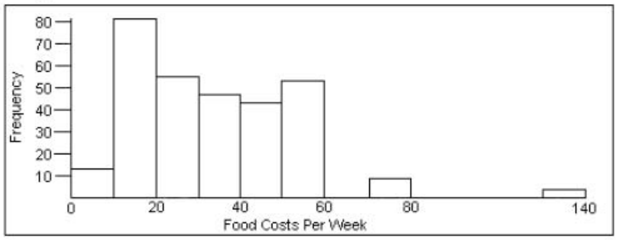
   a. Range and mean

   b. Mean and median

   C. Median and IQR

   d. Mean and standard deviation

Items 10 and 11 refer to the following situation:
This is a distribution of how much money was spent per week for a random sample of college students.



10. The range for this distribution is \$132.50. Indicate your agreement or disagreement with the following statement: The range is not a useful summary of the variability of this data set.

    a. Agree, it is too vague.

    B. Agree, it is too easily influenced by outliers.

    c. Agree, it does not use information on the center of the data.

    d. Disagree, a range of \$132.50 is a good measure of variability because students are apt to spend any amount of money between \$0 and \$132.50.

11. What is the best measure to use to summarize the variability of this data set?

    a. Range, because it tells you the overall spread of the data.

    b. Standard deviation, because it is based on all the information in the data set.

    c. Standard deviation, because it is the most commonly used measure of variability.

    D. Interquartile range, because it is resistant to outliers.

12. A random sample was taken to determine the left foot length of female bears based on measuring their tracks. The following statistics were calculated for this sample: Mean = 12.8 inches, median = 12.7 inches, standard deviation = 1.4 inches, interquartile range = 2 inches. The distribution is mound-shaped and symmetric. Based only on this information, choose the best estimates for the minimum and maximum values of the distribution.

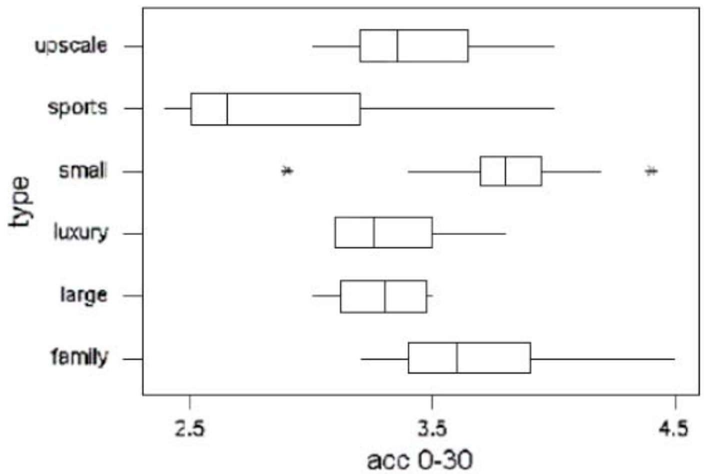
    a. min = 11.4 and max = 14.2

    b. min = 10.7 and max = 14.7

    C. min = 8.6 and max = 17.0

    d. min = 5.0 and max = 20.0

Items 13 and 14 refer to the following situation:
The 1999 Consumer Reports new Car Buying Guide reported on the number of seconds required for a variety of cars to accelerate from 0 to 30 mph. The cars were also classified into six categories according to type. The following boxplots display the distribution of acceleration times for each type of car:



13. Which type of car has the smallest interquartile range for number of seconds to accelerate?

    a. Upscale.
    b. Sports.
    C. Small.
    d. Luxury.
    e. Large.
    f. Family.

14. If the outliers were removed from the dataset of Small cars, which of the following statistics would be least affected?

    a. Range.
    B. IQR.
    c. Standard Deviation.
    d. None of the above.

ARTIST SCALE: **PROBABILITY**

1. Bob and Bill each bought one ticket for a lottery each week for the past 100 weeks. Bill has not won a single prize yet. Bob just won a $20 prize last week. Who is more likely to win a prize this coming week if they each buy only one ticket?

   a. Bill.

   b. Bob.

   C. They have an equal chance of winning.

   d. Not enough information to tell.

2. Two containers, labeled A and B, are filled with red and blue marbles according to the quantities listed in the table below. Each container is shaken vigorously. After choosing one of the containers, you will reach in and, without looking, draw out a marble. If the marble is blue, you win $50. Which container gives you the best chance of drawing a blue marble?

Container	Red	Blue
A	6	4
B	60	40

   a. Container A (with 6 red and 4 blue)

   b. Container B (with 60 red and 40 blue)

   C. Equal chances from each container

   d. Not enough information to tell.

3. When two fair six-sided dice are simultaneously thrown, these are two of the possible results that could occur: Result 1: a 5 and a 6 are obtained in any order. Result 2: a 5 is obtained on each die. Which of the following statements is correct?

   a. The probability of obtaining each of these results is equal.

   B. There is a higher probability of obtaining Result 1 (a 5 and a 6 in any order).

   c. There is a higher probability of obtaining Result 2 (a 5 on each die).

   d. It is impossible to give an answer.

4. Suppose you read on the back of a lottery ticket that the chances of winning a prize are 1 out of 10. Select the best interpretation.

   a. You will win at least once out of the next 10 times you buy a ticket.

   b. You will win exactly once out of the next 10 times you buy a ticket.

   C. You might win once out of the next 10 times but it is not for sure.

   d. You will not win even once.

5. You are about to roll 2 fair six-sided dice, hoping to get a double. (A double = both dice show the same value on top). Which double will occur the least often?

   a. 6 - 6.

   b. 1 - 1.

   c. 1 - 1 and 6 - 6 are both least likely to occur.

   D. All doubles are equally likely.

6. Colin is flipping a fair coin. Heads has just come up 5 times in a row! The chance of getting heads on the next throw is

   a. less than the chance of getting tails since we are due for a tails.

   B. equal to the chance of getting tails since the flips are independent and the coin is fair.

   c. greater than the chance of getting tails since these data prove Colin is more likely to obtain heads with this coin.

7. A game company created a little plastic dog that can be tossed in the air. It can land either with all four feet on the ground, lying on its back, lying on its right side, or lying on its left side. However, the company does not know the probability of each of these outcomes. They want to estimate the probabilities. Which of the following methods is most appropriate?

   a. Since there are four possible outcomes, assign a probability of 1/4 to each outcome.

   B. Toss the plastic dog many times and see what percent of the time each outcome occurs.

   c. Simulate the data using a model that has four equally likely outcomes.

   d. None of the above.

8. Two containers, labeled A and B, are filled with red and blue marbles according to the quantities listed in the table below. Each container is shaken vigorously. Which of the following outcomes has the smallest probability?

Container	A	B
Red	80	40
Blue	20	60

a. Obtaining a blue marble from container A.

B. Obtaining a blue marble from container A and a blue marble from container B.

c. Not enough information to tell.

9. The local Meteorological claims that there is a 70% probability of rain tomorrow. Provide the best interpretation of this statement.

a. Approximately 70% of the city will receive rain within the next 24 hours.

B. Historical records show that it has rained on 70% of previous occasions with the same weather conditions.

c. If we were to repeatedly monitor the weather tomorrow, 70% of the time it will be raining.

d. Over the next ten days, it should rain on seven of them.

ARTIST SCALE: **DATA COLLECTION**

1. In a survey people are asked "Which brand of toothpaste do you prefer?" The data gathered from this question would be what type of data?

   A. categorical

   b. quantitative

   c. continuous

Items 2 and 3 refer to the following situation:
A student is gathering data on the driving experiences of other college students. One of the variables measured is the type of car the student drives. These data are coded using the following method: 1 = subcompact, 2 = compact, 3 = standard size, 4 = full size, 5 = premium, 6 = mini van, 7 = SUV, and 8 = truck.

2. What type of variable is this?

   A. categorical

   b. quantitative

   c. continuous

3. The student plans to see if there is a relationship between the number of speeding tickets a student gets in a year and the type of vehicle he or she drives. Identify the response variable in this study.

   a. College students

   b. type of car

   C. number of speeding tickets

   d. average number of speeding tickets last year

4. A researcher is studying the relationship between a vitamin supplement and cholesterol level. What type of study needs to be done in order to establish that the amount of vitamin supplement causes a change in cholesterol level?

   a. Correlational study

   B. Randomized experiment

   c. Time Series study

   d. Survey

5. An instructor is going to model an experiment in his statistics class by comparing the effect of 4 different treatments on student responses. There are 40 students in the class. Which is the best way for the instructor to distribute the students to the 4 treatments for this experiment?

   a. Assign the first treatment to the first 10 students on the class list, the second treatment to the next 10 students, and so on.

   B. Assign a unique number to each student, then use random numbers to assign 10 students to the first treatment, 10 students to the second treatment, and so on.

   c. Assign the treatment as students walk into class, giving the first treatment to the first 10 students and the second treatment to the next 10 student, and so on.

   d. All of these are equally appropriate methods.

   e. None of these is an appropriate method.

Items 6 and 7 refer to the following situation:
Suppose two researchers wanted to determine if aspirin reduces the chance of a heart attack.

6. Researcher 1 studied the medical records of 500 randomly selected patients. For each patient, he recorded whether the person took aspirin every day and if the person had ever had a heart attack. Then he reported the percentage of heart attacks for the patients who took aspirin every day and for those who did not take aspirin every day. What type of study did Researcher 1 conduct?

   A. Observational

   b. Experimental

   c. Survey

   d. None of the above

7. Researcher 2 also studied 500 patients that visited a regional hospital in the last year. He randomly assigned half (250) of the patients to take aspirin every day and the other half to take a placebo everyday. Then after a certain length of time he reported the percentage of heart attacks for the patients who took aspirin every day and for those who did not take aspirin every day. What type of study did Researcher 2 conduct?

   a. Observational

   B. Experimental

   c. Survey

   d. None of the above

8. The dean of a college would like to determine the feelings of students concerning a new registration fee that would be used to upgrade the recreational facilities on campus. All registered students would pay the fee each term. Which of the following data collection plans would provide the best representation of students' opinions at the school?

   a. Survey every 10th student who enters the current recreational facilities between the hours of 1:00 and 5:00 pm until 100 students have been asked.

   B. Randomly sample fifty student ID numbers and send a survey to all students in the sample.

   c. Place an ad in the campus newspaper inviting students to complete an online survey. Collect the responses of the first 200 students who respond.

   d. All of the above would be equally effective.

9. A team in the Department of Institutional Review at a large university wanted to study the relationship between completing an internship during college and students' future earning potential. From the same graduating class, they selected a random sample of 80 students who completed an internship and 100 students who did not complete an internship and examined their salaries 5 years past graduation. They found that there was a statistically higher mean salary for the internship group than for the non-internship group. Which of the following interpretations do you think is the most appropriate?

   a. More students should take internships because having an internship produces a higher salary.

   B. There could be a confounding variable, such as student major, that explains the difference in mean salary between the internship and no intership groups.

   c. You cannot draw any valid conclusions because the samples are not the same size.

ARTIST SCALE: **CONFIDENCE INTERVALS, ONE-SAMPLE**

1. Two different samples will be taken from the same population of test scores where the population mean and standard deviation are unknown. The first sample will have 25 data values, and the second sample will have 64 data values. A 95% confidence interval will be constructed for each sample to estimate the population mean. Which confidence interval would you expect to have greater precision (a smaller width) for estimating the population mean?

   A. I expect the confidence interval based on the sample of 64 data values to be more precise.

   b. I expect both confidence intervals to have the same precision.

   c. I expect the confidence interval based on the sample of 25 data values to be more precise.

2. A 95% confidence interval is computed to estimate the mean household income for a city. Which of the following values will definitely be within the limits of this confidence interval?

   a. The population mean

   B. The sample mean

   c. The standard deviation of the sample mean

   d. None of the above

3. Each of the 110 students in a statistics class selects a different random sample of 35 Quiz scores from a population of 5000 scores they are given. Using their data, each student constructs a 90% confidence interval for μ the average Quiz score of the 5000 students. Which of the following conclusions is correct?

   a. About 10% of the sample means will not be included in the confidence intervals.

   B. About 90% of the confidence intervals will contain μ .

   c. It is probable that 90% of the confidence intervals will be identical.

   d. About 10% of the raw scores in the samples will not be found in these confidence intervals.

4. A 95% confidence interval for the mean reading achievement score for a population of third grade students is (43, 49). The margin of error of this interval is:

   a. 5

   B. 3

   c. 6

5. Justin and Hayley conducted a mission to a new planet, Planet X, to study arm length. They took a random sample of 100 Planet X residents and calculated a 95% confidence interval for the mean arm length. What does a 95% confidence interval for arm length tell us in this case? Select the best answer:

   a. I am 95% confident that this interval includes the sample mean ($\overline{x}$) arm length.

   b. I am confident that most (95%) of all Planet X residents will have an arm length within this interval.

   c. I am 95% confident that most Planet X residents will have arm lengths within this interval.

   D. I am 95% confident that this interval includes the population mean arm length.

6. Suppose that a random sample of 41 state college students is asked to measure the length of their right foot in centimeters. A 95% confidence interval for the mean foot length for students at this university turns out to be (21.709, 25.091). If instead a 90% confidence interval was calculated, how would it differ from the 95% confidence interval?

   A. The 90% confidence interval would be narrower.

   b. The 90% confidence interval would be wider.

   c. The 90% confidence interval would be the same as the 95% confidence interval.

7. A pollster took a random sample of 100 students from a large university and computed a confidence interval to estimate the percentage of students who were planning to vote in the upcoming election. The pollster felt that the confidence interval was too wide to provide a precise estimate of the population parameter. What could the pollster have done to produce a narrower confidence interval that would produce a more precise estimate of the percentage of all university students who plan to vote in the upcoming election?

   A. Increase the sample size to 150.

   b. Increase the confidence level to 99%.

   c. Both a and b

   d. None of the above

8. A newspaper article states with 95% confidence that 55% to 65% of all high school students in the United States claim that they could get a hand gun if they wanted one. This confidence interval is based on a poll of 2000 high school students in Detroit. How would you interpret the confidence interval from this newspaper article?

   a. 95% of large urban cities in the United States have 55% to 65% high school students who could get a hand gun.

   b. If we took many samples of high school students from different urban cities, 95% of the samples would have between 55% and 65% high school students who could get hand guns.

   C. You cannot use this confidence interval to generalize to all teenagers in the United States because of the way the sample was taken.

   d. We can be 95% confident that between 55% and 65% of all United States high school students could get a hand gun.

9. The Gallup poll (August 23, 2002) reported that 53% of Americans said they would favor sending American ground troops to the Persian Gulf area in an attempt to remove Hussein from power. The poll also reported that the "margin of error" for this poll was 4%. What does the margin of error of 4% indicate?

   a. There is a 4% chance that the estimate of 53% is wrong.

   b. The percent of Americans who are in favor is probably higher than 53% and closer to 57%.

   C. The percent of Americans who are in favor is estimated to be between 49% and 57%.

10. Suppose two researchers want to estimate the proportion of American college students who favor abolishing the penny. They both want to have about the same margin of error to estimate this proportion. However, Researcher 1 wants to estimate with 99% confidence and Researcher 2 wants to estimate with 95% confidence. Which researcher would need more students for her study in order to obtain the desired margin of error?

   A. Researcher 1.

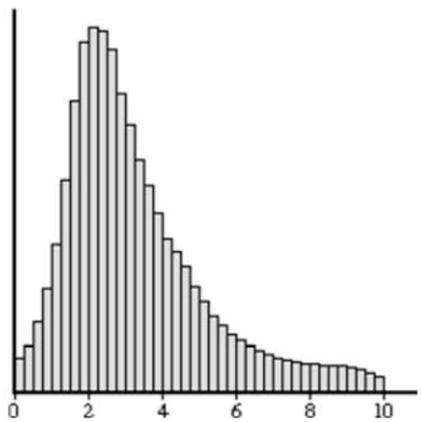
   b. Researcher 2.

   c. Both researchers would need the same number of subjects.

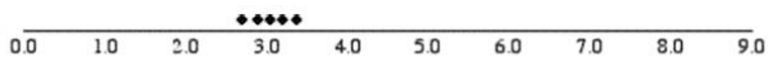
   d. It is impossible to obtain the same margin of error with the two different confidence levels.

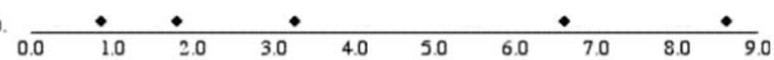
ARTIST SCALE: **SAMPLING VARIABILITY**

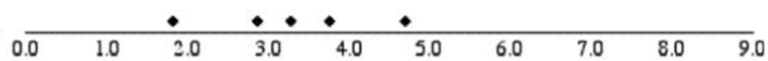
The distribution for a population of measurements is presented below. The mean is 3.2 and the standard deviation is 2. Suppose that five students each take a sample of ten values from the population and each student calculates the sample mean for his or her ten data values. The students draw a dotplot of their five sample means on the classroom board so that they can compare them.



1. Which of the following dotplots do you think is the most plausible for the one they drew on the board?

a.


b.


c.


a. a.

b. b.

C. c.

2. Consider the distribution of average number of hours that college students spend sleeping each weeknight. This distribution is very skewed to the right, with a mean of 5 and a standard deviation of 1. A researcher plans to take a simple random sample of 18 college students. If we were to imagine that we could take all possible random samples of size 18 from the population of college students, the sampling distribution of average number of hours spent sleeping will have a shape that is

   a. Exactly normal.

   B. Less skewed than the population.

   c. Just like the population (i.e., very skewed to the right).

   d. It's impossible to predict the shape of the sampling distribution.

3. Imagine you have a huge jar of candies that are a generic version of M&Ms. We know that 40% of the candies in the jar are brown. Imagine that you create a sample by randomly pulling 20 candies out of the jar. If you repeated this 10 times to create 10 samples, each with 20 candies, about how many browns would you expect to find in each of the 10 samples?

   a. Each sample would have exactly 8 brown candies.

   b. Most of the samples would have 0 to 8 brown candies.

   c. Most of the samples would have 8 to 20 brown candies.

   D. Most of the samples would have 6 to 10 brown candies.

   e. You are just as likely to get any count of brown candies between 0 and 20.

4. A random sample of 25 college statistics textbook prices is obtained and the mean price is computed. To determine the probability of finding a more extreme mean than the one obtained from this random sample, you would need to refer to:

   a. the population distribution of all college statistics textbook prices.

   b. the distribution of prices for this sample of college statistics textbooks.

   C. the sampling distribution of textbook prices for all samples of 25 textbooks from this population.

5. In a geology course, students were learning to use a balance scale to make accurate weighings of rock samples. One student plans to weigh a rock 20 times and then calculate the average of the 20 measurements to estimate her rock's true weight. A second student plans to weigh a rock 5 times and calculate the average of the 5 measurements to estimate his rock's true weight. Which student is more likely to come the closest to the true weight of the rock he or she is weighing?

   A. The student who weighed the rock 20 times.

   b. The student who weighed the rock 5 times.

   c. Both averages would be equally close to the true weight.

   d. It is impossible to predict which is more likely to be closer to the true weight.

6. Suppose half of all newborns are girls and half are boys. Hospital A, a large city hospital, records an average of 50 births a day. Hospital B, a small, rural hospital, records an average of 10 births a day. On a particular day, which hospital is less likely to record 80% or more female births?

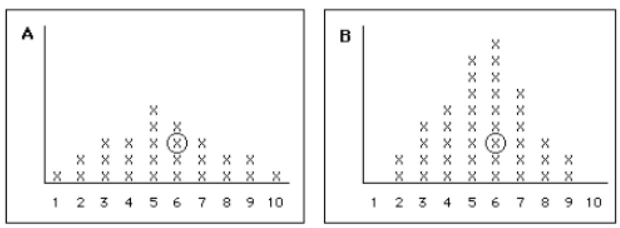
   A. Hospital A (with 50 births a day), because the more births you see, the closer the proportions will be to .5.

   b. Hospital B (with 10 births a day), because with fewer births there will be less variability.

   c. The two hospitals are equally likely to record such an event, because the probability of a boy does not depend on the number of births.

   d. There is not enough information to determine which hospital is more likely to have 80% or more female births.

7. Figure A represents the weights for a sample of 26 pebbles, each weighed to the nearest gram. Figure B represents the mean weights of a random sample of 3 pebbles each, with the mean weights rounded to the nearest gram. One value is circled in each distribution. Is there a difference between what is represented by the X circled in A and the X circled in B? Please select the best answer from the list below.

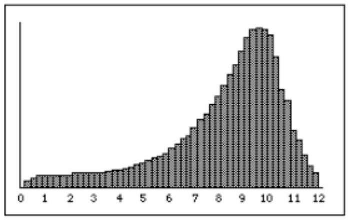


   a. No, in both Figure A and Figure B, the X represents one pebble that weights 6 grams.

   b. Yes, Figure A has a larger range of values than Figure B.

   C. Yes, the X in Figure A is the weight for a single pebble, while the X in Figure B represents the average weight of 3 pebbles.

Items 8 and 9 refer to the following situation:
The distribution for a population of measurements is presented below.



8. A sample of 10 randomly selected values will be taken from the population and the sample mean will be calculated. Which of the following intervals is MOST likely to include the sample mean?

   a. 4 to 6

   B. 7 to 9

   c. 10 to 12

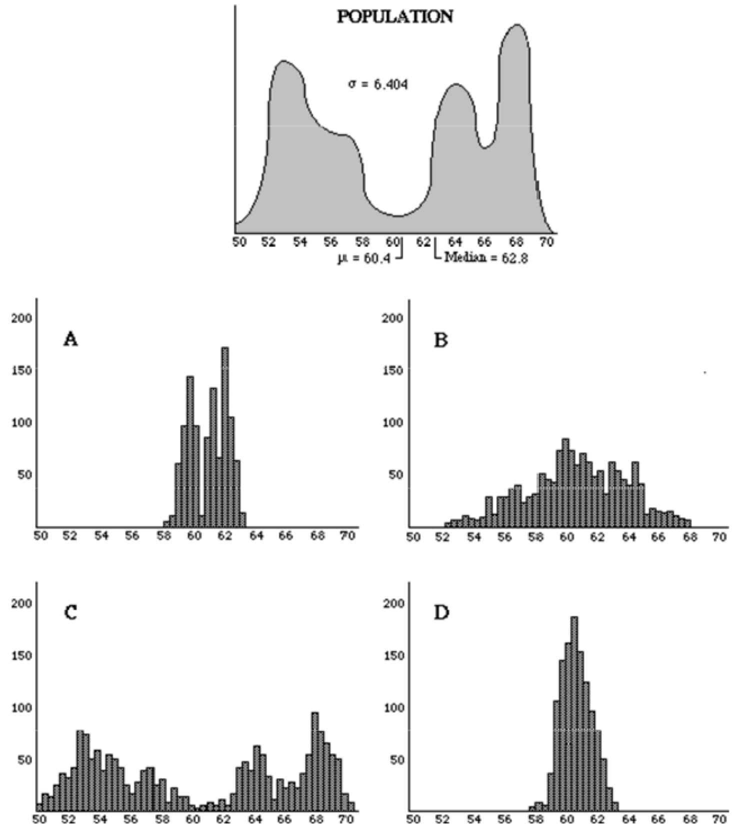
9. Another sample of 10 randomly selected values will be taken from the population and the sample mean will be calculated. Which of the following intervals is LEAST likely to include the sample mean?

   A. 0 to 3

   b. 4 to 7

   c. 8 to 11

Items 10 to 15 refer to the following situation:
A hypothetical distribution for a population of test scores is displayed below. The population has a mean of 60.4, a median of 62.8, and a standard deviation of 6.404. Each of the other four graphs labeled A to D represent possible distributions of sample means for random samples drawn from the population.



10. Which graph best represents a distribution of sample means for 1000 samples of size 4?

   a. A

   B. B

   c. C

   d. D

11. What do you expect for the shape of the sampling distribution (the distribution of sample means for all possible samples of size n = 4)?

A. Shaped more like a normal distribution then like the population distribution.

b. Shaped more like the population distribution then like a normal distribution.

c. Shaped like neither the population or the normal distribution.

12. What do you expect for the variability (spread) of the sampling distribution?

a. Same as the population.

B. Less variability than the population (a narrower distribution).

c. More variability than the population (a wider distribution).

13. Which graph best represents a distribution of sample means for 1000 samples of size 50?

a. A

b. B

c. C

D. D

14. What do you expect for the shape of the sampling distribution (the distribution of sample means for all possible samples of size n = 50)?

A. Shaped more like a normal distribution.

b. Shaped more like the population.

c. Shaped like neither the population or the normal distribution.

15. What do you expect for the variability (spread) of the sampling distribution?

a. Same as the population.

B. Less variability than the population (a narrower distribution).

c. More variability than the population (a wider distribution).

JMU

James Madison University
Center for Assessment and Research Studies [MSC 6806]

Memorandum

To: Researchers Requesting the *Attitudes Toward Statistics* Scale

From: Steven Wise

RE: Instructions for using the *ATS*

You have requested a copy of the *Attitudes Toward Statistics* scale to use in your research. I have attached a copy of the scale that should be suitable for photocopying.

The 29-item ATS has two subscales. The *Attitudes Toward Field* subscale consists of the following 20 items, with reverse-keyed items indicated by an "(R)":

1, 3, 5, 6(R), 9, 10(R), 11, 13, 14(R), 16(R), 17, 19, 20(R), 21, 22, 23, 24, 26, 28(R), 29

The *Attitudes Toward Course* subscale consists of the following 9 items:

2(R), 4(R), 7(R), 8, 12(R), 15(R), 18(R), 25(R), 27(R)

To score the *ATS*, simply sum the appropriate item scores for the subscales and/or total scale.

The original reference for the *ATS* is

Wise, S. L. (1985). The development and validation of a scale measuring attitudes toward statistics. *Educational and Psychological Measurement, 45*, 401-405.

Another useful (and more recent) reference regarding the scale is:

Schultz, K. S., & Koshino, H. (1998). Evidence of reliability and validity for Wise's Attitudes Toward Statistics scale. *Psychological Reports, 82*, 27-31.

In exchange for permission to use my scale, I'd appreciate your sending me a copy of any manuscripts that result from your use of the *ATS*, as I am always interested in seeing studies that use the scale. I thank you in advance for your cooperation, and I wish you success with your research.

Appendix B

The Attitudes Toward Continuing Professional Development in Statistics Instrument

Attitudes Toward Continuing Professional Development in Statistics Instrument (ATPDS)

<Lancaster, S. M., 2006>

There are now special workshops which have been started in certain areas of the United States. These workshops are designed to help teachers to improve their ability to teach statistics to elementary students. The workshops are not classes where teachers learn statistics as they would in some college course. Instead, the workshops help teachers to recognize and use methods known to help improve students' learning of statistical investigation and analysis. A typical workshop lasts 3 weeks with follow up contacts occurring during the school year.

Do you believe that you would enjoy such a workshop?
(choose one)

____ very much ____ I am not sure

____ somewhat ____ probably not

Would you be interested in attending such a workshop if some colleagues with whom you are friends attended as well?
(choose one)

____ very much ____ I am not sure

____ somewhat ____ probably not

Do you believe that such a program would help you in your classroom teaching?
(choose one)

____ very much ____ I am not sure

____ somewhat ____ probably not

Would you be interested in attending such a workshop if the cost was covered for you?
(choose one)

____ very much ____ I am not sure

____ somewhat ____ probably not

Would you be interested in attending such a workshop if the cost was covered for you and you received a small stipend (say $300)?
(choose one)

____ very much ____ I am not sure

____ somewhat ____ probably not

Would you attend such a workshop even if you had to pay for it yourself?
(choose one)

____ very much ____ I am not sure

____ somewhat ____ probably not

Appendix C

The Grading Project

NAME ___Write your name on the BACK of this sheet of paper.

Project for the Grading of Mathematical Problems, 4th grade level.

On the following pages, there are several "lessons". For each lesson, there are samples of problems which have been completed. There are not any indicators of the teaching style used to arrive at the work which is shown. This work may have been completed during group work in class, it may have been completed at home individually, the circumstances surrounding the work may not be either of those previously mentioned.

Using a pen of any color except black (your marks need to be distinguishable from the marks on the original copy), make corrections to the paper. Assume for each problem that it is worth ten points. If the problem has parts (a), (b), and so on, then count all parts added together as ten points.

Before you make any marks, determine how you intend to mark the paper. Will you write the correct "step" where errors have occurred?, will you only give hints or suggestions?, will you only mark the area of concern and put the number of points deducted?, will you have a system differing from any of those suggested above? Remember that one important issue is that the students should know how much has been deducted (or positively valued) for each problem.

Following each problem, there are specific questions for you to answer. The questions may not be the same for each problem. As you grade each "student problem", think about each question and how you might respond. Write down answers or ideas about the questions. If you participate in the short (10 minute)voluntary interview, then we may discuss these questions again at that time.

Once you have graded all of the problems on the papers and considered all of the questions on the word document, then answer the general questions at the end of the document.

1

Grading Assignment **Table of Contents**

2

Topic: Dividing by Two Digit Divisors

Question 1: Andre has 128 pieces of candy. His mom will not let him eat that much candy. She explained to him that when his 14 guests (counting Andre) come over for his party, he can share the candy with them. He wants to give everyone the same amount of candy.
(a) How many pieces of candy will each person have if they each have an equal amount?
(b) Is there any candy left over after dividing by 14? If so, how many pieces?

Student 1:

(a)

$$\begin{array}{r} 9 \\ 14\overline{)128} \\ \underline{126} \\ 2 \end{array}$$

Each person can have 2 pieces of candy.

(b) There are 9 pieces left over.

Student 2:

(a)

14 times 8 is 112
14 times 9 is 126
14 times 10 is 140

Since 126 is under 128 each person can have 9 pieces of candy.

(b)

$$\begin{array}{r} 128 \\ -126 \\ \hline 2 \end{array}$$

There are 2 pieces left over.

3

Questions to answer regarding the dividing by two digit divisors problems.

In the grading of question 1, which student was deducted the most points (or did they get equal deductions)?

Write a short sentence explaining what student 1 did wrong.

How might you explain to student 1 the correct way to solve the problem?

Can you use part of what student 2 did to help student 1 understand their error? Write a short explanation of how this might be done.

4

Topic: Prime and Composite Numbers

Question 2: Determine which of the following numbers are prime and which are composite. (a) 21 (b) 91 (c)17

Student:

(a) composite since 3 goes into 21 with no remainder

(b) prime nothing divides 91

(c) prime nothing divides 17

Topic: Number Patterns

Question 3: For each list presented, determine the next value in the sequence. You do not need to show any work.

(a) 2, 4, 6, __ (b) 1, 3, 7, __ (c) 2, 6, 18, __

Student:

(a) 2, 4, 6, 8

(b) 1, 3, 7, 15

(c) 2, 6, 18, 54

5

Questions to answer regarding the prime and composite numbers and the numerical patterms problems.

In the prime numbers problem, if the student made any errors, then show, for each error, the correct way to VERIFY the correct answer. (ONLY do this for answers with errors.)

This question is to be worked by hand on the student work page. No explanation to give here.

In the numerical patterns problem, did you decide that the student missed one or more problems?

Note that the problem did not specify that the student had to use either an arithmetic or a geometric progression. List first the patterns in which the student used an arithmetic progression. List second the patterns in which the student used a geometric progression. List third, the patterns in which the student used a pattern that was not either one of the above types.

Did you count (b) as not fully correct? (Be honest here.) Do not change your grading of that problem. The student recognized that you can get each subsequent term by doubling the previous term and then adding one. That is how the student obtained the answer written on the paper. [At issue here is the importance of well-written questions. But also, it is important to be flexible and recognize possibilities if you accidentally make a problem more open-ended.] Add any comments you have regarding this problem.

6

Topic: Data Representation

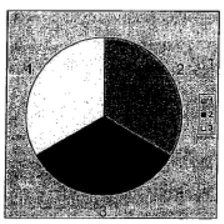
Question 4: Your class has taken a survey and has found that the favorite food of each person in the class is as follows: pizza, 11 people: fish, 4 people, macaroni and cheese, 6 people, hot dogs, 3 people. Make a frequency table for your survey with a frequency column and with a cumulative frequency column.

Student:

Food Item	Frequency	Cumulative Frequency
Pizza	~~||||~~ ~~||||~~ | 11	24
Fish	|||| 4	24
Macaroni and Cheese	~~||||~~ | 6	24
Hot Dogs	||| 3	24

Topic: Probability

Question 5: Look at the spinner below. After two spins, what is the more likely sum of the two spin values, 2 or 3 or 4? (Hint: One of these is more likely than the others.)



Student 1 It cannot be 4 since there is not a 4 on the spinner.
I think 2 looks bigger than 3 so 2 is the likely sum.

Student 2 I can get 1 or 2 or 3 on the first spin.
I can get 1 or 2 or 3 on the next spin.
Sum means add so I need to add the spins together.

1+1=2 1+2=3 2+3=5 Each has 1 way except 4.
2+2=4 1+3=4 3+3=6 It has 2 ways so it is 4.

7

Questions to answer regarding the data representation and probability problems.

In the data representation problem, the student made a small error while doing the cumulative frequency. What is the error the student made.

How would you explain why the student should write the table using the correct method instead of writing the table as they did?

In the probability problem, which student was farther from the correct solution?

Even though student 1 did not correctly answer this question (question 5) there is one single concept that they were correct about. This has to do with the word "bigger". Can you explain the notion for which the student had correct understanding?

Student 2 arrived at the correct answer, but they actually had an incorrect method of deciding on this answer. Explain the correct way to work the last step (the only step this student did incorrectly) to show that **4** is the correct answer.

Topic: Mean, Median, Mode, and Outliers

Question 6: (a).Find the mean, median, and mode of the following list. 6,7,7,9,11

(b) Add the number 17 to the list. Recalculate the median in part (a).

(c) Do you think mean or median changed more when the 17 was added to the list? Explain.

Student:

(a) mean $\frac{6+7+7+9+11}{5} = \frac{40}{5} = 8$

median It is in order so the middle number is the median. 7

mode 7

(b) 6, 7, 7, 9, 11, 17

median There is a tie so use $\frac{7+9}{2} = \frac{16}{2} = 8$

(c) It affected the median the most since I had to change the way that I found the answer.

Topic: Stem and Leaf Plots

Question 7: For the stem-and-leaf plot given, what is the mode? what is the median?

Stem	Leaves
3	0 0 1
4	1 2 2 5 7
5	0 0 3 3 3 3 8

Student: The mode is 53.

The median is the middle term.

The median is 42 since 4 is the middle stem and 2 is the middle leaf of the 4 stem.

9

Questions to answer regarding the mean, median, mode, and outliers problem.

In the mean, median, mode, and outliers problem, list the specific errors the student made. No comments are needed.

Questions to answer regarding the stem-and-leaf plot problem.

Did the student get the mode correct? What is the next most repeated number? (Hint: There is a three way tie.)

There is an error for the student's median answer. They explained their work. What is the correct answer with an explanation of how it was found.

If you can, explain why the fact that the 3 stem has fewer leaf values than the 5 does affected the student's incorrect answer.

For each problem graded, label each according to the needs that you had.

1 = knew the answer and how to grade without any effort

2 = had to think for a while, but did not need to reference any material

3 = had to make a quick reference to a text or something like that

4 = had to spend a long time with references to figure it out, but it was rewarding

4* = as in 4 but it was more frustrating than rewarding

5 = needed a lot of time to reference and had to give up on a satisfactory review of the material due to being short on time

1. _____
2. _____
3. _____
4. _____
5. _____
6. _____
7. _____

Appendix D

Raw Data for Some of the Data Sets

Summary of values for the3213sp data set including ARTIST scores (SC) and results (WK) from the instrument for predicting interest in CPD in statistic (ATPDS).

ID	ATS sum	CSSE	SELS	SCgr	SCcntr	SCsprd	SCprob	SCsum4	WKen	WKfr	WKhp	WKpd	WKin	WKls	WKttotal
1	114	75	81	7	6	0	0	13	3	3	4	4	4	2	20
2	95	38	58	5	2	2	6	15	2	3	3	2	4	1	15
3	109	47	50	9	3	4	6	22	3	3	3	3	4	1	17
4	70	22	19	2	1	4	5	12	1	1	2	2	3	1	10
5	109	54	77	9	3	9	7	28	2	3	3	3	4	1	16
6	107	54	58	7	3	8	8	26	2	3	3	3	4	1	16
7	105	51	65	5	0	5	8	18	4	4	4	4	4	1	21
8	84	56	53	5	1	2	3	11	3	4	3	4	4	2	20
9	77	14	56	6	1	2	7	16	3	3	2	3	4	2	17
10	101	31	56	5	2	4	6	17	3	3	3	3	4	2	18
11	79	14	56	7	1	4	7	19	4	3	3	4	4	3	21
12	97	63	65	6	1	3	4	14	3	3	3	3	3	1	16
13	93	44	49	9	1	3	4	17	3	4	3	3	4	1	18
14	97	34	70	5	0	7	8	20	2	4	4	4	4	3	21
15	87	46	47	4	2	3	5	14	1	3	2	2	3	1	12
16	113	41	68	7	1	8	4	20	3	3	3	3	4	1	17
17	89	55	69	0	0	0	0	0	1	1	2	1	1	1	7
19	104	58	84	7	2	6	7	22	3	3	3	4	4	3	20
20	100	29	70	7	2	6	7	22	3	3	3	3	4	2	18
21	75	23	21	8	3	4	2	17	2	3	1	3	4	1	14
22	105	48	45	6	4	8	8	26	3	4	3	4	4	2	20

ATS item responses for the first 34 participants from 1473

ID	ATS	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	FIELD	COURSE
1	104	5	4	5	2	5	2	4	2	4	3	1	3	2	2	2	2	3	1	2	3	4	4	4	4	1	5	1	1	1	72	32
2	92	4	5	4	3	4	2	5	1	2	1	4	3	4	2	5	3	5	5	3	1	4	4	4	4	5	4	2	2	1	76	16
3	37	1	4	2	5	1	5	5	1	1	5	1	5	1	5	5	5	5	5	1	5	1	1	1	1	3	1	5	5	1	25	12
4	76	2	3	5	1	2	4	2	1	2	2	4	4	2	3	4	3	4	5	1	1	2	4	3	2	4	1	4	4	1	54	22
5	107	2	3	5	2	5	4	4	2	4	1	2	4	2	3	4	3	5	2	2	1	4	5	5	5	2	5	1	1	5	79	28
6	66	1	4	4	4	2	4	4	1	3	3	2	4	2	3	4	3	3	4	2	4	2	4	3	3	4	2	5	3	1	50	16
7	78	2	4	4	2	3	4	4	2	4	4	2	4	2	4	4	4	4	2	2	5	2	4	4	4	2	2	3	2	1	53	25
8	76	2	4	4	3	4	4	4	2	3	3	2	3	2	4	3	2	3	4	2	3	1	3	2	3	4	4	4	2	2	55	21
9	94	4	3	5	2	4	3	3	2	2	2	3	2	3	3	4	4	4	4	3	2	3	4	4	4	4	4	2	2	1	68	26
10	94	4	3	4	2	3	2	2	2	3	2	3	3	4	3	4	2	2	4	3	2	3	4	3	3	3	4	2	3	2	67	27
11	87	4	5	3	4	4	3	3	2	3	2	3	2	3	3	4	3	4	5	3	3	4	3	3	4	5	5	3	3	3	68	19
12	75	3	4	3	3	3	4	5	2	3	2	3	5	2	3	5	4	3	5	2	2	4	4	3	3	5	3	4	2	3	61	14
13	94	4	4	5	4	4	2	4	3	3	2	4	3	4	2	3	3	4	4	2	2	4	3	3	4	4	3	4	2	3	73	21
14	61	1	4	4	2	2	5	3	1	1	4	1	4	1	5	4	3	4	5	1	5	1	4	4	3	2	2	4	4	1	40	21
15	84	2	4	5	3	4	3	5	1	4	2	3	4	2	2	5	2	4	4	2	4	1	4	4	4	4	3	2	2	3	66	18
16	78	3	4	5	4	4	5	4	2	3	3	2	4	2	3	4	3	5	4	2	2	2	3	3	4	4	2	3	2	1	59	19
17	105	3	1	4	2	3	4	1	2	4	3	3	2	3	2	2	2	3	2	3	1	4	4	4	3	1	3	2	3	3	68	37
18	86	1	1	5	1	4	5	3	1	1	2	1	3	1	2	3	4	4	1	1	3	3	4	4	1	1	3	1	3	1	51	35
19	83	3	4	5	2	3	4	4	1	3	2	2	4	3	4	4	3	5	5	2	2	4	4	4	4	5	2	3	2	2	65	18
20	97	4	4	4	3	4	2	3	3	3	2	2	2	2	3	3	3	4	2	3	2	2	4	4	5	2	3	3	2	2	68	29
21	74	4	2	3	2	2	2	4	1	2	3	1	5	1	2	2	5	4	4	1	5	2	3	2	1	1	1	1	1	1	46	28
22	90	3	2	5	2	3	4	3	2	4	2	1	3	2	3	2	3	4	2	2	3	2	4	3	4	3	2	2	2	1	59	31
23	78	2	4	5	4	5	3	3	1	4	2	2	3	4	5	5	4	4	5	1	3	1	4	4	4	5	2	4	3	4	62	16
24	70	2	5	5	4	1	4	4	1	2	3	2	4	2	4	5	2	4	4	1	3	2	4	5	2	5	5	4	3	2	56	14
25	91	3	3	4	2	4	4	3	2	3	2	2	2	3	3	2	2	3	4	2	2	4	3	3	4	4	4	3	3	2	64	27

ATS item responses for the first 34 participants from 1473, continued

ID	ATS	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	FIELD	COURSE
26	92	1	2	3	3	4	3	2	1	4	2	2	2	4	3	5	4	4	2	2	2	4	3	3	4	1	3	1	4	2	61	31
27	126	5	1	5	1	4	1	1	3	5	1	5	1	5	1	3	1	4	1	5	1	5	2	3	4	1	3	1	3	2	85	41
28	74	4	5	4	4	3	2	4	1	3	2	2	5	2	4	5	3	4	4	2	2	5	4	2	2	5	2	4	3	2	61	13
29	99	4	4	4	2	3	2	3	2	4	2	4	3	4	2	4	2	4	2	3	3	4	3	3	3	2	4	3	3	3	72	27
30	95	4	4	3	3	3	2	4	2	2	2	4	3	4	2	2	2	4	5	2	2	4	5	5	4	4	3	2	3	2	72	23
31	107	3	2	4	2	4	2	2	2	3	2	3	2	4	2	2	2	4	3	3	2	4	4	4	4	3	4	2	2	3	75	32
32	50	1	5	4	5	2	5	5	1	3	3	1	5	1	5	4	3	4	5	1	5	1	2	2	2	5	2	5	3	2	40	10
33	60	1	5	5	3	4	5	5	1	1	2	1	5	1	3	5	2	4	5	1	2	1	3	2	2	5	2	5	2	1	49	11
34	84	3	5	4	4	4	3	5	2	2	2	2	4	3	2	4	2	4	3	2	2	2	4	4	3	3	2	2	2	2	64	20

SELS item responses for the first 34 participants from 1473

ID	SELS	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	64	5	5	5	4	3	3	4	2	4	5	6	6	6	6
2	68	5	5	5	4	5	4	4	4	4	4	6	6	6	6
3	84	6	6	6	6	6	6	6	6	6	6	6	6	6	6
4	55	3	4	4	4	3	3	4	3	3	3	3	6	6	6
5	82	5	6	6	6	6	5	6	6	6	6	6	6	6	6
6	21	2	2	1	1	1	1	2	1	1	1	2	1	4	1
7	53	4	4	4	3	2	4	4	4	4	4	4	4	4	4
8	41	3	4	2	4	2	4	3	2	2	2	1	3	5	4
9	72	5	5	5	5	5	5	5	5	5	5	5	5	6	6
10	58	3	3	3	5	5	4	4	5	4	3	4	5	5	5
11	70	5	5	5	5	5	5	5	5	5	5	5	5	5	5
12	50	5	5	3	3	3	2	4	2	2	2	4	5	5	5
13	72	4	4	5	5	6	5	5	5	5	5	6	5	6	6
14	48	4	3	3	4	4	3	2	4	3	2	2	4	6	4
15	55	5	5	3	2	2	3	5	5	4	3	3	3	6	6
16	54	4	4	3	4	5	5	3	2	3	4	5	3	5	4
17	64	5	5	4	5	5	4	6	4	4	4	4	4	6	4
18	84	6	6	6	6	6	6	6	6	6	6	6	6	6	6
19	21	1	1	1	2	1	2	3	1	1	1	1	1	4	1
20	75	5	5	6	5	5	5	5	5	5	5	6	6	6	6
21	42	5	5	3	3	4	4	2	1	1	1	1	4	4	4
22	80	5	5	6	6	6	5	6	6	5	6	6	6	6	6
23	53	2	2	2	x	3	5	3	4	4	6	5	5	6	6
24	49	5	5	3	4	3	4	5	1	1	2	3	2	6	5
25	58	4	2	4	5	5	4	5	3	3	3	4	5	6	5
26	69	5	6	5	4	5	5	3	6	5	6	4	4	6	5
27	64	3	6	1	6	6	6	6	1	6	1	4	6	6	6
28	45	1	5	4	4	4	1	3	1	1	1	5	5	5	5
29	75	5	5	6	6	6	6	5	2	5	5	6	6	6	6
30	79	5	6	5	6	6	5	6	5	5	6	6	6	6	6
31	55	4	5	5	5	4	3	4	3	3	3	5	3	5	3
32	20	1	2	1	1	1	2	2	1	1	1	1	2	2	2
33	52	4	4	4	4	3	4	3	4	4	3	4	3	4	4
34	64	3	3	5	6	6	4	5	3	3	5	4	5	6	6

Quantitative Results for the Qualitative Analyses of the Two Preservice Teachers Who Participated in the Grading Project

Report

Mean

ID	collstat	ATSfield	ATScours	ATSsum	CSSE	SELS	SCI
1	1.00	80.00	34.00	114.00	75.00	81.00	.00
21	.00	61.00	14.00	75.00	23.00	21.00	.00

Report

Mean

ID	SCgraphs	SCcenter	SCspread	SCprob	SCsum4
1	7.00	6.00	.00	.00	13.00
21	8.00	3.00	4.00	2.00	17.00

Report

Mean

ID	SCdacoll	SCcnfint	SCsmpvar	SCsum7
1	.00	.00	.00	13.00
21	.00	.00	.00	17.00

Report

Mean

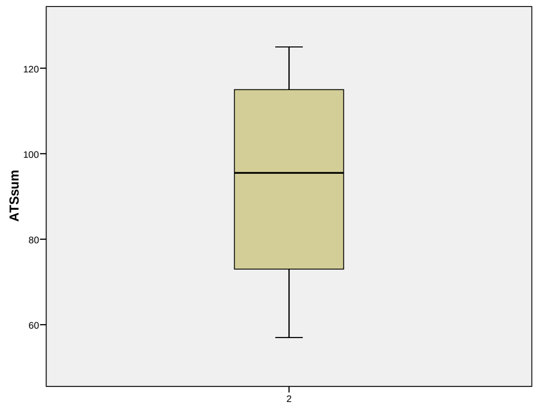
ID	WKen	WKfr	WKhp	WKpd	WKin	WKls	WKtotal
1	3.00	3.00	4.00	4.00	4.00	2.00	20.00
21	2.00	3.00	1.00	3.00	4.00	1.00	14.00

Appendix E

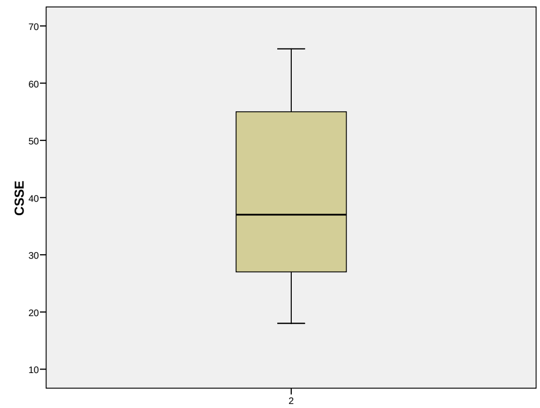
Box Plots, Q-Q Plots, and Kolmogorov-Smirnov Tests of Normality for the 3213fa, 1473, and 3213sp Data

Box plots for the 3213fa Data Set

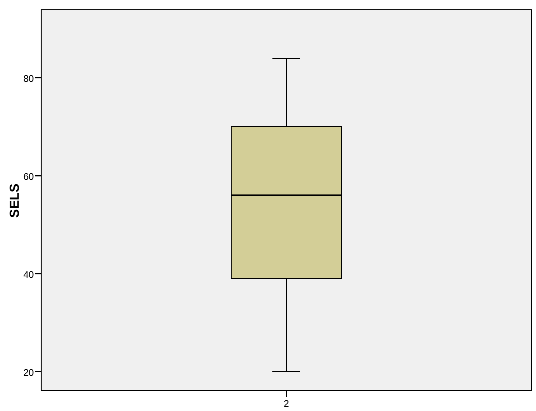
ATSsum



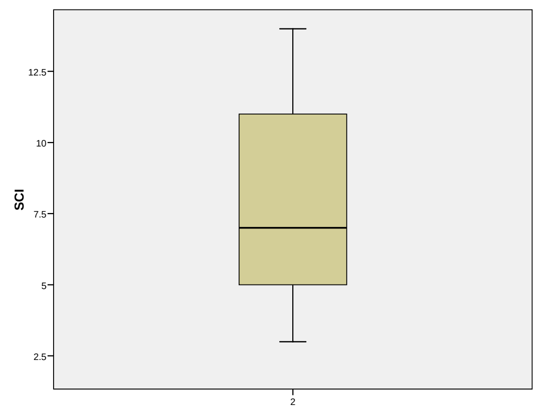
CSSE



SELS

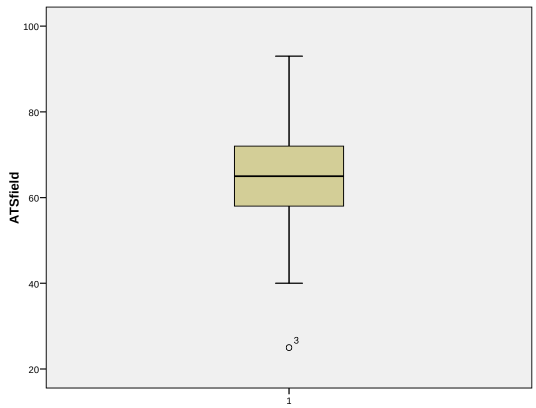


SCI

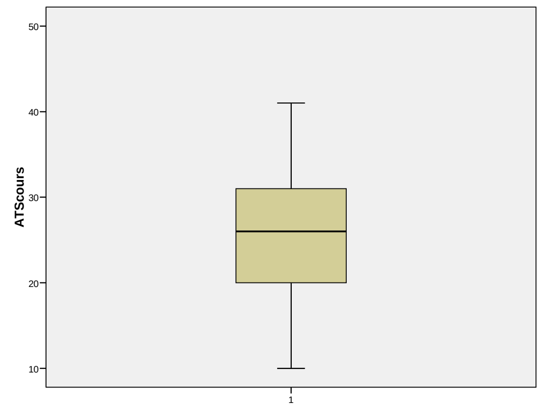


Box plots for the 1473 Data Set

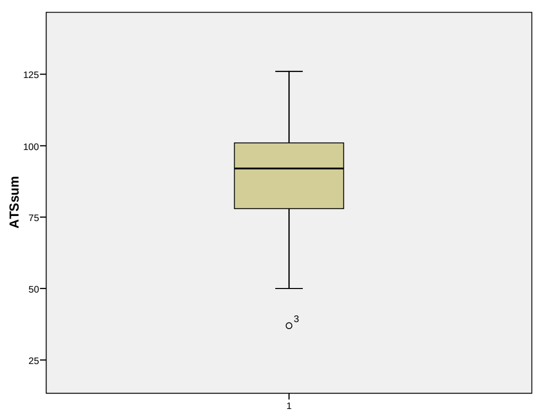
ATSfield



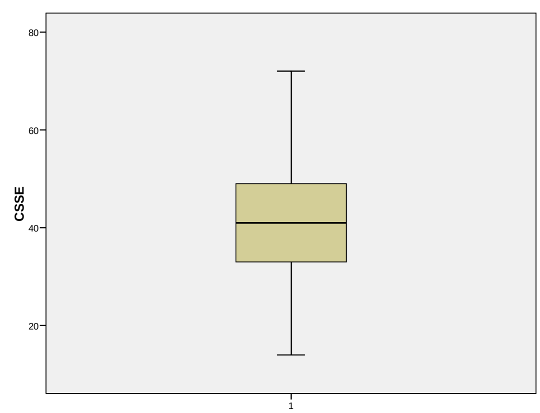
ATScourse



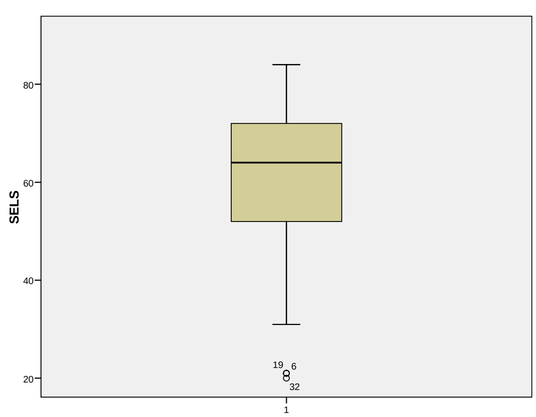
ATSsum



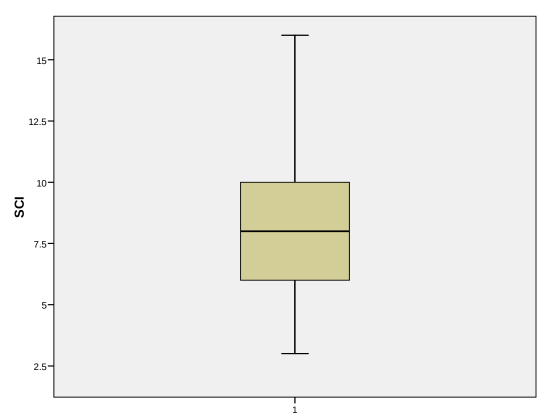
CSSE



SELS

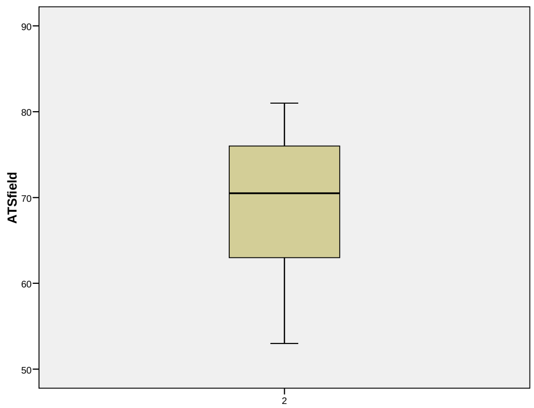


SCI

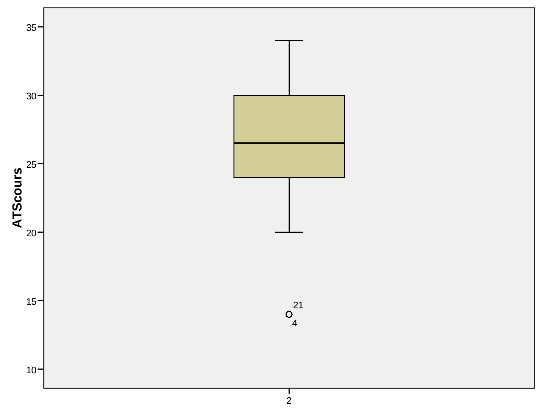


Box plots for the 3213sp Data Set

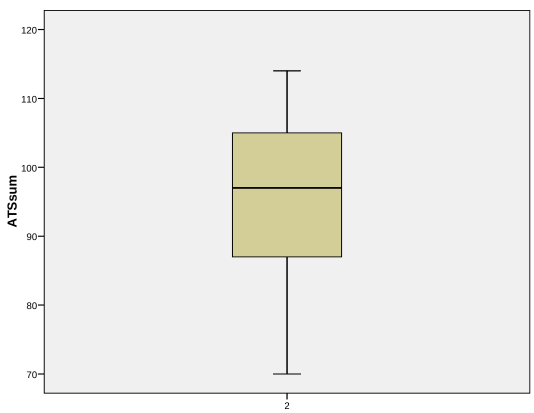
ATSfield



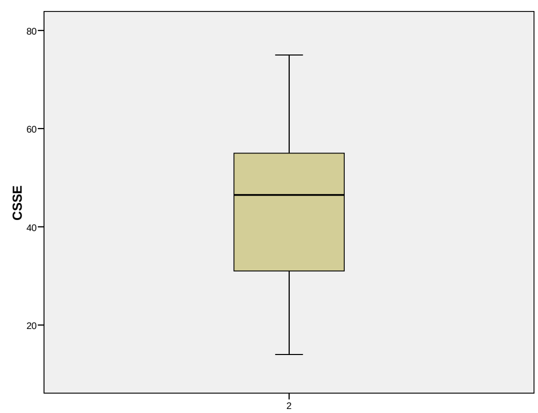
ATScourse



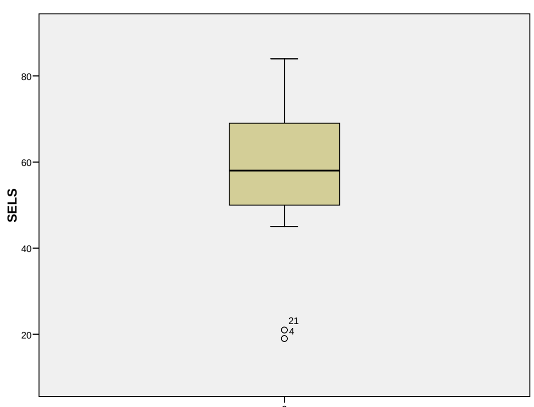
ATSsum



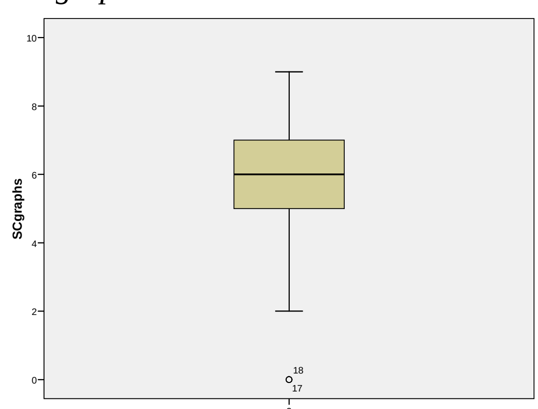
CSSE



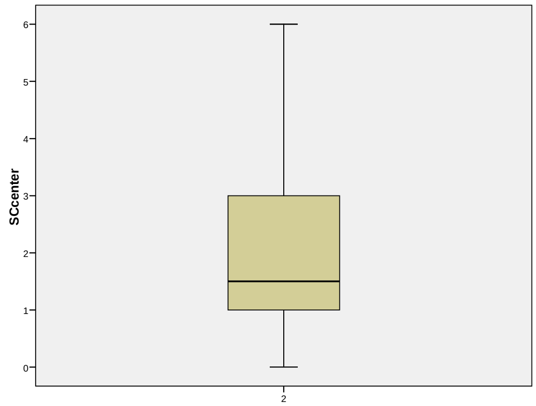
SELS



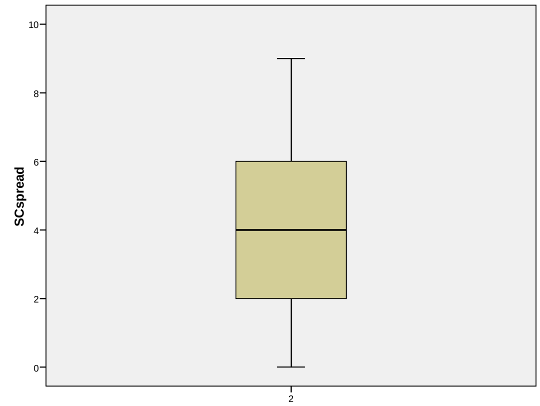
SCgraphs



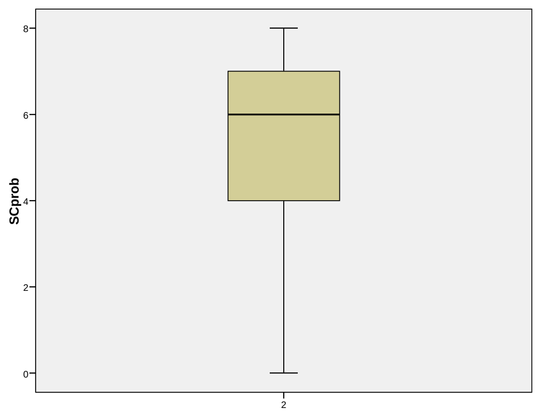
SCcenter



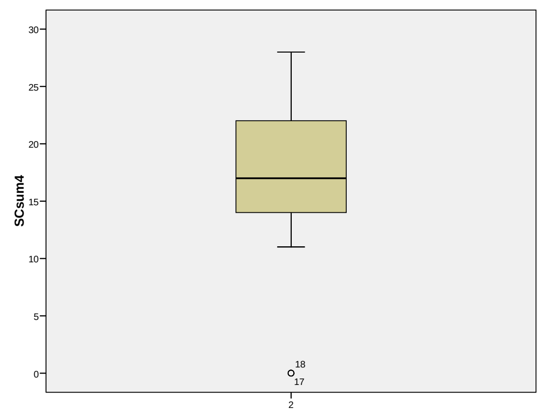
SCspread



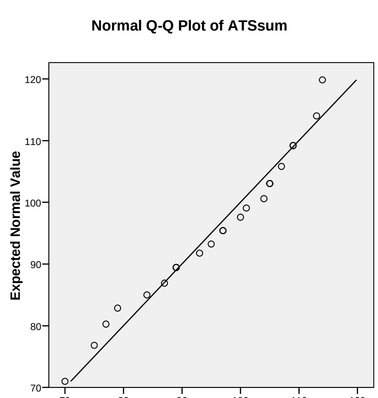
SCprob

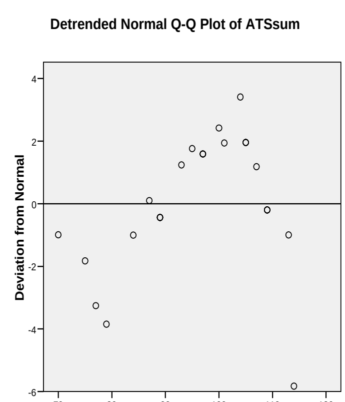


SCsum4

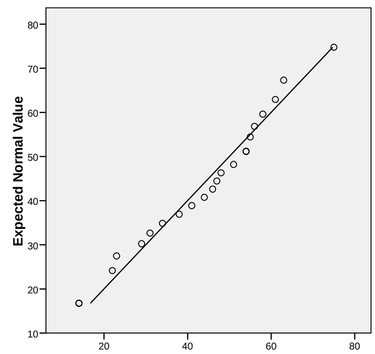


Q-Q plots of Quantitative Variables for the 3213sp Data

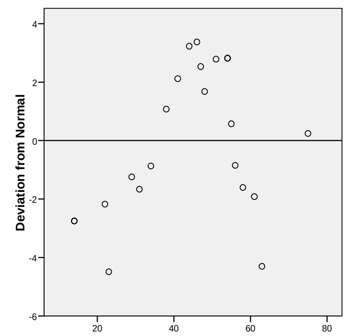




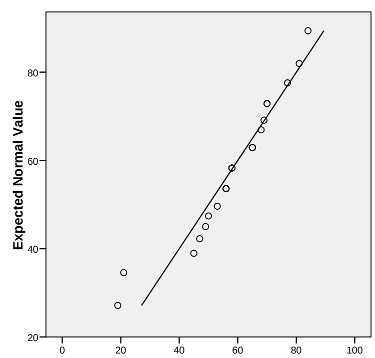
Normal Q-Q Plot of CSSE



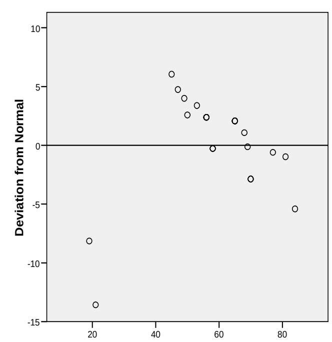
Detrended Normal Q-Q Plot of CSSE



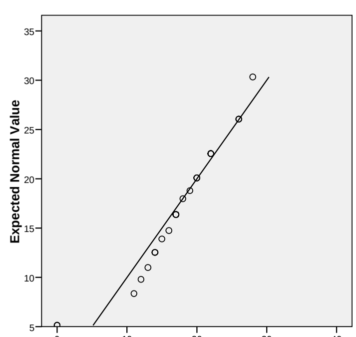
Normal Q-Q Plot of SELS



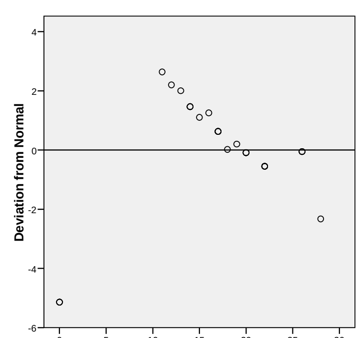
Detrended Normal Q-Q Plot of SELS



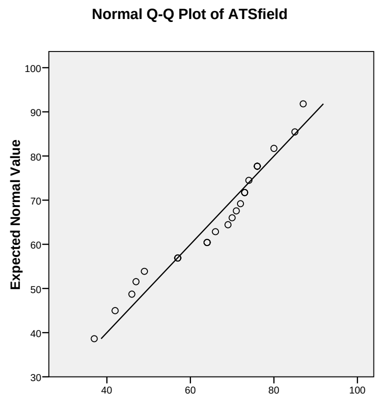
Normal Q-Q Plot of SCsum4

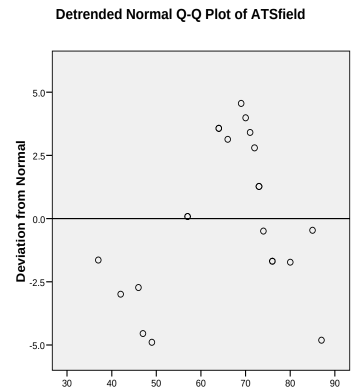


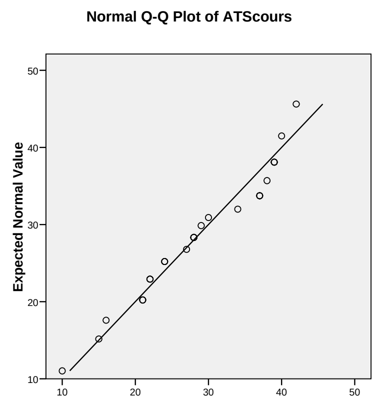
Detrended Normal Q-Q Plot of SCsum4

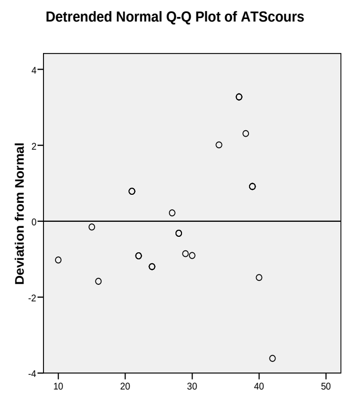


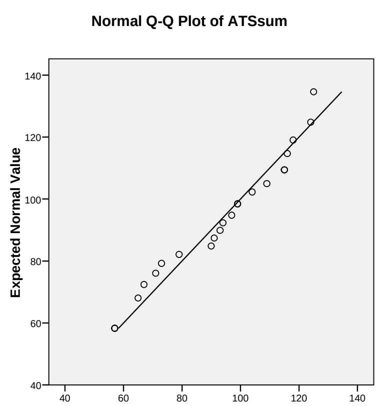
Q-Q plots of Quantitative Variables for the 3213fa Data

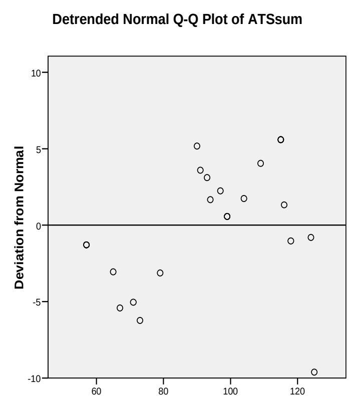




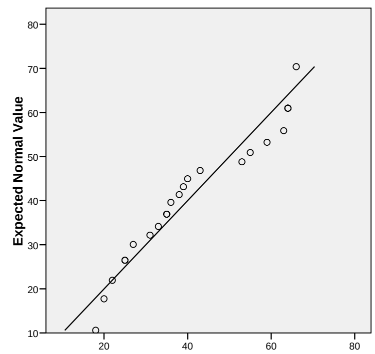




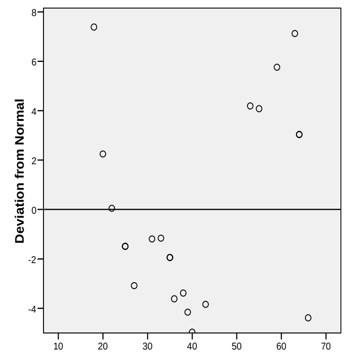




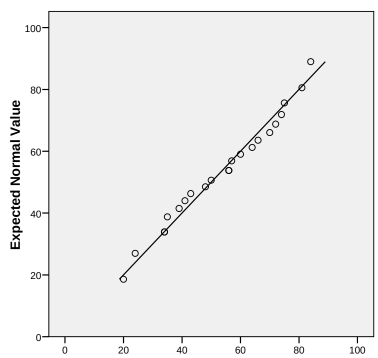
Normal Q-Q Plot of CSSE



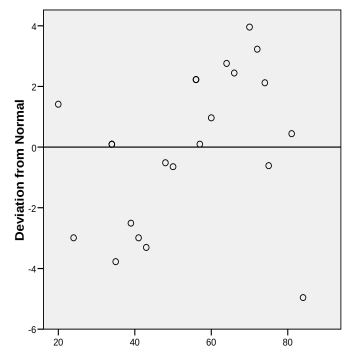
Detrended Normal Q-Q Plot of CSSE



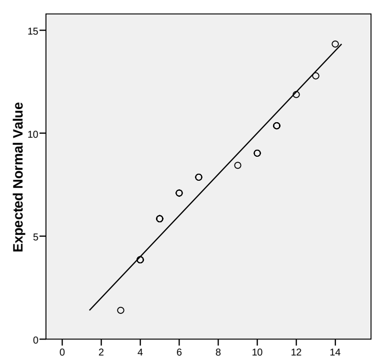
Normal Q-Q Plot of SELS



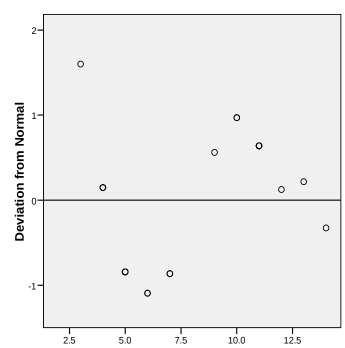
Detrended Normal Q-Q Plot of SELS



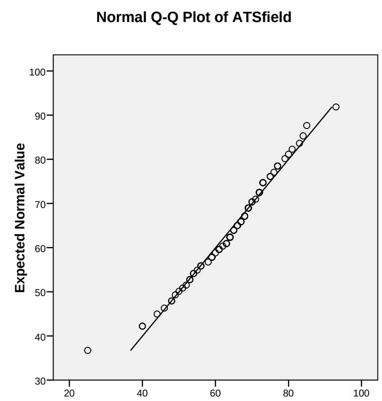
Normal Q-Q Plot of SCI

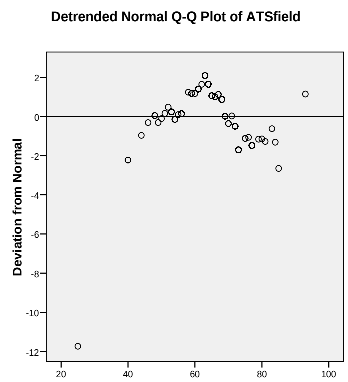


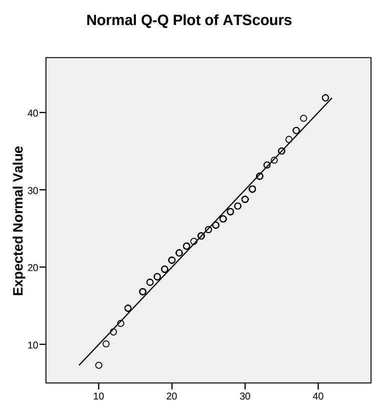
Detrended Normal Q-Q Plot of SCI

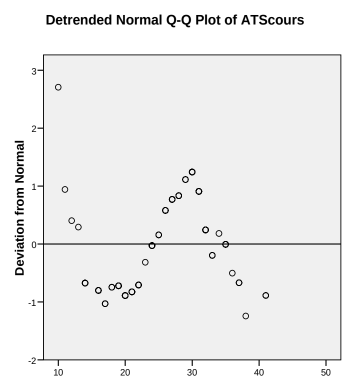


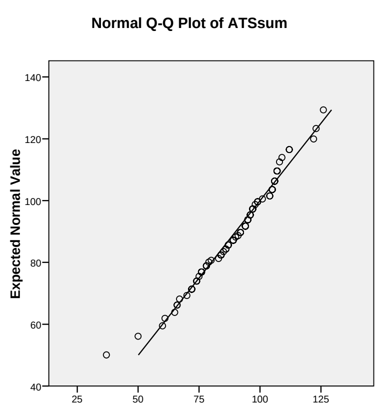
Q-Q plots of Quantitative Variables for the 1473 Data

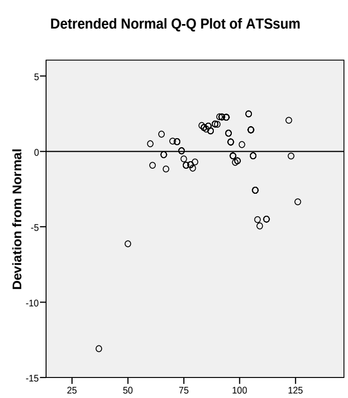


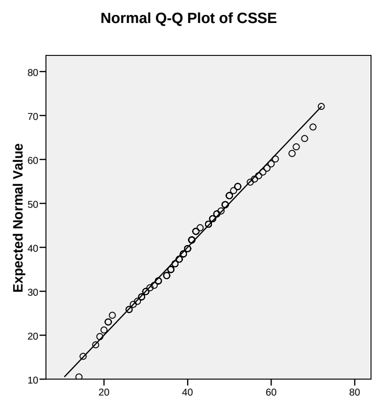


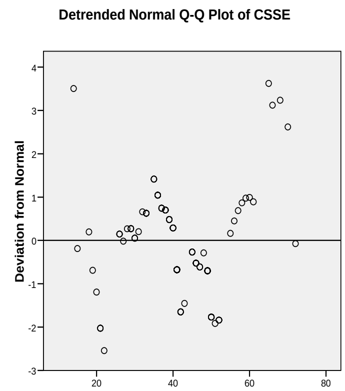


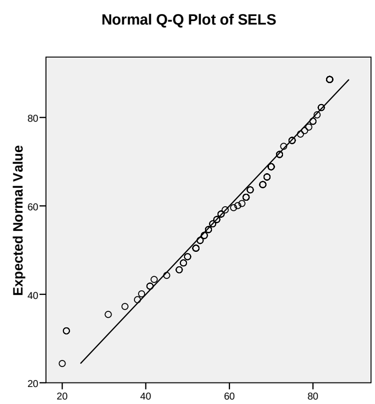


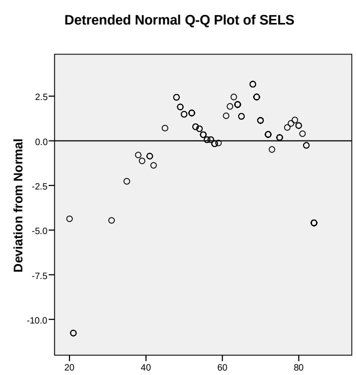


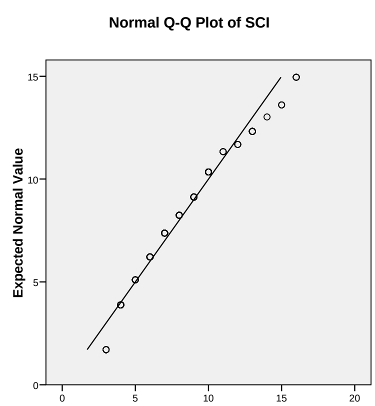


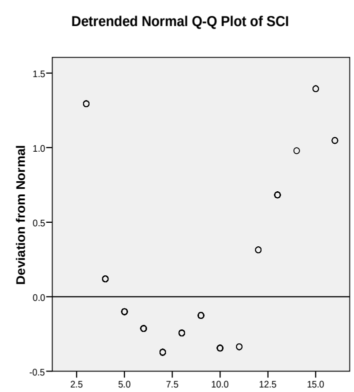

Normal Q-Q Plot of CSSE
Expected Normal Value
80
70
60
50
40
30
20
10
20
40
60
80


Detrended Normal Q-Q Plot of CSSE
Deviation from Normal
4
3
2
1
0
-1
-2
-3
20
40
60
80


Normal Q-Q Plot of SELS
Expected Normal Value
80
60
40
20
20
40
60
80


Detrended Normal Q-Q Plot of SELS
Deviation from Normal
2.5
0.0
-2.5
-5.0
-7.5
-10.0
20
40
60
80


Normal Q-Q Plot of SCI
Expected Normal Value
15
10
5
0
0
5
10
15
20


Detrended Normal Q-Q Plot of SCI
Deviation from Normal
1.5
1.0
0.5
0.0
-0.5
2.5
5.0
7.5
10.0
12.5
15.0

The Kolmogorov-Smirnov tests of normality for all of the variables for 3213fa.

One-Sample Kolmogorov-Smirnov Test

		ATSfield	ATScour	ATSsum	CSSE	SELS	SCI
N		22	22	22	22	22	22
Normal Parameters(a,b)	Mean	65.23	28.32	93.55	40.50	53.77	7.86
	Std. Deviation	13.921	9.057	21.505	15.650	18.426	3.385
Most Extreme Differences	Absolute	.152	.149	.116	.149	.094	.165
	Positive	.105	.092	.103	.149	.084	.165
	Negative	-.152	-.149	-.116	-.109	-.094	-.145
Kolmogorov-Smirnov Z		.714	.700	.546	.699	.439	.773
Asymp. Sig. (2-tailed)		.688	.711	.927	.712	.991	.588

a Test distribution is Normal.
b Calculated from data.

The Kolmogorov-Smirnov tests of normality for all of the variables for 1473.

One-Sample Kolmogorov-Smirnov Test

		ATSfield	ATScour	ATSsum	CSSE	SELS	SCI
N		81	81	81	81	81	81
Normal Parameters(a,b)	Mean	64.30	25.42	89.72	41.28	61.49	8.10
	Std. Deviation	11.374	7.480	16.354	12.707	15.322	3.113
Most Extreme Differences	Absolute	.095	.088	.085	.070	.097	.110
	Positive	.062	.062	.059	.070	.071	.110
	Negative	-.095	-.088	-.085	-.055	-.097	-.057
Kolmogorov-Smirnov Z		.851	.791	.763	.631	.869	.992
Asymp. Sig. (2-tailed)		.464	.559	.605	.821	.437	.279

a Test distribution is Normal.
b Calculated from data.

The Kolmogorov-Smirnov tests of normality for SCgraphs and SCcenter from the 3213sp data

One-Sample Kolmogorov-Smirnov Test

		ATSfield	ATScours	ATSsum	CSSE	SELS
N		22	22	22	22	22
Normal Parameters(a,b)	Mean	69.36	26.05	95.41	43.55	58.27
	Std. Deviation	8.539	5.499	12.786	16.344	16.304
Most Extreme Differences	Absolute	.165	.128	.113	.105	.126
	Positive	.104	.077	.082	.077	.100
	Negative	-.165	-.128	-.113	-.105	-.126
Kolmogorov-Smirnov Z		.773	.599	.529	.493	.593
Asymp. Sig. (2-tailed)		.588	.866	.942	.968	.874

a Test distribution is Normal.
b Calculated from data.

The Kolmogorov-Smirnov tests of normality for SCgraphs and SCcenter from the 3213sp data excluding participants #17 and #18.

One-Sample Kolmogorov-Smirnov Test

		SCgraphs	SCcenter
N		20	20
Normal Parameters(a,b)	Mean	6.30	1.95
	Std. Deviation	1.780	1.432
Most Extreme Differences	Absolute	.153	.196
	Positive	.147	.196
	Negative	-.153	-.154
Kolmogorov-Smirnov Z		.684	.879
Asymp. Sig. (2-tailed)		.738	.423

a Test distribution is Normal.
b Calculated from data.

The Kolmogorov-Smirnov tests of normality for SCspread, SCprob, and SCsum from the 3213sp data excluding participants #1, #17, and #18.

One-Sample Kolmogorov-Smirnov Test

		SCspread	SCprob	SCsum4
N		19	19	19
Normal Parameters(a,b)	Mean	4.84	5.89	18.74
	Std. Deviation	2.267	1.823	4.759
Most Extreme Differences	Absolute	.224	.202	.116
	Positive	.224	.124	.116
	Negative	-.129	-.202	-.094
Kolmogorov-Smirnov Z		.975	.879	.506
Asymp. Sig. (2-tailed)		.297	.423	.960

a Test distribution is Normal.
b Calculated from data.

Appendix F

Descriptive Statistics, Original SPSS™ Outputs

Descriptive Statistics for the 1473 Participants

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
ATSfield	81	25	93	64.30	11.374
ATScours	81	10	41	25.42	7.480
ATSsum	81	37	126	89.72	16.354
CSSE	81	14	72	41.28	12.707
SELS	81	20	84	61.49	15.322
SCI	81	3	16	8.10	3.113
Valid N (listwise)	81				

Descriptive Statistics for the 1473 Participants Who Are in the Elementary Education Program

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
ATSfield	12	25	80	62.33	14.067
ATScours	12	12	36	25.75	7.533
ATSsum	12	37	112	88.08	20.659
CSSE	12	29	72	47.42	10.749
SELS	12	31	84	66.33	16.422
SCI	12	4	15	9.08	3.315
Valid N (listwise)	12				

Descriptive Statistics for the 1473 Participants Who Are Not in the Elementary Education Program

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
ATSfield	69	40	93	64.64	10.926
ATScours	69	10	41	25.36	7.524
ATSsum	69	50	126	90.00	15.654
CSSE	69	14	70	40.22	12.785
SELS	69	20	84	60.65	15.090
SCI	69	3	16	7.93	3.069
Valid N (listwise)	69				

Descriptive Statistics for the 3213fa Participants

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
ATSfield	22	37	87	65.23	13.921
ATScours	22	10	42	28.32	9.057
ATSsum	22	57	125	93.55	21.505
CSSE	22	18	66	40.50	15.650
SELS	22	20	84	53.77	18.426
SCI	22	3	14	7.86	3.385
Valid N (listwise)	22				

Descriptive Statistics for the 3213sp Participants: Variables in Common with 3213fa

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
ATSfield	22	53	81	69.36	8.539
ATScours	22	14	34	26.05	5.499
ATSsum	22	70	114	95.41	12.786
CSSE	22	14	75	43.55	16.344
SELS	22	19	84	58.27	16.304
Valid N (listwise)	22				

432: Descriptive Statistics for Variables in Common with 3213sp

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
ATSfield	6	73	84	78.50	4.764
ATScours	6	30	38	35.00	3.347
ATSsum	6	106	122	113.50	6.156
CSSE	6	32	57	43.50	11.041
SELS	6	58	84	70.33	9.438
SCgraphs	6	6	12	9.17	2.137
SCcenter	6	0	5	3.50	1.761
SCspread	6	0	12	7.67	4.227
SCprob	6	6	9	7.33	1.211
SCsum4	6	23	33	27.67	4.546
WKen	4	2	4	2.75	.957
WKfr	4	3	4	3.25	.500
WKhp	4	3	4	3.50	.577
WKpd	4	3	4	3.25	.500
WKin	4	4	4	4.00	.000
WKls	4	2	3	2.75	.500
WKtotal	4	18	23	19.50	2.380

3213sp: Descriptive Statistics for Variables in Common with 432

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
ATSfield	21	53	81	69.57	8.692
ATScours	21	14	34	26.14	5.615
ATSsum	21	70	114	95.71	13.020
CSSE	21	14	75	42.71	16.264
SELS	21	19	84	57.95	16.636
SCgraphs	21	0	9	6.00	2.214
SCcenter	21	0	6	1.86	1.459
SCspread	21	0	9	4.38	2.598
SCprob	21	0	8	5.33	2.477
SCsum4	21	0	28	17.57	6.177
WKen	21	1	4	2.57	.870
WKfr	21	1	4	3.05	.805
WKhp	21	1	4	2.86	.727
WKpd	21	1	4	3.10	.831
WKin	21	1	4	3.71	.717
WKls	21	1	3	1.57	.746
WKtotal	21	7	21	16.86	3.719

The descriptive statistics outputs from SPSS™ *for 3213sp21.*

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
WKen	21	1	4	2.57	.870
WKfr	21	1	4	3.05	.805
WKhp	21	1	4	2.86	.727
WKpd	21	1	4	3.10	.831
WKin	21	1	4	3.71	.717
WKls	21	1	3	1.57	.746
WKtotal	21	7	21	16.86	3.719
Valid N (listwise)	21				

The descriptive statistics outputs from SPSS™ *for 432*

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
WKen	4	2	4	2.75	.957
WKfr	4	3	4	3.25	.500
WKhp	4	3	4	3.50	.577
WKpd	4	3	4	3.25	.500
WKin	4	4	4	4.00	.000
WKls	4	2	3	2.75	.500
WKtotal	4	18	23	19.50	2.380
Valid N (listwise)	4				

Appendix G

ANOVA, Original SPSS™ Outputs

ANOVA Results for Each of the Variables Measured in the 1473 Course. These Results Compare Elementary Education Majors to All Other 1473 Majors Combined.

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
ATSfield	Between Groups	54.280	1	54.280	.417	.521
	Within Groups	10294.609	79	130.312		
	Total	10348.889	80			
ATScours	Between Groups	1.536	1	1.536	.027	.870
	Within Groups	4474.192	79	56.635		
	Total	4475.728	80			
ATSsum	Between Groups	37.552	1	37.552	.139	.710
	Within Groups	21358.917	79	270.366		
	Total	21396.469	80			
CSSE	Between Groups	529.813	1	529.813	3.379	.070
	Within Groups	12386.656	79	156.793		
	Total	12916.469	80			
SELS	Between Groups	329.928	1	329.928	1.413	.238
	Within Groups	18450.319	79	233.548		
	Total	18780.247	80			
SCI	Between Groups	13.656	1	13.656	1.417	.238
	Within Groups	761.554	79	9.640		
	Total	775.210	80			

ANOVA for the 1473 (all 81) Compared to the 3213fa (all 22) Participants

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
ATSfield	Between Groups	14.995	1	14.995	.105	.747
	Within Groups	14418.753	101	142.760		
	Total	14433.748	102			
ATScours	Between Groups	145.344	1	145.344	2.368	.127
	Within Groups	6198.501	101	61.371		
	Total	6343.845	102			
ATSsum	Between Groups	253.707	1	253.707	.824	.366
	Within Groups	31107.924	101	307.999		
	Total	31361.631	102			
CSSE	Between Groups	10.633	1	10.633	.059	.808
	Within Groups	18059.969	101	178.812		
	Total	18070.602	102			
SELS	Between Groups	1031.404	1	1031.404	4.021	.048
	Within Groups	25910.111	101	256.536		
	Total	26941.515	102			
SCI	Between Groups	.956	1	.956	.095	.758
	Within Groups	1015.801	101	10.057		
	Total	1016.757	102			

ANOVA between the 3213fa (n=22) and the 1473 elem. Ed. Majors (n=12) participants.

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
ATSfield	Between Groups	65.029	1	65.029	.333	.568
	Within Groups	6246.530	32	195.204		
	Total	6311.559	33			
ATScours	Between Groups	51.213	1	51.213	.698	.410
	Within Groups	2347.023	32	73.344		
	Total	2398.235	33			
ATSsum	Between Groups	231.658	1	231.658	.515	.478
	Within Groups	14406.371	32	450.199		
	Total	14638.029	33			
CSSE	Between Groups	371.466	1	371.466	1.853	.183
	Within Groups	6414.417	32	200.451		
	Total	6785.882	33			
SELS	Between Groups	1225.029	1	1225.029	3.883	.057
	Within Groups	10096.530	32	315.517		
	Total	11321.559	33			
SCI	Between Groups	11.551	1	11.551	1.022	.320
	Within Groups	361.508	32	11.297		
	Total	373.059	33			

*ANOVA between the 3213fa (n=22) and the 3213sp (n=22) participants***.**

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
ATSfield	Between Groups	188.205	1	188.205	1.411	.242
	Within Groups	5600.955	42	133.356		
	Total	5789.159	43			
ATScours	Between Groups	56.818	1	56.818	1.012	.320
	Within Groups	2357.727	42	56.136		
	Total	2414.545	43			
ATSsum	Between Groups	38.205	1	38.205	.122	.729
	Within Groups	13144.773	42	312.971		
	Total	13182.977	43			
CSSE	Between Groups	102.023	1	102.023	.398	.531
	Within Groups	10752.955	42	256.023		
	Total	10854.977	43			
SELS	Between Groups	222.750	1	222.750	.736	.396
	Within Groups	12712.227	42	302.672		
	Total	12934.977	43			

ANOVA between the 1473 (all 81) and the 3213sp (all 22) participants.

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
ATSfield	Between Groups	444.253	1	444.253	3.777	.055
	Within Groups	11879.980	101	117.624		
	Total	12324.233	102			
ATScours	Between Groups	6.773	1	6.773	.134	.715
	Within Groups	5110.683	101	50.601		
	Total	5117.456	102			
ATSsum	Between Groups	560.737	1	560.737	2.281	.134
	Within Groups	24829.787	101	245.839		
	Total	25390.524	102			
CSSE	Between Groups	88.484	1	88.484	.482	.489
	Within Groups	18525.924	101	183.425		
	Total	18614.408	102			
SELS	Between Groups	179.506	1	179.506	.744	.390
	Within Groups	24362.611	101	241.214		
	Total	24542.117	102			

ANOVA between the 3213sp (n=22) and the 1473 elem. Ed. Majors (n=12) participants.

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
ATSfield	Between Groups	383.772	1	383.772	3.312	.078
	Within Groups	3707.758	32	115.867		
	Total	4091.529	33			
ATScours	Between Groups	.678	1	.678	.017	.896
	Within Groups	1259.205	32	39.350		
	Total	1259.882	33			
ATSsum	Between Groups	416.706	1	416.706	1.641	.209
	Within Groups	8128.235	32	254.007		
	Total	8544.941	33			
CSSE	Between Groups	116.364	1	116.364	.541	.467
	Within Groups	6880.371	32	215.012		
	Total	6996.735	33			
SELS	Between Groups	504.499	1	504.499	1.888	.179
	Within Groups	8549.030	32	267.157		
	Total	9053.529	33			

ANOVA between the 1473 (Ed, n=12) and all 3213 (n=44) participants.

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
ATSfield	Between Groups	232.156	1	232.156	1.574	.215
	Within Groups	7965.826	54	147.515		
	Total	8197.982	55			
ATScours	Between Groups	19.330	1	19.330	.343	.560
	Within Groups	3038.795	54	56.274		
	Total	3058.125	55			
ATSsum	Between Groups	385.463	1	385.463	1.164	.285
	Within Groups	17877.894	54	331.072		
	Total	18263.357	55			
CSSE	Between Groups	274.320	1	274.320	1.222	.274
	Within Groups	12125.894	54	224.554		
	Total	12400.214	55			
SELS	Between Groups	1002.338	1	1002.338	3.404	.071
	Within Groups	15901.644	54	294.475		
	Total	16903.982	55			

ANOVA between the 1473 (n=81) and all 3213 (n=44) participants.

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
ATSfield	Between Groups	256.464	1	256.464	1.955	.165
	Within Groups	16138.048	123	131.204		
	Total	16394.512	124			
ATScours	Between Groups	88.526	1	88.526	1.580	.211
	Within Groups	6890.274	123	56.018		
	Total	6978.800	124			
ATSsum	Between Groups	646.346	1	646.346	2.299	.132
	Within Groups	34579.446	123	281.134		
	Total	35225.792	124			
CSSE	Between Groups	15.562	1	15.562	.081	.777
	Within Groups	23771.446	123	193.264		
	Total	23787.008	124			
SELS	Between Groups	853.448	1	853.448	3.310	.071
	Within Groups	31715.224	123	257.847		
	Total	32568.672	124			

ANOVA comparison of Common Variables Between the 3213sp and 432 participants

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
ATSfield	Between Groups	372.024	1	372.024	5.725	.025
	Within Groups	1624.643	25	64.986		
	Total	1996.667	26			
ATScours	Between Groups	366.095	1	366.095	13.331	.001
	Within Groups	686.571	25	27.463		
	Total	1052.667	26			
ATSsum	Between Groups	1476.214	1	1476.214	10.309	.004
	Within Groups	3579.786	25	143.191		
	Total	5056.000	26			
CSSE	Between Groups	2.881	1	2.881	.012	.913
	Within Groups	5899.786	25	235.991		
	Total	5902.667	26			
SELS	Between Groups	715.344	1	715.344	2.990	.096
	Within Groups	5980.286	25	239.211		
	Total	6695.630	26			
SCgraphs	Between Groups	46.796	1	46.796	9.682	.005
	Within Groups	120.833	25	4.833		
	Total	167.630	26			
SCcenter	Between Groups	12.595	1	12.595	5.422	.028
	Within Groups	58.071	25	2.323		
	Total	70.667	26			
SCspread	Between Groups	50.381	1	50.381	5.616	.026
	Within Groups	224.286	25	8.971		
	Total	274.667	26			
SCprob	Between Groups	18.667	1	18.667	3.590	.070
	Within Groups	130.000	25	5.200		
	Total	148.667	26			
SCsum4	Between Groups	475.598	1	475.598	13.722	.001
	Within Groups	866.476	25	34.659		
	Total	1342.074	26			

Appendix H

Regression Analysis Outputs for CSSE to Predict SCI for 1473ed

1473ed regression model for CSSE to predict SCI scores.

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.024(a)	.001	-.099	3.476

a Predictors: (Constant), CSSE

ANOVA(b)

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	.072	1	.072	.006	.940(a)
	Residual	120.844	10	12.084		
	Total	120.917	11			

a Predictors: (Constant), CSSE
b Dependent Variable: SCI

Coefficients(a)

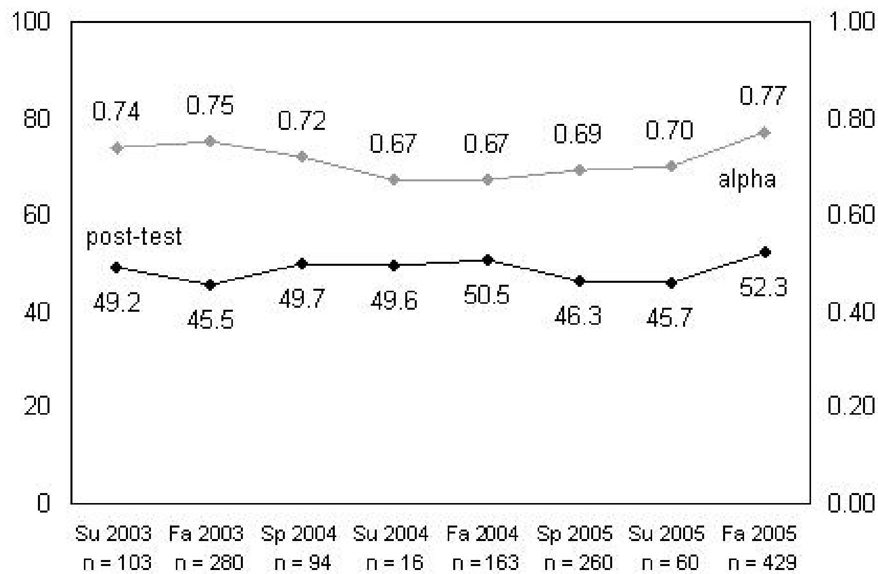
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta	B	Std. Error
1	(Constant)	8.726	4.731		1.844	.095
	CSSE	.008	.098	.024	.077	.940

a Dependent Variable: SCI

Appendix I

Percentage Scores for the Participants in the Original SCI Study

Results of the total administration of the SCI (engineers, etc),



(Rhoades, 2006)

Appendix J

Original Model Summaries, ANOVA and Standard Error of the Coefficients for Each Regression Model Generated

Model Summary, ANOVA, and Coefficients of the Stepwise Regression to Predict SCI for the 3213fa Participants

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.507(a)	.257	.220	2.990

a Predictors: (Constant), ATSsum

ANOVA(b)

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	61.789	1	61.789	6.911	.016(a)
	Residual	178.802	20	8.940		
	Total	240.591	21			

a Predictors: (Constant), ATSsum
b Dependent Variable: SCI

Coefficients(a)

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta	B	Std. Error
1	(Constant)	.402	2.909		.138	.891
	ATSsum	.080	.030	.507	2.629	.016

a Dependent Variable: SCI

Excluded Variables(b)

Model		Beta In	t	Sig.	Partial Correlation	Collinearity Statistics: Tolerance
1	collstat	-.219(a)	-.954	.352	-.214	.705
	ATSfield	.148(a)	.212	.834	.049	.080
	ATScours	-.096(a)	-.212	.834	-.049	.190
	CSSE	-.145(a)	-.519	.610	-.118	.498
	SELS	-.277(a)	-1.028	.317	-.230	.510

a Predictors in the Model: (Constant), ATSsum
b Dependent Variable: SCI

Model Summary, ANOVA, and Coefficients of the Stepwise Regression to Predict SCI for the 3213fa Participants

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.525(a)	.276	.248	3.606

a Predictors: (Constant), ATSsum

ANOVA(b)

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	128.831	1	128.831	9.906	.004(a)
	Residual	338.133	26	13.005		
	Total	466.964	27			

a Predictors: (Constant), ATSsum
b Dependent Variable: SCI

Coefficients(a)

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta	B	Std. Error
1	(Constant)	-1.266	3.321		-.381	.706
	ATSsum	.105	.033	.525	3.147	.004

a Dependent Variable: SCI

Excluded Variables(b)

Model		Beta In	t	Sig.	Partial Correlation	Collinearity Statistics: Tolerance
1	collstat	.078(a)	.397	.694	.079	.743
	ATSfield	-.009(a)	-.014	.989	-.003	.072
	ATScours	.006(a)	.014	.989	.003	.182
	CSSE	-.022(a)	-.100	.921	-.020	.598
	SELS	-.111(a)	-.440	.664	-.088	.452

a Predictors in the Model: (Constant), ATSsum
b Dependent Variable: SCI

Model Summary, ANOVA, and Coefficients for the Quadratic Regression Model to Predict SCsum4 scores for the Combined 3213fa and 432 Participants Using the Variable ATSsum

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.577(a)	.333	.280	3.529

a Predictors: (Constant), ATSsumSq, ATSsum

ANOVA(b)

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	155.635	2	77.818	6.249	.006(a)
	Residual	311.329	25	12.453		
	Total	466.964	27			

a Predictors: (Constant), ATSsumSq, ATSsum
b Dependent Variable: SCI

Coefficients(a)

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta	B	Std. Error
1	(Constant)	-21.767	14.346		-1.517	.142
	ATSsum	.578	.324	2.904	1.782	.087
	ATSsumSq	-.003	.002	-2.390	-1.467	.155

a Dependent Variable: SCI

Model Summary, ANOVA, and Coefficients of the Stepwise Regression to Predict SCsum4 scores for the 3213sp19 Participants

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.687(a)	.472	.441	3.558

a Predictors: (Constant), ATSfield

ANOVA(b)

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	192.472	1	192.472	15.204	.001(a)
	Residual	215.212	17	12.660		
	Total	407.684	18			

a Predictors: (Constant), ATSfield
b Dependent Variable: SCsum4

Coefficients(a)

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta	B	Std. Error
1	(Constant)	-7.133	6.685		-1.067	.301
	ATSfield	.374	.096	.687	3.899	.001

a Dependent Variable: SCsum4

Excluded Variables(b)

Model		Beta In	t	Sig.	Partial Correlation	Collinearity Statistics: Tolerance
1	collstat	-.169(a)	-.952	.355	-.232	.986
	ATScours	.131(a)	.582	.569	.144	.637
	ATSsum	.303(a)	.582	.569	.144	.119
	CSSE	-.276(a)	-1.288	.216	-.306	.649
	SELS	.047(a)	.211	.836	.053	.653

a Predictors in the Model: (Constant), ATSfield
b Dependent Variable: SCsum4

Model Summary, ANOVA, and Coefficients for the Quadratic Regression Model to Predict SCsum4 scores for 3213sp19 Participants Using the Variable ATSfield

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.789(a)	.623	.575	3.101

a Predictors: (Constant), ATSfldSq, ATSfield

ANOVA(b)

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	253.821	2	126.910	13.197	.000(a)
	Residual	153.864	16	9.616		
	Total	407.684	18			

a Predictors: (Constant), ATSfldSq, ATSfield
b Dependent Variable: SCsum4

Model Summary, ANOVA, and Coefficients of the Stepwise Regression to Predict SCsum4 scores for the Combined 3213sp19 and 432 Participants

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.744(a)	.553	.534	4.123

a Predictors: (Constant), ATSsum

ANOVA(b)

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	483.688	1	483.688	28.456	.000(a)
	Residual	390.952	23	16.998		
	Total	874.640	24			

a Predictors: (Constant), ATSsum
b Dependent Variable: SCsum4

Coefficients(a)

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta	B	Std. Error
1	(Constant)	-10.924	6.019		-1.815	.083
	ATSsum	.320	.060	.744	5.334	.000

a Dependent Variable: SCsum4

Excluded Variables(b)

Model		Beta In	t	Sig.	Partial Correlation	Collinearity Statistics: Tolerance
1	collstat	.076(a)	.517	.610	.110	.919
	ATSfield	-.103(a)	-.245	.809	-.052	.115
	ATScours	.075(a)	.245	.809	.052	.216
	CSSE	-.112(a)	-.681	.503	-.144	.736
	SELS	-.026(a)	-.127	.900	-.027	.476

a Predictors in the Model: (Constant), ATSsum
b Dependent Variable: SCsum4

Model Summary, ANOVA, and Coefficients for the Quadratic Regression Model to Predict SCsum4 scores for the Combined 3213sp19 and 432 Participants Using the Variable ATSsum

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.762(a)	.581	.543	4.079

a Predictors: (Constant), ATSsumSq, ATSsum

ANOVA(b)

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	508.511	2	254.255	15.278	.000(a)
	Residual	366.129	22	16.642		
	Total	874.640	24			

a Predictors: (Constant), ATSsumSq, ATSsum
b Dependent Variable: SCsum4

Coefficients(a)

		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
Model		B	Std. Error	Beta	B	Std. Error
1	(Constant)	34.064	37.315		.913	.371
	ATSsum	-.643	.790	-1.497	-.814	.425
	ATSsumSq	.005	.004	2.247	1.221	.235

a Dependent Variable: SCsum4

Model Summary, ANOVA, and Coefficients of the Stepwise Regression to Predict SELS Results for the 3213fa and 1473ed Participants Combined

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.500(a)	.250	.227	14.565

a Predictors: (Constant), CSSE

ANOVA(b)

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	2265.510	1	2265.510	10.680	.003(a)
	Residual	6788.020	32	212.126		
	Total	9053.529	33			

a Predictors: (Constant), CSSE
b Dependent Variable: SELS

Coefficients(a)

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta	B	Std. Error
1	(Constant)	35.562	8.209		4.332	.000
	CSSE	.569	.174	.500	3.268	.003

a Dependent Variable: SELS

Excluded Variables(b)

Model		Beta In	t	Sig.	Partial Correlation	Collinearity Statistics: Tolerance
1	collstat	.124(a)	.802	.429	.142	.996
	ATSfield	-.007(a)	-.044	.965	-.008	.922
	ATScours	.216(a)	1.306	.201	.228	.840
	ATSsum	.074(a)	.447	.658	.080	.880
	SCI	.089(a)	.574	.570	.103	.985

a Predictors in the Model: (Constant), CSSE
b Dependent Variable: SELS

Model Summary, ANOVA, and Coefficients of the Stepwise Regression to Predict SELS Results for the Combined 3213fa and 1473ed Participants

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.671(a)	.450	.432	13.954

a Predictors: (Constant), CSSE

ANOVA(b)

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	5090.564	1	5090.564	26.143	.000(a)
	Residual	6230.995	32	194.719		
	Total	11321.559	33			

a Predictors: (Constant), CSSE
b Dependent Variable: SELS

Coefficients(a)

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta	B	Std. Error
1	(Constant)	21.014	7.658		2.744	.010
	CSSE	.866	.169	.671	5.113	.000

a Dependent Variable: SELS

Excluded Variables(b)

Model		Beta In	t	Sig.	Partial Correlation	Collinearity Statistics: Tolerance
1	collstat	.002(a)	.011	.991	.002	.783
	ATSfield	-.039(a)	-.258	.798	-.046	.763
	ATScours	.217(a)	1.512	.141	.262	.799
	ATSsum	.068(a)	.441	.662	.079	.749

a Predictors in the Model: (Constant), CSSE
b Dependent Variable: SELS

Model Summary, ANOVA, and Coefficients of the Stepwise Regression to Predict SELS Results for the Combined 3213sp, 3213fa, and 1473ed Participants

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.612(a)	.375	.363	13.991
2	.664(b)	.440	.419	13.360

a Predictors: (Constant), CSSE
b Predictors: (Constant), CSSE, ATScours

ANOVA(c)

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	6334.332	1	6334.332	32.362	.000(a)
	Residual	10569.650	54	195.734		
	Total	16903.982	55			
2	Regression	7444.093	2	3722.046	20.853	.000(b)
	Residual	9459.890	53	178.488		
	Total	16903.982	55			

a Predictors: (Constant), CSSE
b Predictors: (Constant), CSSE, ATScours
c Dependent Variable: SELS

Coefficients(a)

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta	B	Std. Error
1	(Constant)	27.372	5.738		4.770	.000
	CSSE	.715	.126	.612	5.689	.000
2	(Constant)	15.799	7.181		2.200	.032
	CSSE	.561	.135	.481	4.165	.000
	ATScours	.677	.271	.288	2.494	.016

a Dependent Variable: SELS

Excluded Variables(c)

Model		Beta In	t	Sig.	Partial Correlation	Collinearity Statistics: Tolerance
1	collstat	.127(a)	1.139	.260	.155	.922
	ATSfield	.082(a)	.667	.508	.091	.770
	ATScours	.288(a)	2.494	.016	.324	.792
	ATSsum	.183(a)	1.480	.145	.199	.742
2	collstat	.087(b)	.804	.425	.111	.899
	ATSfield	-.157(b)	-1.060	.294	-.145	.478
	ATSsum	-.235(b)	-1.060	.294	-.145	.215

a Predictors in the Model: (Constant), CSSE
b Predictors in the Model: (Constant), CSSE, ATScours
c Dependent Variable: SELS

Model Summary, ANOVA, and Coefficients for the Quadratic Regression Model to Predict SELS scores for the Combined 3213sp, 3213fa, and 1473ed Participants Using the Variable CSSE

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.617(a)	.380	.357	14.061

a Predictors: (Constant), CSSEsqrd, CSSE

ANOVA(b)

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	6425.812	2	3212.906	16.251	.000(a)
	Residual	10478.170	53	197.701		
	Total	16903.982	55			

a Predictors: (Constant), CSSEsqrd, CSSE
b Dependent Variable: SELS

Coefficients(a)

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta	B	Std. Error
1	(Constant)	35.702	13.537		2.637	.011
	CSSE	.274	.660	.235	.415	.680
	CSSEsqrd	.005	.008	.385	.680	.499

a Dependent Variable: SELS

Model Summary, ANOVA, and Coefficients of the Stepwise Regression to Predict SELS Results for the 3213fa Participants

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.740(a)	.547	.525	12.705
2	.826(b)	.683	.649	10.912

a Predictors: (Constant), ATScours
b Predictors: (Constant), ATScours, CSSE

ANOVA(c)

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	3901.575	1	3901.575	24.171	.000(a)
	Residual	3228.288	20	161.414		
	Total	7129.864	21			
2	Regression	4867.359	2	2433.679	20.437	.000(b)
	Residual	2262.505	19	119.079		
	Total	7129.864	21			

a Predictors: (Constant), ATScours
b Predictors: (Constant), ATScours, CSSE
c Dependent Variable: SELS

Coefficients(a)

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta	B	Std. Error
1	(Constant)	11.157	9.081		1.229	.234
	ATScours	1.505	.306	.740	4.916	.000
2	(Constant)	4.958	8.098		.612	.548
	ATScours	.955	.326	.469	2.927	.009
	CSSE	.538	.189	.457	2.848	.010

a Dependent Variable: SELS

Excluded Variables(c)

Model		Beta In	t	Sig.	Partial Correlation	Collinearity Statistics: Tolerance
1	collstat	.288(a)	1.839	.082	.389	.822
	ATSfield	.115(a)	.507	.618	.115	.452
	ATSsum	.178(a)	.507	.618	.115	.190
	CSSE	.457(a)	2.848	.010	.547	.649
	SCI	-.137(a)	-.811	.427	-.183	.808
2	collstat	.095(b)	.547	.591	.128	.577
	ATSfield	-.212(b)	-.951	.354	-.219	.339
	ATSsum	-.327(b)	-.951	.354	-.219	.142
	SCI	-.153(b)	-1.067	.300	-.244	.807

a Predictors in the Model: (Constant), ATScours
b Predictors in the Model: (Constant), ATScours, CSSE
c Dependent Variable: SELS

Model Summary, ANOVA, and Coefficients of the Stepwise Regression to Predict ATSfield Results for the 3213sp Participants

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.643(a)	.413	.384	6.704

a Predictors: (Constant), ATScours

ANOVA(b)

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	632.321	1	632.321	14.071	.001(a)
	Residual	898.770	20	44.938		
	Total	1531.091	21			

a Predictors: (Constant), ATScours
b Dependent Variable: ATSfield

Coefficients(a)

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta	B	Std. Error
1	(Constant)	43.372	7.075		6.130	.000
	ATScours	.998	.266	.643	3.751	.001

a Dependent Variable: ATSfield

Excluded Variables(b)

Model		Beta In	t	Sig.	Partial Correlation	Collinearity Statistics: Tolerance
1	collstat	.123(a)	.702	.491	.159	.979
	CSSE	.286(a)	1.446	.164	.315	.713
	SELS	.289(a)	1.344	.195	.295	.610

a Predictors in the Model: (Constant), ATScours
b Dependent Variable: ATSfield

Model Summary, ANOVA, and Coefficients of the Stepwise Regression to Predict WKen Results for the 3213sp19 Participants

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.477(a)	.228	.182	.751

a Predictors: (Constant), ATScours

ANOVA(b)

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	2.829	1	2.829	5.015	.039(a)
	Residual	9.592	17	.564		
	Total	12.421	18			

a Predictors: (Constant), ATScours
b Dependent Variable: WKen

Coefficients(a)

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta	B	Std. Error
1	(Constant)	.799	.836		.955	.353
	ATScours	.071	.032	.477	2.239	.039

a Dependent Variable: WKen

Excluded Variables(b)

Model		Beta In	t	Sig.	Partial Correlation	Collinearity Statistics: Tolerance
1	collstat	-.080(a)	-.367	.719	-.091	.999
	ATSfield	-.297(a)	-1.122	.278	-.270	.637
	ATSsum	-.439(a)	-1.122	.278	-.270	.293
	CSSE	-.308(a)	-1.236	.234	-.295	.710
	SELS	.170(a)	.614	.548	.152	.614
	SCsum4	-.126(a)	-.501	.623	-.124	.753

a Predictors in the Model: (Constant), ATScours
b Dependent Variable: WKen

Model Summary, ANOVA, and Coefficients of the Stepwise Regression to Predict WKhp Results for the 3213sp19 Participants

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.688(a)	.473	.442	.514

a Predictors: (Constant), SELS

ANOVA(b)

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	4.035	1	4.035	15.276	.001(a)
	Residual	4.491	17	.264		
	Total	8.526	18			

a Predictors: (Constant), SELS
b Dependent Variable: WKhp

Coefficients(a)

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta	B	Std. Error
1	(Constant)	1.217	.432		2.817	.012
	SELS	.029	.007	.688	3.908	.001

a Dependent Variable: WKhp

Excluded Variables(b)

Model		Beta In	t	Sig.	Partial Correlation	Collinearity Statistics: Tolerance
1	collstat	.252(a)	1.433	.171	.337	.947
	ATSfield	.224(a)	1.028	.319	.249	.653
	ATScours	.328(a)	1.513	.150	.354	.614
	ATSsum	.336(a)	1.468	.162	.344	.554
	CSSE	.154(a)	.779	.447	.191	.812
	SCsum4	.011(a)	.056	.956	.014	.810

a Predictors in the Model: (Constant), SELS
b Dependent Variable: WKhp

Model Summary, ANOVA, and Coefficients for the Quadratic Regression Model to Predict WKhp scores for the 3213sp19 Participants Using the Variables SELS and SELSsqrd

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.738(a)	.544	.487	.493

a Predictors: (Constant), SELSsqrd, SELS

ANOVA(b)

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	4.642	2	2.321	9.559	.002(a)
	Residual	3.885	16	.243		
	Total	8.526	18			

a Predictors: (Constant), SELSsqrd, SELS
b Dependent Variable: WKhp

Coefficients(a)

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta	B	Std. Error
1	(Constant)	.108	.815		.132	.897
	SELS	.079	.033	1.880	2.432	.027
	SELSsqrd	-.001	.000	-1.221	-1.580	.134

a Dependent Variable: WKhp