The Pennsylvania State University

The Graduate School

College of Education

STATISTICAL LITERACY IN ADULT COLLEGE STUDENTS

A Dissertation in

Adult Education

by

Barbara Ann Wade

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Education

May 2009

The dissertation of Barbara Ann Wade was reviewed and approved* by the following:

Edward Taylor
Associate Professor of Adult Education
Dissertation Advisor
Chair of Committee

Patricia A. Cranton
Professor of Adult Education

Pamela Preston
Associate Professor of Criminal Justice

John Urenko
Assistant Professor of Mathematics

Ian E. Baptiste
Associate Professor of Education
In Charge of Graduate Programs in Adult Education

*Signatures are on file in the Graduate School.

ABSTRACT

Currently statistics is used in society as social indicators that measure, for example crime rates, or research concerning health and medical issues, which are reported by the media. Educated citizens need to be able to understand statistics; it is assumed that adult students who graduate from college will have this ability. Knowing how important statistical literacy is, the purpose of this research was to measure statistical literacy in adult learners before and after they have completed a statistics class, or a research methods class with no prior statistics, or a research methods class with prior statistics.

Participants were 110 adult students, 74 females and 37 males, 72% Caucasian, 26% African American, 1% Native American and 1% Asia/Pacific Islander; ages ranged from 18 to 40 years old (M = 21, S.D. = 3.25), with most reporting marital status as single and being full-time students.

Using a quasi-experimental research design, adult students completed pre- and post-test surveys, which measured statistical literacy, based on Gal's (2004) Model of Statistical Literacy that embraced both knowledge and dispositional elements. Fisher's LSD post-hoc test results showed a statistically significant difference among class types for the knowledge elements, statistical thinking, reasoning, and literacy, but no statistically significant differences for the critical questions. Results from a MANOVA on post-test scores among class types showed no significant differences on the dispositional elements, affect, cognitive competence, difficulty and value. For beliefs, a Chi-square analysis showed that 90% of adult students in the research methods class with prior statistics believed that statistics was relevant in their lives. More importantly, further tests that examined pre- to post-test scores revealed significant differences among class types.

Implications for policy include a re-examination of prerequisites for statistics courses, and requiring completion of a statistics course before a research methods course. Many teaching implications are indicated in this research; however, the most important is the use of constructivist perspectives in the classroom; accordingly, future research should examine teaching methodologies. In addition, statistical literacy needs to be examined again; however, using modifications to the model that separates statistics into its two main branches, descriptive and inferential.

TABLE OF CONTENTS

Chapter 5: DISCUSSION, IMPLICATIONS, LIMITATIONS, CONCLUSION

LIST OF TABLES

LIST OF CHARTS

CHAPTER 1

Introduction

This first chapter of this dissertation lays the foundation for this quantitative research study on statistical literacy. It describes the background of the problem, the teaching of statistics, an overview of the conceptual framework, the problem and purpose statements, hypothesis, significance of the research, the précis of the research design, assumptions and limitations of the research, the organization of the dissertation, and definitions of terms.

Background to the Problem

This section examines why it is important for higher education and adult education to place an emphasis on the teaching of statistics. In the not-so-recent past, the teaching of statistics was not considered to be important in the college curricula, and when it was taught, it was taught in the abstract—it had no application to everyday life. However, a change in mathematics and statistics education was imminent, and during the 1990s a movement toward improving statistical education began to unfold.

Currently and historically, numerical data is used and produced by our political system, but unfortunately, most Americans are unable to understand the reported data. The government uses data extensively in a variety of ways to monitor the social economy. For example, poverty is measured by numbers, and through these numbers individuals may claim benefits. Crime rates are measured as an indicator of public order, divorce rates are used to examine private morality and family life, and opinion polls measure the feelings of a country on a given issue at a given time (Rose, 1991). And as shown in the past, statistics were used for political gain. Take, for example, the American

Census of 1840—unbeknownst to many this Census, like others, was riddled with errors. For example:

> The 1840 Census added a count of the insane and idiots, distinguished by race and mode of support, to the counts of the blind, deaf and dumb, that had been included in 1830. When the results of the Census were published in 1841, the total number of those reported as insane or feeble-minded in the United States was over 17,000. More to the point, perhaps, nearly 3,000 were black, and the rate of insanity amongst free blacks was 11 times higher than that of slaves and six times higher than that of the white population. For those who opposed abolition, like U. S. Vice President John C. Calhoun, these Census figures proved that blacks were congenitally unfit for freedom. (Rose, p. 685)

But more importantly, statistical literacy is more problematic today, because most information is governed by numbers. Numbers underlie everyday decisions, from quantitatively based proposals that shape public policy in education (Steen, 2001), to decisions regarding political candidates, medicines and health, which are based on logic and quantitative information (Moreno, 2002). Many research studies concerning these health and medicine issues are reported regularly in newspapers and magazines with conflicting results, and leave many to become cynical about research in general (Utts, 2003). Also, in this age of information, much of what is reported to citizens is unregulated, unrestricted, and difficult for many to interpret. Because of the inability to understand statistical information, many individuals become either skeptical or believe anything that is printed (Rumsey, 2002). This leaves those who are not mathematically or statically literate outsiders in a society whose inability to participate in public dialogue leads to mistrusting the technocratic elite, resulting in a retreat from public issues, or possibly an aggressive opposition to change.

Especially important to the discussion on statistical literacy is the global economy we live in and the global competition that ensues. Compared to other countries, the United States is behind in mathematics education. According to the Third International Mathematics and Science Study (TIMSS), which examined 38 countries, the United States was significantly behind 14, even with 6 and doing better than 17. One reason the United States remains competitive now is due to our sheer size, which allows us to select the best for our workforce. Labor markets are minimally regulated, giving employers the flexibility to pay based on performance, leaving no guarantees for jobs, wages or benefits (Carnevale & Desrochers, 2003).

Respectively, all citizens in society need to be able to think quantitatively, from farmers to lawyers, from consumers to manufacturers and even jurors to the accused (Moreno, 2002). In fact, as Steen (1997) warns, "an innumerate citizen today is as vulnerable as the illiterate peasant of Gutenberg's time" (p. xxvii). Most schools do not provide a bridge from arithmetic in elementary education to mathematics in high school, to the world of data and statistics. One reason for this disjuncture is the perception about statistics by the public that it is not as rigorous or prestigious as some of the mathematics courses that are in place—for example calculus, which is associated with the hard sciences of physics and engineering. But in fact, the reasoning involved in data-based statistical inference "is harder for students to grasp and explain than the comparable symbol-based problems and proofs in a typical calculus course" (Steen, p. 62).

Traditionally, algebra and calculus courses have been the dominant goals of high school mathematics. These are often the standard courses students need to complete their high school diplomas, and great importance is placed on them, because of college entrance exams, such as the ACT or SAT. These exams often determine if and where high school students will go to college. However, passing the requirements to get accepted into a college does not necessarily prepare students for higher education or later employment. Many students who are accepted into college find themselves required to take remedial courses in intermediate algebra, but unfortunately, the skills that they are required to master are rarely encountered in adult life. “From school to college, mathematics follows an isolated trajectory of increasing difficulty and abstraction whose implicit purpose is to select and prepare the best mathematics students for graduate education in mathematically intensive fields” (Carnevale & Desrochers, 2003, p. 21). On the other hand, courses that promote statistical literacy could help adult learners become better prepared to function in the real world, because in the real world, “mathematical activity begins not with formulas, but with data” (Steen, 2003, p. 59). Our educational system is only beginning to promote the need to understand mathematical reasoning or quantitative reasoning.

How mathematics is being taught—and what is missing from the curricula—has been the impetus for change, and was addressed by the National Council of Teachers of Mathematics (NCTM) in the 1990s to reform teaching at all levels. The call for change—the *Democratization of Mathematics*—is the remedy to close the cultural, political and economic gap between those who are literate and illiterate in mathematics. It is not the dumbing down of mathematics courses, but rather making them more accessible to all students, who later become citizens in a data-driven society (Carnevale & Desrochers, 2003). It also asserts mathematics is no longer a curriculum designed for the elite, as it was in the past, but as a necessary part of general education for all citizens (Moore, 1997). Data has shown education in mathematics "has always been about separation—of rich from poor, of boys from girls, of elites from plebeians…the biggest barrier to upward mobility in educational attainment" (Steen, 2004, n.p.).

Moore further argues that democratization also states mathematics studies need to move away from the esoteric toward the immediately useful. Reformers in mathematics education urge:

> A change of culture toward the concrete, toward applications, toward ability to use mathematical concepts and tools over rigor of detail…they are responding to the pressures of democratization. This is an opportunity for statistics; as mathematical studies shift toward a more utilitarian approach, a larger place for statistics opens up. (p.124)

Hence, the movement, the Democratization of Mathematics, not only places emphasis on the teaching of mathematics, but also the teaching of statistics, especially at the college level.

During some of the changes that were being implemented in mathematics education at the beginning of the Democratization of Mathematics, a concern about adults and mathematics education surfaced. While the teaching of reading and writing to adults has been a primary focus of adult educators, the teaching of mathematics has not. Most research on the teaching of mathematics and adults has been focused on math anxiety or comparing the mathematical abilities of women to men (Rose, 1997). In adult education and literacy, quantitative skills have been viewed as a basic skill adults need to possess. However, defining the types of mathematical skills necessary to function in society have been vague. As defined by the United States National Literacy Act in 1991,

> adult literacy is an individual's ability to read, write and speak in English, and to compute and solve problems at levels of proficiency necessary for functioning on the job and in society, achieving one's goals and developing one's knowledge and potential. (Gal, 1997, p. 13)

From this definition it is clear that computation and problem-solving abilities are of chief concern within adult literacy; however, to specifically outline them remains a challenge. What are the "mathematics-related skills, knowledge, abilities and dispositions adults require to meet the goals of literacy and numeracy" (Gal, 1997, p. 13)? Gal informs us that this is a complicated question to answer, due to the multiple and diverse types of situations adults encounter in their daily lives "involving numbers, measurements, mathematical ideas, patterns, probabilities and events that unfold over time" (p. 13). One area where mathematical skills are interwoven with statistical literacy is the community, as it is of the utmost importance of the community to have an informed citizenship. Here, an individual, who is statistically literate, can comprehend poll results or crime figures, or promote social action by interpreting surveys that may have important environmental implications (Gal).

Others who join the discussion on adult literacy, define mathematical skills adults require to include "computational skills and number sense, statistical and probabilistic knowledge and reasoning, including data representation and interpretation of graphics" (Weist, Higgins, & Frost, 2007, p. 48). These skills are considered life skills for adults by some in that they enable individuals to assess claims, detect fallacies, evaluate risk and weigh evidence (Steen, 1997), which adults encounter in the media in their daily lives.

As important as statistical literacy is, it is surprising to find that it is not included as part of the most commonly used instrument to measure adult literacy in the United States, the National Assessment of Adult Literacy (NAAL). The NAAL is a nationally representative assessment of English literacy among American adults age 16 and older, and examined over 19,000 adults' literacy skills in the United States in 2003. This scale measures adult literacy on three literacy dimensions of adult learning, prose, document and quantitative. Prose literacy refers to skills needed to search, comprehend and use continuous texts—for example, interpreting editorials, news stories, brochures and instructional materials. Document literacy is the knowledge and skills required to search, comprehend and use non-continuous texts in various forms such as job applications, transportation schedules, or drug and food labels. Quantitative literacy embraces the skills required to perform quantitative tasks requiring an adult to identify and perform computations by using numbers embedded in text. Examples of quantitative literacy would include demonstrating the ability to balance a checkbook, compute a tip in a restaurant, or complete an order form (National Center for Education Statistics [NCES], 2006). The scale does not examine adults' ability to interpret statistical data that is presented in multiple ways and in various forms to the general population and further, it does not examine the related abilities necessary to critique various types of numerical or statistical based information (NCES, 2006). However, these measures of literacy will probably not remain stagnant, as various projects are under way to establish content standards for what adults should be able to do after completing their education (Gal, 1997).

Changes on other scales have occurred. Similar to the NAAL, the renamed Adult Literacy and Lifeskills Survey (ALL) measures adult literacy in the United States and in 20 other countries, for example in Canada and Italy. The new instrument has modified its assessment from prose, document and quantitative literacy to prose and document literacy, numeracy and problem solving. The new numeracy scale is replacing the old quantitative scale and is "designed to be broader than the quantitative literacy scale, going beyond applying arithmetic skills to a wider range of mathematical skills such as the use of number sense, estimation and statistics" (NCES, 2006, p. 1).

Results from the ALL surveys suggest adults' numeracy skills are deficient in specific areas. In fact, 25 to 50% of American adults are unable to complete a range of tasks such as, "where the numbers to be used have to be located in different types of forms or texts; where mathematical operations to be performed have to be inferred; or where quantitative information has to be gleaned from graphs or tables" (Gal, 2002, p. 23). While these are important concepts in adult numeracy, missing from the instrument was the understanding of basic statistical concepts that may be relevant to issues in the media or work contexts (Gal).

In comparing numeracy scores with other countries the ALL survey showed that "U.S. adults outperformed adults in Italy in numeracy skills, in 2003, but were outperformed by adults in Switzerland, Norway, Bermuda and Canada" (NCES, 2006, n.p.). Despite the fact that the United States has not included statistical literacy as a national assessment of adults' literacy skills, the Democratization of Mathematics movement has included statistical literacy within its goals, whose purpose is to make children and adults proficient in mathematics and statistics.

As Moore (1997) previously stated, the focus in teaching of mathematics is “a change of culture toward the concrete, toward applications, toward ability to use mathematical concepts and tools over rigor of detail…” (13), and as Gal (1997) has stated, “literacy is an individual’s ability to read, write and speak in English, and to compute and solve problems at levels of proficiency necessary for functioning on the job and in society” (p. 13), with both the Democratization of Mathematics and Adult Literacy complementing each other. The teaching of mathematics in higher education has embraced the importance of teaching statistics and, as adult literacy becomes redefined to reflect the purpose of literacy, what is taught will be reflected in adult literacy measurements throughout the United States. This is essential if the United States is to remain competitive in a global society.

Teaching of Statistics

In higher education, the Democratization of Mathematics has emphasized the necessity of including statistics in various non-mathematical majors, such as the behavioral and social sciences. With many disciplines agreeing and incorporating introductory statistics into their course curricula, an interest in examining how these courses are being currently taught has resulted. Many have been shown to be inclusive of active learning techniques in the statistics classrooms (for example, see Aberson, et al., 2000; 2002; 2003). The three predominant ways to incorporate active learning of statistics into the classroom has been through computer simulations, data analysis programs and Web-based classes.

First, computer-based simulations demonstrated specific statistical concepts and allowed students to analyze data, create charts and graphs, and receive immediate feedback about the choices they have made (Aberson, et al., 2000, 2002, 2003; Morris 2001; Morris, et al., 2002). This type of active learning resulted in higher scores on exams for students who used the computer simulation than students who received instruction in a lecture-only classroom. And students who used computer simulations valued them—they believed they helped them learn specific statistical concepts (Aberson, et al., 2000, 2002, 2003). On the contrary, similar research completed by Morris, et al., found students who attended lecture-only classes in introductory statistics scored higher on assessments than students who participated in using computer simulations.

Second, computer programs such as Excel or SPSS, integrated into the statistics classrooms, allow students to engage hands-on in completing data entry and data analyses (Proctor, 2002; Raymondo & Garrett, 1998; Warner & Meehan, 2001). Similar to the computer simulation programs, learning outcomes from these teaching methods were mixed. Raymondo and Garrett found no statistically significant differences in students' final grades between students who used the computer programs as active learning and students who did not. Differently, Proctor found students who used the data analysis program Excel had a deeper understanding of statistical concepts than the group who used SPSS. Further, when examining students' beliefs about the data analyses programs, Warner and Meehan found that "although the students agreed that the assignments improved their knowledge of statistical concepts, they rated the assignments more highly in terms of improving their computer skills" (p. 297).

Third, are the Web-based instructional statistics programs, which use a combination of a Web-based statistics program and minimal student-instructor face-to-face contact (Bushway & Flowers, 2002; Johnson & Dasgupta, 2005; Symanzik & Vukasinovic, 2006; Utts, et al., 2003; Ward, 2004). Students who attend these types of statistics classes often have hand-in homework assignments (Bushway & Flowers; Symanzik & Vukasinovic; Utts, et al.; Ward), engage in activities such as interactive worksheets, data analyses, and applet demonstrations of statistical concepts (Johnson & Dasgupta; Symanzik & Vukasinovic; Utts, et al.; Ward).

Evaluations of these Web-based instructional statistics courses were based on two criteria, students' preference between a traditional lecture-based statistics classroom and students' learning. Some studies showed students preferred non-traditional teaching methods (i.e., hybrid programs) more than traditional lecture-based teaching methods (Bushway & Flowers, 2002; Johnson & Dasgupta, 2005), while another showed students in the non-traditional statistics classroom felt the workload was excessive, whereas students in the traditional statistics class did not (Utts et al., 2003). No difference in students' preference for Web-based instruction versus a traditionally taught classroom was found in research by Symanzik and Vukasinovic (2006).

More consistent were evaluations of students' learning in Web-based instructional statistics courses and traditional statistics classrooms. Utts et al. (2003) and Ward's (2004) results showed no significant differences in students' knowledge of statistical concepts between students in the traditional (i. e., lecture) and non-traditional (Web-based instruction) classrooms. When final grades were used to evaluate students' learning of statistics, Symanzik and Vukasinovic (2006) and Bushway and Flowers (2002) found no significant differences in grades between students in the Web-based and traditional statistics classrooms.

A few teaching methodologies that incorporated active learning into the statistics classroom used real data sets for statistical analyses (Morris, 2001) or like-real data sets (Cralley & Ruscher, 2001; Proctor, 2002). Unfortunately, a comparison of students' learning was only examined between students who used shared and individual data sets, with students who analyzed individual data sets scoring higher than students who used shared data sets.

Other active learning techniques in the classroom included incorporating everyday problems into the course curricula (Lawson, et al., 2003; Vanderstoep & Shaughnessy, 1997), or by adding a writing component to the course (Rajecki, 2002; Vanderstoep & Shaughnessy). Both methods resulted in students demonstrating deeper knowledge of statistical concepts on assessments.

Important to the research on the teaching of statistics, and hence statistical literacy, is the course context in which the teaching of statistics occurs. Most of the studies examined the teaching of statistics in introductory statistics classes (Bushway & Flower, 2002; Cralley & Ruscher; Johnson & Dasgupta, 2005; Kvam, 2000; Morris, 2001; Morris et al., 2002; Proctor, 2002; Proctor, 2006; Raymondo & Garrett 1998; Symanzik & Vukasinovic, 2006; Ward, 2004), albeit a few studies employed students from research method classes for their study (Morris; Vanderstoep & Shaughnessy, 1997), and it is unclear if students already had taken a statistics course, or were taking it concomitantly with their research method class, or had previously taken a statistics class in a different semester.

This is an important distinction to articulate. There are no empirical studies that have examined statistics and/or research methods course requirements for the behavioral or social sciences. However, it is easy to see how course requirements vary within the same disciplines, by scanning college Web pages and examining course descriptions. And more importantly, when measuring adult learners' gains in statistical learning, it is necessary to distinguish between a statistics course and a research methods course, because they cover similar, but different topics.

Consider the majors of psychology and sociology. Some colleges require both statistics and research methods courses (e.g., Penn State Harrisburg, 2007a; Penn State Schuylkill, 2007; 2007a); however, there are some that require a combined statistics and research methods for an applied behavior science degree (e.g., Penn State Harrisburg, 2007b), or only research methods (e.g., Lebanon Valley College, 2007), albeit some colleges recommend that students also take statistics (University of Pittsburgh, 2007) or have an unspecified quantification course requirement (e.g., Penn State Harrisburg, 2007).

Second, to examine students' proficiency in statistics, it is necessary to encompass elements from both statistics and research methods courses. Hence, the term statistical literacy becomes an important concept, and can be succinctly defined as "the study of statistics used in everyday life. Statistical literacy helps citizens in a democracy read and interpret numbers in the news to make intelligent decisions" (Statistical Literacy, 2007, n.p.). Statistical literacy includes not only the understanding of statistical concepts and the mathematics behind it, but also the research methodology that provides legitimacy to a study. A poorly designed study can lead to false results about any phenomena, while purporting to be empirically valid, leaving the reader with false information. To be accurate in measuring students' statistical literacy, it is crucial to know which course students have completed, in order to understand which type of content students have mastered to identify the gaps in their literacy and to provide recommendations to close the gaps. Statistical literacy is not a singly defined term, but a multifaceted concept.

Overview of the Conceptual Framework

Statistical literacy cannot be measured using a single instrument due to the various elements that make up this concept. In order to encapsulate and examine these elements empirically, Gal's Model of Statistical Literacy (2004) will be the conceptual framework for this study, because this model has the ability to incorporate various elements for measurement and analyses. This will inform my study, because this model will allow me to examine the seven elements that comprise statistical literacy, which can be measured together or separately. In essence, it will allow me to examine gaps in students' learning due to the type of courses they have completed.

Importantly, statistical literacy is a term that is comprised of various statistical phraseologies—in which each element comprises a necessary part of this multifaceted term. These are numeracy (Brown, Askew, Baker, Denvir, & Millett, 1998; Steen, 1990; Watson, 2002), adult numeracy (Gal, 2002; NCES, 2006), innumeracy (Cerrito, 1999; Paulos, 1989), quantitative literacy (Manaster, 2001; Rosen, Weil & Van Zastrow, 2003; Steen, 1990; Steen, 2001), mathematical literacy (Rosen, et al.), mathematical illiteracy (Paulos), statistical literacy (Moore, 1997; Schield, 1999; Rumsey, 2002a; Wallman, 1993), and statistical reasoning (Garfield & Chance, 2000; Garfield, 2003). Some terms define it within arithmetic or mathematics, or mathematics and statistics, while others combine arithmetic, mathematics and statistics; some incorporate social implications of being statistically literate and some the importance of research methodology. Accordingly, to capture the essence of this multifaceted term, a guiding framework is essential to examine statistical literacy and is provided by Gal's (2004) Model of Statistical Literacy, whereby seven elements comprise the term relative to the various statistical phraseologies that are found in various empirical literatures. As he reminds us, the "elements in the proposed model should not be viewed as fixed and separate entities, but as a context-dependent dynamic set of knowledge and dispositions that together enable statistically literate behavior" (p. 51).

Succinctly, these elements are divided into two categories, knowledge and dispositional elements. Five knowledge elements are literacy skills, statistical knowledge, mathematical knowledge, context knowledge and critical questions. These are the elements, according to Gal (2004), that examine how individuals "interpret and critically evaluate statistical information and data-related arguments" (p. 49), which are encountered in different mediums in everyday life. Dispositional elements in Gal's Model of Statistical Literacy are beliefs and attitudes, and critical stance, and are related to individuals' ability to "discuss or communicate their reactions to such statistical information, such as their understanding of the meaning of the information, their opinions about the implications of this information, or their concern regarding the acceptability of given conclusions" (Gal, p. 49).

Even though Gal's Model of Statistical Literacy purports to examine statistical literacy, there are a few assumptions the model makes, which should be taken into consideration. These are: (a) the model purports that to examine statistical literacy a holistic approach should be taken; (b) it is not necessary for the model's elements to be completely mutually exclusive; and (c) because the model was developed from various statistical phraseologies within the mathematics realm, it can measure statistical literacy within the various behavioral and social sciences disciplines.

Problem Statement

As demonstrated both historically and currently, an understanding of statistics is an important part of adult life and hence, the educational process in higher education. Spurred by the labor market and globalization, the NCTM began the movement to reform mathematics education and it expanded to include statistics, because it is recognized as a crucial part of higher education. Its importance is reflected in the data that is used to sell everything from medicines to automobiles, to promote political agendas or to influence our opinions on important issues. This new emphasis on the teaching of statistics has led to research in the last decade that informs us about current teaching practices in introductory statistics classes. Teaching methods in many statistics classes now include a variety of active learning techniques that include computer simulations, Web-based statistics courses, computer data analyses programs, an integration of real-world problems into the course curricula, and the use of real data sets. Evaluations of these courses used students' opinions about the course, final grades and/or single exam scores.

Although these types of evaluations can provide us with important information about the success of these new teaching methodologies, they fail to inform us if students have become statistically literate. By using active learning techniques in the statistics classrooms, students should be able to interpret data in a proficient manner; however, no research has examined if they are statistically literate after completing their required statistics and/or research methods courses. And importantly, some research failed to examine if courses were statistics or research methods class with prior statistics or research methods class with no prior statistics. Further, there is no research that examines gender or first-generation adult students and statistical literacy. This leads to unanswered questions: Are students statistically literate after completing their discipline's requirement of statistics and/or research methods course? Will they be able to become an informed citizen in a society that is constantly influenced by data? Are there gaps in students' statistical literacy, and if so, where are the gaps in students' statistical literacy? And more importantly, what recommendations can be made to help students to be proficient in statistical literacy, after completing this study?

Most noteworthy, past research has left a void in the research that embraces the learning of statistics; this is different than course evaluations, final grades or single exams scores, as it is applying the knowledge learned in a statistics and/or research method course to real-world situations. And because statistical literacy embraces 7 elements, it should be measured by an instrument that can incorporate these elements. No research has measured statistical literacy using Gal's (2004) Model of Statistical Literacy.

Purpose Statement

With knowing how important statistical literacy is our data-driven society, the purpose of this research is to measure statistical literacy in adult learners before and after they have completed a statistics class, a research methods class without prior statistics, and a research methods class with prior statistics; and further to examine if there are learning differences in statistical literacy between gender, and first-generation adult student status.

Hypotheses

1. Adult learners who have completed a research methods class with prior statistics will be more proficient in their knowledge of statistics than learners who have only completed a statistics or research methods class with no prior statistics.

    1a. Adult learners who are not first-generation learners will be more proficient in their knowledge of statistics than learners who are first-generation adult learners.

    1b. Adult learners who are male will be more proficient in their knowledge of statistics than learners who are female.

2. Adult learners who have completed a research methods class with prior statistics will have more of a positive disposition toward statistics than learners who have only completed a statistics or research methods class with no prior statistics.

    2a. Adult learners who are not first-generation learners will have more of a positive disposition toward statistics than learners who are first-generation learners.

    2b. Adult learners who are male will have more of a positive disposition toward statistics than learners who are female.

Significance of the Study

"In God we trust, all others bring data" (Cerrito, 1999), was echoed by W. Edwards Demings, a statistician, college professor, author, and lecturer many years ago; however, it is more relevant in today's world. In our contemporary times, numbers constantly permeate society and are constantly used in political dialogue (Cerrito), in the workforce (Pugalee, 1999; Rumsey, 2002), in higher education for promotion and tenure, and embrace health issues from our baby immunizations to treatments for medical afflictions, just to mention a few. Statistical information, which is produced by an increasing number of public agencies, non-profit organizations and commercial companies, “has a special role in the information fabric of modern societies, as it enables people to be aware and capable of reacting to phenomena of social, political, economic, and personal importance” (Murray & Gal, 2002, p. 1). And importantly, one duty of a responsible government is to provide statistical information about the welfare of its citizens and should be studied by all who aspire to improve the state of the nation (Schaeffer, 2001). This is reflected in the growth of the collection and dissemination of data contained in the U. S Census Bureau.

Much statistical information is disseminated in various media contexts, which include television, newspapers, or Internet sites, through written or oral text, numbers, symbols, and graphical or tabular displays (Murray & Gal, 2002). Hence, the information age has made the world quantitative (Schaeffer, 2001), but unfortunately, in many instances numbers are misleading and taken for granted by a society that is mostly innumerate (Cerrito, 1999; Schaeffer), which can be profoundly disabling in many spheres of human endeavors. These include “home, private life, work, or public and professional pursuits” (Orrill, 2001, p. xvi). This inability to understand data continues to occur even though “conceptions of statistics and probability have steadily advanced within scientific and mathematical communities, adults in mainstream American society cannot think probabilistically or statistically about important societal issues” (Derry, Levin, & Schauble, 1995, p. 51).

Currently, many adults who may be statistically illiterate are presently enrolled as adult learners in degree-granting institutions of higher learning. In fact, data from the National Center for Educational Statistics (NCES) shows there are approximately 5.4 million adult students enrolled in college out of a total population of 17 million, approximately one-third of the current college population. Because one of the tasks of higher education is to develop an informed citizenry, and in light of the emphasis on data available in mainstream society, an important change in higher education has been to incorporate statistics courses into a variety of disciplines (Schaeffer, 2001).

However, to incorporate statistics courses into course curriculums does not inform us if adult learners have achieved statistical literacy. No studies have examined whether adult learners have achieved statistical literacy during the course of their college experience. And importantly, adult learners who have been out of school for a number of years may have insufficient math skills and/or anxiety, which could impede their learning of statistics. Therefore, not only will this study examine adult students' learning gains in statistics, but it can also examine where there are deficiencies in their learning, because Gal's Model of Statistical Literacy breaks down this multifaceted concept into 7 elements, which includes a scale that can examine, for example, their attitudes and beliefs about statistical data. With a focus on statistical literacy as the outcome of learning in statistics and research methods courses, learning gaps can be identified and future research can address these gaps, enabling adult learners to become critical consumers of data reports that surround their daily lives and to become responsible citizens in a civic society.

As previously discussed, this research brings a much broader perspective to the field of adult education through adult literacy. Adult literacy is measured on three literacy dimensions of adult learning, prose, document and quantitative. Prose literacy refers to skills needed to search, comprehend and use continuous texts—for example, interpreting editorials, news stories, brochures and instructional materials. Document literacy is the knowledge and skills required to search, comprehend and use non-continuous texts in various forms such as job applications, transportation schedules, or drug and food labels. And quantitative literacy embraces the skills required to perform quantitative tasks requiring an adult to identify and perform computations by using numbers embedded in text. Examples of quantitative literacy would include demonstrating the ability to balance a checkbook, compute a tip in a restaurant, or complete an order form (National Center for Education Statistics [NCES] 2006).

Missing from the discourse on adult literacy is statistical literacy, and like prose literacy it embraces important aspects of empowering individuals to become responsible citizens in a civic society. Being illiterate in either prose or statistics alienates individuals to "the culture of silence, the masses are mute; that is they are prohibited from creatively taking part in the transformations of their society…" (Freire, 1970; 1998, p. 486). And most importantly as stated by Freire:

> Illiterates know that they are concrete men. They know that they do things. What they do not know in the culture of silence—in which they are ambiguous, dual beings—is that men's actions as such are transforming, creative and re-creative. Overcome by the myths…of their own natural inferiority, 'they do not know that their action upon the world is transforming.' Prevented from having a 'structural perception of the facts involving them, they do not know that they cannot have a

> voice,' that is, they cannot exercise the right to participate consciously in the socio-historical transformation of their society…(p. 486)

Succinctly, there are four important reasons to examine statistical literacy. First, it is significant to examine statistical literacy because our society is, and has been, immersed in data that rules many aspects of our lives. It is important for adults to have the ability to interpret and critically question statistical results in order to be an informed citizen in a civil society. Second, to the field of adult education it is significant because it adds to the knowledge base regarding adults' preparation to become an informed citizen in a civic society. Third, by examining adult students' level of statistical literacy before and after completing their required courses of statistics and/or research methods, gaps in their statistical literacy can be identified and addressed within the teaching of adults in higher education. And fourth, because statistical literacy is now being addressed in higher education for adults, it is likely that it will become part of the broader adult literacy scale that currently fails to address this important need in society.

As a first-generation college student whose entire college life was experienced as an adult learner, I soon began to understand the role statistics played in constructing social policies, for example auto insurance and health insurance rates, and the data that allows pharmaceutical companies to sell new drugs in the public sphere. Somewhere between my experiences as an adult and my learning I realized that some data analyses were fictitious—they were manipulated for monetary or political gain, but no one questioned the research methods or the data analyses, except me. This silence, the inability to critically examine or question the research, fueled my interest in statistics, leading me to examine why students had statistical anxiety when I was a professional tutor for statistics students. This became the topic of my master's research. After completing my master's degree, and because I enjoyed tutoring, I became an instructor in higher education. And now as I work toward my Ed.D in adult education, I realize that my purpose in teaching is to promote social change, though empowering students to understand that the data reported in any media needs to be critically questioned and reflected upon before it is accepted as a fact. Succinctly, this research is significant to me because it will enable me to examine statistical literacy in adult learners and address the deficits in their learning, in order to promote social change.

Précis of Research Method

In contemplating which type of research design to use, it is important to match the design to the purpose of the study; and hence, to examine statistical literacy, a quantitative research design was chosen. Quantitative research is based on the philosophy of positivism, and later, logical positivism. Positivism is a philosophical doctrine "that recognizes only natural phenomena or facts that are objectively observable…and not debatable" (Schultz & Schultz, 2000, pp. 39-40) can constitute knowledge. Later, this philosophy was furthered through logical positivism, which focused on the verifiability of meaning and logical analysis. Verification to logical positivists means that "a statement is meaningless if verification is not possible or the criteria for verification are not clear" (Ho Yu, 2006, p. 28), and logical analysis adds an emphasis of language, as complex phenomena could be expressed in terms of mathematics, and mathematics could be further reduced to logic (Russell, 1963). And further, positivism purports that "objective information about human behavior cannot be obtained from subjective meanings, beliefs and explanations since human beings are capable of placing any number of interpretations upon their own behavior" (James, 2005, p. 2). Therefore, to examine human behavior it is necessary to measure it with some type of instrument. And likewise, to measure adult students' learning of statistical concepts an instrument will be used, Gal's (2004) Model of Statistical Literacy, which incorporates the 7 elements of statistical literacy.

Accordingly, a quasi-experimental, pre-post test research design was chosen, as a true experimental design would not be feasible. But similar to an experimental design, a quasi-experimental design allows an examination to determine if treatment groups are equal before treatment, "then pre-test selection differences could not be a cause of post-test differences" (Shaddish, et al., 2002, p. 249), and temporal precedence remains in check, as cause precedes effect. The use of statistical analysis can check to see whether cause co-varies with effect, and the remaining task to examine causality is to eliminate any alternative explanations for the results (Shaddish, et al.).

The method for choosing research participants for this study will entail non-probability sampling—purposive sampling. Both groups, the treatments and control, must be adult college students, with the treatment groups currently or recently enrolled in a statistics or a research methods class. The data in this study will be collected through surveys that measure statistical literacy. Both the pre- and post-test instrument will consist of seven sections and can be described as consisting of : (a) 6 questions regarding statistical knowledge, (b) 6 questions regarding statistical reasoning, (c) 6 questions regarding statistical thinking, and (d) 4 questions regarding an individual's belief about statistics. The fifth and sixth sections of the instrument use a 7-point Likert scale (strongly disagree to strongly agree) to measure an individual's (a) critical stance, and

(b) attitudes toward statistics. The seventh section examines the worry questions about statistics and research methods by using a real research example, which will allow an individual to write an open-ended response to it. In addition, there are 8 demographic questions.

Adult students who were enrolled in a statistics class, or a research method class with prior a prior statistics course, or a research methods course with no prior statistics, were asked if they would like to participate in a study examining statistical literacy, and simply filled out the survey online or filled out a take home version. Statistical analyses primarily consisted of (a) MANOVAs, ANOVAs, independent t-tests, and Fisher's (LSD) post-hoc tests, used to examine post-test scores among groups; (b) mixed between-within subjects ANOVAs and paired t-tests were used to examine pre- to post-test scores; and (c) an ANCOVA was used to examine for pre-test effects on post-test scores.

Assumptions of the Research

The following are the assumptions the researcher has in initiating this study.

1. It is assumed that all the statistics and/or research methods courses basically covered the same statistical and methodological concepts.
2. It is assumed that adult students who participated in this study answered the survey instruments to the best of their ability.
3. It is assumed that students completed the instrument without the help of others or referred to a textbook to help them answer the questions.
4. It is assumed that students were interested in achieving statistical literacy instead of just completing an introductory statistics course to complete their college degree.
5. It is assumed that adult learners had adequate computer skills to complete the survey

online.

Limitations of the Research

1. This research used a quasi-experimental design; therefore, it was limited in its ability to generalize the results of statistical literacy to all adults who are enrolled in college.
2. Adult students who participated in this research were from small rural colleges on the east coast of the United States. This limits the generalizability to all adult college students.

Organization of the Writing

This dissertation on statistical literacy is divided into five chapters. Chapter one explains the background of the problem in its relation to the teaching of statistics, an overview of the conceptual framework, a problem and purpose statement, the hypotheses, the significance of the research, précis of the research design, assumptions and limitations of the research and a section that defines the terms used in this study.

Chapter two provides a literature review on research on the teaching of statistics, the details of the conceptual framework, along with various statistical phraseologies used to define statistical literary, which led to the Model of Statistical Literacy. This is followed by a brief history of the teaching of mathematics and statistics.

Chapter three provides the research design and methodology used to examine statistical literacy, including the participants, procedure, and instrumentation. It also provides a glimpse at the underlying ideas of quantitative research. Chapter four provides the details of the data analyses, and chapter five provides an in-depth discussion on the results and recommendations for future research.

Definition of Terms

Active Learning: a variety of teaching activities that is promoted by learner-centered activities (McKeachie, 2002).

Adult Learner (formerly referred to as a non-traditional student) consists of a variety of definitions, including a person who has assumed major life responsibilities and commitments; one who is no longer dependent upon parents, can operate independently in society (Mancuso, 2001); one who has assumed care of another, for example a child or elderly relative, is employed or has experienced a delay after high school in enrolling in college (Belcastro & Purslow, 2006).

Computer Simulation Programs: computer-based simulations that demonstrate specific statistical concepts, for example population sampling distributions, the central limit theorem or statistical power, by allowing students to analyze data, create charts and graphs, and receive immediate feedback about the choices they have made (Aberson, et al., 2000, 2002, 2003; Morris et al., 2002; Morris, 2001).

Data Analyses Programs**:** Computer programs that are used in business, industry and education that can analyze large data sets, and create graphs and charts.

First-Generation Student: Students whose parents never attended college (Lee, Sax, Kim, & Hagedom, 2004).

Gal's Model of Statistical Literacy: A model of statistical literacy that incorporates seven elements; 5 knowledge components, literacy skills, statistical knowledge, mathematical knowledge, context knowledge and critical questions; and 2 dispositional elements, beliefs and attitudes, and critical stance.

Non-probability Sampling: a type of sampling in "which not everyone has an equal chance of being selected from the population" (Neuman, 2000, p. 196).

Purposive Sampling: a method non-probability sampling that entails getting all possible cases that fit particular criteria.

Quasi-Experimental Design: A research methodology in which one or more independent variables are manipulated to observe their effects on one or more dependent variables; and an experiment in which units are not assigned to conditions randomly (Shaddish, et al., 2002).

Research Method Courses: An introduction to methods of psychological research, with special attention to hypothesis formation and testing, threats to validity, and data presentation (Penn State University, 2007).

Elementary Statistics Courses**:** Courses that examine frequency distributions and graphs; measures of central tendency and variability; normal probability curve; elementary sampling and reliability; correlations; simple regression equations (Penn State University, 2007).

Statistical Literacy: statistical literacy is the study of statistics used in everyday life. Statistical literacy "helps citizens in a democracy read and interpret numbers in the news to make intelligent decisions" (Statistical Literacy, 2007, n.p.).

T Test: A statistical technique to compare groups.

Traditional Student**:** students who have entered college immediately following high school graduation.

Web-based Courses: These courses are often referred to as *hybrid* programs because they use a combination of a Web-based statistics program and have students have face-to-face, albeit less contact with the instructor than students who attend a traditionally taught statistics classroom (Bushway & Flowers, 2002; Johnson & Dasgupta, 2005; Symanzik & Vukasinovic, 2006; Utts, et al., 2003; Ward, 2004).

CHAPTER 2

LITERATURE REVIEW

Introduction

The purpose of this research is to measure statistical literacy in adult learners before and after they have completed a statistics course, or a research methods class with no prior statistics, or a research methods class with prior statistics. This chapter contains three main sections. The first section provides a discussion on constructivism and a review of the empirical literature on the teaching of statistics—on how statistics is currently being taught and assessed in higher education today. The second contains the conceptual framework and statistical phraseologies, which have been used to define statistical literacy. These phraseologies, through their definitions, are the underpinnings of the model of statistical literacy, which is also discussed in this section. The third section contains a brief history of the teaching of mathematics and statistics, which elucidates how, historically, mathematics and statistics have been taught in the United States. This is essential to understand, because the discipline of statistics emerged from mathematics, albeit statistics is not mathematics. It continues by elucidating why the learning of statistics is crucial to adult students in higher education today.

Constructivism in the Statistics Classroom

This section details constructivism in-depth, by defining it broadly, followed by a discussion that details its early philosophical roots. Next, the discussion will focus on three specific types of constructivism, cognitive, sociocultural and radical, and will proceed with a critique on some of the tenets that can be ascertained from constructivism. The final part of this paper will then explain how constructivism has informed the teaching of statistics.

Constructivism has multi-meanings—it is considered a philosophy, and it also refers to a set of views about how individuals learn. Phillips (2000) offers a basic distinction between meanings; as a philosophy, it embraces a thesis "about the disciplines or bodies of knowledge that have been built up during the course of human history" (p. 6). These bodies of knowledge are made from human constructs, formed from knowledge taken from a variety of fields, such as politics, ideologies, values, religious beliefs, the exertion of power, and self-interest. Further, this thesis denies that these disciplines are "an objective reflection of the objective world" (Phillips, p. 6). Therefore, the origin of human knowledge, and its standing as knowledge, is to be examined by using sociological tools, instead of epistemological ones. This area of constructivism is known as social constructivism or sometimes, social constructionism (Philips).

As Tobin and Tippins (1993) explain, constructivism is a form of realism—that is, an existence of a reality is acknowledged from the beginning and individuals come to know it in a personal and subjective way. For example, for constructivists, gravity does exist, because of our experiences with it. Therefore, knowledge is individual and social, and through negotiation with others in the social system, an agreement is met that the concept of gravity has numerous verifiable properties; hence, the model exists through the processes of negotiations and consensus building. And as our experiences change, the model is updated, as constructions are constrained by experience, as objects fall downward, not upward. And because there is no objective account of what gravity really is, "we cannot tell whether the model for gravity gets closer and closer to an absolute reality…we can only know gravity in a personally, socially mediated way" (Tobin &

Tippins, pp. 3-4). Succinctly, constructivism is a theory about "knowledge and learning; it describes both what 'knowing is' and how one 'comes to know'" (Fosnot, 1996, p .ix).

Likewise, Kivinen and Ristela (2003) purport that constructivism is about how "an individual learner constructs knowledge in his or her mind; whereas for others the main concern is with the construction of human knowledge in general or with the sociopolitical construction of knowledge" (p. 363). Similarly, with individuals in mind, Merriam, Caffarella and Baumgartner (2007) define constructivism as a learning theory in which "learning is a process of constructing meaning; it is how people make sense of their experience" (p. 291). Learners construct their own sets of meanings or understandings, "knowledge is not a mere copy of the external world; nor is knowledge acquired by passive absorption or by simple transference from one person to another…Knowledge is made, not acquired" (Phillips, 2000, p. 7). In this sense, knowledge for individuals may be different as they may not construct the same understandings, even if they use the same language to express what they have learned, as their deep understandings may be different (Phillips).

Accordingly, to pin down an exact meaning of what constructivism is, is a great challenge; because of its popularity in the education literature, its meaning has different purposes and continually changes. So in order to explain what it is, it is necessary to examine some of the philosophical precursors in order to give context to present-day meanings.

The principle originator of constructivist thought was the philosopher Immanuel Kant, who sought to resolve competing claims of knowledge, rationalism and empiricism, under vigorous debate during his time. Rationalists (i.e., those who believe that reason is the only valid knowledge of reality), like Rene Descartes, viewed knowledge "as derived from intuitively clear and indubitable ideas," and the empiricists (i.e., those who believe the only source of knowledge is experience), such as John Locke, who "viewed knowledge as synthesized from elementary sensory experience" (Bredo, 2000, p. 128). Kant believed that both "mental organization and sensory input are involved in knowing" (Bredo, p. 129). Basic categories, such as spatial, temporal, and causal relations that give form to the flux of experiences, are provided by the mind. Therefore, sensory experience provides concrete particulars that give specific content to the mind's categories. In this perspective "we can never know the 'things in themselves' that cause perceptual experiences, because even the phenomena of experience are shaped by mental relationships. The most basic experiences are constructs, since they have been given form by mental categories and relationships" (Bredo, p. 129). Further, there is no escape from our a priori assumptions (i. e., knowledge independent of experience), but many of these implicit categories are universal, since we live in a commonly constructed world (Bredo).

Other philosophers, such as Georg Hegel, also influenced the constructivist perspective as he attempted to synthesize opposing forms of thought by viewing them as phases in a process of sociocultural evolution. He believed that different forms of consciousness developed in different ages and this resulted in different qualitative character of subjects, objects or methods of representation; hence, collective thought and reality evolved together. To Hegel, history demonstrated this, as the mind evolved toward increased self-awareness and freedom; hence, his approach toward constructivism was an evolutionary approach to thinking about relationships between mind and nature, different than Kant's static approach. Hegel's work influenced Karl Marx's materialistic interpretation of social evolution, which then influenced Lev Vygotsky's thinking on the social formation of mind. For Hegel and Marx, individual thinkers did play a role on the process of social revolution, but their emphasis was on a collection evolution focused on an ultimate state of permanent harmony and consensus (Bredo, 2000).

An evolutionary approach was evident in Charles Darwin's work, albeit Darwin placed an emphasis on individual uniqueness and within-group variation, and rejected the notion that evolution had an ultimate goal. In fact, he challenged "conventional notions of essential sameness and rationality in nature while introducing individual diversity and contingency" (Bredo, 2000, pp. 129-130). From here, other thinkers like William James and John Dewey looked for a synthesis between neo-Hegelian and Darwinian views. With Hegel, they believed in the "mind as a social product and as a factor within nature and social life, helping to alter the course of its evolution" and with Darwin, "they saw the mind as a practically adaptive function, rather than an aspect of Absolute Spirit, saw every individual as unique, and conceived of no fixed end to national or social evolution" (Bredo, p. 130). Accordingly, the mind is not a spectator, but is a partial and limited participant in the course of social and natural evolution (Bredo). And lastly, Rudolf Carnap contributed to constructivist thought during a time when formal logic and physics became the models of knowledge adopted by philosophers, rather than biology, and viewed

> formal relationship to logical systems as defining distinctive worlds. A set of logical primitives creates the basis for a world, not unlike the way a computer programming language creates a micro-world. Systems based on different primitives then form different 'worlds'. (Bredo, 2000, p. 130)

Succinctly, the principal claim in the many views of the evolution of constructivism is that knowledge is made, rather than found, as “the objects and properties that we experience and know are themselves in some manner products of human (i.e., mental or physical) activity” (Bredo, 2000, p. 131). This is not surprising in contemporary times, as our lives are constantly affected by new discoveries as knowledge is controversial and subject to change.

Thus, constructivism “remains a theory about learning, not a description of teaching” (Fosnot, 1996, p. 29), and a variety of types of constructivism have evolved from its basic tenets, for example, individual idealist and individual realist (Bredo, 2000), and philosophical constructivism (Matthews, 2000). But for brevity and for the purpose of this discussion on constructivism, three types of constructivism related to education will be discussed next. These are Piagetian or cognitive, sociocultural, and radical.

Jean Piaget’s name is commonly associated with constructivism. His early background in biology later influenced his learning in psychology (Phillips, 1997; Wadsworth, 1996). He was also influenced by the Kantian philosophy, aforementioned that both “mental organization and sensory input are involved in knowing” (Bredo, p. 129), as he believed that “biological acts are acts of adaptation to the physical environment and help organize the environment” (Wadsworth, p. 13). Further, he believed that mental activity is subject to the same laws of development as biological activity, in the sense that concepts of biological development are useful and valid in viewing intellectual development (Phillips, 1997; Wadsworth).

Hence, the mind and body do not operate independently of one another, and the intellectual organization and adaptation, according to Piaget, contain four basic cognitive concepts—schema, assimilation, accommodation and equilibration. As explained by Piaget, schemas are “cognitive mental structures by which individuals intellectually adapt to and organize their environment” (Wadsworth, 1996, p. 14). To illustrate, in biology, the stomach is a biological structure that animals use to adapt to their environment, because as the environment changes, their diet changes. Similarly, schemas are psychological structures that adapt and change during mental development—for example, if the only animal a child saw was a cow, he might confuse another animal, a dog, as a cow. He has limited schemata, which will grow and change. But because we cannot observe them, Piaget called these called hypothetical constructs (Wadsworth).

Second, assimilation is “the cognitive process by which a person integrates new perceptual motor or conceptual matter into existing schemata or patterns of behavior” (Wadsworth, 1996, p. 17). This would occur if a child saw a cow and stated that it was a dog, because the cow has the same characteristics of a dog, it for example, also having four legs. Thus, assimilation is the cognitive process of putting new stimuli into existing schemata, and allows for the schema to grow (Wadsworth).

Third, accommodation occurs when a new stimulus does not assimilate into an existing schema. This results in the creation of a new schema, or modifying an existing schema so that a stimulus fits, with both of these processes resulting in the development of cognitive structures. And fourth, the final concept is equilibrium. Equilibrium is the balance between assimilation and accommodation, and allows an external experience to be incorporated into an individual’s existing schema. This is Piaget’s cognitive constructivism, where individuals construct their own meaning of the world from their experiences.

Further, Piaget believed that "cognitive growth and development proceed in this way at all levels of development. From birth through adulthood, "knowledge is constructed…the schemata of adulthood being built from the schemata of childhood" (Wadsworth, 1996, p. 20). And he believed that an individual interacting with real structures in the real world will come to "construct his or her internal structures that, while not copies of those in the world, will be logically isomorphic with them" (Phillips, 1997, p. 26), as individuals will be exposed to the same stimuli will construct structures having the same logical features (Phillips).

While Piaget's cognitive constructivism focused how knowledge is formed inside the mind of an individual, Vygotsky was concerned on how social and cultural factors influenced intellectual development (Wadsworth, 1996). His most popular constructivist idea was demonstrated by his theory of the zone of proximal development, in which importance is placed on social interactions with more knowledgeable others (Fosnot, 1996), and through these interactions, students can learn things they could not learn on their own (Wadsworth). Vygotsky also purported "basic forms of minds to be socially constructed and constructing" and "viewed symbolically mediated thought as a social process, like a dialogue that is 'internalized' though participation in social action" (Bredo, 2000, p. 133), as higher mental functions originate in social activity (Hausfather, 1996). Further, the type of knowledge one learns from social interaction is dependent on the sociohistorical nature of society. For example, as individuals become educated and experience life, their ways of thinking tend to become more formal-logical. In this sense, modern life creates modern minds (Bredo).

But to what purpose does modern life create modern minds? This is the subject focused on by some sociocultural constructivists, as sometimes “social facts and identities are socially constructed while made to appear natural” (Bredo, 2000, p. 133). Some sociocultural constructivists purport how social and political elites define knowledge in self-serving ways. For example, in the schools, rather than promoting competence in students, educational success is promoted by schools as getting ahead of others (Bredo).

The final type of constructivism that will be discussed is radical constructivism, which redefines the definition of knowledge as an adaptive function. Hence, the results of our cognitive efforts, instead of the “traditional goal of furnishing an objective representative of the world as it might exist apart from us and our experience, has the purpose of helping us to cope in the world of our experience” (von Glasersfeld, 1992, p. xiv). Von Glasersfeld, like other radical constructivists, prefers to use the verb knowing instead of the noun knowledge. Applying this to an individual’s cognition, to know is “not to possess true representations of reality, but to possess ways and means of acting and thinking that allow one to attain the goals one happens to have chosen” (McCarthy & Schwandt, 2000, p. 45). Therefore, we only have access to the world that we ourselves create out of our own experience, “never to a world of reality, and never conclude that our knowledge is in fact, knowledge of the real world” (Goldin, 1990, p. 35).

Because individuals have different experiential worlds and see things differently, some would argue that we could not agree on anything; therefore, we cannot communicate. But as von Glasersfeld (1992) points out, just because we can communicate and can agree on certain things, does not mean that we experience object reality (i.e., a truth about the world), but consensual domains. Hence, "all our experience is subjective, but we manage in communication with those around us, to render our subjective meanings intersubjective and to create consensual domains" (Maturana as cited Schoenfeld, 1992, p. 291). In other words, these consensual domains are constructed out of our "in-context experience of each others' speeches and actions" (Goldin, 1990, p. 35). To communicate, individuals do not necessarily need to have identically shared meaning of things; only compatible meanings are necessary. It is these shared beliefs that become important in communication between the instructor and the learner (von Glasersfeld). It is through these consensual domains that constructivist ideas, which are subjective, can be examined through the tenets of logical positivism, which posits that knowledge occurs from observable phenomena. Logical positivism is discussed in chapter 3.

Albeit these three different views of constructivism put forth different ideas, commonalities can be seen among these various epistemologies. A commonality between cognitive and sociocultural constructivist theorists is that both highlight the critical role that activity plays in learning and development. The difference is that sociocultural theorists typically "link activity to participation in culturally organized practices, whereas cognitive theorists give priority to individual students' sensory-motor and conceptual activity" (Cobb, 1996, p. 36). Moreover, sociocultural theorists believe that in the learning process, cognitive processes are subsumed by social and cultural processes. In effect, the primary directive for sociocultural constructivists is to understand how participation in social interactions and culturally organized activities influence psychological development (Fosnot, 1996). And similarly, radical constructivist von Glasersfeld, who defines learning as self-organization, acknowledges that this "constructive activity occurs as the cognizing individual interacts with other members of a community" (Cobb, p. 37), but further elaborates to state "knowledge refers to conceptual structures that epistemic agents, given the range of present experience within their tradition of through and language, consider viable" (von Glasersfeld, 1992, p. 381). Viable means a concept that works "to the extent that it does what we it need do—to make sense of our perceptions or data, or to make an accurate prediction, to solve a problem" (Simon, 1995, p. 115). Accordingly, each one is a little different than the other, while at the same time each one complements and builds, in a sense, its epistemology.

From these three views of constructivism, cognitive, sociocultural and radical, and from the overview that describes constructivism, some important tenets can be summarized. Many of these are not mutually exclusive of each other's epistemology. Hence, the following six tenets are summarized, and will be critiqued:

> (a) Learning is an active process (cognitive, sociocultural, radical); (b) knowledge is constructed rather than innate or passively absorbed (cognitive); (c) knowledge is invented and not discovered (sociocultural, radical); (d) all knowledge is idiosyncratic and personal (cognitive, radical); (e) all knowledge is socially constructed (sociocultural); and (f) learning is essentially a process of making sense. (Fox, 2001, pp. 24-30)

The first claim, "learning is an active process" (Fox, 2001, p. 24), is criticized because in general, humans do not always gain knowledge by acting upon their world, but by the world acting upon them; for example, light adaptation of the eye to changing levels of brightness. However, Piaget would insist that children are not merely recipients of stimulation, but frequently investigate their world while getting to know it (Fox). Second, “knowledge is constructed rather than innate or passively absorbed” (Fox, p. 25), as our ability to perceive, to learn, to speak and to reason are based on innate capacities; therefore, learning is not an active process. Third, “knowledge is invented and not discovered” (Fox, p. 26), as we will always view the world from some sociocultural or historical viewpoint. Hence, we cannot come to know things in themselves.

Fourth, “all knowledge is idiosyncratic and personal” (Fox, 2001, p. 29). This can lean toward solipsism (i.e., the external world and other minds are not known and might not exist) and contradicts the possibility of sharing and communicating knowledge between people, which then contradicts the idea of social constructivism, the idea in which knowledge is constructed socially. But this claim implies that, in the real world when instructors present the same lesson to a variety of students, their experiences may result in different learning by each student. Fifth, “all knowledge is socially constructed” (Fox, p. 29); all learning then is leaning toward an implausible extreme, in which social factors would determine all learning and all conscious thought. It ignores the fact that learning sometimes depends on independent practice and problem solving. Hence, the fourth and fifth claims contradict each other. Further, as Fox states, “knowledge may be seen as essentially defined in terms of the subjective mental states of each knower,” and also, as Popper states (as cited by Fox), “knowledge may be defined in the terms of publicly communicated and constructed bodies of knowledge that make up academic disciplines, databases, books, theories, works of art and other cultural products” (p. 30).

And finally, “learning is essentially a process of making sense” (Fox, 2001, p. 30), as constructivism purports that the aspect of learning is about understanding and, in doing so, it takes us beyond the conception of rote learning. But sometimes a conscious effort to do so is missing, especially if the context is too removed from students' horizon of expectations, resulting in their feelings of boredom or confusion, or both.

Despite the critiques and the different types of constructivism that exist, a pedagogy of learning can be extrapolated from the epistemology of constructivism. Simply, Bredo (2000) informs us, a constructivist pedagogy has "a concern with students having an active role in learning and their being allowed to redefine or discover new meanings for the objects which they interact" (p. 132). Or as explained by Howe and Berv (2000),

> learning takes as its starting point the knowledge, attitudes and interests student bring to the learning…learning results from the interaction between these characteristics and experiences in such a way that learners construct their own understanding, from the inside, as it were. (p. 31)

More importantly, and immediately germane to this discussion, constructivist perspectives on learning have contributed to the teaching of statistics in the classroom, as the importance of teaching statistics has been part of the educational reform within mathematics education, which was established by the NCTM (Garfield & Chance, 2000). As part of the reform in teaching, a constructivist perspective of pedagogy has replaced the traditional behaviorist teaching method, where the teacher was the sole information-giver to passive students. In this setting, teachers lecture and students listen as teachers transfer their thoughts and meanings to the passive students. These classes relied heavily on textbooks that promote the idea that there is a fixed world of knowledge that the student must come to know (Hanley, 1994). Differently, in the constructivist classroom, the instructor's role is to "facilitate and negotiate meaning-making with the learner" (Merriam et al., 2007, p. 295), with the purpose of learning being to construct knowledge. Sometimes these are manifested in adult learning through experimental learning, transformative learning, or situated learning (Merriam, et al.). Further, Hanley explains:

> The role of the teacher is to organize information around conceptual clusters of problems, questions and discrepant situations in order to engage the student's interest. Teachers assist the students in developing new insights and connecting them with their previous learning…activities are student centered and students are encourage to ask their own questions, carry out their own experiments, make their own analogies and come to their own conclusions. (p. 2)

For the learners, their role is to become actively engaged in their own learning, through willing participation in hands-on activities, and to share thoughts in dialogues in order to pursue topics in depth (Mvududu, 2005).

Most important is how this variety of constructivist thought, cognitive, sociocultural and radical, along with the underlying tenets of constructivism, has impacted the teaching of statistics for adults, as each one has some relevance to the teaching of statistics today.

Going back to Piaget's ideas of cognitive constructivism, it is understood that adults construct their own meanings of the world from their experiences, and they bring this experience into the statistics classroom. These experiences include their ideas, opinions, values and beliefs, as they relate to their abilities to succeed in a statistics course. And for a constructivist, learning can be affected by the beliefs and attitudes of the learner; consequently, if adults had bad experiences learning mathematics during their earlier school days, it can leave them with negative views of their mathematical skills, which, in turn, can affect their ability to handle statistics problems (Gal & Ginsburg, 1994). Statistics instructors need to be sensitive to students' attitudinal dispositions, because these factors can have an impact on the learning and teaching process in statistics (Mvududu, 2005). Additionally, because adults bring with them knowledge from their own experiences, new knowledge is constructed internally by transforming, organizing, and reorganizing previous knowledge (Cobb, 1994), as well as externally through the environment, and social factors that are influenced by culture, language and social interactions (Mills, 2003).

According to cognitive and sociocultural constructivism ideas, "learning is an active process, and knowledge is constructed rather than innate or passively absorbed" (Fox, 2001, pp. 24-25). To create an active learning environment, an instructor needs to create a supportive atmosphere where students feel safe to explore statistical concepts, sometimes through trial and error, and increasingly, this is accomplished through the use of technology in the classroom. One use of technology in the statistics classroom is to incorporate web-based learning tools, which allows students to learn specific topics (i.e., Central Limit Theorem, statistical power, or hypothesis testing) in statistics through computer-based simulations which include analyses, charts, and graphs (Aberson, Berger, Healy, Kyle, & Romero, 2000; Aberson, Berger, Healy, & Romero, 2002; Aberson, Berger, Healy, & Romero, 2003). Instructors can further enhance students' learning by allowing them to complete assignments in the classroom through group work or individual completed assignments.

Additionally, because, as radical and sociocultural constructivism purports, "knowledge is invented and not discovered" (Fox 2001, p. 26), we will always view the world from some sociocultural or historical viewpoint. And to understand how the world is viewed in our contemporary time, it is recommended that real data sets be used in the classroom for analyses and interpretation (Raymondo & Garrett, 1998). Further, because many research studies concerning health and medical issues are reported regularly in newspapers and magazines, and often with conflicting results (Utts, 2003), these can be brought into the statistics classroom in order to enhance students' statistical reasoning ability toward everyday issues (Lawson, Schwiers, Doellman, Grady, & Keinhofer, 2003). The use of real-world problems can bridge the gap between reality and numbers as the context makes the number meaningful; therefore, examples should be presented in the context of real-world problems (Garfield, Hogg, Schau, & Whittinghill, 2002).

It is important for instructors in the statistics classroom to remember that when they present the same lesson to a variety of students, their experiences may result in different learning by each student as, "all knowledge is idiosyncratic and personal" (Fox, 2001, p. 29). Therefore, some students may need more instructional support from teachers or from more skilled peers who can help them bridge the gap between their current skill level and the desired one (Mvududu, 2005). This is much like the idea of sociocultural constructivist Vygotsky's Zone of Proximal Development, in which importance is placed on social interactions with more knowledgeable others. Or, it can be accomplished by an instructor who includes time for interaction and discussion with the students (Mills, 2003).

"Knowledge is socially constructed" (Fox, p. 30) is one tenet of sociocultural constructivism, which can be accomplished in a variety of ways within the classroom. It can be constructed through the use of textbooks, or through students using small groups to promote active learning as they think, discuss and process information (Keeler & Steinhorst, 1995). Or, it can consist of group work to develop a research question, design a collection strategy, collect and analyze data, and give an oral presentation of the results to the class (Chance, 2005).

It is important to note there is one other area of great importance that sociocultural constructivists focus on that has relevance to the teaching of statistics—namely, "social facts and identities are socially constructed while made to appear natural" (Bredo, 2000, p. 133), as some social and political elites define knowledge in self-serving ways. And to really become educated in contemporary society, where "the age of information is an age of numbers" (Steen, 2003, p. 62), adults need to be statistically literate, which should be the outcome of their learning experiences in a statistics course. Because numbers underlie everyday decisions, from quantitatively based proposals that shape public policy in education and health (Steen) to decisions regarding political candidates, medicines and health (Moreno, 2002), there is a need to be statistically literate in order to sort social facts from social fallacies. And this links with the last tenet, "learning is essentially a process of making sense" (Fox, 2001, p. 30), as constructivism purports that learning is about understanding and, in doing so, it takes us beyond the conception of rote learning.

In sum, the many different contributions from the epistemology of constructivism have changed the way statistics is currently being taught today in higher education. These include the cognitive, sociocultural and radical ideas of constructivism and all contribute different perspectives, but complement each other by including factors that need to be considered in the total construct of learning. But more importantly, the statistics classroom has changed from a behaviorist orientation to a more constructivist orientation in learning in adult education, as shown by the current literature, which follows next.

Review of the Empirical Literature on Statistics

The introductory statistics class has become a required part of many different majors in the college curricula, as an interest in statistical education has resulted from the Democratization of Mathematics. As aforementioned, this movement was set in motion by the NCTM, in which they responded "to the changing needs of society by publishing standards for school mathematics" (Steen, 2003, p. 3), and later, leading the way to promote statistics in higher education, which included the social sciences, psychology and sociology. The challenge now is for educators to help students become statistically literate, and this is currently being accomplished through research in statistics education.

A literature review search was conducted in various databases across multiple disciplines, as the goal was to locate numerous articles on statistical literacy through articles that examine the teaching and assessment of introductory statistics classes in the social sciences. Databases searched were EBSCO Host, Education Abstracts, Elsevier Science Direct, ERIC, Kluwer, JSTOR, Psychological Abstracts, PsycINFO, ProQuest, Social Science Citation Index Sociological Abstracts, and Wiley, and articles chosen were published within the last 10 years. Keywords and phrases to locate articles were statistical literacy, quantitative literacy, numeracy, innumeracy, mathematical literacy, teaching statistics, and assessment and statistics. Accordingly, this literature review will examine empirical research concerning introductory statistics courses, and will be arranged in four sections, beginning with the populations and purposes of the research, followed by methodological issues and teaching, and ending with evaluations.

Populations and Purposes

This section will identify who the populations are in the studies and the various purposes for undertaking this empirical research. Populations of students will be identified in three ways. The first is by examining students' enrollment by majors; second, by the type of course they are enrolled in, for example, statistics or research methods; and third, whether students are traditional or non-traditional college students. In addition, various reasons for examining statistics classes are reviewed. These reasons include (a) evaluation of tutorial and data analysis programs, (b) Web-based statistics courses, (c) traditional and non-traditional teaching methods, and (d) differences in students' abilities according to majors.

Nineteen empirical studies on the teaching of introductory statistics were found, with most focused on populations of undergraduate students in the social sciences, namely psychology and sociology. There were eight studies that focused on students who were psychology majors (Aberson, Berger, Healy, & Romero, 2002; Cralley & Ruscher, 2001; Lawson, Schwiers, Doellman, Grady, & Kelnhofer, 2003; Morris, 2001; Morris, Joiner, & Scanlon, 2001; Rajecki, 2002; Vanderstoep & Shaughnessy, 1997; Warner & Meehan, 2001), three on criminal justice (Bushway & Flower, 2002; Proctor, 2002; Proctor, 2006), and one who used a combination of these majors (Raymondo & Garrett, 1998). However, there were some studies that did not focus on students in the social sciences. One study's participants included students from the majors of geography, fisheries and wildlife, animal sciences, history and computer science (Symanzik & Vukasinovic, 2006), business (Ward, 2004), engineering (Kvam, 2000), and four studies did not clarify the discipline of their participants, other than they were college students in introductory statistics courses (Aberson, Berger, Healy, Kyle, & Romero, 2000; Aberson, Berger, Healy, & Romero, 2003; Johnson & Dasgupta, 2005; Utts, Sommer, Acredolo, Maher, & Matthews, 2003). This collection of empirical literature indicates that attention is focused toward examining students enrolled in the social sciences and their learning of statistics, rather than other college majors.

There are two concerns that emerge from this review. First, this is a small collection of empirical literature for an approximate span of 10 years, and no articles were found that examined statistics and learning in some other majors, for example sociology, political science or education. This is a concern, because in these disciplines, like the social sciences, statistics are often used in research to create public policy or examine teachers' effectiveness in the classroom in elementary, secondary and post-secondary education. Clearly, there should be more research focused on college students' learning of statistics.

Although most courses students were enrolled in for these studies were introductory statistics (Bushway & Flower, 2002; Cralley & Ruscher; Johnson & Dasgupta, 2005; Kvam, 2000; Morris, 2001; Morris et al., 2001; Proctor, 2002; Proctor, 2006; Raymondo & Garrett 1998; Symanzik & Vukasinovic, 2006; Ward, 2004), two studies employed students from research method classes for their study (Morris; Vanderstoep & Shaughnessy). And differently, one study was designed for students who already completed a statistics course, as the purpose of the research was to examine if students' learning of statistics could be improved by using computer-enhanced instruction (Morris, et al.). Two concerns emerge from the last two findings. It is not clear if the research method course also included statistics, as some courses are a combined research/statistics course in which students do not have any prior statistical knowledge.

One the other hand, some research method courses only include methodological issues pertinent to designing a research project; statistics is not taught, but sometimes reviewed, because for many students in the social sciences, statistics is a prerequisite for research methods. Therefore, it is not known if the research is measuring students' previous or new learning of statistical concepts. And likewise, if the research is examining students' previous learned statistical skills, does this infer that students did not learn these concepts in their introductory statistics class?

Many participants in the empirical literature were identified as traditional college students (Aberson, et al., 2000; Aberson, et al., 2003; Johnson & Dasgupta, 2005; Raymondo & Garrett 1997; Vanderstoep & Shaughnessy, 1997; Ward, 2004). And some of the studies included a mixed population of traditional and non-traditional students (Bushway & Flowers, 2002; Morris et al., 2001; Proctor 2006); however, others did not identify whether traditional or non-traditional students were the participants (Cralley & Ruscher, 2001; Kvam 2000; Morris 2001; Proctor, 2002; Rajecki, 2002; Lawson, et al., 2003; Symanzik & Vuskasinovic, 2006). Moreover, in the studies that included both traditional and non-traditional students, none of the data was separated and analyzed to examine if there were different results between the two groups concerning their learning of statistical concepts. Accordingly, not much is known about non-traditional students and their learning of statistics, either because the data was not separated for analyses, or due to a lack of information collected in the demographics of these studies. This could be important, because some adult students, who return to college years after completing high school, may have mathematics anxiety, which could impede their learning of statistics.

Many different purposes for examining the teaching of statistics were uncovered in the literature, and are summarized by commonalities in the research. Five studies focused on evaluation of interactive tutorials for teaching statistical concepts, which are statistical power (Aberson, et al., 2002), correlations (Morris, 2001; Morris, et al., 2001), the central limit theorem (Aberson, et al., 2000; Morris, et al., 2001), and hypothesis testing concepts (Aberson, et al., 2003). Other computer data analysis programs were also examined, such as Microsoft Excel (Warner & Meehan, 2001), and SPSS (Proctor, 2002; Raymondo & Garrett, 1998). Similar students' performance of mastering statistical concepts was examined by using Web-based statistics courses (Symanzik & Vukasinovic, 2006; Utts, et al., 2003; Ward, 2004). Two studies focused on students' preference and learning outcomes regarding traditional teaching and non-traditional teaching methods (Johnson & Dasgupta, 2005; Kvam, 2000). Some similar studies focused on using real-life events, by using real data sets or everyday problems from the media in the statistics course in which students could analyze and interpret statistical results (Cralley & Ruscher, 2001; Lawson, et al., 2004; Rajecki, 2002; Vanderstoep & Shaughnessy, 1997). Tutoring and its affects on students' learning was the focus of another article (Bushway & Flowers, 2002), while another compared statistical knowledge between criminal justice majors and non-criminal justice majors (Proctor, 2006). While this research informs us on many different pieces of information on statistical learning, there are no studies that address students' statistical literacy skills. Definitely, this is an area that warrants further research.

Methodological Issues

This section examines methodological issues of the research, as properly designed research is essential to good empirical results. This section includes four parts: first, the research design, the methods section, with a focus on demographic variables, and the purpose statement, which should be inclusive in all types of empirical research.

The most prevalent research design employed in this literature review was a quasi-experimental research design in which some were conducted using pre- and post-tests (Aberson, et al., 2000; Lawson, el al., 2003; Morris, 2001; Morris, et al., 2001; Vanderstoep & Shaughnessy, 1997), while others were conducted using no pre-tests, just comparison groups in their analyses (Aberson, et al., 2002, 2003; Bushway & Flower, 2002; Johnson & Dasgupta, 2005; Kvam, 2000; Proctor, 2002, 2006; Rajecki, 2002; Raymondo & Garret, 1998; Symanzik & Vukasompvoc, 2006; Utts, 2003; Warner & Meehan, 2001; Ward, 2004). Only one study used an experimental design (Cralley & Ruscher, 2001). And while all these types of research can inform us with a variety of information, only the quasi- or experimental designs with pre- and post-tests are able to measure learning gains in students' statistical knowledge, as comparison groups can only find differences between groups.

An important part of any quantitative research is the methods section, in which information can be found about the participants, the instrument and the procedure of the research. A central issue in some of the empirical literature is that many did not include a methods section (Aberson, et al., 2002; Bushway & Flowers, 2002; Johnson & Dasgupta, 2005; Kvam, 2000; Symanzik & Vukasinovic, 2006; Utts, et al., 2003), and the few that had a methods section, did not include the demographics of the participants

(Proctor, 2006), while some only listed the demographics in a chart (Bushway & Flowers; Johnson & Dasgupta; Utts, et al.).

Further, some of the empirical literature that did include a methods section with demographic information about their participants provided only scant information. Some research articles only indicated that participants were undergraduate students (Rajecki, 2002), and many did not include the ages of the participants (Aberson, et al., 2000, 2002, 2003; Cralley & Ruscher, 2001; Lawson, et al., 2003; Morris, 2001; Proctor, 2002; Rajecki, 2002; Utts, et al., 2003). Undergraduate students can be either traditional or non-traditional students, and with no ages reported, it is impossible to know which population the research was completed on. Further, some research included both undergraduate and graduate students jointly as participants when undertaking their research (Aberson, et al., 2000, 2003; Johnson & Dasgupta, 2005), which is problematic, because there is a difference in age and learning accomplishments, and mixing this population can affect the results, especially when examining students' learning of statistics. It can be assumed that graduate students have prior learning in statistics, and undergraduates have none or very little.

Gender was another demographic missing from much of the research articles (Aberson, 2000; Cralley & Ruscher, 2001; Rajecki, 2002; Utts, et al., 2003; Vanderstoep & Shaughnessy, 1997; Ward, 2004; Warner & Meehan, 2001). And likewise, none of the empirical research listed the ethnicity of the participants (Aberson, et al., 2000, 2002, 2003; Bushway & Flower, 2002; Cralley & Ruscher; Johnson & Dasgupta, 2005; Kvam, 2000; Lawson, et al., 2003; Morris, 2001; Morris, et al., 2001; Proctor, 2002, 2006; Rajecki; Raymondo & Garrett, 1998; Symanzik & Vukasinovic, 2006; Utts, et al.;

Vanderstoep & Shaughnessy; Ward; Warner & Meehan). Conversely, all the aforementioned research articles did report the number of participants in the studies, either listed in the method section, or in the text of the articles. While we know how many participants participated in these studies, we do not know if there are gender, ethnic, or type of college students (i.e., traditional or non-traditional) differences that could have impacted the results, because of the lack of this information.

However, a few older studies examined performance in statistics classes and gender (Ware & Chastain, 1991; Brooks, 1987; Buck, 1985). Buck found no differences in final course grades between males and females; similarly, Ware and Chastain (1991) found no gender difference between males and females when comparing scores for statistical concepts. On the other hand, Brooks' research showed females had higher final course grades than men. To try to account for these differences, Buck believed females came into her classroom as mathematically disadvantaged, but because she was a female instructor, she motivated and inspired them, and therefore, their abilities became equal to males. Brooks (i.e., a male instructor) hypothesized his results were due to females attending more classes and seeking help.

Although this may have some merit, another study about gender differences in mathematical abilities may provide us with an explanation. In comparing male and female abilities in mathematics, results showed that females performed better than males on tasks of computation, while males performed better than females in problem solving (Hyde, Fennema, & Lamon, 1990). Overall, gender differences may be due to the type of introductory statistics course, as some contain more mathematical computations than others. Or as Schram (1996) explains, gender differences can be accounted for due to the types of assessment; women perform better than men on final grades, while men perform better when tests are used. Interestingly, gender differences in doing well in mathematics are explained by gender for some. Mendick (2005) informs us that some believe that "doing mathematics is doing masculinity" (p. 237), or in other words, males do better at mathematics because they are male; therefore, females are not expected to perform well in mathematics because they are female.

One final and important demographic to consider in this research is the location where the research took place. Most research on students' learning of statistics took place in 4-year universities, which included state (Aberson, et al., 2000; Raymondo & Garrett, 1997; Symanzik & Vukasinovic, 2006), urban state (Aberson, et al., 2002; 2003; Johnson & Dasgupta, 2005; Proctor, 2002) rural state (Aberson, et al., 2002; 2003), and private universities (Aberson, et al., 2002, 2003; Vanderstoep & Shaughnessy, 1997; Ward, 2004). There was only one that listed that its research included participants at a community college (Aberson, et al., 2000). And some of the empirical literature did not list the location where the research took place (Bushway & Flowers, 2002; Cralley & Ruscher, 2001 Kvam, 2000; Lawson, et al., 2003; Morris, 2001; Morris et al., 2001 Proctor, 2006; Rajecki, 2002; Utts, et al., 2003; Warner & Meehan, 2001). This set of empirical literature indicates that most research is done at state universities, and little research undertaken at private universities, and even less at community colleges. This is disturbing, as it indicates a lack of interest in the teaching of statistical methods in the social sciences; and most types of colleges have some type of degree in the social sciences. Also, it shows a lack of interest in examining how statistics courses are taught and how students do academically in statistics courses, and most importantly, no one has examined whether students have achieved statistical literacy after completing their statistics courses.

There is one final comment concerning methodology—and that is an explicit statement concerning the purpose of the research. This should be clearly stated at the end of the literature review, and just before the methodology section; however, many of these empirical articles did not state a specific purpose of their research, as a result of which the reader had to infer the purpose after reading the text (Aberson, et al., 2000, 2002, 2003; Bushway & Flowers, 2002; Kvam, 2000; Rajecki, 2002). Not having a specific purpose statement makes it hard to interpret the focus of the research, and in some cases, the purpose of the research can be misinterpreted.

Teaching of Statistics Courses

Teaching methods and evaluations of introductory statistics classes are the focus of this section. Most teaching methods in the statistics classroom focus on active learning, aided by the use of technology. While there are a variety of ways in which introductory statistics classes are evaluated, the most common way is to compare traditional (i.e., lecture) with non-traditional classrooms (i.e., active learning) with students' exam scores or final grades. Other ways include students' evaluation of the courses they complete.

Many of the empirical articles that examine teaching in introductory statistics courses reflect active learning techniques in the classroom (Aberson, et al., 2000; 2002; 2003; Bushway & Flower, 2002; Cralley & Ruscher, 2001; Johnson & Dasgupta, 2005; Kvam, 2000; Lawson, et al., 2003; Morris et al., 2002; Morris, 2001; Proctor, 2002; Rajecki, 2002; Raymondo & Garrett, 1998; Symanzik & Vukasinovic, 2006; Utts, et al., 2003; Vanderstoep & Shaughnessy, 1997; Ward, 2004; Warner & Meehan, 2001). And often to engage students in active, rather than passive learning, many statistics classes integrate technology into their statistics classroom in three ways—computer simulations, data analyses programs and Web-based classes.

The first way technology is used in a statistics classroom is through computer-based simulations that demonstrate specific statistical concepts, for example population sampling distributions, the central limit theorem or statistical power. These simulations allow students to analyze data and create charts and graphs and receive immediate feedback about the choices they have made (Aberson, et al., 2000, 2002, 2003; Morris et al., 2002; Morris, 2001). This type of teaching methodology was evaluated by comparing students' learning in traditional lecture-only and those in computer-based simulation classrooms. Aberson, et al. (2000, 2002, 2003), found that students who used the computer simulations scored higher on exam questions related to the concepts contained in the computer based simulations than students who received instruction in a lecture-only classroom. In addition, these active learning techniques were evaluated by the students who reported the simulations were easy to use, concept explanations were clear, and overall, a useful aid that helped them learn a specific statistical concept (Aberson, et al., 2000, 2002, 2003). However, mixed research results were found in measuring students' learning of correlations and measures of central tendency. Students who used the computer simulation programs scored higher on their knowledge on the measure of central tendency, but not on correlations. And the students who attended a traditional lecture-only classroom scored higher than students who participated in using the computer simulation (Morris; Morris, et al.).

Second, computer programs are incorporated into the statistics classrooms, which allow students to work with data sets in Excel or SPSS. Learning is hands-on, as they enter and analyze data and interpret results (Proctor, 2002; Raymondo & Garrett, 1998; Warner & Meehan, 2001). Similar to the computer simulation programs, learning outcomes from these teaching methods were mixed. Raymondo and Garrett found no statistically significant difference in students' final grades between students who used the computer program SPSS to analyze data sets and students who did not. Proctor compared students who used Excel and SPSS and found that students' knowledge of statistical concepts was higher in the group who used Excel. This could be due to the repetitive task of writing up computational formulas and then entering the formulas into Excel.

Succinctly. students are exposed to a replication of hand calculation procedures, whereas in SPSS, no hand calculations are used, because the program is more automated. Differently, Warner and Meehan examined the outcome of using Excel in a statistics classroom, but did not measure students' learning. Instead, their research examined students' perceptions of their learning and found that "although the students agreed that the assignments improved their knowledge of statistical concepts, they rated the assignments more highly in terms of improving their computer skills" (Warner & Meehan, p. 297).

Third are the Web-based instructional statistics programs, which are not to be confused with distance learning classes. These are called *hybrid* programs because they use a combination of a Web-based statistics program and have students conduct face-to-face, albeit less, contact with the instructor than students who attend a traditionally taught statistics class (Bushway & Flowers, 2002; Johnson & Dasgupta, 2005; Symanzik &

Vukasinovic, 2006; Utts, et al., 2003; Ward, 2004). Students who attend these hybrid statistics classes often have hand-in homework assignments (Bushway & Flowers; Symanzik & Vukasinovic; Utts, et al.; Ward), and engage in activities such as interactive worksheets, data analyses, and applet demonstrations of statistical concepts (Johnson & Dasgupta; Symanzik & Vukasinovic; Utts, et al.; Ward).

Evaluations of these Web-based instructional statistics courses were based on students' preference between a traditional (i.e., lecture) and non-traditional (i.e., active learning) statistics classroom. Some results showed students preferred non-traditional more than traditional teaching methods (Bushway & Flowers, 2002; Johnson & Dasgupta, 2005), while others showed students in the non-traditional statistics classroom felt the workload was excessive, whereas students in the traditional statistics class did not (Utts et al., 2003). And no difference in students' preference for Web-based instruction versus a traditionally taught classroom was found in research completed by Symanzik & Vukasinovic (2006).

More consistent were evaluations of students' learning in Web-based instructional statistics courses and traditional statistics classrooms. Utts et al. (2003) and Ward's (2004) results showed no significant differences in students' knowledge of statistical concepts between students in the traditional (i. e., lecture) and non-traditional (i.e., active learning) classrooms. And when final grades were used to evaluate students' learning of statistics, Symanzik and Vukasinovic (2006) and Bushway and Flowers (2002) found no significant differences in grades between students in the Web-based and traditional statistics classrooms.

Other teaching methodologies incorporate technology into the statistics classroom by using real data sets for statistical analyses (Morris 2001) or like-real data sets, which allowed students to develop an understanding of statistical concepts and data analyses (Cralley & Ruscher, 2001; Proctor, 2002). However, only one study evaluated students' learning after using common or individual data sets. Results showed students scored higher on exams that used individual data sets than students who had analyzed shared data sets. This could indicate that some of the students' analyses could have been copied from other students, and some may have not analyzed the data themselves.

Two studies did not incorporate technology into their statistics classroom, but tried to engage students in their learning of statistical concepts by incorporating everyday problems into the course curricula (Lawson, et al., 2003; Vanderstoep & Shaughnessy, 1997), or by adding a writing element to the course (Rajecki, 2002; Vanderstoep & Shaughnessy, 1997). Both studies that added a writing component to their statistics course showed students who participated in these classes scored higher on evaluations of statistical concepts than students who were not in statistics classes that incorporated writing into the curriculum. And although many classes that incorporated technology into their statistics classrooms included the use of collaborative projects and discussion groups in their course curriculum (Johnson & Dasgupta, 2005; Kvam, 2000; Utts, et al., 2003; Ward, 2004), none of the research examined whether these techniques aided students in their learning.

Non-Cognitive Factors and Learning Statistics

Adult students' attitudes and beliefs can either impede or assist them in achieving statistical literacy. Their attitudes about statistics courses are often negative, as some "view statistics as the worst course taken in college" (Hogg, 1991, p. 342). Many become anxious because they believe that "statistics is a difficult subject, involves lots of math, and is not relevant to their career goals" (Kirk, 2002, p. 2). Hence, adult students' attitudes and beliefs are important to include as part of the literature review for two main reasons: First, "their role in influencing the teaching/learning process; and second, their role in influencing students' statistical behavior after they leave the classroom" (Gal, Ginsburg, & Schau, 1997, p. 2).

Five studies examined students' attitudes toward statistics using the SATS (Wiberg, 2009; Carnell, 2008; Froelich, Stephenson, & Duckworth, 2008; Alldredge, Johnson & Sanchez, 2006; Chadjipadelis & Andeadis, 2006; Nassser, 1999). However, there were differences in research methodologies. These studies compared pre- to post-test scores for groups in multiple ways. Treatment/control groups were used by Alldredge et al., (i.e., treatment: video clips of statistics in real world settings), Froelich, et al., (i.e., groups separated by mathematical ability) Carnell, Chadjipadelis and Andeadis, (i.e., treatment, special projects), and Wiberg (i.e., treatment, revised course).

And results were mixed. Carnell (2008), Alldredge et al. (2006), and Chadjipadelis and Andeadis (2006) showed no significant differences between the treatment and the control groups on scores from the SATS elements of affect, cognitive competence, value and difficulty.

However, for Alldredge et al. (2006) there was a significant interaction effect between the treatment group and the preliminary algebra test score, with these students scoring higher on the element of affect. As their algebra test scores increased, their feeling concerning statistics grew more positive. As the authors suggest, this may have occurred because students who are more capable mathematically are more receptive to the video message showing positive uses of statistics.

Carnell (2008), whose research involved the inclusion of a single project, suggested one possible explanation for the lack of significance for affect, which was that one project may not have been “sufficient to impact global attitudes” (p. 8), and likewise for cognitive competence. And importantly, test performance and past experiences in different types of quantitative classes could impact cognitive competence, negatively or positively. Further Carnell reports, students with a “more extensive math background might feel differently about statistics than students with a more limited background” (p. 7). And for difficulty, no changes many indicate that the course may have turned out to be more difficult than students originally anticipated. There was no interpretation for the element, value.

Conversely, Wiberg (2009) showed increased scores for SATS elements, affect, cognitive competence, and value, but no differences for difficulty on pre- to post-test scores, for the treatment group, the revised course. These increases were attributed to the different types of teaching methods used in the revised course, which includes data-driven problems and student-centered learning, Similar to Wiberg, Froelich et al. (2008) found differences in attitudes on two elements of the SATS, affect and cognitive competence, on pre- to-post-test scores, albeit the post-test scores decreased for the group with the lowest mathematical abilities. A possible explanation would be that students view an introductory statistics course as a mathematics course, and because they have poor mathematical abilities, they have negative feelings about it, and are less confident in their abilities to learn the course materials.

Like many of the current studies on the teaching of statistics, there was no gender or adult student status reported for Wiberg (2009) and Froelich et al. (2008). Alldredge et al. (2006) reported gender but did not include it in the analyses, and no adult student status was reported. Carnell (2008), and Chadjipadelis and Andeadis' (2006) results showed no differences between gender on any of the elements, affect, cognitive competence, difficulty and value on the SATS scale.

Students' beliefs about statistics are the internal feelings that underlie their attitudes toward statistics; however, as explained by Gal, Ginsburg and Schau (1997), "other than the commonly-held belief that statistics is heavily mathematical and that statistics is a somewhat difficult discipline, students' beliefs about statistics as a domain remain mostly unexplored" (p. 4). Further, many students do not come into a statistics class ready to learn statistics, but rather, they carry baggage from past experiences that can include negative beliefs about themselves in relation to mathematical issues (McLeod, 1992). Some students have already formed beliefs about the value or lack of value that statistics has for their future careers.

Gal and Ginsburg (1994) explain that attitudinal factors become important to students once they begin to experience difficulties with statistics class material. These attitudinal factors are similar to students' learning of mathematics. Tobias (1994) suggests that adults' early memories of learning mathematics, which are often negative, are triggered by becoming confused in mathematics classes now, when they fail to understand some mathematical concept. This leads to losing a sense of confidence in their abilities and a loss of control over their comprehension. These losses lead to students becoming bored or becoming disengaged from learning, as they perceive it as a futile attempt to learn. Often students have developed negative views about their abilities to do mathematics. Gal and Ginsburg suggest that a similar process happens in a statistics class, as for most students, statistics is similar to mathematics.

Gal and Ginsburg (1994) strongly suggest that when examining students' attitudes toward statistics, which is accomplished by using a Likert-type scale, it should be used in conjunction with open-ended questions that reflect some of the elements from the scale. This will allow students to "describe the intensity and frequency of specific emotional responses, and elaborate on their source" (n.p). This will give insight into the beliefs that underlie attitudes toward statistics.

From this literature review. it is evident that many introductory statistics course instructors incorporate active learning into their course curriculum aided by the use of technology, but many classes are still taught by traditional methods. Nevertheless, it is not evident which type of teaching method can help students learn important statistical concepts. In some studies, students who used computer simulations to learn statistical concepts scored higher on exam questions, while in other studies they did not. Likewise were the results when comparing students' scores who used computer data analyses programs. Some students who used the data analyses programs scored higher on exams, while other research showed no significant difference in students' final grades between those who used data analyses programs and those who did not. Similar were the results of research examining students' final grades who completed Web-based statistics courses and those who completed the course in a traditional statistics classroom—there was no significant difference in grades. Differently, other teaching methods, analyzing individual data sets and completing a writing element, resulted in students with higher scores on their exams. Collaborate work and discussion as a way of teaching was not used as an independent variable in measuring students' learning.

Hence, the ideas of constructivism are reflected in many of the studies discussed in this review of the current literature on the teaching of statistics. From how statistics is currently taught, it is necessary to examine how the term statistical literacy came into being to understand how to measure statistical literacy in adult students. And to understand the conceptual framework for this research, it is necessary to discuss the various terms that imply statistical literacy, because these terms help us to understand that it is a multifaceted concept, and necessitates a model that encompasses many elements.

Conceptual Framework

In the late 19th century the British statesman Benjamin Disraeli proclaimed: "There are three kinds of lies: lies, damned lies, and statistics" (Andrews, Biggs & Seidel, 1996, n.p.). Perhaps he was one who was not fooled by misleading graphs or by flawed research. Although the importance of statistical literacy in this era was important, it has become increasingly so in contemporary society. We are constantly surrounded by statistics that are presented to add credibility and marketability to products from drugs to automobiles, to promote political agendas, and to set standards for health and safety issues, thereby affecting the personal welfare of the nation's citizens. The concern for statistical literacy is ever-increasing for each person's "quality of life and for our collective well-being" (Steen, 2004, p. 27). Succinctly, "statistical literacy is the study of statistics used in everyday life. Statistical literacy helps citizens in a democracy read and interpret numbers in the news to make intelligent decisions" (Statistical Literacy, 2007, n.p.); however, the term statistical literacy has an abundance of conceptual definitions, and there is an array of different words that imply statistical literacy. The statistical phraseology (which will be discussed next) shows many similarities and differences in defining statistical literacy. Many denotations are neither mutually inclusive nor exclusive in their overall meanings; however, all are important elements that help conceptualize statistical literacy. Accordingly, it is necessary to explicate how the various conceptual elements become operational in order to measure statistical literacy.

In order to join the conceptual definitions of statistical literacy into an operational one, a conceptual framework will be used that incorporates these numerous definitions into elements, which can later be quantitatively measured to examine levels of statistical literacy in adult college students. Because the research is *cutting edge* (i.e., new research), only one model of statistical literacy has been created, Gal's (2004) Model of Statistical Literacy. Appropriately, the first section of this conceptual framework demonstrates the variety of definitions of statistical literacy, which are used in multiple disciplines within the research literature. The second section elucidates how these definitions fit the seven elements that define statistical literacy on Gal's Model of Statistical Literacy, which is the conceptual framework for this research.

Statistical Phraseologies

Statistical phraseologies refer to the different terms and words that are used to define statistical literacy; these are numeracy **(**Brown, Askew, Baker Denvir & Millett, 1998; Steen, 1990; Watson, 2002), adult numeracy (Gal, 2002; NCES, 2006), innumeracy (Cerrito, 1999; Paulos, 1989), quantitative literacy (Manaster, 2001; Rosen, Weil & Van Zastrow, 2003; Steen; 1990; Steen, 2001), mathematical literacy (Rosen, et al.), mathematical illiteracy (Paulos), statistical literacy (Moore, 1997; Schield, 1999; Rumsey, 2002a; Wallman, 1993), and statistical reasoning (Garfield & Chance, 2000;

Garfield, 2003). This variety embraces the multiple meanings of statistical literacy, with some defining it within arithmetic or mathematics, or mathematics and statistics, while others combine arithmetic, mathematics and statistics, with some incorporating social implications of being statistically literate and some the importance of research methodology. These will be discussed in detail next.

Numeracy/Adult Numeracy. The term numeracy was conceptualized in the British study, the Crowther Report in 1959 (Brown, et al., 1998), which focused on elementary school children's achievement in numeracy (Passow, 1962). Numeracy, as defined in the report, focused on proficiency of basic numerical skills, such as mental and written calculations, or the multiplication tables. One issue in the British school systems, according to the Crowther Report, was that number problems being taught in elementary schools had no relevance to real-life mathematical skills—the teaching of mathematics was based on artificial contexts (Brown, et al.).

Conversely, another definition of numeracy is much more complex. It is defined through worldly dimensions of life deeply embedded within societal constructs. Steen (1990) asserts that "numeracy is to mathematics as literacy is to language" (p. 211), because each represents a type of communication that is indispensable to civilized life. Compared with the Yin and Yang, "numeracy and literacy are the entwined complements of human communication" (Steen, p. 212).

Accordingly, Steen (1990) defines numeracy through four societal dimensions that reflect ideas from both mathematics and statistics. First, practical numeracy is the importance of numbers to individuals' functioning in their everyday life, from being able to compare loans, to being a savvy consumer, or even to understand their chances of winning the daily lottery. Consider those who lack the ability to employ basic rules of probability and play lotteries. This allows others to "take in disproportionate revenue from less well-educated citizens in part because few people with minimal education understand chance" (p. 217). Second, to explain civic numeracy, the focus turns toward society where inferences are drawn from analyzed data, which determine major public policies. Third, professional numeracy relates to the use of mathematics skills used in various jobs, from the medical field to assembly-line operations. New drugs are approved for the public only after going through rigorous testing, and in manufacturing, statistical processes are crucial to quality control. Finally, there is leisure numeracy, which is related to leisure activities in the American culture, which include illegal numbers games, casino gambling and horse betting; all are based on theories of probability (Steen). However, numeracy described as adult numeracy has a different connotation.

Literacy, according to the 1991 National Literacy Act, is broadly defined and encompasses multiple skills. It is an individual's ability to read, write and speak in English and to compute and solve arithmetic problems at levels of proficiency necessary to function on the job and in society, to achieve one's goals, and develop one's knowledge and potential. (Gal, 2002, p. 20) By leaving the definition broad, it leaves the adult education community (i.e., practitioners, program administrators and policy makers) to engage in ongoing dialogue to clarify the goals and appropriate methods to develop adults' numeracy skills necessary in a civic society. However, numeracy has received less dialogue than other skills such as reading, writing and speaking (Gal).

In fact, the term numeracy is less often used in the adult education community, and has no exact meaning. Some views of numeracy are related with basic computational skills, similar to basic reading, or writing skills in adults (NCES, 2006). This is similar to the definition of numeracy as defined in by the Crowther Report, which primarily focused on proficiency of basic numerical skills, such as mental and written calculations, or the multiplication tables in the British school systems (Passow, 1962). However, a better approach in defining numeracy within adult education would be to focus on the quantitative aspects of the adult world, because “one key declared goal of educational programs…is preparing learners to become more informed citizens and workers who can effectively function in an information-laden society” (Gal, 2000, p. 135).

Innumeracy. The opposite of numeracy is innumeracy, and it is used in reference to statistical literacy by Paulos (1989), who also uses the term mathematical literacy. He describes innumeracy by explaining there are social costs to contemporary society when citizens are innumerate. For example, to understand medical or drug testing, to be savvy to pseudo-medical treatments, the chance occurrence of an airplane or automobile accident, or casino gambling, an understanding of probabilities is necessary. Also, the public is shown information on elections and political polls that consistently use confidence intervals to describe data collected on specific candidates. To understand these polls, it is necessary to understand how the idea of randomness can affect the results. Further to understand published research in mainstream society, the difference between statistical and practical significance is warranted; but, unfortunately, this is often misunderstood by the general public. Probability is the foundation of statistics, with significance an important element of statistics, while randomness is a part of research methodology, but both are crucial to understand if citizens are to be informed citizens in society (Paulos).

Similar to Paulos's conceptualization of innumeracy, Cerrito (1999) describes it by referring to statistics that are related to the political milieu and medical studies. For individuals to become a part of the political process, it is necessary that they become informed about how the issues are being debated, and to be informed is to be literate in understanding the data supporting the issue. Especially in medical trials, it is essential to understand, for example how the pertussis vaccine for infants and hepatitis vaccine for high-risk groups may affect individuals who receive them. Elaborating further on the conception of innumeracy, Cerrito describes it in relation to an individual's functioning in society as "numbers permeate society and are constantly referred to in political dialogue. Those with credentials that label them as experts are free to expound on numbers without fear of challenge" (p. 2). Hence, an innumerate society allows for others to deceive those who do not understand.

Quantitative/Mathematical Literacy. Quantitative literacy is a broad term that can include many mathematical topics, such as arithmetic, geometry and algebra. However, to understand its meaning from the perspective of statistical literacy, an examination of the skills that are embedded in the elements and expressions of quantitative literacy are warranted, because statistics as a whole is a distinct discipline, albeit inclusive of numerous elements from mathematics. These skills include arithmetic skills, such as (a) calculations; (b) having the ability to use information conveyed as data, graphs or charts; (c) being proficient in using a computer to record data and perform calculations; (d) modeling of exponential, multivariate, and simulation models; (e) statistical concepts, to understand the importance of variability, the difference between correlation and causation, between randomized experiments and observational studies; and (f) chance, to be able to evaluate risks from available evidence and reasoning that is being able to use logical thinking and exercising causation in making generalizations (Steen, 2001).

Similar to Steen's definition, Rosen, et al. (2003) define quantitative literacy or mathematical literacy (terms are interchangeable) through two social structures, business and education. A business description of quantitative literacy may span a continuum of mathematical skills, from basic arithmetic skills to more complex skills of identifying "metrics for gathering data, and understand how to utilize data to take action to improve performance" (Rosen, et al., p. 45). From an education perspective the definition is more concrete, and often referred to as mathematical literacy. Nevertheless, the National Council of Education and the Disciplines provided us with the seven elements that define quantitative literacy. These are: (a) arithmetic, the use of simple calculations for numbers; (b) data, using data to draw inferences, understanding graphs and charts; (c) computers, to record data, create graphic displays of fitting line or curves to a data set; (d) modeling, the ability to understand linear, exponential, multivariate and simulation models; (e) statistics, to understand the importance of variability in a data set, recognizing the differences between correlation and causation, the difference between experiments and non-experiments, statistical significance and practical significance; (f) chance, to evaluate risks, understand the value of random samples and to understand that improbable coincidences are not uncommon; and (g) reasoning, to exercise caution in making generalizations, checking hypotheses and using logical thinking (Steen, 2001).

Statistical Literary/Reasoning. Wallman (1993) combines the ideas of mathematics and statistics alongside the importance of political dialogue to succinctly define statistical literacy. It is "the ability to understand and critically evaluate statistical results that permeate our daily lives—coupled with the ability to appreciate the contributions that statistical thinking can make in public and private, professional and personal decisions" (Wallman, p. 1).

Expanding on Wallman's definition of statistical literacy, Rumsey (2002) describes a basic element of it as the "chain of statistical information and the people who participate in it" (p. 33). There are people in society who are the data producers, those who are the data consumers and those who are the data communicators. Often the communicators incorrectly disseminate information from the producers to the consumers through their own translations, which are incorrectly interpreted (Rumsey). This miscommunication is rooted in society's lack of statistical literacy and occurs through misunderstanding about the sources of statistical data, and misgivings about the value of statistics in public and private spheres (Wallman, 1993).

To offset miscommunication and the expand on the definition of statistical literacy, individuals proficient in it need to understand the importance of asking the *worry questions* when interpreting statistical messages. These questions should include, "where the data came from, how reliable is it, whether the data is summarized correctly, the validity of conclusions, and the completeness of information" (Rumsey, 2002, p. 33). Further strengthening this definition, the Royal Statistical Society believes statistical literacy "involves the ability to critically evaluate the use of statistical data by others, in media and elsewhere. This refers to the use of official statistics, both in providing 'snapshots' of current situations and in showing important changes over time" (Goodall, 2005, p. 96).

Consider an operationalized definition of statistical literacy that provides much more details to explicate its meaning based on specific skills needed to become statistically literate. For Schield (1999), “statistical literacy is the ability to read and interpret data, and the ability to use statistics as evidence in arguments. Statistical literacy is a competency: the ability to think critically about statistics” (p. 1). It is a competency much like reading, because it involves comprehension and interpretation in order to make decisions using statistics as evidence. Further, he explains, to be statistically literate, an individual must be able to accomplish the following tasks:

1) be able to distinguish statements of association from statements of causation
2) be able to distinguish a sample statistic from a population parameter
3) be able to distinguish between the target population and the sampled population
4) be able to distinguish the quality of a test from the predictive power of a test
5) be able to interpret what a statistic means
6) be able to distinguish an observational study from an experiment
7) know the various sources of problems in interpreting a measurement or an association
8) be able to ask the following questions: Is this statistic true? Is this statistic representative? Is this association spurious? (Schield, pp. 2-6)

Another term used to define statistical literacy is statistical reasoning, which states it is “the way people with statistical ideas make sense of statistical information” (Garfield & Chance, 2000, p. 101). Interpretations often combine ideas about chance and data leading to inferences and interpreting statistical results. Important conceptual ideas, for example, distributions, center, spreads, association, and sampling are necessary to understand, in order to be able to reason in statistics. Further, Garfield (2003), in defining statistical reasoning, cautions that statistical reasoning is not mathematical reasoning. Mathematical reasoning is the understanding why assertions based on assumptions are true or about the assertions are about relationships among abstractions (Manaster, 2001).

Differences in Statistical Phraseologies. Within the discipline of statistics, mathematics is used to calculate various statistical equations that enable researchers to come to conclusions about phenomena. But as Steen (2001) notes, mathematical literacy is not the same as statistical literacy, because mathematical literacy stresses the traditional tools and vocabulary of mathematics to solve problems. By the same token, Pugalee (1999) purports that mathematical literacy embraces five processes which are different than statistical literacy. Students must value mathematics, become confident in one's ability to do math, become problem solvers, communicate mathematically, and reason mathematically. Or as Manaster (2001) explains, mathematics is the science of "numbers and their operations, interrelations, combinations, generalizations, and abstractions; space configurations and their structure, measurement, transformations and generalizations…" (pp. 67-68). Differently, statistical literacy includes an understanding of research methodology. Moore (1997) explains, "a student who emerges from a first statistics course without an appreciation of the distinction between observation and experiments and of the importance of randomized comparative experiments…has been cheated" (p. 127).

It is important to note, according to Steen (1990), there is a difference between numeracy and quantitative literacy. He asserts that quantitative literacy is "more of a habit of mind, an approach to problems that employ and enhance both statistics and mathematics, unlike statistics, which is primarily about uncertainty, numeracy is often about the logic of certainty" (Steen, p. 5). Conversely, numeracy is defined under the umbrella of quantitative literacy as it "requires inferences based on estimates and approximation, on incomplete or sometimes inaccurate data" (Manaster, 2001, p. 68). The term quantitative literacy also purports to allow "people a quantitative perspective for understanding the world" (Manaster, p. 68). Examples include charts of income distributions, effects of medical treatments, and differences between social programs, or policies, through descriptive statistics. Numbers in this sense help compare features of real-world situations and help us make decisions based on the numbers. Hence, sometimes the distinction between quantitative literacy and statistical literacy is small, as both terms blend meaning in the social sciences to include the understanding of the meaning and sense of data, as it is used in political campaigns, to inform government decisions and by businesses to inform marketing strategies (Manaster).

Despite these differences in definitions, numeracy is often defined by the different dimensions in which both mathematics and statistical ideas operate (Steen, 1990). And in some schools quantitative literacy is simply defined as an informal synonym for elementary statistics, because in the real world, most mathematical activity does not begin with formulas, but with data (Steen, 2003).

Similarities in Statistical Phraseologies. Aside from the differences in statistical phraseologies, there are many similarities in the definitions, as many terms are neither exclusive nor mutually inclusive. Numeracy, (Brown et al., 1998; Steen, 1990) adult numeracy, (Gal, 2002), quantitative (Rosen, et al., 2003; Steen, 2001), innumeracy (Cerrito, 1999; Paulos, 1989) and mathematical literacy (Rosen et al.; Steen); all encompass basic computation skills of mathematics and statistics.

Sometimes quantitative and mathematical literacy are defined through social structures such as business and education (Rosen, et al, 2003.; Steen, 2001) while others, adult numeracy, innumeracy, statistical literacy, statistical reasoning, include the necessity of understanding statistics to be a fully functioning individual in a civic society (Cerrito, 1999; Gal, 2002; Paulos, 1999; Rumsey, 2002; Schield, 1999; Steen, 1990; Wallman, 1993) in their definitions. This also includes the ability to interpret statistical information as explained by in Wallman's and Rumsey's definition of statistical literacy and reasoning.

Underlying the ability to understand statistics is the ability to understand research methods, which was included in the definition of statistical literacy and reasoning by Rumsey (2002), Schield (1999) Garfield and Chance (2002), Paulos (1999) and Cerrito (1999) in innumeracy, Rosen, (2003) et al., and Steen (2001) in quantitative and mathematical literacy. This is necessary in order to have an understanding of how political policy is created by data analysis or medical research for drug approval (Cerrito).

And importantly, as Garfield and Chance (2002), Rumsey (2002) and Schield (1999) noted in their definition of statistical literacy, it includes the ability to not only interpret statistics, but to argue with assumptions made by the data and to critically challenge the results of the data based on their individual knowledge and attitudes toward statistics.

Accordingly, these similarities can be summed up and classified as mathematical, statistical and literacy skills, context knowledge, critical skills, attitudes toward statistics and critical stance. From these conceptual definitions a more precise model has been developed by Gal (2004), which operationalizes these conceptual definitions. His model of statistical literacy will be discussed next.

A Model of Statistical Literacy

Because of all the various conceptual definitions that define statistical literacy in the literatures, a model to examine statistical literacy should contain the various elements that make up this multifaceted concept. Accordingly, Gal's (2004) Model of Statistical Literacy was chosen not only because it is the only model of statistical literacy, but because it encompasses all seven conceptual elements that describe statistical literacy in the literature. Each element is discussed in detail, but as Gal reminds us, the "elements in the proposed model should not be viewed as fixed and separate entities but as a context-dependent dynamic set of knowledge and dispositions that together enable statistically literate behavior" (p. 51).

Knowledge Elements of Statistical Literacy

Gal's model is composed of five knowledge elements; literacy skills, statistical knowledge, mathematical knowledge, context knowledge and critical questions. These are the elements that, according to Gal (2004), examine how individuals "interpret and critically evaluate statistical information and data-related arguments" (p. 49), which are encountered in different mediums in everyday life. Accordingly, these knowledge elements are part of statistical literacy and will be explained next.

Literacy Skills. The first knowledge element, literacy skills, pertains to the understanding of statistical messages either in written text (which may be long or short), or may involve a graph with a few words. Readers have to be able to understand certain terms related to statistics that are used by the message originators such as randomness, representative, percentage, average or reliable, and understand what they mean according to the context in which they are embedded. Many times**,** terms that are used by the message originators are not explained—for example, sampling error or margin of error that is commonly used in discussing poll results. Hence, in order to be statistically literate, one must be capable of general literacy first, as both are intertwined (Gal, 2004). Also individuals must be able to identify, interpret and use information given in lists, tables, indexes, schedules, charts and graphical displays. These displays often include explicit quantitative information, such as numbers or percentages. And these graphical displays can vary in degrees of complexity, as they can be a simple bar graph or pie chart or in a graph that combine multiple elements (Gal, 2004).

Statistical Knowledge. The second element, statistical knowledge, consists of five parts. These are (a) knowing why data are needed and how data can be produced, (b) familiarity with basic terms and ideas related to descriptive statistics, (c) familiarity with basic terms and ideas related to graphical and tabular display, (d) understanding the basic notions of probability, and (e) knowing how statistical conclusions or inferences are reached (Gal, 2004).

Knowing the origins of the data collection and why the data was produced allows an understanding of the logic behind the research design. Adults who have this knowledge can understand the research design used in order to refute or acknowledge a claim of causality that the data may be purporting (Moore, 1998). Research design methods include the (a) experimental method and the use of experimental and control groups, (b) pilot studies (c) the logic of sampling and the need to infer from samples to populations, and (d) the notions of representativeness (Cobb & Moore, 1997).

Additionally, the type of sampling the research used is also important; for example, did the research use convenience sampling, or probability sampling (Gal, 2004)?

Familiarity with basic terms and ideas related to descriptive statistics are also important to the statistical knowledge base. Descriptive statistics includes percentages (Parker & Leinhardt, 1995), and measures of central tendency, the mean (i.e., average as used by the media), mode and median. Gal (1995) argues that adults need to know that:

> Means and medians are simple ways to summarize a set of data and show its center; that means are affected by extreme values, more so than medians, and that measures of center can mislead when the distribution or shape of the data of sample from which they are calculated is not representative of the whole population. (p. 59)

Adults should also be familiar with graphical and tabular displays and their interpretations, and be able to detect when the relative length of the bars is not proportional to the actual data, whether given as percentages or as whole numbers, which can make the graphical or tabular displays misleading (Gal, 2004). And adults need to be aware that graphs can be intentionally misleading to show a specific trend or difference (Huff, 1954).

Another part of the knowledge base for statistics is the ability of adults to understand the basic norms of probability, because chance and random events are very common in the types of messages adults encounter. These messages can include estimates made by forecasters, genetic counselors and physicians. Adults should have an understanding about the various ways in which probability is communicated by the various methods, for example, percentages or verbal estimates. And adults should be able to understand and critically evaluate probabilistic claims, and recognize the source of probability estimates, because some can be based on data modeling or on subjective claims (Clemen & Gregory, 2000).

The final element of statistical knowledge is the ability for adults to understand how conclusions or inferences are reached, as most adults are data consumers and not producers. There are various types of research designs in which data is collected, and adults need to be aware of different errors that might be evident in the sampling methods, or in how the phenomenon was measured. Accordingly, specific types of errors can be controlled through a proper research design. One type of estimation that may require interpretation by an adult is the margin of error, because it is frequently used by the media (Gal, 2004).

Mathematical Knowledge. Mathematical knowledge is the computations that underpin how statistical analysis is computed. For example, adults need to understand how an arithmetic mean is computed, in order to understand how a mean may be influenced by extreme values in a data set, and may not be truly representative of the middle of that data set. The media often reports statistical information by percentages, and often how they report them is different than what adults have encountered in the classroom; for instance, some percentages are reported larger than 100%. Percentages have different mathematical meanings and also statistical uses. They may represent a number, an expression of a relationship, a statistic, a function, or an expression of likelihood. Also, percentages may represent complex relationships, such as conditional probabilities of an event, or may be linked to concepts such as 15% below average, or the 2% margin of error that is commonly used by the polls (Gal, 2002). It is important to note, there is an ongoing debate about the amount of mathematics adults need to know to understand more sophisticated concepts such as a statistically significant difference. To understand these concepts, a solid understanding of underlying statistical ideas such as the quantification of variance, repeated sampling, sampling distributions, curves, and logic of statistical inference are necessary (Cobb & Moore, 1997).

Context/World Knowledge Base. Context—how and where the data was collected—is important to properly interpret statistical messages. It is necessary for adults to place messages in a context and to tap into their world knowledge, which allows them to make sense of the various messages that statistics represent (Gal, 2004). In statistics, the data needs to be viewed as numbers within its context, as the context is the source of meaning and the basis for the interpretation of the results (Moore, 1990). Adults need to be familiar with data generation processes, such as the research methodology used, and the processes used to analyze the data. Context knowledge is "the main determinant of the reader's familiarity with sources for variation and error" (Gal, p. 64), because without this information it is difficult to imagine why group differences occur, or what alternative explanation may exist or how a study could be completed incorrectly. In the media this is a common problem, as ads shown by the media can easily mask or distort information to the reader. Many times reporters use the term *experiment* in such a way to enhance the validity of a study, when in fact the study may be non-experimental in nature (Gal, 2002).

Critical Skills. A critical evaluation of research results is a necessary skill because general media information are produced by various sources, such as politicians, manufacturers or advertisers and depending on their needs and goals, they might not present a balanced and objective report of their research findings. In fact, they may be intentionally biased and used to create media hype. For example, in 1992 a national magazine attempted to create a public image of a drug plague by reporting only some of data collected as part of a multiyear survey project (Orcutt & Turner, 1993). Gal (2002) recommends that adults need to ask the worry questions (i.e., questions that challenge how the data was analyzed and gathered) about statistical messages when interpreting any type of research. In asking and answering these questions, a critical evaluation of statistical information will lead to a more informed consumer of research.

Dispositional Aspects of Statistical Literacy

The dispositional elements in Gal's model of statistical literacy are beliefs and attitudes, and critical stance, and are related to individuals' ability to "discuss or communicate their reactions to such statistical information, such as their understanding of the meaning of the information, their opinions about the implications of this information, or their concern regarding the acceptability of given conclusions" (Gal, 2004, p. 49). These elements will be explained in detail next.

Critical Stance, Beliefs and Attitudes. Dispositional aspects of statistical literacy are critical stance, beliefs and attitudes. Taking a critical stance can be a covert or an overt process. It may be an internal process where adults may think about the meaning in a particular research report and raise some critical questions in their minds. Or it can be an overt process where a discussion of the research findings would take place with family members or with co-workers. Individuals' beliefs and attitudes about statistics are intertwined with their ability to take a critical stance. Accordingly, dispositional aspects—critical stance, beliefs and attitudes—are discussed separately, but in reality are interconnected (Gal, 2002). Taking a critical stance involves an adult's ability to have a questioning attitude toward quantitative messages without external cues. They should be able to invoke their list of worry questions when reading and interpreting results or conclusions from various types of research (Gal).

Underlying adults' ability to take a critical stance and their willingness to challenge research results is their beliefs and attitudes. There is a fine distinction between beliefs and attitudes; beliefs are individually held ideas or opinions, about oneself or about a social context (Wallman, 1993); they are stable and less resistant to change than attitudes (Gal, 2002). On the other hand, attitudes are "relatively stable intense feelings that develop through gradual internalization of repeated positive or negative emotional responses over time" (Gal, p. 69).

Beliefs, attitudes and critical stance mesh together in statistical literacy; for adults to maintain a critical stance, they need to develop "a belief in the legitimacy of critical action" (Gal, 2002, p. 70). They need to hold on to the idea that it is legitimate to be critical about statistical messages or arguments, no matter what sources report the data. And the worry questions should be raised even if they do not have access to all needed background details (Gal).

Strength and Weakness of the Model

Gal's (2004) Model of Statistical Literacy incorporates the numerous elements discussed in the literature that encompass statistical literacy. As previously stated, these elements are the knowledge elements, literacy skills, mathematical, statistical and context knowledge and critical questions. The dispositional elements are beliefs and attitudes, and critical stance. The ability to incorporate these conceptual elements into one model of statistical literacy is the strength of the model. Additionally, Gal's model allows the conceptually defined terms to be incorporated into operationally defined elements that can be used to examine adult college students' statistical literacy skills (see Chapter 3 for a complete discussion of creating the instruments to examine statistical literacy). However, the weakness of the model is that it has not been empirically tested, as there is no known instrument that encompasses all the elements of the model. In addition, the elements of the model are not mutually exclusive.

Accordingly, because achieving statistical literacy has become an important part of the education process, it has been carefully defined to encompass seven elements. However, statistical literacy has not always been considered important in the educational curricula. To understand why it has become important in education today, it is necessary to understand the early teachings of mathematics and statistics in America. This section will be discussed next.

A Brief History of the Teaching of Mathematics and Statistics

This section of the literature review elucidates how mathematics was taught from colonial times to the early 1990s, because statistics, albeit not mathematics**,** emerged from within it. The pedagogical nature of teaching mathematics during this time in America is examined through teaching resources and practices, as there is a severe paucity in the literature in these eras (i.e., 1700s, 1800s and 1900s). This is followed by a section on the emergence of the teaching of mathematical statistics in colleges in the 1800s, to the separation of mathematics and statistics as they became separate, but intertwined disciplines. The last section follows by examining the teaching of statistics in the 1900s to the end of this era, when important changes begin to emerge.

Colonial Era

Historically, in America before the 1790s, simple counting, adding, and subtracting in Arabic numerals below one hundred was commonly taught by parents to children, so they could pay taxes, sell eggs or measure lumber; essentially, arithmetic skills were related to the static colonial economy (Cohen, 2003; Smith & Ginsburg, 1934). In most colonies, higher arithmetic skills were limited to those who were going into the mercantile trades, because arithmetic was identified with commerce. Further, teaching at this level required children to be 11 or 12 years of age (Monroe, 1912; Cohen), but most children abandoned school before their 12th birthday; consequently, very few learned higher arithmetic skills. In addition, though some schools in this era, in particular those in the state of New England, supported literacy and numeracy, (i.e., reading, and basic adding and subtraction), most mercantile operations were predominately located in the South, which further discouraged those from the North from learning more arithmetic. Unlike New England, in the southern states education was sparse, comprising of a few private schools, and tutors educated the small number of the gentry's children (Cohen). Classes for these children were held in private homes, or in special buildings and private schools supported by tuition fees, which were numerous in many larger towns. These schools instructed boys who aspired to careers such as ship's officers, surveyors, clerks, and merchants. Some private schools had classes in the evening, which offered courses in mathematics or surveying (Kiefer, 1948). This left most of the population with little arithmetic skills, other than the ones needed to survive in their environment (Cohen).

Another important issue related to the learning of arithmetic was books.

Arithmetic books during the 1700s were scarce in colonial America, and most were imported from England (Cohen, 2003). The few that were printed in America focused mostly on elementary arithmetic such as *Arithmetick* written by James Hodder. His book was originally printed in London in 1661, but the 25th edition (Monroe, 1912) was reprinted in Boston in 1719. Another popular work used extensively in the early colonial schools, but never printed in America, was *Arithmetic,* written by Edward Cocker (Carpenter, 1963; Cremin, 1970). Some consider him to be the father of modern arithmetic, because his book went through 100 editions in England within a century (Carpenter).

Eventually, an arithmetic book was written and published in America. In 1729, *Arithmetic, Vulgar and Decimal: with the Application thereof, to a Variety of Cases in Trade, and Commerce* was published (Monroe, 1912; Nietz, 1961); however, anonymously. Its authorship was established by a newspaper advertisement in the *Boston News Letter* (Carpenter, 1963; Cremin, 1970; Smith & Ginsburg, 1938). There was no reason given to why the author, Isaac Greenwood, announced his book to the world in this manner, as he was a professor of mathematics and philosophy at Harvard University. However, it is known that there was no second edition, most likely due to his frequent lapses of sobriety (Carpenter).

More popular than Greenwood's book, and considered the most popular textbook in America, was Nicholas Pike's A *New and Complete System of Arithmetic.* It was supposedly "composed for the use of the citizens of the United States" (Nietz, 1961, p. 156), but interestingly, the tables used in it represented English money, and there were no problems in the book that involved the monetary system in the United States. And similar to the other texts published at this time, the problems were always preceded by a rule (Nietz). So in essence, education was not focused on educating its citizens on the necessary arithmetic skills that would be essential in their everyday lives, and what was taught was rote learning.

A distinctive quality of textbooks at this time was that they organized "knowledge of the arithmetic arts into a catechism-like set of rules that relied on memory rather than reasoning" (Cohen, 2003, p. 10). Further, as Carter (1827), explained, "the plan of all arithmetics [ibid]…has been to state the principle or rule to be taught in the most concise manner possible, and then arrange under it, examples of its application" (p. 30). Further, textbooks were disorganized in the way they presented material; hence, no section was dependent on a previous section to build upon previous arithmetic learning. This was the accepted pedagogy of the times as parents were satisfied by the education of their children. Carter explained:

> And he is hardly better prepared for the business of life, for he can neither remember the rule, nor the application of it. But the parent is satisfied because the child has been through the book and can repeat all the rules it contains; and moreover, he can flourish in the application of any rules to the examples, which are put under it, and which his instructor has probably led him through again and again. The instructor is satisfied because the parent is; and the pupil is doubly satisfied, on both accounts. (p. 31)

Further, textbooks were definitely confusing. Hodder's *Arithmetick* consisted of 216 pages and was fairly well printed (Carpenter, 1963), but contained numerous errors shown by revised editions. The 25th edition was revised by Hodder's successor, Henry

Mose, who stated, "there are above a thousand faults amended" (Smith & Ginsburg, 1938, p. 37). In Hodder's book, the *scratch* system of division was used which involved "the method of writing first in a column the beginning multiples of the divisor, the remainders being put in a vertical column without preserving the decimal order" (Carpenter, p. 29). Equally confusing was his definition of subtraction; "subtraction teacheth (sic) to take any lesser number out of a greater and to know what remains" (Sleight, 1943, p. 256).

A common type of arithmetic in this era was denominate arithmetic, taught for the acquisition for numeracy, which was described as pages and pages of texts that presented equivalencies in gallons and pints. Students struggled with denominations of volume and size, because these were specific to the item being measured; for example, apothecary ounces totaled 12 to the pound equivalencies, or avoirdupois (i.e., a system of weights in the United States used for goods other than gems, precious metals or drugs) ounces that equaled 16 per pound. Arithmetic with this mercantile focus was often taught with one textbook for the teacher, and students wrote the explanations for specific problems in their copybooks. Instruction did not provide explanations for problems; there were minimal examples as they treated "each problem as a universe unto itself" (Cohen, 2003 p. 10).

Accordingly, students memorized problem after problem, instead of learning abstract rules of mathematics capable of generalization to many types of problems, as the dominant pedagogy of this era was to "state a rule, give an example, and set exercises to be followed or in geometry, give proofs to be memorized" (Jones & Coxford, 1970, p. 17). Interestingly, different fields of study, such as navigation, surveying or gunnery, used this same style of problem-based memory learning for arithmetic (Jones & Coxford; Cohen, 2001). And, the rote learning commercial arithmetic of the 1800s "was so completely context-specific that it probably retarded the development of quantitative literacy" (Cohen, p. 26).

Noteworthy, in some colonial schools students learned arithmetic without a book. The *masters of common schools* (i.e., teachers) taught students with a *sumbook,* which was made during the teachers' own school days and used in place of a textbook (Keifer, 1948; Monroe, 1912). After the master told the pupil the rule he would dictate sums (i.e., problems) to the students, and they worked on solving them on scraps of paper. If the student had his work approved by the master, he would carefully copy the problem, the solution, and the rule into his *Cyphering Book* (i.e., a blank book made of a quire of paper folded and sewed together; today, similar to a student's notebook).

It is also important to note, the status of grammar school teachers in this era was not high; "it was about that of skilled laborers" (Jones & Coxford, 1970, p. 21). And the teaching of arithmetic was not required by many colonial schools, but masters who had a sum book were considered *more learned and good at figures* (i.e., very intelligent) and acquired teaching positions easier, because they were held in higher esteem than their less talented colleagues (Keifer, 1948). This is evident in the latter part of the 1800s, when a master did not have a rule to figure out a simple computation on a teachers' exam in, for example Indiana. Monroe (1912) quotes Hobbs from an interview which was published in *Early School Days:*

> In the late 1800s as the only question they asked me at my first examination was what is the product of 25 cents by 25 cents...we only had Pike's arithmetic (i.e., a

textbook with rules) which gave the sums and the rules…How could I tell the product of 25 cents by 25 cents when such a problem could not be found in the book? The examiner thought it was 6 ¼ cents but was not sure. I thought just as he did, but this looked too small to both of us. We discussed its merits for an hour or more, when he decided that he was sure I was qualified to teach school, and a first-class certificate was given me. (p. 23)

Division between the genders was also apparent in the colonial school era, especially in the teaching of arithmetic. Girls were educated according to their prescribed roles in society, while males were educated according to the purpose of education and intelligence (Keifer, 1948). Girls generally did not learn much arithmetic (Smith & Ginsburg, 1934) beyond the skills necessary to manage a house or purchase necessities. Rather, a girl's education should consist of knowledge of English, a little French, albeit not to perfection, and music, drawing, and geography—nothing that would involve deep thought (Nelson, 1763). In fact, Nelson believed that a woman "should not aim at more deep or learned studies, which would only make her affected or pedantic; make her a pain to herself and disgustful to all who converse with her, particularly her own sex" (p. 316).

Unlike girls, boys, in general were educated in the practical rules of arithmetic, while higher arithmetic teaching focused on those who were thought to be geniuses. Most boys received a *Merchant Account Book* (i.e., a book that recorded sales) in which they would become familiar with the simplest entries of debtor and creditor (Keifer, 1948). Most occupations for males required simple computations of numbers, because they would grow up to be farmers or artisans and they would only need to understand how to add and subtract numbers; rarely did they need to multiply or divide. Knowledge of simple computations included the measures that were common to colonial life such as halves, fourths, and eighths (Smith &Ginsburg, 1934).

After the War of 1812, there was a rapid expansion of commerce in the United States, and a takeoff period of early capitalism, leading many citizens to a market economy characterized by banking, economics, wage labor and urbanization. This led to the support of education and the development of public education in the early 1820s. With more schools and more teachers, an inductive approach to arithmetic instruction came into play, spearheaded by Warren Colburn. His method was to train the minds of children to reason with numbers, not to do problems with rote formulas (Cohen, 2001). His method of teaching was evident in his first textbook, *First Lessons in Arithmetic,* which was followed by *Intellectual Arithmetic* and *Sequel to the First Lessons.* These books were noted for "rescuing school arithmetic from ciphering, an old rut bound method" (Carpenter, 1963, p.141). Colburn's teaching methods were based on the teaching of Pestalozzi, who was a Swiss educational reformist in the early 1800s. He argued that children should learn though activities and through things, not abstraction (Smith, 1997). Hence, children should not learn mathematics through abstraction, but first learn the idea of numbers through observing sensible objects. For example, Colburn would teach the idea of numbers wholly as an observing process, not as a ciphering process, by asking children how many thumbs or how many hands they had (Carpenter). However, Colburn's method that children could create arithmetic in their heads was severely criticized by others, leaving his method short-lived. Soon his method was replaced by a deductive approach where axiom and definitions were to be memorized and applied, and his inductive method was remembered as a failure (Cohen).

The Emergence of Statistics and Higher Education: The 1800s

The emergence of the discipline of statistics in higher education was not evident until the 1800s, and similar to the early teachings of arithmetic in colonial schools, not much importance was placed on the learning of mathematics, or later statistics. However, when it was taught, the instructor had more credentials, and most likely had studied in England before teaching in the colonies (Jones & Coxford, 1970). But nevertheless, higher education at first did not place an emphasis on mathematics, and only later was an emphasis was placed on statistics, albeit at first the emphasis was political. This will be discussed next.

Early American colleges in colonial times were often small and mirrored the universities in Great Britain. One that reflected Great Britain and became the intellectual center of the Puritan movement in the American colonies was Cambridge University. Here the curriculum of studies "emerged from the medieval tradition *trivium*, (i.e., grammar, rhetoric, logic) and *quadrivium* (i.e., arithmetic, geometry, astronomy, music) and from three philosophies, natural, moral, mental" (Kraus, 1961, p. 64). However, interest in the teaching of mathematics dwindled, and more emphasis was placed on a student learning rhetoric, logic, and philosophy (Kraus).

By the mid 1760s, the purpose of a college education was to train leaders of the state through a liberal education. And it was understood, in order to be a leader of state it would be necessary that all should be taught to read the scriptures and understand the highest branches of literature. Those who completed this type of education would be qualified to be the future judges or senators (Robson, 1983). By 1795, the subjects an individual should learn to become judges or senators focused more on the subjects of

English, history and the natural sciences. But more importantly, an emphasis was placed on moral and political philosophy through the works of Hume and Millot, which suggested an emphasis to teach both piety and rationality, both tenets of a liberal philosophy of education (Robson). The teaching of mathematics in higher education was rare at this time, and when it was taught, it was no more advanced than a contemporary 7th grade curriculum. Because of the paucity of instruction in mathematics at the college level, private classes in mathematics were often advertised in the newspaper (Smith & Ginsburg, 1934).

In that the purpose of a college education was at this time (e.g., 1800s) to train future preachers and politicians, many professors of higher education were preachers, or doctors of Divinity, or politicians, not mathematicians. Preachers would often teach at universities in order to teach the tenets of their faith, and worked for very small salaries and would teach other courses, including mathematics, for which they were ill prepared. Often they had long and numerous titles, for example, F. A. P. (sic) was a professor of mathematics, natural philosophy, astronomy, and chemistry at Alabama University, then simply a professor of mathematics and natural philosophy at the University of Mississippi (Smith & Ginsburg, 1934).

Because of the paucity of research in the teaching of mathematics, no articles explaining how it was taught in the colleges was published until 1822. And similar to the methods of teaching mathematics at the elementary level, higher education used the synthetic approach. In this approach, a mathematical rule to be taught is stated, and then under it were examples of its application (Carter, 1827; Mathematics, 1822). Most likely this method of teaching had been passed down from great ancient scientists and mathematicians, for example Sir Isaac Newton, who "delivered his splendid discoveries by the synthetic method and that authority has influenced the greatest part of mathematicians who have written in his native language" (Mathematics, p. 315). This type of teaching was evident in the early history of Cambridge College (Mathematics), whereas the student would commit a mathematical rule to memory, and solve a question related to it. The student had never been called upon to "exercise any discrimination, judgment or reasoning" (Carter, p. 30).

Related to the subject of mathematics, is statistics, as it requires the use of mathematical computations. The first institution to offer statistics as a course was the University of Virginia, in 1845. It was offered by the Department of Moral and Philosophy and interestingly, it was taught by two professors, William H. McGuffey and Professor Tucker, both of whom had the title of Reverend. The only other institution of higher learning to offer statistics as a course at this time was the University of Louisiana. It was to be taught by James D. B. De Bow, and funded by Manuel White, once a poor immigrant, who had recently become wealthy and a New Orleans merchant. He had secured an endowment for a Chair of Commerce, Public Economy and Statistics, and had De Bow appointed; however, the course did not attract any students, and later the college closed its doors (Fitzpatrick, 1955).

Despite the lack of interest in college students to learn statistics, census data started to reach the American public through publication of descriptive statistics. Data had been collected in earlier years, but it was not shown to the public until the 1830s, because it was riddled with errors. Data published in statistical almanacs was purported to be, according to Joseph Worcester (1930), "an account of whatever influences the condition of the inhabitants or the operations of government on the welfare of men in promoting the needs of social beings, and the best interest of communities" (as cited by Cohen, 2001). Other statistical information began to flow to civilians as newspaper reporters started to publish data on various topics, for example the total number of people who traveled on stagecoaches and railways. Or some groups published descriptive statistics in order to gain attention and legitimacy for political or social gains. For example, a group of moral reformers published the number of prostitutes as a way to give an analysis to the problem with an aura of scientific results (Cohen), to further promote their cause.

From descriptive statistics, new ideas about statistics grew from professionally trained statisticians, who pushed for an increasingly sophisticated mathematical methodology. This allowed the Federal Census to be used as a social indicator, which would be useful not only to legislators, but to businesses as well. However, this growing sophistication of data was not matched by a corresponding improvement in quantitative literacy by the general public—most people could not comprehend the data, despite major attempts to reform the mathematics curriculum to enable people to understand the new data reports (Cohen, 2001).

Most people could not comprehend the publication of data in the mid-1800s because at this time, mathematics had been and was still poorly taught in the elementary schools, with most focusing on only basic arithmetic skills (e.g., addition, subtraction, multiplication and division). And at this time, many citizens opposed the idea of a public education. They felt that it would make boys lazy and lead to dissatisfaction with farm life (Smith & Ginsburg, 1934). In fact, colleges did not include mathematics or statistics in their early curricula at Harvard. To earn a four-year degree only basic arithmetic, along with geometry in a student's second year**,** was required. Later, mathematics was included into the course curriculum, when John Ward's *Young Mathematician's Guide* became a standard college text at Harvard and Yale. This text included arithmetic, algebra and geometry (Kraus, 1961). Perhaps the focus on arithmetic was due to the fact that students who were to enter higher learning in this period lacked the necessary training in arithmetic, and consequently had to be taught in college, before they could engage in the study of mathematics.

It was not until 1873 that Yale College offered a course in statistics. The focus of the course was public finance and statistics of industry, and was taught by A.M. Walker, who was a Professor of Political Economy and History. When he left the school to take a position with the Federal Census Bureau, statistics courses were not offered at the college until 1887. A handful of other colleges in the 1880s did start to offer courses in statistics, and by the 1890s, 16 other colleges did. Some described their courses in statistics as statistics of population, with these courses offering instruction on populations, religion, education, births, deaths and marriages. Others listed them under Moral Statistics and examined suicide, vice, crime, and effects of penalties (Fitzpatrick, 1955).

Similar to Yale, the University of Michigan offered a course in statistics in 1887, titled the *Principles of the Science of Statistics,* taught by Henry C. Adams, a professor of political economy and finance. Like Walker, he also held another position—his was with the Interstate Commerce Commission. The course was dropped in 1888 and not offered again until 1891. And similar to both Yale and the University of Michigan, the University of Indiana offered an *Introductory Course in Statistics* in 1890, which was dropped in

1891 because of the departure of Professor Jeremiah Jenks. When it was offered again in 1893, it was described as a course in economics and statistics—an introduction to the science of the political economy. Likewise, other colleges in the late 1800s offered courses in statistics, with titles of these courses indicating topics of economics or the political economy (Fitzpatrick, 1955).

Courses in the political economy were the first courses that taught statistics, and albeit there were few articles published describing how statistics was taught during this era, one gives us some details to the construction of the course, and a little bit of information about its pedagogy. For example, at the Massachusetts Institute of Technology, two courses were offered in statistics around 1888, which included graphical methods (i.e., the ability to read charts and graphs) to illustrate statistics, and were to be taken in connection with a course in United States finance. An advanced course was also offered in the statistics of sociology, which included all those facts of life "which admit of mathematical determination to express the 'average man'" (Wright, 1888, pp. 13-14). This gives a description of the course, but not so much of how it was taught. Similar was the course in statistics offered by Columbia College, in which the teaching of statistics was described as a two-hour conceptual lecture on the history of statistics, statistical methods and the connection of statistics with political and social science. In this sense, teaching was most likely rote learning of facts, but included teaching students how to use statistics to interpret and explain social phenomena. For example, statistics showed that an alarming amount of illiteracy was present in the state of Massachusetts, and "statistical inquiry shows that by far the greater number of these literates is of foreign birth, so that the fault is not with the public school system, but the evil is due to a temporary cause, namely, immigration" (Wright, p. 16).

The type of statistics courses offered reflected the zeitgeist. As previously mentioned, the Federal Census was collecting data to be used as social indicators, useful to legislators and businesses. Most professors who taught statistics at this time had a background in public finance, economics, or population statistics. And the history seems to indicate that professors who taught statistics at the university often left for other jobs with the government. This could be due to the lower salaries that the universities offered during this era. At this time, no professors who taught statistics were designated to be a professor of statistics, but this started to change (Fitzpatrick, 1955). So in a sense, the teaching of statistics was not important within the college environment.

Important changes occurred at the end of the 1800s. The teaching of statistics at the college level was mostly focused on undergraduate students, while graduate studies in statistics were rare. However, in 1897, the University of Minnesota began a graduate course in statistics, taught by William Folwell, who was a Professor of Political Science and Librarian of the university. Soon others followed, and one university, the Catholic University of America, in 1895 created the graduate school of Social Sciences in the Department of Economics. It offered three graduate level courses in statistics, and like the University of Minnesota, it was taught by individuals who were not professors. These courses were taught by the Honorable Carroll D Wright, LL.D**.** (i.e., a doctorate level academic degree in law), United States Commissioner of Labor. Previously, the university lost professors because they went to work for the government, but in this case the university recruited instructors from the government to teach courses in statistics (Fitzpatrick, 1955).

No professor who taught statistics, even at the college level, had the distinction of being a Professor of Statistics until Elgin Gould appeared and taught statistics courses at Johns Hopkins University in 1894. He was the first American professor to achieve this title. This is not surprising, due to the paucity of statistics classes offered in universities in this era, as some would have to travel to Germany to achieve a doctorate degree in statistics and later immigrate to the United States (Fitzpatrick, 1955; Wright, 1888).

Mathematics and Statistics Education: The 1900s

Significant changes in mathematics education occurred during the 1910s and 1920s, when the importance of teaching mathematics diminished due to the opinions of the professional mathematics educators in universities. Normally, schools would have courses in higher mathematics that included algebra, geometry and trigonometry to teach to students who were primarily white and middle to upper class. However, when there was an influx of immigrants, and the school's population started to include children of immigrants, emancipated slaves and industrial workers, the teaching of higher mathematics was questioned. Mathematics educators felt these students would be unable to comprehend courses in higher mathematics; therefore, teaching them was seen to be impossible (Cohen, 2003).

Perhaps this was further promoted by Edward Thorndike, who as one the leading educational and behaviorist theorists, argued that "mathematics did not encourage mental discipline" (Cohen, 2003, p. 15), which was quite the opposite in the early 19th century. At this time, the purpose of education changed; it was decided instruction should be geared toward teaching abilities that would likely help job placement. So widely believed was this ideology that some middle and high schools withdrew algebra and geometry as a requirement for graduation. Additionally, it was thought this type of curriculum did not have any practical value for workers' children, and it should be replaced with a two-year course that would teach mathematics that would correspond to the type of job they would be employed in as adults (Cohen).

This had a disastrous effect on students who pursued a college education. As Baten (1950) reported, "college freshmen do not know how to place the decimal point in multiplication and division problems and do not know how to add, multiply, and divide simple fractions" (p. 24). This left colleges to offer courses in mathematics without college credit in an attempt to bring students' knowledge of mathematics to the college standard. However, because students entered college with a deficit of mathematical knowledge, courses in college algebra became a mere reflection of what students should have learned in high school. Perhaps, because of the lack of interest in teaching mathematics at this time, there is no literature on the pedagogical nature of mathematics.

Likewise, there was no literature on the pedagogical aspects of teaching statistics on the college campus, except one article describing who would be qualified to teach statistics. For example, Hotelling (1940) suggested only those who have "made comprehensive studies of the mathematical theory of statistics and have been in active contact with applications in one or more fields" (p. 472) are qualified to teach it. What is evident is that courses in statistics at this time were not taught independent of mathematics. Statistics was taught in the discipline of mathematics, because the underlying concepts of statistics are mathematical. However, these courses did not include the elements of research methodology, which are an important part of understanding the outcomes of statistical inference. These courses, known as mathematical statistics, included concepts of algebra, geometry and probability, with probability being the most popular course taught within it. Its importance to the understanding of statistics was discussed numerous times with examples of mathematical formulas (Bailey, 1941; Baten, 1932, 1934; Camp, 1932; Copeland, 1937; Jackson, 1917), which does suggest statistics was taught with only precise mathematical formulas, and as previously stated, without research methodology.

More changes in mathematics education occurred in the 1950s to 1960s, spurred by the Cold War competition between the United States and the Soviet Union over scientific brain power. A new way of teaching mathematics called the *new math* attempted to introduce set theory and discovery methods into the elementary curriculum. However, this attempt failed, as teachers were not excited about the new ways and not ready to abandon their old ways of instruction. And the new method tended to the abstract, resulting in not promoting the kind of quantitative literacy related to political or civil life (Cohen, 2003). It was at this time that a preliminary discussion began on the teaching of statistics to undergraduates and graduate students at the university level. This discussion was started by the National Research Council (i.e., part of the United States National Academy of Science) with its publication, *Personnel and Training Problems Created by the Recent Growth of Applied Statistics in the United States.* It focused on the growing need for workers to have an understanding of statistics, and suggested that college students in multiple disciplines should learn statistics because it is used by the government, businesses, and in everyday life (Dutka & Fafka, 1950; Hotelling, Bartkly, Deming, Friedman, & Hoel, 1948). This renewed interest in the training of students in statistics is a result of statistical control as quality control, in research and development in business, research in the biological and psychological sciences, collection and analysis of government statistics, and market research (Wilks, 1947). Hotelling, et al., also made a point that students will be future consumers of statistics; however, future consumers were defined as students who would have careers as business executives, government administrators, research workers and teachers. There was no mention of how statistics would be necessary in everyday life and promote a civil society. And some recommendations for what to teach all college students in a statistics course included the fundamental logic and philosophy of statistics with no mathematics. A different approach in learning statistics was suggested for those who will be research workers and teachers of statistical methods—these students would be required to take advanced mathematics. And for adults in the workforce, there were more recommendations. Because previously statistics had been taught with inadequate methods, adults who worked in the research profession needed to be retrained. Suggestions were made for implementing programs in the evening for these adults through professional statistical organizations (Hotelling, et al.).

Included in the discussion initiated by the National Research Council were some problems in the teaching of statistics at the university level. For example, courses on probability theory were left out of many mathematical curricula. Probability, as previously discussed, is the backbone of statistics courses; without it, statistics courses lose the meaning of the data. As Hotelling et al., (1948) concluded, "the whole foundation of descriptive statistical methods, of inductive inference, and of the design of experiments, rests upon probability theory" (p. 105). Another problem was that many statistics courses were taught department-wide without any collaboration among departments, and there was a wide variation in the selection of the topics according to the abilities of individual instructors. This resulted in students being ill-prepared for the workforce, whether it was government service, or business statistics positions (Hotelling et al.). Many business students were taught inappropriate content in statistics courses that was not applicable to the real world. This resulted from the educational curriculum of the 1930s that remained stagnant in the universities, which required students to take courses of fragmented practical business curricula. These were “offered at the expense of broad educational experience in the arts and sciences as the basic tool courses” (Lee, 1960, p. 16), much a part of the liberal philosophy of education. It is important to note, during this time, there were apparently no pedagogical concerns on how mathematical statistics was taught, as discussions included who should be taught, and what should be taught within the context of employment after college.

During the 1970s, more discussion emerged on mathematical education, but this time it was focused toward mathematics for undergraduates in the social sciences. Social sciences at this time included the disciplines of anthropology, political science, psychology and sociology. Although many schools in the social sciences use mathematics, only a small number of them required college-level mathematics, resulting in a small number of social scientists with strong mathematical skills. In fact, many students selected social sciences as their major because they disliked mathematics and understood that sometimes mathematics courses are minimally required or not required at all. And those that do require statistics, do not mandate a prerequisite in particular mathematics courses (Darcy, 1971). Darcy believes this is due to the diversity of educational objectives of a particular discipline within the social sciences and the fact that social science instructors are not trained in mathematical statistics.

Because of the lack of a prerequisite in mathematics courses before taking a statistics course, these courses became known as non-calculus service courses in statistics, which focused less on the mathematics that formed the theoretical foundation, and more on statistical concepts. In fact, these courses were described as more of an *artsy* type, where students would “learn about the subject of statistics and statistical ideas, but would not attain proficiency in the technical aspects of the subject” (Federer, 1978, p. 119). But even though the mathematics requirements were lifted, many students entered the statistics classroom unprepared, unmotivated and left without an understanding of basic statistical concepts (Carlson, 1978). The problem as explained by Carlson;

> Is the failure to confront the pedagogical problems of statistics. Pedagogy, or the psychology of learning, needs to be taken seriously, and especially in mathematics where the logical flow of an idea is clearly different from the path of learning it. (p. 140)

Further, in 1989, the debate of integrating statistics courses into a curriculum focused on the liberal arts curriculum. A liberal arts education is different than other types of education by its concern for the general over the specific, and long-term versus short-term perspective.

> It takes the view that the undergraduate degree prepares the student for a lifetime of learning. Rather than training a student for some specific career or occupation, a liberal education emphasized methods of inquiry that are appropriate in whatever future career the graduate undertakes. (Moore & Roberts, 1989, p. 80)

Accordingly, the purpose of a liberal arts education is to prepare a student to become a functioning and contributing citizen in society. It encourages students to try a variety of subjects in order to expand their view of the world, rather than the more functional aspects of a vocational degree. And a liberal arts college is a place where statistics could easily be incorporated into the curriculum, because the focus of the college places more emphasis on teaching than on research. Here, class sizes are usually kept small, which can help foster a data-driven statistics education and promote citizenry though the understanding of statistics that are used in everyday life (Moore & Roberts, 1989).

Also, in the early 1990s, ideas for improving the teaching of mathematics/ statistics education had begun, with some purporting to re-examine the curriculum for statistics majors (Gaudard & Hahn, 1991) with the idea that statistics should be "taught with statistical rigor—data driven explorations of the discipline of statistics as opposed to dry presentations of formulas (Moore & Witmer, 1991, p. 433). Likewise is Hogg's (1991) recommendation that students should gather and work with real data and the beginning focus in statistics courses should be less on mathematics and more on careful thinking.

Summary

Historically, in the colonial era, the teaching of arithmetic was riddled with problems. Textbooks were often unavailable, and those that were available were published in England and imported to the colonies. When they were finally published in the colonies, they still had application for the British economy, instead of the colonies. Textbooks had arithmetic rules that were oriented toward rote learning, leaving students unable to apply the rules of arithmetic to other similar problems, often had numerous errors in them, and were difficult to understand; hence, memorization and rote learning were the pedagogical methods of this era. Teachers were poorly trained or never received training, and taught by the same methods in which they had learned arithmetic, and when Colburn tried to change the pedagogy of arithmetic to a more hands-on approach, his efforts were defeated.

Government census data flowed into mainstream America through the media in the mid-1800s, albeit most Americans were unable to understand it, as most only had basic arithmetic skills. Accordingly, data could be used for political or social gain with little questioning of the validity of the data. Also, colleges in this era still taught arithmetic in college, but most focused on the training of individuals to become leaders of the state, and believed that studying moral and political philosophy could achieve this goal. The teaching of mathematics at this time was mostly done by private tutors; hence, only those who could afford the tutoring had an opportunity to study mathematics. When mathematics was finally offered at the college level, it was taught not by mathematicians, but rather by preachers, who gained pseudo titles due to the subjects they taught. And when the teaching of statistics was finally offered by colleges around the 1870s there were only a few individuals who could teach it, and those who did were professional statisticians who often worked for the government and taught in the colleges as their second job, because of the low salaries offered by the universities. In a sense, the lower salaries de-emphasized the importance of statistics as a discipline.

In the United States during the 1920s, due to the influx of immigrants, who were of the lower class and thought to be inferior in intelligence, the importance of teaching mathematics diminished in universities. And within a few years the purpose of an education began to become focused toward vocational pursuits, instead of education for the purpose of producing preachers or leaders of the state. The Cold War brought the topic of mathematics education to the forefront again, but this time it focused on secondary education through the teaching of the new mathematics. This attempt toward reforming mathematics education failed, as it did not promote the kind of quantitative literacy necessary for a civil society. Because of the increase in using statistics in businesses during the 1950s, industries started to demand employees who had knowledge of statistics, and they looked to the universities to train them. The purpose in teaching statistics at this point was purely business in manner, as missing from this discussion was the important role statistics plays in a civil society. And by the 1970s introductory statistics courses became part of the curriculum for numerous disciplines in universities, but there were issues, especially in the discipline of social sciences. In many universities, the social sciences required little or no mathematical courses and students often pursued a social science degree because of this. However, many of the social science disciplines required an introductory course in statistics, but without a prerequisite of probability theory.

During the late 1980s and early 1990s some suggested that statistics courses should be part of a liberal arts education. Statistics courses began to be seen as an important part of the educational curriculum in colleges. After all, the purpose of a liberal arts education is to prepare students to become functioning and contributing citizens in society, and statistics is essential in this type of education.

Efforts to reform the mathematics curriculum had failed in the past; however, the new movement, the Democratization of Mathematics*,* spurred by the NCTM, started to help improve mathematics education in the United States through its curriculum reform in K-12, and undergraduate college education. Some argue that statistics is a branch of mathematics, while others say it is a distinct discipline that stands alone; nevertheless, it uses computations from mathematics, and is “a powerful tool of political and civic functioning” (Schaeffer, 2003, p. 147). Historically, there was no support by social institutions to change the way mathematics was taught or to bring statistics into the school’s curriculum. But now, because we live in an age of fast-growing technology and data analysis, support has been growing from industry leaders who need statistically literate workers. This support from outside academia is part of the reason that this movement is gaining momentum, as businesses and industries need individuals who are statistically and/or quantitatively literate. It is easily seen that “democratization is driven in part by the quantification of society” (Moore, 1997, p. 124). And it has moved mathematics and statistics studies “away from the esoteric toward the immediately useful” (Moore, p. 124). This fundamental change in the teaching and learning of statistics can be seen as the abandonment of the information transfer model (i.e., rote learning) to a constructivist view of learning, as more importance on learning mathematics and statistics has occurred over the last decade.

Chapter Summary

To understand the current movement to proliferate the teaching of statistics in higher education, three main topics emerged from the literature regarding the teaching of statistics in higher education—populations and purposes, methodological issues, and the teaching of statistics. And interestingly, no literature that specifically examines students’ statistical literacy was uncovered. A summary of the literature follows.

First, it is evident that much of the research in the last decade on statistics education has focused on students in the social sciences, albeit the amount of research completed is disappointing; it fails to include other majors within the social sciences, for example education, and it has either excluded non-traditional students or fails to identify them. Second, some empirical research mixed research methods courses with statistics courses in their analysis; these should be analyzed separately, because the classes focus on similar, but different concepts. Third, albeit there were a variety of purposes in examining statistics classrooms, none of the research has examined statistical literacy, which should be the outcome of successfully completing a statistics course.

Fourth, albeit comparison groups in a research design have their place in empirical research, more research on statistical education needs to examine students' gains in statistical literacy through quasi-experiments using pre- to post-test research designs. Fifth, the methods section of the research paper needs to be clearly written and be inclusive of important demographic variables that could affect the research results, especially age, gender and type college student. And sixth, a clear purpose statement should be stated in order to guide the reader through the study.

And importantly, as the literature shows, statistics instructors used both traditional and non-traditional teaching methodologies in their classroom and measured students' achievement through grades or final projects. However, many courses were evaluated by students' evaluations that measured their opinions to whether or not they liked or disliked their statistics courses. While these studies are important and measured students' learning, or students' opinions of their statistics classroom, there is no research that specifically examines students' statistical literacy at the completion of their statistics course. And missing from the literature review is an examination of students' learning in research methods courses, either with or without a prior statistics course.

Historically, efforts to change the way mathematics was taught has failed. But due to renewed interest in the teaching of mathematics and statistics, spurred by the NCTM, curriculum reform has begun. Some argue that statistics is a branch of mathematics, while others say it is a distinct discipline that stands alone; nevertheless, it uses computations from mathematics, and is "a powerful tool of political and civic functioning" (Schaeffer, 2003, p. 147). This is important as we are constantly surrounded by statistics that are presented to add credibility and marketability to products from drugs to automobiles, to promote political agendas, and to set standards for health and safety issues, thereby affecting the personal welfare of its citizens. Statistical literacy is crucial to our "quality of life and for our collective well-being" (Steen, 2004, p. 27). Succinctly, "statistical literacy is the study of statistics used in everyday life. Statistical literacy helps citizens in a democracy read and interpret numbers in the news to make intelligent decisions" (Statistical Literacy, 2007, n.p.). Accordingly, being statistically literate is similar to having statistical knowledge, but it is the ability to apply this knowledge to the real world that is different, and the most important outcome of learning in a statistics class. This has never been examined in the research literature.

It is evident from the numerous statistical phraseologies discussed in this literature review, that statistical literacy is a multifaceted concept. And because of this, a model that embraces these phraseologies is necessary to examine statistical literacy. Gal's (2004) Model of Statistical Literacy embraces these various statistical phraseologies in the knowledge and dispositional elements. This is the model that was used for this research on statistical literacy, and is discussed in more detail, along with the research methods, in chapter 3.

CHAPTER 3

METHODOLOGY

Introduction

The purpose of this research was to measure statistical literacy in adult learners before and after they have completed a statistics course, a research methods class with no prior statistics, and a research methods class with prior statistics. To measure their statistical literacy skills, a quantitative research design was used. Accordingly, to lay the foundation for this chapter, it begins by elucidating the early beginnings of quantitative research. From here, to explain what quantitative research is, a discussion is provided on the classic experimental design—the foundation of quantitative research, to a more specific discussion on the quasi-experimental design, which is used in this study. Included in this section is (a) a discussion of internal and external validity issues that pertain to this design, (b) information on the background of the researcher, and (c) the methods section, which provides detailed information concerning participant selection, instrumentation, and the procedures of the research.

The Foundation of Quantitative Research

The foundation of quantitative research is rooted in the philosophies of positivism and logical positivism, which emerged during the scientific revolution. This made extensive changes in the way knowledge was acquired by societies. No longer was knowledge to be credited to the metaphysical, which was subjective—the creation of knowledge was to occur from observable phenomena.

Positivism

Developed through the works of Auguste Comte and Henri Saint-Simon (Bryant, 1985; Hart, 1964), positivism is a philosophical doctrine "that recognizes only natural phenomena or facts that are objectively observable…and not debatable" (Schultz & Schultz, 2000, pp. 39-40) can constitute knowledge. Succinctly, the philosophy of positivism makes five major claims:

> 1) There is a reality out there that has an independent existence outside human consciousness.
>
> 2) Human beings can accumulate knowledge of this reality through observation and measurement, provided that they always proceed from observation to theory, and not vice versa.
>
> 3) Systematic data collection reveals that reality works in certain ways, and that these ways can be analyzed.
>
> 4) The purpose of theory is to produce empirically testable causal explanations that build knowledge about why things, including people, behave as they do.
>
> 5) The social world can be investigated in the same way as the natural world. (James, 2005, p. 3)

From these early tenets of positivism, the building blocks of logical positivism emerged, and it is from this philosophy we can see the emergence of quantitative methodology take shape, through the use of mathematics and statistics.

Logical Positivism

Emerging in the early part of the 1900s, logical positivism soon became the dominant philosophical perspective on science. It originated in Austria, by a group of theorists who became known as the Vienna Circle, with most of the founders being physicists and mathematicians (Bechtel, 1988). Included in logical positivism were more philosophical topics that included philosophy of language, symbolic logic, philosophy of science, and philosophy of mathematics, whereas in positivism, only sociology was inclusive. Further, logical positivism focused on the verifiability of meaning and logical analysis. Verification to logical positivists means that “a statement is meaningless if verification is not possible or the criteria for verification are not clear” (Ho Yu, 2006, p. 28), and logical analysis adds an emphasis of language, as complex phenomena could be expressed in terms of mathematics, and mathematics could be further reduced to logic (Russell, 1963). And within these ideas are the tenets of logical positivism.

From the Vienna Circle, four tenets of logical positivism related to quantitative methods emerged. First, “there is only empirical knowledge” (Ho Yu, 2006, p. 27), means that the pursuit of knowledge can only occur through observation and experimentation (Schultz & Schultz, 2000), as science is the only genuine form of knowledge (Benton & Craib, 2001), and an empirical theory is one that can be tested quantitatively (Abercrombie, Hill, & Turner, 2000). Related to the first, the second tenet, “metaphysics is meaningless” (Ho Yu, p. 27), derives from the idea that metaphysicians have been unable to achieve any concrete, universally agreed-upon, result. In fact, they have not been able to contribute to the understanding of the world (Bushkovitch, 1970). For logical positivists the metaphysical world is the other world, beyond our physical realm, and is denied to exist (Ho Yu).

Language is the core of the third tenet, “a proposition is only meaningful if it can be verified" (Ho Yu, 2006, p. 27), as a proposition is the content of a sentence that affirms or denies something and is capable of being true or false (Borowski & Borwein, 1991). And for logical positivism, "a proposition is only meaningful if it can be verified" (Ho Yu, 2006, p. 27), hence theological statements, for example, God is love, or peace is good, are viewed as confused discourse. In their view, language similar to this is often found in literature and poetry, and could arouse emotional responses or inspire action in others, and thereby serves another function in society. But when it comes to science, science is concerned with the truth; therefore, the discourse is restricted to propositions of meaningfulness (Bechtel, 1988; Ho Yu).

Logical positivists believe sentences and words are the basic vehicles of meaning, as linguistics are the criterion of verification to explain appropriately related experiences. Accordingly, "the meaning of a sentence was the set of conditions that would show that the sentence was true" (Bechtel, 1988, p. 20), and these conditions would not occur if a sentence was false; however, the proposition could state what would be the case if it were true (e.g., the idea of hypothesis testing). Sentences, not individual words, could be true or false, because they are analyzed by the meaning of words in terms of their roles in it. This type of meaning became known as the "verifiability theory of meaning" (Bechtel, p. 20). Within verification theory, the logic of statistical hypothesis testing is not to verify whether the hypothesis is right, as conclusive verification of a hypothesis is not possible, but conclusive falsification is possible within a finite sample (Ho Yu, 2006).

Therefore, originating from the verifiability theory of meaning, quantitative research uses synthetic sentences when forming research hypotheses. These are propositions that can be either true or false, but are written in the positive. For example, “abused children have lower self-esteem” (Smith & Davis, 2007, p. 49), can either be confirmed or disconfirmed. Also, emerging from this theory is the *general implication form,* the statement of a research hypothesis in the if-then form, which can include a conditional sentence as part of the proposition. For example, “If x is placed in water, then x will dissolve if and only if x is soluble” (Bechtel, 1988, p. 22).

The fourth tenet of logical positivism, “mathematics can be reduced to logic” (Ho Yu, 2006, p. 27), encompasses the ideas of logical analysis and reductionism. That is, mathematics is a tool to take a phenomenon that can be expressed in terms of manageable variables, numbers, and mathematical equations, and through statistical analysis this data can be reduced into a single determining factor (Abercrombie, et al., 2000; Ho Yu, 2006). Nonetheless, it does not necessarily mean that it is only reducing events to numeric data to mathematical models, but rather events, data, and theory form a positive feedback loop. Consider the concept, construct validity, which is “the degree to which a measurement device accurately measures the theoretical construct it is designed to measure” (Cozby, 2004, p. 369), as these drive the nature of the data collection, and the resulting data from the administration of an instrument can sometimes revise the theory itself (Ho Yu, 2006). From these early philosophies of positivism and logical positivism emerged the experimental and quasi-experimental research designs.

Experimental and Quasi-Experimental Research Designs

Some educational researchers incorrectly link causal inferences to logical positivism. Cause, according to logical positivists, is something that cannot be observed or measured; therefore, according to vertificationism, a statement that cannot be verified has no content—causal statements are non-verifiable statements (Ho Yu, 2006). Russell (cited by Ho Yu), explained relationships according to functions, where Y = a + bX, which can be rewritten as X=(Y-a)/b; therefore, X could not be interpreted as a cause of Y because the position of X and Y are interchangeable in the equation. Likewise, even though the principal goal of statistics in experimental and non-experimental data analysis is to find causation, the cause and effect relationship remains undetermined even when X is seen to cause Y (Ho Yu). To have causal inference, three elements are needed in an experimental design.

First, temporal precedence—did X (i.e., independent variable) come before Y (i.e., dependent variable)?; second, co-variation of the cause and effect—when the cause is present the effect occurs, and likewise, when it is not, the effect does not; and third, have all alternative explanations been eliminated? Could there be another causal variable responsible for the results (Cozby, 2004)? These are the three essential elements that make up a true experimental research design; however, in the social sciences, it is extremely unlikely that a research design can encompass all three elements to initiate a true experimental design, so a quasi-experimental design is used instead. In order to understand the purpose of a quasi-experimental design, it is necessary to briefly discuss and explain an experimental design first. Campbell and Stanley (1963) state it best. The experiment is

> the only means for settling disputes regarding educational practice, as the only way of verifying educational improvements and as the only way of establishing a cumulative tradition in which improvements can be introduced without the danger of a faddish disregard of old wisdom in favor or inferior novelties. (p. 2)

A true experimental design is the only research methodology that can, as previously stated, control for temporal precedence, co-variation, and eliminate other plausible explanations for the results of the study. Also, it can assure the researcher of good internal and external validity. Internal validity "is the basic minimum without which any experiment is uninterpretable [ibid]" and can answer the question "did in fact the experimental treatment make a difference in this specific experimental instance?" (Campbell & Stanley, 1963, p. 5). External validity is focused on generalizability. "To what populations, settings, treatment variables, and measurement variables can this effect be generalized" (Campbell & Stanley, p. 5)? Both questions are important in educational research.

One of the most important elements in a true experimental design is randomization of participants. This is "a process occurring at a specific time, and is the all-purpose procedure for achieving pretreatment equality of groups, within known statistical limits" (Campbell & Stanley, 1963, p. 6). Further, when treatment groups are equated before treatment, "then pre-test selection differences could not be a cause of post-test differences" (Shaddish, Cook, & Campbell, 2002, p. 249). Because of an experimental design, temporal precedence remains in check, as cause precedes effect. The use of statistical analysis can verify whether cause co-varies with effect, and the remaining task—to examine causality—is to eliminate any alternative explanations for the results (Shaddish, et al.). There may be numerous alternative explanations for research results, and by eliminating them, "varying degrees of 'confirmation' are conferred upon a theory…the fewer remaining…the greater degree of confirmation" (Campbell & Stanley, p. 36). These are the general underlying assumptions in a pure experimental design. Further, as explained by Stanley and Campbell, the most efficient model in research methodology is the experimental model. However,

> in a very fundamental sense, experimental results never confirm or prove a theory—rather the successful theory is tested and escapes being disconfirmed. The word 'prove' by being frequently employed to designate deductive validity, has acquired in our generation a connotation inappropriate both to its older uses and to its application to inductive procedures such as experimentation. The results of an experiment 'probe' but do not 'prove' a theory. (p. 35)

Likewise, this applies to a quasi-experiment, which is similar to an experimental design, except with one major difference—participants are not randomly assigned to either a treatment or control group. Hence, the purpose of using a quasi-experimental design is "well worth employing when more efficient probes are unavailable" (Campbell & Stanley, 1963, p. 35). And to examine statistical literacy in adult college students, an experimental design would not be appropriate, because it would be impossible to randomly assign participants to either a control or treatment group. But like an experimental design, treatment groups are equated before treatment, "then pretest [sic] selection differences could not be a cause of posttest [sic] differences" (Shaddish, et al., 2002, p. 249), and temporal precedence remains in check, as cause precedes effect. The use of statistical analysis can check to see whether cause co-varies with effect, and the remaining task to examine causality is to eliminate any alternative explanations for the results (Shaddish, et al.).

Research Design for Examining Statistical Literacy

A quantitative rather than a qualitative methodology was chosen to examine statistical literacy in adult college students. A qualitative approach would not be appropriate because it collects data through the form of text, written words, phrases or symbols, which results in thick, rich descriptions of data. Hence, qualitative research focuses more on an understanding of a social phenomenon from participants' perspectives (Neuman, 2000). Further, a qualitative approach allows a researcher to construct social reality, become involved, and use a thematic analysis. On the other hand, a quantitative methodology allows a researcher to measure objective facts, focus on key variables, remain detached from the research, and perform statistical analyses to secure results (Neuman).

In essence, to measure statistical literacy in adult college students, a quantitative methodology was used, which reflects the philosophical underpinning of positivism and logical positivism. These philosophies purport that knowledge can be accumulated through observation, measurement, and systematic data collection, and can help build knowledge to answer questions about why individuals behave the way they do. Hence, the social world can be investigated in the same way as the natural world (i.e., biology or physics) (James, 2005). Likewise, to measure learning outcomes of adult students in statistics, data was collected, analyzed and interpreted to examine adults' statistical literacy.

And as previously stated, logical positivism purports that there is only empirical knowledge (i.e., a theory which can be tested by some kind of evidence drawn from experience); metaphysics is meaningless, because metaphysicians have been unable to achieve any concrete, universally agreed-upon knowledge, and have not been able to contribute to the understanding of the world. Further, verification theory purports the logic of statistical hypothesis testing is not to verify whether the hypothesis is right, as conclusive verification of a hypothesis is not possible, but conclusive falsification is possible within a finite sample (Ho Yu, 2006). In the same manner, to measure adult students' statistical literacy is to measure their knowledge drawn from experiences in their statistics class, through hypothesis testing.

Appropriately, a quantitative approach has been employed to examine statistical literacy in adult college students, as objective instrumentation was necessary to examine their learning gains in statistics. Adult students were invited to participate in this study by completing a pre-test at the beginning of the semester, and a post-test at the end. Because different disciplines have various course requirements regarding statistics and research methods, students were enrolled in either a statistics class, a research methods class with no prior statistics, or a research methods class with prior statistics. A fourth group, a control group, was added, consisting of adult students who have not completed any type of a research or statistics courses. The independent variable was the type of course students have completed, and the main dependent variables measured students' knowledge and dispositional elements concerning statistics before and after completion of their courses. The main dependent variables were measured with instruments that reflected Gal's (2004) Model of Statistical Literacy. Hence, there are 2 main hypotheses and 4 sub-hypotheses that were examined empirically in this study:

1. Adult learners who have completed a research methods class with prior statistics will be more proficient in their knowledge of statistics than learners who have only completed a statistics, or research methods course, with no prior statistics.

    1a. Adult learners who are not first-generation learners will be more proficient in their knowledge of statistics than learners who are first-generation adult learners.

1b. Adult learners who are male will be more proficient in their knowledge of statistics than learners who are female.

2. Adult learners who have completed a research methods class with prior statistics will have more of a positive disposition toward statistics than learners who have only completed a statistics, or research methods class with prior statistics.

2a. Adult learners who are not first-generation learners will have more of a positive disposition toward statistics than learners who are first-generation learners.

2b. Adult learners who are male will have more of a positive disposition toward statistics than learners who are female.

Some concerns in a quasi-experimental design could affect internal validity; this is important because good internal validity has “the ability to eliminate alternative explanations of the dependent variable” (Neuman, 2000, p. 236). Further, as explained by Babbie (2004), “validity is the extent to which an empirical measure adequately reflects the real meaning of the concept under consideration” (p. 143), as it is the “truthfulness of a measure...a valid measure of a concept is one that measures what it claims to measure” (Shaughnessy, Zechmeister, & Zechmeister, 2003, p. 25). Relative to this quasi-experimental research design to measure statistical literacy, there were five threats to internal validity. These are participant selection, history, maturation, attrition, and testing. Most of these are controlled for through a quasi-experimental design, and will be discussed next.

First, selection of participants was a concern, because assignment to groups was not controlled by the researcher; rather, in this research, assignment to a treatment group was in concordance with participants' class types. This could have resulted in groups that were not equivalent, which could have affected the post-test results, because "selection bias is a confounding of treatment effects with population differences" (Shaddish et al., 2002, p. 56). To control for this, a statistical procedure, Levene's test for homogeneity of variances, was employed to test whether the variance of scores between the groups is the same, in the pre-test. Results showed no significant differences among groups for this research.

Second, history, refers to all events that "occur between the beginning of the treatment and the post-test that could have produced the observed outcome in the absence of that treatment" (Shaddish, et al., 2002, p. 56). This is important to consider, because students were exposed to other relevant sources of information beyond those under the researchers' control in the classroom However, this threat was minimized because participants came from the same general location (Shaddish et al.) whereas, in this situation, they were exposed to similar media sources that could increase their statistical knowledge, rather than the course itself.

Third, maturation is important to consider because this research took place over a college semester, resulting in participants growing older or wiser, which could threaten internal validity, if this "could have produced the outcome attributed to the treatment" (Shaddish, et al., 2002, p. 57). However, this threat was reduced by having groups of adult students who were similar in age so that their maturation status was akin; in addition, they were grouped according to their class enrollment.

Fourth, attrition refers to the occurrence in which participants in research sometimes fail to complete the outcome measure (Shaddish, et at., 2002). And unfortunately, there were some participants who did not complete the study, as some of the participants dropped the class or simply did not want to complete this study. There is no control for this, as participants are free to discontinue being part of the research at any time. And last, are testing effects. Because this research design used pre-and post-tests, practice or familiarity with the test could be mistaken for treatment effects. However, this threat was reduced because there was an interval of time between tests, with the pre-test being given at the beginning of the semester and the post-test at the end. In addition, the data was examined by an Analysis of Covariance (ANCOVA), which is a useful statistical technique when there is a two-group, pre-test/post-test design. Scores on a pre-test are treated as a covariate to control for pre-existing differences between groups. This is accomplished by using a SPSS (i.e., a computer data analysis program), which "uses regression procedures to remove the variation in the dependent variable that is due to a covariate, and then perform the normal analysis of the variance techniques on the corrected scores" (Pallant, 2006, p. 363). These results are detailed in chapter 4 and 5.

In addition, there were limitations in this study I wish to briefly discuss concerning external validity. First, this research used a quasi-experimental design; therefore, it was limited in its ability to generalize the results of statistical literacy to all adults who are enrolled in college. And second, adult students who participated in this research were from small rural colleges on the east coast of the United States. This limits the generalizability to all adult college students.

Background of the Researcher

I am instructor in the behavioral/social sciences, who teaches a variety of college courses including statistics and research methods. However, my journey into academia, teaching, and my interest into statistical analyses began very differently than most people with whom I have become acquainted in higher education.

My entire college experience was undertaken as an adult learner, as I began college after becoming a widow with four small children, and after building (literally) a house for my family. Like most college students, I had a difficult time deciding what type of career I would choose after completing my college education, which inspired me to complete dual bachelor degrees, in Criminal Justice and Psychology. During my undergraduate studies, I was a peer tutor to students in virtually any course within these disciplines, and this included statistics in the social and behavioral sciences. Deciding that I wanted to further my education, I completed my Master's degree in Community Psychology and Social Change. My master's thesis was *Statistical Anxiety in Undergraduate Students and Anxiety Reduction Techniques.*

My thesis was related to my current work at the time. While a graduate student in my master's program, I became a professional tutor, with a specialty in tutoring statistics courses. I noticed that many of the students I tutored were very anxious about taking their statistics classes. This inspired my research into statistical anxiety, and my results revealed a variety of ways I could help my tutees overcome their anxiety and succeed in their statistics classes. Some of these anxiety reducing methods I incorporated into the statistics classes I now teach.

While working on the completion of my master's degree, I decided I was going to stay in education and teach at the college level. In my first semester of teaching, I found myself in a unique position. After tutoring students for psychology statistics the previous semester, I had the opportunity to have the same students in my research methods class the next semester. Many of the students I had tutored the previous semester were very good students, who received A's and B's at the end of the semester. However, when they began their research methods course, they did not remember any statistical concepts that were applicable in a research methods course. They were statistically illiterate.

This has fueled my interest in statistical literacy; I believe it is an extremely important part of a college education, as data governs and rules our lives on a daily basis. And students' learning in college should be applicable later to their daily lives. As *learning is life*, an understanding of statistics will be applicable in adults' lives as they continue to learn about political and social issues underpinned by statistics long after their formal college education is complete.

Research Methods

This section will describe the research methods that were employed to examine statistical literacy in adult learners. First, the participation section elucidates how the selection of participants was accomplished; second, a detailed discussion ensues that describes how the instrumentation used to measure statistical literacy in this study incorporated the seven elements of Gal's (2004) Model of Statistical Literacy; and third is a description of how the research was undertaken and analyzed.

Participant Selection

Selection of participants was accomplished by using purposeful samplings, which is a widely used sampling method when the researcher wants to "get all cases that fit particular criteria, using various methods" (Neuman, 2003, p. 211). Hence, students who were enrolled in a statistics class, a research methods class with prior statistics, a research methods class with prior statistics, and students who had never taken classes in statistics and research methods were invited to participate in this research. Because this research is focused on statistical literacy in adult learners in higher education, and because it is difficult to define an adult, the following criteria were used in the selection method. Inclusion criteria for an adult learner was someone; (a) who has assumed major life responsibilities and commitments; (b) one who is no longer dependent upon parents (Manacuso, 2001); or (c) one who assumed care of another, for example a child or elderly relative; (d) is employed; or (e) has experienced a delay after high school in enrolling in college (Belcastro & Purslow, 2006). And within the adult student status, a first-generation college student was defined as those adult students whose parents never attended college (Lee, et al., 2004).

Initially, 167 students volunteered to participate in this research and completed the pre-test on statistical literacy; however, 33 students did not complete the post-test, realizing an attrition rate of 20%. Students dropped out of the study for various reasons, with the most common being—they dropped the class during the semester. Inclusive criteria were applied to the remaining 134 participants who completed the pre- and post-test surveys, which disqualified 24 participants because they did not meet the criteria for being an adult student.

For this reason, participants were 110 adult students, 74 females and 37 males. The sample was 72% Caucasian, 26% African American, 1% Native American and 1% Asia/Pacific Islander, ranging in age from 18 to 40 years old (M = 21, S.D. = 3.25).

Marital status as reported by participants was single (89%), married (4%), and living with a partner (7%). Most participants reported full-time student status (96%), with a few part-time (4%), and reported GPAs ranging between 1-2 (3%), 2-3 (42%), and 3-4 (55%).

Procedure

Instructors at various campuses, who taught statistics or research methods and those who taught introductory courses, were asked by the researcher for their permission to invite their students to participate in a study examining statistical literacy. Accordingly, students who were currently enrolled in a statistics course, a research methods course with prior statistics, a research methods course with no prior statistics, and students who did not take a statistics or a research method class were invited to participate in this study. Because this research was multi-campus, and some campuses were not in the locale of the researcher, students' invitation to participate in this research was accomplished in three ways.

At the beginning of the semester, students in classes who were in the locale of the researcher were invited by the researcher in their classrooms to participate in research that examines statistical literacy. Second, students in classrooms that were not in the locale of the researcher were invited by the instructors of those classes by reading a script prepared by the researcher. And for those classes in which the instructors preferred not to read a script, an invitation to participate in the research on statistical literacy was sent via an e-mail to students.

Regardless of the method of invitation, students were informed of the purpose of the study, procedures, duration, confidentiality, their right to ask questions, reminded that their participation was voluntary, and their payment for participation. Some students received extra credit, or research participation credit, or were entered in a drawing for an iPod. This study was in compliance with the Office of Research Protection at Penn State University.

Adult students who participated in the research completed a demographic survey and four instruments that measured statistical literacy, both of which were completed twice during the semester. The pre-test instrument was completed at the beginning of the semester and the post-test at the end of the semester. Instruments were completed in two ways.

Each test, the pre- and post-, took approximately 1 hour for adult students to complete, and were available for most to access on Angel (i.e., a course management web-based system) for a specific period of time—10 days at the beginning of the semester and 10 days at the end of the semester. The 10-day period allowed adult students who worked or were involved in other campus activities ample time to participate. In addition, by being available on Angel, adult students completed the research either at home on a personal computer, or in the lab. Those who did not complete the instruments on Angel were given a paper and pencil version, and had 10 days to complete it. Even though there are differences between completing instruments on the web and via a pencil and pen version, both sets of adult students were able to use the 10 days to complete the survey, and on the web, they could save their answers and come back the next day to complete it, similar to a paper and pencil version. At the end of the pre- and post-test periods, the instruments were downloaded and printed, and hand-tabulated into SPSS, a data analysis program.

Instrumentation

There were five instruments used in this study. One instrument collected information on adult students' demographics, and four instruments collected data to examine statistical literacy in adult learners. The demographic survey collected data regarding adult students' age, ethnicity, sex, GPA, class standing, and the type of course the student was enrolled in (e.g., statistics, research methods class with no prior statistics, research methods class with prior statistics). In addition, they were asked four questions to determine whether they were considered an adult learner.

The four instruments used in this study to examine statistical literacy encompassed the seven elements of Gal's (2004) Model of Statistical Literacy. The elements are broken down into two categories, knowledge and dispositional elements. There are five knowledge elements: literacy skills, statistical, mathematical and context knowledge, and critical questions; and two dispositional elements, beliefs and attitudes, and critical stance.

The first instrument examined four knowledge elements, literacy skills, statistical, mathematical and context knowledge, and were constructed from questions regarding statistics chosen from the Assessment Resource Tools for Improving Statistical Thinking (ARTIST) web. Some questions were slightly modified with wording changes.

The second instrument examined the fifth knowledge component, critical questions. This instrument was constructed by using a combination of published research from mainstream media (e.g., newspaper, magazines, Internet news) with Gal's worry questions and measured students' ability to critically question published research.

The third instrument was designed to examine the first of two dispositional elements of statistical literacy, beliefs and attitudes, and was accomplished by using the *Survey of Attitudes Toward Statistics* (Schau, Dauphinee, DelVecchio & Stevens, 1994), in combination with an extension of students' open-ended responses, to further probe students' beliefs about statistics.

And finally, the fourth instrument examined the second dispositional element, critical stance, by using the Scale of Critical Stance (SCS), a new instrument containing 10 statements to measure individuals' perception of how they respond to statistical messages in the media. Each of the four instruments were designed to measure the seven elements of statistical literacy, according to Gal's (2004) Model of Statistical Literacy, and will be discussed in detail to explain how they were constructed.

The first instrument, as previously discussed, examined four knowledge elements, literacy skills, statistical, mathematical, and context knowledge, and were constructed from questions regarding statistics chosen from the ARTIST web. The ARTIST web is a collaborative project to improve statistical assessment in higher education, nationally and internationally. The principal investigators are Joan Garfield and Robert delMas, from the University of Minnesota, and Beth Chance from the California Polytechnic State University. Numerous educators make up the Advisory Group, while other educators contributed test items to the project. Funding was provided in part by a grant from the National Science Foundation for the group to develop a website to provide an online assortment of resources to assess statistical literacy.

The project followed after Garfield (2001) conducted an assessment that explored how the reform movement in statistics education has affected the teaching of statistics. Her study indicated that many instructors are incorporating technology and designing courses that allow students to do more than use formulas or computations in their introductory statistics courses. But more importantly, it focused on how the reform movement has affected assessment. Results from her study indicated that most assessments required students "to recall or recognize definitions, perform calculations, and carry out procedures correctly…this is problematic…students may not be able to reason about statistical information or to apply what they have learned in other courses or contexts" (delMas, Chance, & Garfield, 2003, p. 3). Accordingly, the project focused on specific types of assessment items, which could examine statistical literacy, thinking and reasoning (delMas et al.).

The ARTIST web project was designed to collect high-quality assessment items and tasks, which are coded according to statistical literacy, reasoning or thinking. These items can be used in a variety of assessment tasks, such as online quizzes or offline exams, and depending on how it is used, results can be compiled for research purposes. The collection of test items includes items developed by the three principal investigators and their advisory board, and from other educators who submitted high-quality test questions relevant to research studies. Questions that are submitted to the ARTIST project are first reviewed before inclusion into the database, which gives the questions face validity (delMas et al., 2003).

From the various test items collected by the ARTIST database, a Comprehensive Assessment of Outcomes in Statistics (CAOS) was developed after three years of acquiring, writing, and revising items. The CAOS included a set of items that students who completed their introductory statistics course would be expected to understand. During this time many items were revised, and the first set of items were evaluated by the

ARTIST team in order to acquire content validity of the items, and to identify any important concepts that were missed in the test version. New items were added, and the CAOS was piloted-tested in August 2004. Data from the pilot study was used to make additional revisions (delMas, Garfield, Ooms, & Chance, 2006).

A second version, the CAOS 2, soon emerged and was piloted-tested in January, 2005. The test was given to 800 college level students on-line. Data was collected by the instructor in two ways: one was a copy of the test with percentages filled in for the responses of his or her students with the correct answers highlighted, and the second copy contained a spreadsheet with the total percentage correct for each student. Many instructors, who participated with their students in the second version of the CAOS, administered a pre- and post-test for their assessment, and some used the instrument to assign a grade for their course. Like the first version, the second was a pilot test, and data gleaned from the results was used to make a third version, the CAOS 3 (delMas, et al., 2006).

To further improve the validity the CAOS, the third version was given to a group of 30 statistics instructors. Albeit the test showed that it was measuring what it was designed to measure, some suggestions for changes were made, and this input was used to add or delete items, and to make extensive revisions to produce the final version of the test, the CAOS 4. The fourth version consisted of 40 multiple choice items and was administered during the fall of 2005 (delMas, et al., 2006).

Before administering the CAOS, a group of 18 members of the advisory and editorial boards of the Consortium for the Advancement of Undergraduate Statistics Education (CAUSE) were used as expert raters. Agreement was high, with 94% of the raters in agreement that the fourth version measured important learning outcomes. And all the raters further agreed that the "CAOS measures outcomes for which I would be disappointed if they were not achieved by students who succeed in my statistics courses" (delMas, et al., 2006, p. 8).

The latest version of the CAOS 4 was administered on-line and via hard copies to a total of 1,028 students, with the strict criterion that those who took the exam on-line and out of the classroom had spent at least 10 minutes, but not more than 60 minutes, in completing the test. This left a total of 817 students in the study for the data analyses of the test. Most of the students in this population were enrolled in a four-year college, with less than a fifth of the students enrolled in a two-year or a technical college. An analysis of the internal consistency of the 40-item post-test produced a Cronbach's alpha coefficient of .77 (delMas, et al., 2006).

Accordingly, because the database of questions is quite extensive, ongoing research is still being conducted to examine issues of validity. However, results so far indicate that the questions have content and face validity, and the CAOS 4, which was constructed from the ARTIST questions, has a high Cronbach's alpha coefficient. Therefore, questions for the instrument to examine the four knowledge components, literacy skills, statistical, mathematical and context knowledge, were taken from the larger set of questions available on the ARTIST web.

Questions in the ARTIST web bank are classified into three categories, statistical literacy, reasoning, and thinking. Albeit there is some overlap, these can be used as three distinct categories to examine four knowledge components on Gal's model, literacy skills, statistical, mathematical and context knowledge. I will further elucidate how Gal's components match operationally with the questions in the ARTIST web test bank.

Knowledge Elements

Knowledge elements, as previously discussed, consist of literacy, mathematical and statistical skills, context knowledge and critical questions. The first four elements of Gal's (2004) Model of Statistical Literacy were examined by using questions from the ARTIST web that are classified on the site as statistical literacy, reasoning and thinking. The fifth element, critical questions, was examined by using Gal's worry questions in combination with research that is reported in the general media. Each element will be discussed next.

Literacy Skills. Literacy skills are the interpreting and understanding of prose text to "derive meaning from the stimulus presented to the readers" (Gal, 2004, p. 52). For example, readers of statistical information need to be aware of and understand the meaning of certain statistical terms, such as average or representative. Literacy skills also include what Gal describes as document literacy—the ability to understand the reading of various non-prose texts, for example, graphs, tables or symbols. Likewise is the group of questions classified on the ARTIST web bank as statistical literacy. Under this category, questions of statistical literacy are those that examine "understanding and using the basic language and tools of statistics, knowing what statistical terms mean, understanding the use of statistical symbols and recognizing and being able to interpret representations of data" (ARTIST, 2006, n.p.). Accordingly, questions classified by the ARTIST web as statistical literacy were used to examine the knowledge component, literacy, from Gal's model.

Statistical and Mathematical Knowledge. Albeit the components that make up statistical and mathematical knowledge, such as knowing why data are needed and how data can be produced, familiarity with basic terms and ideas related to descriptive statistics, familiarity with basic terms and ideas related to graphical and tabular display, understanding the basic notions of probability, and knowing how statistical conclusions or inferences are reached (Gal, 2004), it is the combination of these elements that make up statistical reasoning. Further, because even at the elementary level, adults need mathematical knowledge to understand the sum of "a large number of observations by a concise quantitative statement," (Gal, p. 63) such as percentages or means, and they need to understand how to apply certain mathematical tools and procedures. In essence, it is the connection of these abilities that enable statistical reasoning; therefore, to examine statistical and mathematical knowledge, questions from the category of statistical reasoning were used. More precisely defined by the ARTIST (2006) web site,

> statistical reasoning is the way people reason with statistical ideas and make sense of statistical information. Statistical reasoning may involve connecting one concept to another…or may combine ideas about data and chance. Reasoning means understanding and being able to explain statistical processes and being able to fully interpret statistical results. (n.p.)

C*ontext Knowledge.* The last knowledge component, context, according to Gal (2004) is the proper interpretation of statistical messages, which depends on adults' ability to place the messages in context. As further explained by Moore (1990) "data are not merely numbers, but numbers with a context" (p. 96), and therefore need to be understood within that context. The study of data and chance need to be understood with a coherent whole from “the progression of ideas from data analysis to data production to probability to inference” (Moore, p. 102). Likewise is the meaning of statistical thinking. As elucidated by the ARTIST (2006) website,

> statistical thinking involves an understanding of why and how statistical investigations are conducted. This includes recognizing and understanding the entire investigative process, from question posing to data collection to choosing analyses to testing assumptions…recognizing how, when, and why existing inferential tools can be used, and being able to understand and utilize the context of a problem to plan and evaluate investigation and to draw conclusions. (n.p.)

Accordingly, questions to examine context knowledge in Gal’s Model of Statistical Literacy were chosen from the collection of questions categorized as statistical thinking.

In sum, the four knowledge elements from Gal’s Model of Statistical Literacy were examined by the following categories from the ARTIST web on one instrument. Literacy skills were measured by statistical literacy, statistical and mathematical knowledge by statistical reasoning, and context knowledge by statistical thinking (see Appendix A).

Critical Questions. It is recommended by Gal (2004) that adults need to ask the worry questions about statistical messages when interpreting any type of research. In asking and answering these questions, a critical evaluation of statistical information will lead to a more informed consumer of research. To evaluate participants’ critical questioning skills, they received a short research article that was published in national newspapers and asked to examine the list of worry questions posed by Gal. After examining the worry questions posed by Gal, they were asked if these have any application to the research they had just read, and if so, what would these be.

Dispositional Elements

As mentioned previously, Gal's Model of Statistical Literacy includes two dispositional elements, attitudes and beliefs, and critical stance. Individuals' attitudes toward statistics shape their behavior toward statistics (e.g. studying habits, importance of, etc.), which affect their belief systems. Beliefs can include how individuals perceive themselves as learners of statistics. Together, attitudes and beliefs can affect an individual's critical stance—their ability to challenge statistical messages without external cues. Because attitudes and beliefs are an interrelated constructs, attitudes were examined by using the instrument, *Survey of Attitudes Toward Statistics (SATS)*. And to examine students' beliefs, extended questions on the attitude phrases were added to the instrument. A separate instrument was developed to examine students' critical stance.

Attitudes and Beliefs. The SATS developed by Schau, Dauphinee, Del Vecchio and Stevens (1991), was constructed using a variation of the nominal group technique (Moore, 1994). This method employed a panel comprised of two graduate and two undergraduate students, who were enrolled in introductory statistics classes, along with two introductory statistics instructors not affiliated with the students, who together generated words and phrases that represented students' attitudes toward statistics. From this collaboration, 92 words and phrases were collected, along with 21 others gathered from existing instruments. These were sorted into four categories: (a) affect, to measure positive and negative feeling concerning statistics; (b) cognitive competence, to measure attitudes about intellectual knowledge and skills when applied to statistics; (c) value, to measure attitudes about the usefulness, relevance, and worth of statistics in personal and professional life; and (d) difficulty, to measure attitudes about the difficulty of statistics as a subject (Schau, Stevens, Dauphinee, & Del Vecchio, 1995).

From this initial undertaking, 113 words and phrases were sorted into 80 potential survey items and again resorted by the group into a four-category structure, resulting in 60 items, 15 per category. This pilot version was administered to 132 introductory statistics classes in both undergraduate and graduate courses at the University of New Mexico and the University of South Dakota. After the pilot test, each item's contribution to its hypothesized dimension was evaluated statistically, resulting in some items being eliminated from the scale. This resulted in 32 items being selected for the SATS, with 7 measuring Affect, 7 Cognitive Competence, 10 Value, and 8 measuring Difficulty (Schau, et al., 1995).

To further establish concurrent validity, that is, a correlation with a criterion measure obtained at the same time (Cherulnik, 2001), Schau et al. (1995) conducted a comparative study between the 32-item version of the SATS and Wise's (1985) Attitude Toward Statistics (ATS) survey. Wise's ATS has been one of the most frequently used instruments to measure attitudes toward statistics, and is similar to the SATS. The SATS was administered to 1,203 students who were enrolled in 33 introductory courses, with a subset of 230 participants taking both the SATS and the ATS. "The ATS scale correlated positively and significantly with all four SATS scales: Affect = .79, Cognitive Competence = .76, Value = .40 and Difficulty = .42" (Schau et al., p. 873).

An item-based confirmatory factor analysis (CFA) on the SATS indicated two items with problems on the Affect and Cognitive Competence scales, and the traditional item analysis indicated two more, one each on the Value and Difficulty scales. All four items were excluded from the scale, resulting in the final form of the SATS, which contains 28 items. Cronbach's Alpha scores for the final version of the SATS ranged "from .81 to .85 for Affect (six items), .77 to .83 for Cognitive Competence (six items), .80 to .85 for Value (nine items), and .64 to .77 for Difficulty (seven items)" (Schau et al., 1995, p. 872). Wade's (2003) research reports similar values for item reliability, Affect = 79, Cognitive Competence =.87, Value = .85, and Difficulty = .65.

The final version of the SATS is a 28-item attitude survey designed to measure students' attitudes toward statistics with pre- and post-tests. Responses on the SATS are measured on a 7-point Likert scale, with one meaning *strongly disagree*, four *neither disagree nor agree*, and seven, *strongly agree.* Some of the items are positively worded, while others are negatively worded and have to be reverse coded before scoring for data analysis. A higher score on the SATS indicates a more positive attitude toward statistics (Schau, et al., 1995).

As previously discussed, attitudes and beliefs are interrelated constructs; therefore, the SATS can be extended to examine beliefs by asking participants to respond separately to open-ended questions, or to sentence completions. For example, one item on the Value element on the SATS, *I can learn statistics,* would be extended in a separate section, and participants could be asked to respond to one of the following questions: Why did you respond as you did? What experiences form the basis for your response? Or, what aspect(s), if any, of statistics make you feel this way? (Gal, Ginsburg & Schau, 1997). Or sentence completions, as suggested by Gal and Ginsburg (1994) could include, for example "I think statistics is… (e.g., useful, boring, frightening) because…” (n.p.). Accordingly, the belief section was an extension of the SATS, but a separate section to examine participants’ beliefs about statistics. Accordingly, this section was constructed of four open-ended statements, one from each element of the SATS scale, Affect, Cognitive Competence, Value and Difficulty.

Critical Stance. Critical stance can be defined by an individual’s “willingness to invoke action” (Gal, 2004, p. 69), when they encounter statistical messages from the media. It is the idea that individuals do not remain passive when they interpret statistical information, as they have developed a questioning attitude toward these statistical messages. Therefore, the scale is designed to examine if adults are willing to challenge statistical messages they encounter from the media (Gal, 2002).

Currently, because this research is on the cutting edge, there is no established instrument to examine an individual’s critical stance toward statistics; therefore, the Scale of Critical Stance (SCS) (Wade, 2007) that will be used in this research is new, and has only undergone pilot testing. The initial pilot test showed Cronbach’s Alpha reliability statistics of .60 for the pre-critical stance scale and .78 for the post-critical stance scale. Scores of .70 are considered to have good internal consistency reliability; and, albeit the pre-test had a score of .60, this could be due to the sampling size of the pilot test (Pallant, 2007). The scale is constructed of 10 statements and uses a Likert scale ranging from one, meaning *strongly disagree,* to four, *neither disagree nor agree*, and to seven, *strongly agree.* Some of the items are positively worded, while others are negatively worded and have to be reverse coded before scoring for data analysis. As advised by Kline (2005), when constructing a scale it is necessary to include items that are negative in value to “ensure that the respondent is paying attention to the items. It prevents respondents from always selecting a particular response category without really attending to the item" (p. 65). Higher scores on this scale will indicate a higher level of critical stance in adults (see Appendix E). However, for this current research Cronbach's Alpha reliability statistics resulted in scores of .31 for the pre-critical stance scale and .37 for the post-critical stance scale. This could be due to the number of items on the scale, as there were only 10, or due to the sample size.

Succinctly, there were four independent variables and six dependent variables. Independent variables were the four different groups of students, which were represented by the type of class they were enrolled in—statistics, a research methods class with no prior statistics, a research methods with prior statistics and a control group. The six independent variables represented the knowledge and dispositional elements from Gal's Model (2004) (two elements have been combined in order to measure them). The knowledge elements from Gal's Model are (a) literary skills, and were measured using the instrument of statistical literacy; (b) statistical and mathematical knowledge (two elements combined), measured by the instrument statistical reasoning; (c) context knowledge measured by the instrument statistical thinking; and (d) critical questions, measured by a survey created by using Gal's worry questions. The dispositional elements, beliefs and attitudes were measured using Schau's SATS instrument and an extension of this instrument, which contained open-ended responses regarding some statements from the SATS. Critical stance, the last dispositional element, was measured by using Wade's SCS scale. (See complete pre- and post-test instruments in appendices B and C)

Analyses

Data analyses were complied according to the type of scale used to collect data in both the knowledge and dispositional elements. And to review, questions from the ARTIST website were used to examine the first four components literacy, statistical and mathematical, and context knowledge, from Gal's Model (2004). Hence, literacy skills were represented by statistical literacy, statistical and mathematical knowledge by statistical reasoning and context knowledge by statistical thinking. Critical questions were analyzed by using a scale that represented a score in response to Gal's worry questions regarding reported research in the news. A summary of the statistical tests used for the analyses of the knowledge and dispositional elements follows.

There are eight types of statistical tests used in this research, which are an ANOVA, MANOVA, Fisher's (LSD), a mixed between-within subjects analysis of variance, paired-samples t-tests, independent-samples t-tests, an ANCOVA and a Chi-Square. These will be discussed briefly.

First, an analysis of variance (ANOVA) is a statistical test procedure used to compare mean scores in more than three groups. It compares "the variance between the different groups, with the variability within each of the groups" (Pallant, 2006, p. 214). If a significant F statistic is obtained, there is a statistically significant difference among the means scores and the null hypothesis can be rejected. However, an ANOVA does not indicate in which groups there is statistically significant differences; to do this, post-hoc comparisons were used.

Second, similar to an ANOVA is a MANOVA, which tests whether mean

differences among groups on a combination of DVs are likely to have occurred by

> chance. In a MANOVA, a new DV that maximizes group differences is created from the set of DVs. The new DV is a linear combination of measured DVs, combined so as to separate the groups as much as possible. Hypotheses about means in MANOVA are tested by comparing variances—hence multivariate analysis of variance. (Tabachnick & Fidell, 2007, p. 243)

Thus, a MANOVA is used to examine "one or more categorical independent variables" (i.e., type of class) and two or more related continuous dependent variables" (i.e., statistical knowledge, statistical thinking, statistical reasoning) (Pallant, 2007, p. 117). A required sample size to complete an MANOVA recommended by Tabachnick and Fidel is to have more cases in each cell than DVs, and this study has met that criterion. Because ANOVAs and MANOVAs can only inform us if there are group differences, a post-hoc test needs to be performed to distinguish where group difference occurs. Accordingly, the third test is Fisher's (LSD) which,

> is a statistical procedure that determines if the difference found between two treatments is due to the treatment or if the difference is simply due to random chance. For each set of data a value ($LSD_{0.05}$) is calculated at a chosen level of significance. If the differences between two treatments means is greater than this calculated value, then it is said to be a 'significant difference' or a difference not due to random change. (University of Illinois, 2009, n.p.)

The fourth statistical test, a mixed between-within subjects analysis of variance, is a statistical procedure used to investigate the impact of the independent variable on the dependent variable by using pre- to post-test scores. This type of analysis "tests whether there are main effects for each of the dependent variables and whether the interaction between the two variables is significant” (Pallant, 2006, p. 241).

The fifth and sixth statistical tests were paired-samples t-tests and independent-samples t-tests. A paired-samples t-tests is used when testing the same people on more than one occasion, for example to examine a difference between a pre- and a post-test for males. Similar is a independent-samples t-test, which is used when one intends to “compare the mean score on some continuous variable for two different groups of participants” (Pallant, 2006, p. 205).

Seventh, is an analysis of covariance (ANCOVA). which is an extension of analysis of variance that allows the exploration of differences “between groups while statistically controlling for an additional variable…called a covariate” (Pallant, 2006, p. 263). This is a variable that may be responsible for influencing scores on the dependent variable. A data analysis program, for example SPSS, uses regression procedures to remove the variation in the dependent variable, then performs the analysis.

And last, a Chi-Square Test for Independence is used to explore relationships among independent and dependent categorical variables (i.e., class type and beliefs). Each variable must have two or more categories and is used to “compare the observed frequencies or proportions of cases that occurs in each of the categories, with the values that would be expected if there was no association between the two variables being measured” (Pallant, p. 214). Open-ended statements were also used with the belief statements, and analyzed according to themes.

Data analyses were first completed on the knowledge elements, then on the dispositional elements, which made up the dependent variables, while class types, gender and first-generation status were the independent variable. The first three knowledge elements, statistical reasoning, thinking, and literacy, were examined in a number of ways. First, a MANOVA was performed to examine each of the knowledge elements on post-test scores among class types. Further analyses were completed using a post-hoc test, Fisher's (LSD), on each of the knowledge elements, statistical thinking, reasoning, and literacy with each of the class types. Mixed between-within subjects ANOVAs were completed to examine if gender (i.e., male or female) or first-generation status (i.e., first-generation or not first-generation) had an effect on learning gains. In addition, separate paired-samples t-tests were completed on each of the knowledge elements (pre- and post-test scores), statistical thinking, reasoning, and literacy with the independent variables gender and first-generation status. Differences between gender and first-generation status were examined by multiple independent-samples t-tests on post-test scores for each group. Statistical analyses completed on variables of gender and first-generation status do not include data from the control group. Finally, to examine for pre-test influence on post-test scores on the knowledge elements, three one-way between groups analyses of covariance (ANCOVA) were conducted on each of the knowledge elements, statistical thinking, reasoning, and literacy.

The fourth knowledge element in Gal's (2004) Model of Statistical Literacy is critical questions, and this data was analyzed by using a mixed between-within subjects ANOVA to assess the impact of class types across two time periods (pre- and post-test scores). Further analyses were completed using a one-way ANOVA to examine post-test scores on the critical questions and class types. Gender and first-generation status were also examined by using separate independent-samples t-tests to compare post-test scores on critical questions. Because of the lack of statistically significant differences on the pre-and post-test scores, an ANCOVA was conducted to examine if the pre-test scores had affected the post-test scores.

The dispositional elements are attitudes and beliefs, and critical stance. Attitudes were measured by using Schau's SATS scale that incorporates four elements, affect, cognitive competence, value and difficulty. These were first examined by a MANOVA, then by four mixed within-between subjects ANOVAs to examine pre- to post-test scores on each of the elements of the SATS, affect, cognitive competence, value and difficulty, among class types. Next, each element of the SATS, affect, cognitive competence, value, and difficulty, were examined by separate independent and paired t-tests with gender and first-generation status as the independent variables. All statistical analyses completed with the variables gender and first-generation statuses did not include data from the control group, to prevent skewing of the data. A follow-up ANCOVA was completed to examine if pre-test scores affected post-test scores.

The dispositional element, beliefs about statistics, was analyzed by using a chi-square to examine differences among class types and students' beliefs on four statements. These statements are categorical variables and consist of dichotomous responses, insecure or secure, will or will not, easy or not easy, and relevant or not relevant. These responses are part of statements that address beliefs, which were further examined by open-ended responses that asked students to respond to the word *because* at the end of each of these statements: *I feel insecure or secure when doing statistics problems because…; I will or will not make a lot of math errors in statistics because…;statistics formulas are easy or not easy to understand because…; and statistics is or is not relevant in my life because…* And because these are categorical data, they were examined according to themes.

An ANOVA was used to examine the final dispositional element, critical stance, by examining class types and post-test scores. To examine pre-to post- test gains or losses on critical stance among groups, a mixed between-within subjects ANOVA was conducted. Pre- to post-test gains or losses were also analyzed with gender and first-generation status as the independent variable by using four paired-samples t-tests. Next, to examine group differences between gender and first-generation status, two separate independent-samples t-tests were conducted on post-test scores of critical stance. Because of a lack of significant results, an ANCOVA was conducted to examine if the post-test scores may have been influenced by pre-test scores on the critical stance scale.

Chapter Summary

To understand the importance of quantitative research, this chapter briefly examined the emergence of quantitative data analyses through its early roots in positivism, and logical positivism. As empirical knowledge emerged to replace the former method of obtaining knowledge, metaphysics, language in the form of propositions became eminent. And as propositions are a part of language, sentences could be analyzed by the meaning of words in terms of their roles in it, which became known as verification theory (Bechtel, 1988). This theory led to forming research hypotheses, which later could utilize mathematics as a tool to turn a phenomenon that can be expressed in terms of manageable variables, numbers and mathematical equations, into something that could be statistically analyzed (Abercrombie, et al., 2000; Ho Yu, 2006).

From the early tenets of positivism and logical positivism emerged the experimental research design. And even though the purpose of a true experimental design is to examine a causal relationship between some phenomena, research has varying degrees of confirmation, “confirmed upon a theory” (Campbell & Stanley, 1963), as most research cannot encompass a true experimental design, especially in the social sciences. A true experimental design is one that shows temporal precedence, co-variation of the cause and effect, and can control for all alternative explanations (i.e., extraneous variables).

And similar to an experimental design, but different due to participant selection, is a quasi-experimental design. It can be used when a random selection of participants to a treatment or a control group is not possible. As explained by Campbell and Stanley (1963), a quasi-experimental design is “well worth employing when more efficient probes are unavailable” (p. 35), and is the research design used in examining adult students’ statistical literacy.

Participants were adult students enrolled in 4 different types of college classes, statistics, a research methods class with no prior statistics, a research methods class with prior statistics, and those who never had a class in statistics or research methods. The adult students who participated in this research completed a pre-test at the beginning of the semester, and a post-test at the end, which consisted of questions that reflected Gal’s (2004) Model of Statistical Literacy. Questions and statements on the instruments reflected both the dispositional and knowledge elements from the model.

Eight types of statistical tests were used to examine the data and were an ANOVA, MANOVA, Fisher’s (LSD), a mixed between-within subjects analysis of variance, paired-samples t-tests, independent-samples t-tests, an ANCOVA and a Chi-Square. These statistical tests analyzed the independent variables, class types, first-generation status and gender, with the dependent variables, the knowledge elements, statistical thinking, reasoning, and literacy, and the critical questions, and dispositional elements, attitudes and beliefs, and critical stance. Next, chapter 4 gives detailed statistical analyses of the results.

CHAPTER 4

RESULTS OF THE STUDY

Introduction

This chapter contains the results of the data analyses on statistical literacy, which was broken down into two sections according to Gal's Model of Statistical Literacy. The first examined knowledge elements, statistical thinking, reasoning, literacy, and critical questions; the second examined dispositional elements, attitudes and beliefs, and critical stance. As previously stated, the purpose of this research is to measure statistical literacy in adult learners before and after they have completed a statistics class, a research methods class with no prior statistics, or a research methods class with prior statistics. There were 2 main hypotheses and 4 sub-hypotheses.

1. Adult learners who have completed a research methods class with prior statistics will be more proficient in their knowledge of statistics than learners who have only completed a statistics or research methods class with no prior statistics.

    1a. Adult learners who are not first-generation learners will be more proficient in their knowledge of statistics than learners who are first-generation adult learners.

    1b. Adult learners who are male will be more proficient in their knowledge of statistics than learners who are female.

2. Adult learners who have completed a research methods class with prior statistics will have more of a positive disposition toward statistics than learners who have only completed a statistics or research methods class with no prior statistics.

2a. Adult learners who are not first-generation learners will have more of a positive disposition toward statistics than learners who are first-generation learners.

2b. Adult learners who are male will have more of a positive disposition toward statistics than learners who are female.

Knowledge Elements: Statistical Thinking, Reasoning, Literacy and Critical Questions

The first three knowledge elements, statistical thinking, reasoning, and literacy, were examined by multiple tests. First, MANOVA was performed to examine each of the knowledge elements on post-test scores among class types. Since the MANOVA yielded significant results, further analyses were completed using a post-hoc test, Fisher's (LSD), on each of the knowledge elements, statistical thinking, reasoning, and literacy.

To examine learning gains between pre- and post-test scores on each of the knowledge elements and each of the class types, three mixed between-within subjects ANOVAs were completed. And to examine if gender (i.e., male or female) or first-generation status (i.e., first-generation or not first-generation) had an effect on learning gains, separate paired-samples t-tests were completed on each of the knowledge elements (pre- and post-test scores) statistical thinking, reasoning, and literacy. Differences between gender and first-generation status were examined by multiple independent-samples t-tests on post-test scores for each group. Statistical analyses completed on variables of gender and first-generation status did not include data from the control group. Finally, to examine for pre-test influence on post-test scores on the knowledge elements, three ANCOVAs were conducted on each of the knowledge elements, statistical thinking, reasoning and literacy.

The fourth knowledge element in Gal's Model of Statistical Literacy was critical questions, and this data was analyzed by using a mixed between-within subjects ANOVA to assess the impact of class types across two time periods (pre- and post-test scores). Further analyses were completed using a one-way ANOVA to examine post-test scores on the critical questions and class types. Gender and first-generation status were also examined using separate independent-samples t-tests to compare post-test scores on critical questions. Because of the lack of statistically significant differences on the pre- and post-test scores, an ANCOVA was conducted to examine if the pre-test scores had affected the post-test scores.

Statistical Thinking, Reasoning and Literacy: Post-test Scores and Class Type (MANOVA)

A MANOVA was performed to investigate differences in post-test scores on the knowledge elements among class types. There were three dependent variables, statistical thinking, reasoning, and literacy, and the independent variable was class types. Preliminary assumption testing was conducted to check for normality, linearity, univariate and multivariate outliers, homogeneity of variance-covariance matrices, and multicollinearity, with no serious violations noted. There were statistically significant differences among class types, statistics, research methods class with prior statistics, research methods class with no prior statistics and the control group (i.e., neither previous statistics or research methods) on the combined dependent variables, $p = 0$, as displayed in Table 1.

Table 1

Knowledge Elements

Combined, Statistical Thinking, Reasoning and Literacy

Multivariate Tests

Effect		Value	F	Hypothesis df	Sig.	Partial Eta Squared
classtyp	Wilks' Lambda	.603	6.505	9.000	.000	.155

When the results for the dependent variables were considered separately, statistical thinking (p = 0), reasoning (p = 0), and literacy (p = 0) reached statistical significance using an adjusted Bonferroni adjusted alpha level of .017, as summarized in Table 2. Because significant results were found with the MANOVA, follow-up comparisons were conducted with a post-hoc test, Fisher's LSD, with the adjusted Bonferroni alpha level of .017.

Table 2

Knowledge Elements

Separate, Statistical Thinking, Reasoning and Literacy

Tests of Between-Subjects Effects

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
classtyp	Post statistical thinking questions total	3.184	3	1.061	17.575	.000	.332
	Post statistical reasoning questions total	.960	3	.320	6.696	.000	.159
	Post statistical literacy questions total	1.107	3	.369	9.586	.000	.213

Statistical Thinking Post-Hoc: Fisher's LSD. As explained previously by the ARTIST web (2006), statistical thinking involves an understanding of why and how statistical investigations are conducted. This includes recognizing and understanding the entire investigative process, from question posing to data collection to choosing analyses to testing assumptions…recognizing how, when, and why existing inferential tools can be used, and being able to understand and utilize the context of a problem to plan and evaluate investigation and to draw conclusions. (n.p.)

As illustrated in Table 3, post-hoc comparisons using Fisher's LSD, with an adjusted Bonferroni alpha level of .017, for the dependent variable statistical thinking and the independent variable, class types, showed four significant results. First, a statistically significant difference between the control group and the statistics class; second, between the control group and the research methods class with no prior statistics; third, between the control group and the research methods class with prior statistics; and fourth, between the statistics class and the research methods class with prior statistics. The only classes whose scores did not reach significance were the research methods class with no prior statistics and the research methods class with prior statistics.

Table 3

Knowledge Element: Statistical Thinking

Post-hoc Comparisons

Multiple Comparisons

LSD

Dependent Variable	(I) Class type	(J) Class type	Sig.
Post statistical thinking questions total	Statistics	Control	.013
		Research methods w/o statistics	.021
		Research methods after statistics	.000
	Control	Statistics	.013
		Research methods w/o statistics	.000
		Research methods after statistics	.000
	Research methods w/o statistics	Statistics	.021
		Control	.000
		Research methods after statistics	.102
	Research methods after statistics	Statistics	.000
		Control	.000
		Research methods w/o statistics	.102

Mean post-test scores among class types for the dependent variable, statistical thinking, showed higher scores for the statistics class when compared to the control group; higher scores for the research methods class with no prior statistics when compared to the control group; higher scores for the research methods class with prior statistics when compared to the control group; and higher scores for the research methods class with prior statistics when compare to the statistics class. A summary of the mean scores and standard deviations are shown in Table 4.

Table 4

Knowledge Element: Statistical Thinking

Means and Standard Deviations

Descriptive Statistics

	Class type	Mean	Std. Deviation	N
Post statistical thinking questions total	Statistics	.4810	.28230	35
	Control	.3205	.22072	26
	Research methods w/o statistics	.6481	.26127	18
	Research methods after statistics	.7681	.20926	31
	Total	.5513	.29655	110

Statistical Reasoning Post-Hoc: Fisher's LSD. Statistical reasoning is the combination of knowing why data are needed and how data can be produced, familiarity with basic terms and ideas related to descriptive statistics, familiarity with basic terms and ideas related to graphical and tabular display, understanding the basic notions of probability, and knowing how statistical conclusions or inferences are reached (Gal, 2004).

As illustrated in Table 5, post-hoc comparisons using Fisher's LSD, with an adjusted Bonferroni alpha level of .017, for the dependent variable statistical reasoning and the independent variable, class types, showed two significant results. First, a statistically significant difference between the statistics class and the research methods class with prior statistics; and second, between the research methods class with prior statistics and the control group.

Table 5

Knowledge Element: Statistical Reasoning

Post-hoc Comparisons

Multiple Comparisons

LSD

Dependent Variable	(I) Class type	(J) Class type	Sig.
Post statistical reasoning questions total	Statistics	Control	.153
		Research methods w/o statistics	.810
		Research methods after statistics	.002
	Control	Statistics	.153
		Research methods w/o statistics	.151
		Research methods after statistics	.000
	Research methods w/o statistics	Statistics	.810
		Control	.151
		Research methods after statistics	.019
	Research methods after statistics	Statistics	.002
		Control	.000
		Research methods w/o statistics	.019

Mean post-test scores among class types for the dependent variable, statistical reasoning, showed higher scores for the research methods class with prior statistics when compared to the statistics class, and higher scores for the research methods class with prior statistics when compared to the control group. A summary of the mean scores and standard deviations are provided in Table 6.

Table 6

Knowledge Element: Statistical Reasoning

Means and Standard Deviations

Descriptive Statistics

	Class type	Mean	Std. Deviation	N
Post statistical reasoning questions total	Statistics	.3829	.21968	35
	Control	.3013	.20558	26
	Research methods w/o statistics	.3981	.23666	18
	Research methods after statistics	.5522	.21742	31
	Total	.4138	.23517	110

Statistical Literacy Post-Hoc: Fisher's LSD. Statistical Literacy is the "understanding and using the basic language and tools of statistics, knowing what statistical terms mean, understanding the use of statistical symbols and recognizing and being able to interpret representations of data" (ARTIST, 2006, n. p.).

As illustrated in Table 7, post-hoc comparisons using Fisher's LSD, with an adjusted Bonferroni alpha level of .017, for the dependent variable statistical literacy and the independent variable, class types, showed three significant results. First, a statistically significant difference between the research methods class with prior statistics and the statistics class; second, between the research methods class with prior statistics and the control group; and third between the research methods class with prior statistics and the research methods class with no prior statistics.

Table 7

Knowledge Element: Statistical Literacy

Post-hoc Comparisons

Multiple Comparisons

LSD

Dependent Variable	(I) Class type	(J) Class type	Sig.
Post statistical literacy questions total	Statistics	Control	.639
		Research methods w/o statistics	.244
		Research methods after statistics	.000
	Control	Statistics	.639
		Research methods w/o statistics	.135
		Research methods after statistics	.000
	Research methods w/o statistics	Statistics	.244
		Control	.135
		Research methods after statistics	.009
	Research methods after statistics	Statistics	.000
		Control	.000
		Research methods w/o statistics	.009

Mean post-test scores among class types for the dependent variable, statistical literacy, showed higher scores for the research methods class with prior statistics when compared to the statistics class; higher scores for the research methods class with prior statistics, when compared to the control group; and higher scores for the research methods class with prior statistics, when compared to the research methods class with no prior statistics. A summary of the mean scores and standard deviations are provided in Table 8.

Table 8

Knowledge Element: Statistical Literacy

Means and Standard Deviations

Descriptive Statistics

	Class type	Mean	Std. Deviation	N
Post statistical literacy questions total	Statistics	.2667	.22579	35
	Control	.2428	.15333	26
	Research methods w/o statistics	.3333	.15125	18
	Research methods after statistics	.4876	.21370	31
	Total	.3342	.21814	110

Statistical Thinking, Reasoning and Literacy: Mixed Between-Within Subjects ANOVAs

To examine learning gains between the pre- and post-test scores on the knowledge elements, three mixed between-within subjects ANOVAs were completed on each the dependent variables, statistical thinking, reasoning, and literacy with the independent variable class types.

Statistical Thinking. Tables 9 and 10 illustrate the results of mixed between-within subjects analyses of variance conducted to assess the impact of four different types of classes, statistics, research methods class with prior statistics, research methods class with no prior statistics and a control group on participants' scores of statistical thinking across two time periods. There was no significant interaction between class type and time, $p = .068$, and no substantial effect for time, $p = .83$. The main effect comparing the types of classes was significant, $p = .0$. Because the main effect was significant, post-hoc comparisons followed.

Table 9

Knowledge Element: Statistical Thinking

Mixed Between-Within Subjects Analysis of Variance

Multivariate Tests

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
time	Wilks' Lambda	1.000	.049	1.000	106.000	.825	.000
time * classtyp	Wilks' Lambda	.935	2.446	3.000	106.000	.068	.065

Table 10

Knowledge Element: Statistical Thinking

Between Subject Effects

Tests of Between-Subjects Effects

Measure: MEASURE_1
Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Intercept	62.578	1	62.578	571.262	.000	.843
classtyp	3.833	3	1.278	11.665	.000	.248
Error	11.612	106	.110			

Post-hoc comparisons using Fisher's LSD showed five significant results for the dependent variable, statistical thinking, and are summarized in Table 11. First, results indicated a statistically significant difference on pre- and post-test scores between the statistics class and the control group, $p = .02$, with both classes showing a decrease in their scores on statistical thinking, albeit the statistics class had higher post-test scores than the control group.

Second, there was a statistically significant difference on pre- and post-test scores between the research methods class with prior statistics and the control group, $p = 0$, indicating an increase in post-test scores for the research methods class with prior statistics and a decrease in post-test scores for the control group.

Third, a statistically significant difference on pre- and post-test scores was found between the research methods class with no prior statistics and the control group, $p = .009$, indicating an increase in post-test scores for the research methods class with no prior statistics and a decrease in post-test scores for the control group.

Fourth, a statistically significant difference on pre- and post-test scores was found between the statistics class and the research methods class with prior statistics, $p = 0$, indicating that scores increased for the research methods class with prior statistics and scores decreased for the statistics class.

And fifth, a statistically significant difference on pre- and post-test scores was found between the research methods class with prior statistics and the research methods class with no prior statistics, $p = .017$, indicating higher pre- and post-test scores for research methods class with prior statistics. Estimated marginal means for statistical thinking and class types can be found in Table 12.

Table 11

Knowledge Element: Statistical Thinking

Post-hoc Comparisons

Multiple Comparisons

Measure: MEASURE_1
LSD

(I) Class type	(J) Class type	Sig.
Statistics	Control	.024
	Research methods w/o statistics	.436
	Research methods after statistics	.000
Control	Statistics	.024
	Research methods w/o statistics	.009
	Research methods after statistics	.000
Research methods w/o statistics	Statistics	.436
	Control	.009
	Research methods after statistics	.017
Research methods after statistics	Statistics	.000
	Control	.000
	Research methods w/o statistics	.017

Table 12

Knowledge Element: Statistical Thinking

Estimated Marginal Means

2. Class type * time

Measure: MEASURE_1

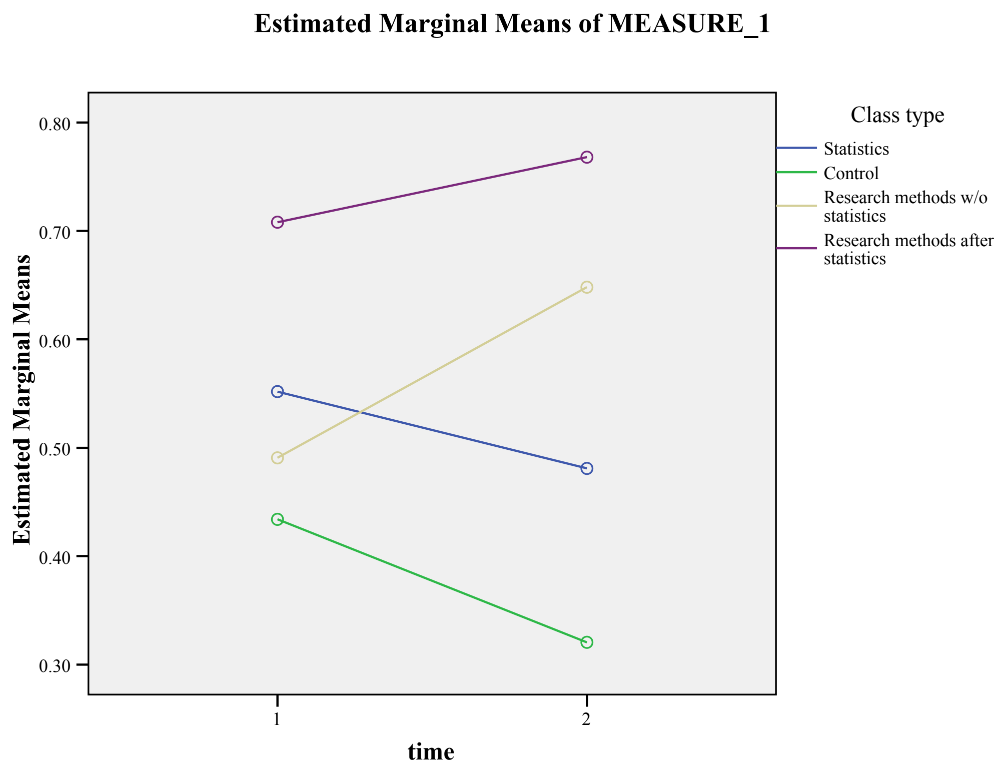
				95% Confidence Interval	
Class type	time	Mean	Std. Error	Lower Bound	Upper Bound
Statistics	1	.552	.059	.435	.669
	2	.481	.042	.399	.563
Control	1	.434	.068	.298	.570
	2	.321	.048	.225	.416
Research methods w/o statistics	1	.491	.082	.328	.654
	2	.648	.058	.533	.763
Research methods after statistics	1	.708	.063	.584	.832
	2	.768	.044	.681	.856

Increases and decreases in pre- to post-test scores among class types, for the dependent variable, statistical thinking, are displayed in Chart 1.

Chart 1

Knowledge Element: Statistical Thinking

Pre- to Post-Test Scores



Statistical Reasoning. Statistical reasoning is the combination of knowing why data are needed and how data can be produced, familiarity with basic terms and ideas related to descriptive statistics, familiarity with basic terms and ideas related to graphical and tabular display, understanding the basic notions of probability, and knowing how statistical conclusions or inferences are reached (Gal, 2004).

Tables 13 and 14 organize the results for mixed between-within subjects analyses of variance conducted to assess the impact of four different types of classes (i.e., statistics, research methods class with prior statistics, research methods class with no prior statistics, and a control group) on participants' scores of statistical reasoning across two time periods. There was no significant interaction between class type and time, p = .335; however, there was a substantial effect for time, p = .003, with all groups showing an increase from pre- to post-test scores. The main effect comparing the type of classes was significant, p = .0. Because the main effect was significant, post-hoc comparisons followed.

Table 13

Knowledge Element: Statistical Reasoning

Mixed Between-Within Subjects Analysis of Variance

Multivariate Tests

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
time	Wilks' Lambda	.921	9.115	1.000	106.000	.003	.079
time * classtyp	Wilks' Lambda	.969	1.144	3.000	106.000	.335	.031

Table 14

Knowledge Element: Statistical Reasoning

Tests of Between-Subjects Effects

Measure: MEASURE_1

Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Intercept	29.160	1	29.160	457.669	.000	.812
classtyp	1.324	3	.441	6.928	.000	.164
Error	6.754	106	.064			

Post-hoc comparisons using Fisher's LSD showed three significant results for the dependent variable statistical reasoning and are summarized in Table 15. First, results showed a statistically significant difference on pre- and post-test scores between the research methods class with prior statistics and the statistics class, $p = .0$, indicating the research methods class with prior statistics scored higher on both pre- and post-test scores than the statistics class.

Second, results showed a statistically significant difference on pre- and post-test scores between the research methods class with prior statistics and the control group, $p = 0$, indicating the research methods class with prior statistics scored higher on both the pre- and post-test scores than the control group.

Third, results showed a statistically significant difference on pre- and post-test scores between the research methods class with prior statistics and the research methods class with no prior statistics, $p = .019$, indicating that the research methods class with prior statistics scored higher on both the pre- and post-test scores than the research methods class with no prior statistics. The estimated marginal means for statistical reasoning and class types are illustrated in Table 16.

Table 15

Knowledge Elements: Statistical Reasoning

Post-hoc Comparisons

Multiple Comparisons

Measure: MEASURE_1
LSD

(I) Class type	(J) Class type	Sig.
Statistics	Control	.377
	Research methods w/o statistics	.528
	Research methods after statistics	.000
Control	Statistics	.377
	Research methods w/o statistics	.181
	Research methods after statistics	.000
Research methods w/o statistics	Statistics	.528
	Control	.181
	Research methods after statistics	.019
Research methods after statistics	Statistics	.000
	Control	.000
	Research methods w/o statistics	.019

Table 16

Knowledge Element: Statistical Reasoning

Estimated Marginal Means

Class type * time

Measure: MEASURE_1

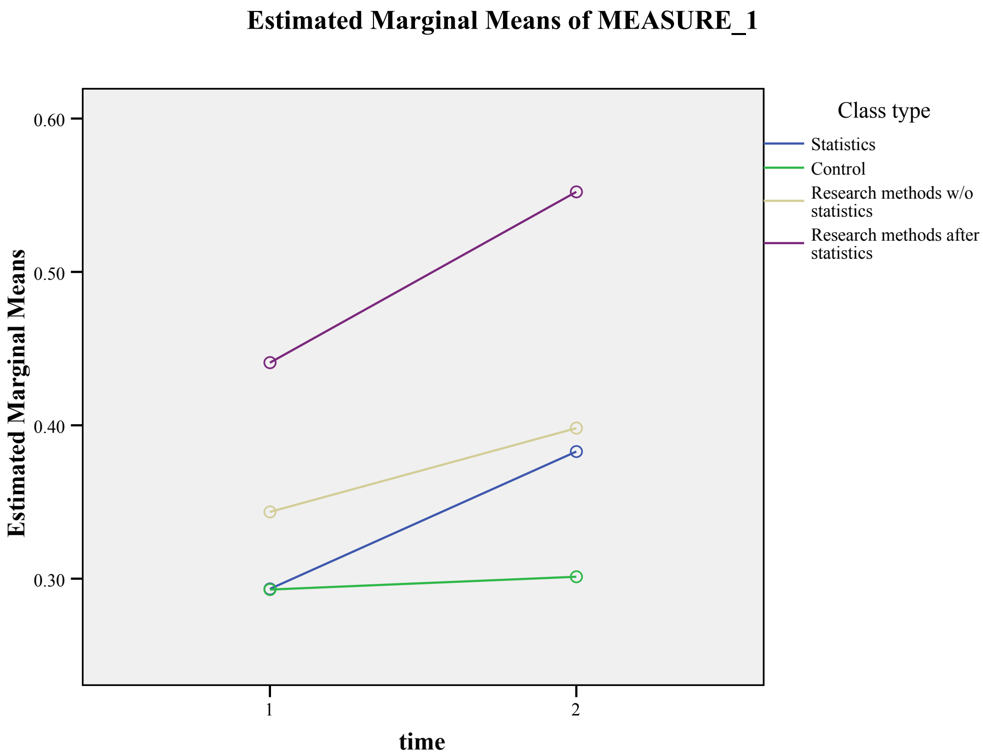
Class type	time	Mean	Std. Error	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
Statistics	1	.293	.034	.226	.361
	2	.481	.042	.399	.563
Control	1	.293	.039	.215	.371
	2	.321	.048	.225	.416
Research methods w/o statistics	1	.344	.047	.249	.438
	2	.648	.058	.533	.763
Research methods after statistics	1	.441	.036	.369	.513
	2	.768	.044	.681	.856

Increases and decreases among class types for pre- to post-test score, for the dependent variable, statistical reasoning, are displayed in Chart 2.

Chart 2

Knowledge Element: Statistical Reasoning

Pre- to Post-Test Scores



Statistical Literacy. Statistical Literacy is the "understanding and using the basic language and tools of statistics, knowing what statistical terms mean, understanding the use of statistical symbols and recognizing and being able to interpret representations of data" (ARTIST, 2006, n. p.).

Tables 17 and 18 organize the results for mixed between-within subjects analyses of variance conducted to assess the impact of four different types of classes (i.e., statistics, research methods class with no prior statistics, research methods class with prior statistics, and a control group) on participants' scores of statistical literacy across two time periods. There was no significant interaction between class type and time, p = .88, and no substantial effect for time, p = .85. However, the main effect comparing the types of classes was significant, p = 0. Because the main effect was significant, post-hoc comparisons followed.

Table 17

Knowledge Element: Statistical Literacy

Mixed Between-Within Subjects Analysis of Variance

Multivariate Tests

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
time	Wilks' Lambda	1.000	.034	1.000	106.000	.854	.000
time * classtyp	Wilks' Lambda	.994	.227	3.000	106.000	.877	.006

Table 18

Knowledge Element: Statistical Literacy

Tests of Between-Subjects Effects

Measure: MEASURE_1
Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Intercept	22.605	1	22.605	425.295	.000	.800
classtyp	1.959	3	.653	12.285	.000	.258
Error	5.634	106	.053			

Post-hoc comparisons using Fisher's LSD showed three significant results, for the dependent variable statistical literacy, and are summarized in Table 19. First, results showed a statistically significant difference on pre- and post-test scores between the research methods class with prior statistics and the statistics class, p = .0, indicating that the research methods class with prior statistics scored higher on both the pre-test and post-test scores, while the scores on the post-test for the statistics changed only slightly.

Second, results showed a statistically significant difference on pre- and post-test scores between the research methods class with prior statistics and the control group, $p = 0$, indicating that the research methods class with prior statistics had higher pre-test scores and increased post-test scores, while the control group had lower pre-test scores and decreased post-test scores.

Third, results showed a statistically significant difference on pre- and post-test scores between the research methods class with prior statistics and the research methods class with no prior statistics, $p = .006$, indicating that research methods class with no prior statistics had lower pre-test scores than research methods class with prior statistics and decreased post-test scores, while the research methods class with prior statistics had increased scores between the pre- and post-tests. The estimated marginal means for statistical reasoning and class types are illustrated in Table 20.

Table 19

Knowledge Element: Statistical Literacy

Post-hoc Comparisons

Multiple Comparisons

Measure: MEASURE_1
LSD

(I) Class type	(J) Class type	Sig.
Statistics	Control	.784
	Research methods w/o statistics	.107
	Research methods after statistics	.000
Control	Statistics	.784
	Research methods w/o statistics	.080
	Research methods after statistics	.000
Research methods w/o statistics	Statistics	.107
	Control	.080
	Research methods after statistics	.006
Research methods after statistics	Statistics	.000
	Control	.000
	Research methods w/o statistics	.006

Table 20

Knowledge Element: Statistical Literacy

Estimated Marginal Means

2. Class type * time

Measure: MEASURE_1

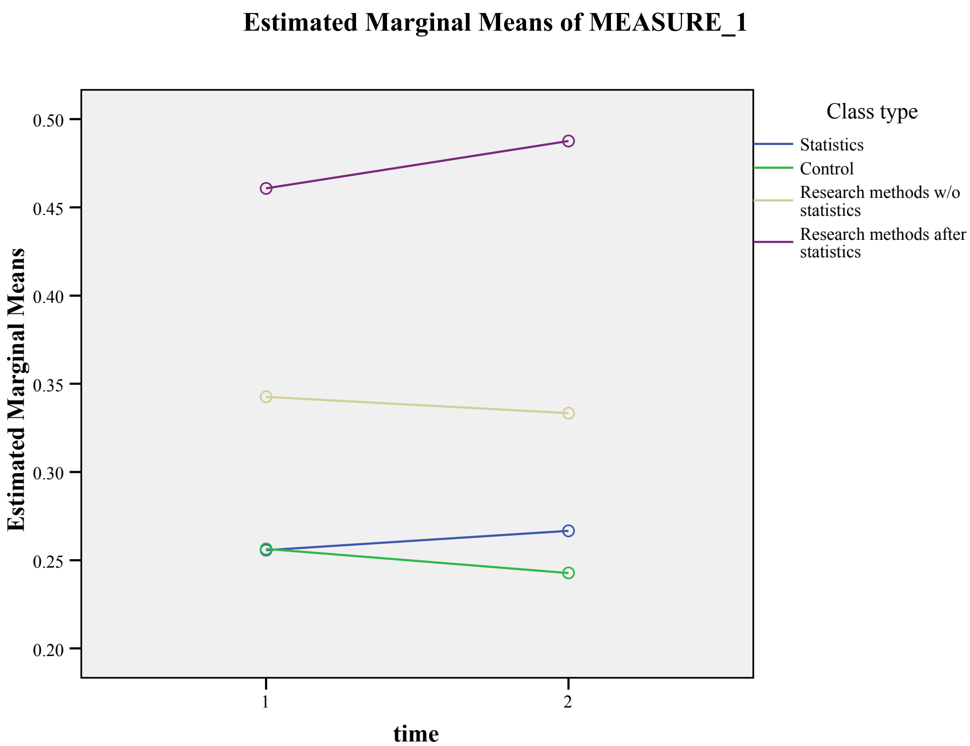
Class type	time	Mean	Std. Error	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
Statistics	1	.256	.032	.192	.319
	2	.267	.033	.201	.332
Control	1	.256	.037	.183	.330
	2	.243	.038	.166	.319
Research methods w/o statistics	1	.343	.045	.254	.431
	2	.333	.046	.242	.425
Research methods after statistics	1	.461	.034	.394	.528
	2	.488	.035	.418	.557

Increases and decreases in pre- to post-test scores among the class types for the dependent variable, statistical literacy, are displayed in Chart 3.

Chart 3

Knowledge Element: Statistical Literacy

Pre- to Post-Test Scores



Gender: Paired-Samples T-Tests. To further explore learning gains between pre- and post-test scores on the knowledge elements, statistical thinking, reasoning, and literacy, the data was analyzed by gender on six paired-samples t-tests.

Learning gains between pre- and post-tests of the dependent variables statistical thinking, reasoning and literacy were analyzed with females as the independent variable through three paired-samples t-tests, summarized in Table 21. For females, no significant difference was found between the pre- and post- test scores on statistical thinking, p =

.96, and statistical literacy, p = .95, but, a statistically significant difference resulted between the pre- and post-scores and on statistical reasoning, p = 0. And as illustrated in Table 22, an increase can be seen for females between pre- to post-test scores on their statistical reasoning skills.

Table 21

Knowledge Elements: Paired-Samples T-Test Females

Paired Samples Test

			Paired Differences				
What is your sex?			Mean	Std. Deviation	t	df	Sig. (2-tailed)
female	Pair 1	Pre statistical thinking questions total - Post statistical thinking questions total	-.00282	.46225	-.047	58	.963
	Pair 2	Pre statistical reasoning questions total - Post statistical reasoning questions total	-.12006	.23885	-3.861	58	.000
	Pair 3	Pre statistical literacy questions total - Post statistical literacy questions total	.00198	.22686	.067	58	.947

Table 22

Knowledge Elements: Females, Means and Standard Deviations

Paired Samples Statistics

What is your sex?			Mean	N	Std. Deviation
female	Pair 1	Pre statistical thinking questions total	.6324	59	.43672
		Post statistical thinking questions total	.6352	59	.28924
	Pair 2	Pre statistical reasoning questions total	.3322	59	.21945
		Post statistical reasoning questions total	.4523	59	.25296
	Pair 3	Pre statistical literacy questions total	.3571	59	.19839
		Post statistical literacy questions total	.3551	59	.21872

Likewise, learning gains between pre- and post-test scores for the dependent variables statistical thinking, reasoning, and literacy were analyzed with males as the independent variable through three paired-samples t-tests, summarized in Table 23. For males, there was no statistically significant difference between the pre- and post- test scores on statistical thinking, $p = .083$, reasoning, $p = .56$ and literacy, $p = .26$. Means and standard deviations for males are summarized in Table 24.

Table 23

Knowledge Elements: Paired-Samples T-Tests, Males

Paired Samples Test

What is your sex?			Paired Differences Mean	Paired Differences Std. Deviation	t	df	Sig. (2-tailed)
male	Pair 1	Pre statistical thinking questions total - Post statistical thinking questions total	-.08200	.22628	-1.812	24	.083
	Pair 2	Pre statistical reasoning questions total - Post statistical reasoning questions total	-.01933	.16200	-.597	24	.556
	Pair 3	Pre statistical literacy questions total - Post statistical literacy questions total	-.04667	.20138	-1.159	24	.258

Table 24

Knowledge Elements: Males, Means and Standard Deviations

Paired Samples Statistics

What is your sex?			Mean	N	Std. Deviation
male	Pair 1	Pre statistical thinking questions total	.5113	25	.21184
		Post statistical thinking questions total	.5933	25	.26387
	Pair 2	Pre statistical reasoning questions total	.4207	25	.18741
		Post statistical reasoning questions total	.4400	25	.18559
	Pair 3	Pre statistical literacy questions total	.3333	25	.22567
		Post statistical literacy questions total	.3800	25	.25240

First-Generation Status: Paired-Samples T-Tests. To further explore learning gains between pre-and post-test scores on the knowledge elements, statistical thinking, reasoning, and literacy, the data was analyzed by first-generation status on six paired-samples t-tests.

As illustrated in Table 25, there was no statistically significant differences for first-generation college students on pre- to post-test scores for statistical thinking, $p = .92$ and literacy, $p = .18$. But, results showed a significant difference between the pre- and post-test scores on statistical reasoning, $p = .001$. And as shown on Table 26, an increase can be seen for first-generation adult college students between pre- to post-test scores on their statistical reasoning skills.

Table 25

Knowledge Elements: Paired-Samples T-Tests, First-Generation Student

Paired Samples Test

First generation college student			Paired Differences Mean	Paired Differences Std. Deviation	t	df	Sig. (2-tailed)
yes	Pair 1	Pre statistical thinking questions total - Post statistical thinking questions total	-.00794	.51998	-.099	41	.922
	Pair 2	Pre statistical reasoning questions total - Post statistical reasoning questions total	-.12897	.23074	-3.622	41	.001
	Pair 3	Pre statistical literacy questions total - Post statistical literacy questions total	-.04762	.22624	-1.364	41	.180

Table 26

Knowledge Elements: Means and Standard Deviations, First-Generation

Paired Samples Statistics

First generation college student			Mean	N	Std. Deviation
yes	Pair 1	Pre statistical thinking questions total	.5908	42	.50000
		Post statistical thinking questions total	.5987	42	.30107
	Pair 2	Pre statistical reasoning questions total	.3476	42	.21816
		Post statistical reasoning questions total	.4766	42	.25050
	Pair 3	Pre statistical literacy questions total	.2806	42	.18086
		Post statistical literacy questions total	.3282	42	.21122

For those who were not first-generation college students no statistically significant differences resulted on scores between the pre- and post-test scores on statistical thinking, p = .13, reasoning, p =.09, and literacy, p = .64. Table 27 provides a summary of the results.

Table 27

Knowledge Elements: Paired-Samples T-Tests, Not First-Generation

Paired Samples Test

			Paired Differences				
First generation college student			Mean	Std. Deviation	t	df	Sig. (2-tailed)
no	Pair 1	Pre statistical thinking questions total - Post statistical thinking questions total	-.05813	.24052	-1.548	40	.130
	Pair 2	Pre statistical reasoning questions total - Post statistical reasoning questions total	-.05650	.20966	-1.726	40	.092
	Pair 3	Pre statistical literacy questions total - Post statistical literacy questions total	.01504	.20608	.467	40	.643

Gender: Independent-Samples T-Tests. Gender differences between males and females were examined on post-test scores of the knowledge elements, statistical thinking, reasoning, and literacy by three independent-samples t-tests.

Independent-samples t-tests were conducted to compare each of the knowledge elements, between males and females. No statistically significant differences were found in post-test scores between males and females on statistical thinking, p = .13, reasoning, p = .83 and literacy, p = .65. Table 28 summarizes the results.

Table 28

Knowledge Elements: Independent-Samples T-Test, Gender

Independent Samples Test

		Gender		
		t	df	Sig. (2-tailed)
Post statistical thinking questions total	Equal variances assumed	.622	82	.535
Post statistical reasoning questions total	Equal variances assumed	.218	82	.828
Post statistical literacy questions total	Equal variances assumed	-.456	82	.650

First-Generation Status: Independent-Samples T-Tests. First-generation status differences between those who were a first-generation adult student and those who were not a first-generation adult student were examined on post-test scores of the knowledge elements, statistical reasoning, thinking and literacy, by three independent-samples t-tests.

An independent-samples t-test was conducted to compare the knowledge elements first-generation and not first-generation adult college students. As shown on Table 29, there were no statistically significant differences on scores between first-generation and not first-generation adult college students on statistical thinking, $p = .33$, reasoning, $p = .38$ and literacy, $p = .14$.

Table 29

Knowledge Elements: First-Generation Status, Independent-Samples T-Test

Independent Samples Test

		First-Generation Status		
		t	df	Sig. (2-tailed)
Post statistical thinking questions total	Equal variances assumed	-.980	81	.330
Post statistical reasoning questions total	Equal variances assumed	.904	81	.369
Post statistical literacy questions total	Equal variances assumed	-1.492	81	.139

ANCOVA: Statistical Thinking, Reasoning and Literacy

To examine for pre-test influence on post-test scores on the knowledge elements, three one-way between-groups of analyses of covariance were conducted to compare the effectiveness of four different types of classes, the independent variable (i.e., statistics, research methods class with prior statistics, research methods class with no prior statistics and a control group) and the dependent variable, participants' post-test scores on each of the knowledge elements (i.e., statistical thinking, reasoning and literacy). Participants' scores on the pre-test scores of each of the knowledge elements were used as the covariate in these analyses.

Preliminary checks were conducted to ensure that there were no violations of the assumptions of normality, linearity, homogeneity of variances, homogeneity of regression slopes, and reliable measurement of the covariate.

After adjusting for pre-test scores on each of the knowledge elements, there were statistically significant differences among the four types of classes, statistics, research methods class with prior statistics, research methods class with no prior statistics and the control group. Results showed the following results for post-test scores; thinking, $p = 0$, reasoning, $p = .014$, and literacy, $p = .018$.

In addition there was a small relationship between the pre- and post-test scores on thinking as indicated by a partial eta-squared value of .047; a moderate relationship between pre-and post-test scores on reasoning as indicated by a partial eta-squared value of .197; a small relationship between pre- and post-test scores on literacy as indicated by a partial eta-squared value of .019. These results are summarized in Tables 30, 31 and 32.

Table 30

Knowledge Element: Statistical Thinking, ANCOVA

Tests of Between-Subjects Effects

Dependent Variable: Post statistical thinking questions total

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
PRESTQTOTA	.300	1	.300	5.161	.025	.047
classtyp	2.554	3	.851	14.651	.000	.295

Table 31

Knowledge Element: Statistical Reasoning, ANCOVA

Tests of Between-Subjects Effects

Dependent Variable: Post statistical reasoning questions total

Source	Type III Sum of Squares	df	F	Sig.	Partial Eta Squared
PRESRQTOTAL	.997	1	25.727	.000	.197
classtyp	.433	3	3.724	.014	.096

Table 32

Knowledge Element: Statistical Literacy, ANCOVA

Tests of Between-Subjects Effects

Dependent Variable: Post statistical literacy questions total

Source	Type III Sum of Squares	df	F	Sig.	Partial Eta Squared
PRESKQTOTAI	.768	1	24.358	.000	.188
classtyp	.333	3	3.524	.018	.091

Estimated marginal means for post-test scores on thinking showed the following results: the research methods class with prior statistics had the highest means in each component—statistical thinking, reasoning, and literacy. Estimated marginal means for the knowledge element statistical thinking were, from the highest to lowest, the research methods class with prior statistics, followed by the research methods class with no prior statistics, statistics and the control group. These are displayed in Table 33.

Estimated marginal means for the knowledge element statistical reasoning, from highest to lowest, were; the research methods class with prior statistics, followed by the research methods class with no prior statistics, statistics, and the control group. These results are displayed in Table 34.

Estimated marginal means for the knowledge element statistical literacy, from highest to lowest, were; the research methods class with prior statistics, followed by research methods class with no prior statistics, statistics, and the control group. Results are displayed in Table 35. These results indicate that after controlling for the effect of the pre-test scores on the post-test scores, means for all three knowledge elements, statistical thinking, reasoning, and literacy, the research methods class with prior statistics scored higher than the other groups.

Table 33

Knowledge Element: Statistical Thinking

Estimated Marginal Means

Class type

Dependent Variable: Post statistical thinking questions total

Class type	Mean	Std. Error	95% Confidence Interval Lower Bound	Upper Bound
Statistics	.482	.041	.401	.563
Control	.339	.048	.244	.435
Research methods w/o statistics	.658	.057	.545	.771
Research methods after statistics	.745	.044	.657	.833

Table 34

Knowledge Element: Statistical Reasoning

Estimated Marginal Means

Class type

Dependent Variable: Post statistical reasoning questions total

Class type	Mean	Std. Error	95% Confidence Interval Lower Bound	Upper Bound
Statistics	.407	.034	.340	.473
Control	.325	.039	.248	.403
Research methods w/o statistics	.398	.046	.306	.490
Research methods after statistics	.505	.037	.433	.578

Table 35

Knowledge Elements: Statistical Literacy

Estimated Marginal Means

Class type

Dependent Variable: Post statistical literacy questions total

			95% Confidence Interval	
Class type	Mean	Std. Error	Lower Bound	Upper Bound
Statistics	.299	.031	.238	.360
Control	.275	.035	.205	.345
Research methods w/o statistics	.327	.042	.244	.410
Research methods after statistics	.428	.034	.360	.495

Critical Questions

The fourth knowledge element in Gal's model is critical questions. It is recommended by Gal (2004) that adults need to ask the worry questions about statistical messages when interpreting any type of research. To evaluate participants' critical questioning skills, participants read a short research article that was published in national newspapers and were asked to examine the list of worry questions posed by Gal, and then were asked if these have any application to the research they have just read, and if so, what would these be.

Critical Questions: Mixed Between-Within Subjects ANOVAs. Mixed between-within subjects analyses were conducted to assess the impact of four different types of classes (i.e., statistics, research methods class with prior statistics, research methods class with no prior statistics, and a control group) across two time periods (pre- and post-test scores of the critical questions scale). There was no significant interaction between class type and time, $p = .10$, and no substantial effect for time, $p = .59$, as shown on Table 36.

The main effect in comparing class types was significant, p = .03, as shown on Table 38, which suggests that the class types, statistics, research methods class with prior statistics, research methods class with no prior statistics and the control group, contributed to differences in the scores on the critical questions.

Table 36

Knowledge Element: Critical Questions

Mixed Between-Within Subjects Analysis of Variance

Multivariate Tests

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
time	Wilks' Lambda	.997	.297	1.000	106.000	.587	.003
time * classtyp	Wilks' Lambda	.943	2.140	3.000	106.000	.100	.057

Table 37

Knowledge Element: Critical Questions

Between Subjects

Tests of Between-Subjects Effects

Measure: MEASURE_1
Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Intercept	1072.485	1	1072.485	93.668	.000	.469
classtyp	109.792	3	36.597	3.196	.026	.083
Error	1213.690	106	11.450			

Critical Questions: Post-Hoc Fisher's LSD. Because the main effect in comparing class types was significant, post-hoc comparisons using Fisher's LSD followed, and results showed three significant results for the dependent variable critical questions and are summarized in Table 38. First, results showed a statistically significant difference on pre- and post–test scores between the research methods class with no prior statistics and the control group, $p = .047$, indicating that the research methods class with no prior statistics scored higher, and increased scores from pre- to post-test.

Second, results showed a statistically significant difference on pre- and post-test scores between the research methods class with prior statistics and the control group, $p = .007$, indicating that the research methods class with prior statistics scored higher, and increased scores from pre- to post-test. And third, results showed a statistically significant difference on pre- and post-test scores between the research methods class with prior statistics and the statistics class, $p = .037$, indicating that the research methods class with prior statistics scored higher, and increased scores from pre- to post-test. The estimated marginal means for critical questions and class types are illustrated in Table 39.

Table 38

Knowledge Element: Critical Questions

Post-Hoc Comparisons

Multiple Comparisons

Measure: MEASURE_1
LSD

(I) Class type	(J) Class type	Sig.
Statistics	Control	.418
	Research methods w/o statistics	.165
	Research methods after statistics	.037
Control	Statistics	.418
	Research methods w/o statistics	.047
	Research methods after statistics	.007
Research methods w/o statistics	Statistics	.165
	Control	.047
	Research methods after statistics	.696
Research methods after statistics	Statistics	.037
	Control	.007
	Research methods w/o statistics	.696

Table 39

Knowledge Element: Critical Questions

Estimated Marginal Means

2. Class type * time

Measure: MEASURE_1

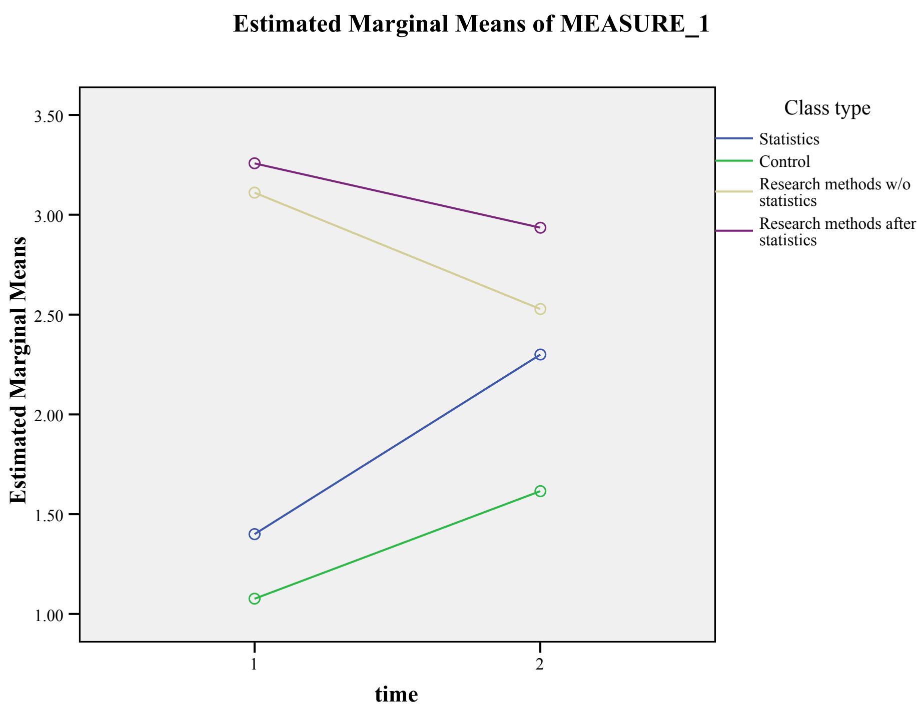
Class type	time	Mean	Std. Error	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
Statistics	1	1.400	.428	.551	2.249
	2	2.300	.481	1.346	3.254
Control	1	1.077	.497	.091	2.062
	2	1.615	.558	.508	2.722
Research methods w/o statistics	1	3.111	.597	1.927	4.296
	2	2.528	.671	1.197	3.858
Research methods after statistics	1	3.258	.455	2.356	4.161
	2	2.935	.511	1.922	3.949

Increases and decreases in pre to post-test scores among class types for the dependent variable, critical questions, are displayed in Chart 4.

Chart 4

Knowledge Element: Critical Questions

Pre- Post-Test Scores



Critical Questions: ANOVA

Further analysis was completed on post-test scores of the critical questions and class types. A one-way ANOVA was conducted to explore the impact of class types on the post-test scores on the critical questions. There were four groups according to the type of classes adult students were enrolled in: statistics, research methods class with no prior statistics, research methods class with prior statistics and a control group. There were no statistically significant differences at the $p < .05$ level in scores for the four groups. A summary of the ANOVA results are shown in Table 40.

Table 40

Knowledge Element: Critical Questions

ANOVA

Post critical question total

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	25.312	3	8.437	1.041	.378
Within Groups	859.361	106	8.107		
Total	884.673	109			

Gender: Paired-Samples T-Tests. Paired-samples t-tests were conducted to evaluate the impact of gender on the critical questions. No significant difference was found between pre- and post-test scores of the critical questions, p = .92, for females as shown in Table 41.

Table 41

Knowledge Elements: Critical Question

Female: Paired-Samples t-Test

Paired Samples Test

			Paired Differences				
What is your sex?			Mean	Std. Deviation	t	df	Sig. (2-tailed)
female	Pair 1	Pre critical question total - Post critical question total	.28814	2.54308	.870	58	.388

However, a statistically significant difference was found for males between pre- and post-test scores on the critical questions, p = .03, as shown in Table 42. A summary of the means and standard deviations for males and females on the critical questions can be seen on Table 43.

Table 42

Knowledge Elements: Critical Questions

Males: Paired-Samples t-Test

Paired Samples Test

What is your sex?			Paired Differences Mean	Paired Differences Std. Deviation	t	df	Sig. (2-tailed)
male	Pair 1	Pre critical question total - Post critical question total	-1.12000	2.41644	-2.317	24	.029

Table 43

Critical Questions: Means and Standard Deviations, Gender

Paired Samples Statistics

What is your sex?			Mean	N	Std. Deviation
female	Pair 1	Pre critical question total	2.6610	59	2.85639
		Post critical question total	2.3729	59	2.79098
male	Pair 1	Pre critical question total	1.9600	25	2.49132
		Post critical question total	3.0800	25	2.91433

First-Generation Status: Paired-Samples T-Tests. Paired-samples t-tests were conducted to evaluate the impact of first-generation status on scores of the critical questions. As shown on Table 44, no statistically significant difference was found between pre- and post-test scores of the critical questions, $p = .38$, for adult students who were first-generation.

Table 44

Knowledge Element: Critical Questions

First-Generation Students: Paired-Samples t-Test

Paired Samples Test

First generation college student			Paired Differences				
			Mean	Std. Deviation	t	df	Sig. (2-tailed)
yes	Pair 1	Pre critical question total - Post critical question total	-.33333	2.42111	-.892	41	.377

Likewise, as displayed on Table 45, no statistically significant difference was found for adult students who were not first-generation between pre- and post-test scores on the critical questions, p = .87. A summary of the means and standard deviations for first-generation status and the critical questions can be found in Table 46.

Table 45

Knowledge Element: Critical Questions

Not First-Generation Students: Paired-Samples t-Test

Paired Samples Test

First generation college student			Paired Differences				
			Mean	Std. Deviation	t	df	Sig. (2-tailed)
no	Pair 1	Pre critical question total - Post critical question total	.07317	2.76713	.169	40	.866

Table 46

Knowledge Element: Critical Questions

First-Generation Status, Means and Standard Deviations

Paired Samples Statistics

First generation college student			Mean	N	Std. Deviation	Std. Error Mean
yes	Pair 1	Pre critical question total	2.2500	42	2.65782	.41011
		Post critical question total	2.5833	42	2.79354	.43105
no	Pair 1	Pre critical question total	2.7195	41	2.87456	.44893
		Post critical question total	2.6463	41	2.90526	.45373

Gender: Independent-Samples T-Tests. As shown in Table 47, an independent-samples t-test was conducted to compare the mean post-test scores on the critical questions scale between males and females. There was no statistically significant difference in scores for males and females, p = .30. Pre- and post-test scores means for males and females were extremely low, as displayed in Table 48.

Table 47

Knowledge Element: Critical Questions

Gender

Independent Samples Test

		Gender		
		t	df	Sig. (2-tailed)
Post critical question total	Equal variances assumed	-1.048	82	.298

Table 48

Knowledge Element: Critical Questions

Means and Standard Deviations, Gender

Paired Samples Statistics

What is your sex?			Mean	N	Std. Deviation
female	Pair 1	Pre critical question total	2.6610	59	2.85639
		Post critical question total	2.3729	59	2.79098
male	Pair 1	Pre critical question total	1.9600	25	2.49132
		Post critical question total	3.0800	25	2.91433

First-Generation College Student: Independent-Samples T-Test. An independent-samples t-test was conducted to compare the post-test scores of the critical questions for adult students who identified themselves as not being first-generation and those who were first-generation. As shown in Table 49, there was no statistically significant difference in scores between first-generation and not first-generation adult college students, $p = .92$. Mean post-test scores for both groups were extremely low, as shown in Table 50.

Table 49

Knowledge Element: Critical Questions

First-Generation Status

Independent Samples Test

		First-Generation Status		
		t	df	Sig. (2-tailed)
Post critical question total	Equal variances assumed	-.101	81	.920

Table 50

Knowledge Element: Critical Questions

Means and Standard Deviations, First-Generation Status

Paired Samples Statistics

First generation college student			Mean	N	Std. Deviation	Std. Error Mean
yes	Pair 1	Pre critical question total	2.2500	42	2.65782	.41011
		Post critical question total	2.5833	42	2.79354	.43105
no	Pair 1	Pre critical question total	2.7195	41	2.87456	.44893
		Post critical question total	2.6463	41	2.90526	.45373

C*ritical Questions: ANCOVA.* Because of a lack of statistically significant differences on the pre- and post-test scores, one more test was conducted to examine if the pre-test had affected the post-test scores. An ANCOVA was conducted to compare the effectiveness of four different class types on participants' scores on the critical questions. The independent variable was the type of class (i.e., statistics, research methods class with no prior statistics, research methods class with prior statistics, and a control group) and the dependent variable consisted of post-test scores from the critical questions. Adult students' scores on the pre-critical questions test were used as the covariate in this analysis.

Preliminary checks were conducted to ensure that there was no violation of the assumptions of normality, linearity, homogeneity of variances, homogeneity of regression slopes, and reliable measurement of the covariate. After adjusting the pre-test scores, there were no statistically significant differences among the four class types on post-test scores on critical questions, $p = .60$. There was a moderate relationship between the pre-test and post-test critical question scores, as indicated by a partial eta squared value of .34. A summary of the data is provided in Table 51.

Table 51

ANCOVA: The Critical Questions

Tests of Between-Subjects Effects

Dependent Variable: Post critical question total

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
PRECQUESTION	288.885	1	288.885	53.171	.000	.336
classtyp	10.245	3	3.415	.629	.598	.018

Dispositional Elements: Attitudes, Beliefs and Critical Stance

Attitudes are participants' feelings that shape their behavior toward statistics (e.g., studying habits, importance of, etc.) and were measured by using Schau's SATS scale that incorporates four elements, affect, cognitive competence, value and difficulty. These were first examined by a MANOVA, then by four mixed within-between subjects ANOVAs to examine pre- to post-test scores on each of the elements of the SATS, among class types. Next, each element of the SATS was examined by separate independent and paired t-tests with gender and first-generation status as the independent variables. All statistical analyses completed with the variables gender and first-generation status did not include data from the control group, to prevent skewing of the data. A follow-up ANCOVA was completed to examine if pre-test scores affected post-test scores.

The dispositional element, beliefs about statistics, was analyzed by using a chi-square to examine differences among class types and students' beliefs on four statements. These statements are categorical variables and consist of dichotomous responses, insecure or secure, will or will not, easy or not easy, and relevant or not relevant. These responses are part of statements that address beliefs, which were further examined by open-ended responses that ask students to respond to the word *because* at the end of each of these statements: *I feel insecure or secure when doing statistics problems because…; I will or will not make a lot of math errors in statistics because…;statistics formulas are easy or not easy to understand because…; and statistics is or is not relevant in my life because*… And because these are categorical data, they were examined according to themes.

An ANOVA was used to examine the final dispositional element, critical stance, by examining class types and post-test scores. To examine pre- to post-test gains or losses on critical stance among groups, a mixed between-within subjects ANOVA was conducted. Pre- to post-test gains or losses were also analyzed with gender and first-generation status as the independent variables by using four paired-samples t-tests. Next, to examine group differences between gender and first-generation status, two separate independent-samples t-tests were conducted on post-test scores of critical stance. Because of a lack of significant results, an ANCOVA was conducted to examine if the post-test scores may have been influenced by pre-test scores on the critical stance scale.

Attitudes: MANOVA

A MANOVA was performed to investigate differences in class types (i.e., independent variables) and attitudes (i.e., dependent variables) toward statistics. Attitudes were measured from each element of the post-test scores from Schau's (1991) SATS scale, which consists of four elements: affect, cognitive competence, value and difficulty.

Preliminary assumption testing was conducted to check for normality, linearity, univariate and multivariate outliers, homogeneity of variance-covariance matrices, and multicollinearity, with no serious violations noted. There were statistically significant differences among class types, statistics, research methods class with prior statistics, research methods class with no prior statistics, and the control group on the combined dependent variables, p = .002, as Table 52 illustrates. However, as Table 53 presents, when the results of the dependent variables were considered separately no significant differences were found on any of the four components, affect, cognitive competence, value and difficulty when using an adjusted Bonferroni adjusted alpha level of .0125.

Table 52

Dispositional Element: Combined Attitudes

Affect, Cognitive Competence, Value and Difficulty

Multivariate Tests

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
classtyp	Wilks' Lambda	.740	2.742	12.000	272.804	.002	.096

Table 53

Dispositional Element: Separate Attitudes

Affect, Cognitive Competence, Value and Difficulty

Tests of Between-Subjects Effects

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
classtyp	Post-test affect	7.118	3	2.373	1.717	.168	.046
	Post-test cognitive competence	10.609	3	3.536	3.439	.020	.089
	Post-test value	10.751	3	3.584	3.052	.032	.080
	Post-test difficulty	1.859	3	.620	.892	.448	.025

Attitudes: Mixed Between-Within Subjects ANOVAs

Four mixed between-within subjects analyses of variance were conducted to assess the impact of four different types of classes (i.e., statistics, research methods class with no prior statistics, research methods class with prior statistics and a control group) on participants' scores on the SATS (Schau, 1991), across two time periods, pre- and post-test scores by the scale elements, affect, cognitive competence, difficulty and value. As stated before, affect examines students' feeling about statistics; cognitive competence examines students' attitudes about their intellectual knowledge when engaging in statistics; value is students' attitudes about the worth, usefulness and relevance of statistics in their professional and personal life; and difficulty examines students' attitudes about how difficult they believe statistics is as a subject (Schau, et al., 1991).

Affect. For the element affect, there was no significant interaction between class type and time, $p = .34$, and no substantial effect for time, $p = .54$, as shown on Table 54. However, the main effect for comparing four types of classes was significant, $p = .045$, as displayed on Table 55. Because the mixed between-within subjects analyses of variance found a significant main effect, further analyses were completed post-hoc.

Table 54

Dispositional Element: Attitudes, Affect

Interaction and Time Effect

Multivariate Tests

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
time	Wilks' Lambda	.996	.377	1.000	106.000	.541	.004
time * classtyp	Wilks' Lambda	.969	1.130	3.000	106.000	.341	.031

Table 55

Dispositional Element: Attitudes, Affect

Main Effect

Tests of Between-Subjects Effects

Measure: MEASURE_1
Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Intercept	3374.801	1	3374.801	1588.221	.000	.937
classtyp	17.732	3	5.911	2.782	.045	.073
Error	225.239	106	2.125			

Post-hoc comparisons from Fisher's LSD test showed there were statistically significant outcomes for the dependent variable, affect, and are summarized in Table 56. First, the test indicated a statistically significant difference on pre- and post-test scores between the statistics and the research methods class with no prior statistics, indicating that adult students in the statistics class changed their feeling more negatively toward statistics at the end of the semester, whereas those in the research methods class with no prior statistics had more positive feelings toward statistics at the end of the semester.

Second, there was a statistically significant difference on pre- and post-test scores between the research methods class with no prior statistics, and the research methods class with prior statistics, indicating that adult students in the research methods class with prior statistics changed their feeling more positively toward statistics at the end of the semester. Likewise, adult students in the research methods class with no prior statistics did change their feeling more positively toward statistics at the end of the semester, but had lower pre- and post-test scores overall.

Third, there was a statistically significant difference on the pre- and post-test scores between the research methods class with no prior statistics and the control group, indicating that albeit both groups changed their feeling more positively toward statistics at the end of the semester, the control group scored higher than the research methods class with no prior statistics. Estimated marginal means for pre- and post-test scores the dispositional element, affect and class types are illustrated in Table 57.

Table 56

Dispositional Element, Attitude: Affect

Post-hoc Comparisons

Multiple Comparisons

Measure: MEASURE_1
LSD

(I) Class type	(J) Class type	Sig.
Statistics	Control	.532
	Research methods w/o statistics	.005
	Research methods after statistics	.392
Control	Statistics	.532
	Research methods w/o statistics	.033
	Research methods after statistics	.853
Research methods w/o statistics	Statistics	.005
	Control	.033
	Research methods after statistics	.041
Research methods after statistics	Statistics	.392
	Control	.853
	Research methods w/o statistics	.041

Table 57
Dispositional Element, Attitude: Affect

Estimated Marginal Means

Class type * time

Measure: MEASURE_1

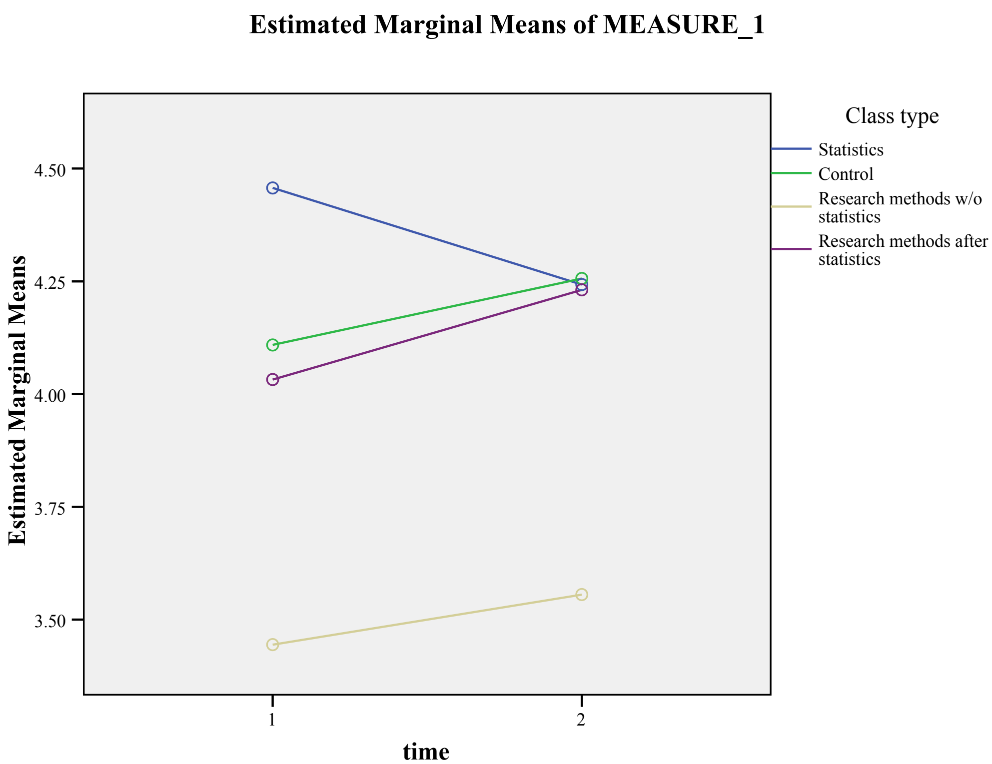
Class type	time	Mean	Std. Error	95% Confidence Interval Lower Bound	Upper Bound
Statistics	1	4.457	.189	4.082	4.832
	2	4.243	.199	3.849	4.637
Control	1	4.109	.219	3.674	4.544
	2	4.256	.231	3.799	4.713
Research methods w/o statistics	1	3.444	.264	2.922	3.967
	2	3.556	.277	3.006	4.105
Research methods after statistics	1	4.032	.201	3.634	4.430
	2	4.231	.211	3.813	4.650

Increases and decreases in pre- to post-test scores among class, for the dependent variable, affect, are graphically displayed on Chart 5.

Chart 5

Dispositional Element: Attitudes, Affect

Pre- to Post-Test Increases and Decreases



Cognitive Competence. For cognitive competence, there was no significant interaction between class type and time, p = .31, and there was no substantial effect for time, p = .96, as shown in Table 58. However, the main effect for comparing four types of classes was significant, p = .002, as illustrated in Table 59. Because the mixed between-within subjects analyses of variance found a significant main effect, further analyses were completed post-hoc.

Table 58

Dispositional Element: Attitudes, Cognitive Competence

Interaction and Time Effect

Multivariate Tests

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
time	Wilks' Lambda	1.000	.002	1.000	106.000	.962	.000
time * classtyp	Wilks' Lambda	.967	1.203	3.000	106.000	.312	.033

Table 59

Dispositional Element: Attitudes, Cognitive Competence

Main Effect

Tests of Between-Subjects Effects

Measure: MEASURE_1
Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Intercept	4716.111	1	4716.111	2738.783	.000	.963
classtyp	28.126	3	9.375	5.445	.002	.134
Error	182.529	106	1.722			

a. Computed using alpha = .05

Post-hoc comparisons from Fisher's LSD test showed three statistically significant outcomes for the dependent variable, cognitive competence, and are summarized in Table 60. First, the test indicated a statistically significant difference on pre- and post-test scores between the statistics and the research methods class with no prior statistics, indicating that adult students in the statistics class had maintained more positive attitudes toward their intellectual knowledge and skills when applied to statistics than students in the research methods class with no prior statistics.

Second, there was a statistically significant difference on pre- and post-test scores between the research methods class with no prior statistics and the research methods class with prior statistics. This indicates that adult students in the research methods class with prior statistics had developed more negative attitudes about their intellectual knowledge and skills when applied to statistics, whereas those in the research methods class with no prior statistics had developed more positive attitudes about their intellectual knowledge and skills when applied to statistics.

Third, there was a statistically significant difference on pre- and post-test scores between the statistics class and the control group. The statistics class had higher pre- and post-test scores than the control group, albeit the pre- and post-test scores were almost exactly the same. Similarly, the pre- and post-test scores for the control group showed almost no change; however, the pre- and post-test scores for the statistics class were higher than the pre- and post-test scores for the control group. This would indicate that both classes did not change their attitudes about their intellectual knowledge and skills when applied to statistics from the beginning of the semester to the end. Estimated marginal means for the dispositional element, cognitive competence, and class types are illustrated in Table 61.

Table 60

Dispositional Element, Attitude: Cognitive Competence

Post-hoc Comparisons

Multiple Comparisons

Measure: MEASURE_1
LSD

(I) Class type	(J) Class type	Sig.
Statistics	Control	.009
	Research methods w/o statistics	.001
	Research methods after statistics	.471
Control	Statistics	.009
	Research methods w/o statistics	.274
	Research methods after statistics	.057
Research methods w/o statistics	Statistics	.001
	Control	.274
	Research methods after statistics	.005
Research methods after statistics	Statistics	.471
	Control	.057
	Research methods w/o statistics	.005

Table 61

Dispositional Element, Attitude: Cognitive Competence

Estimated Marginal Means

Class type * time

Measure: MEASURE_1

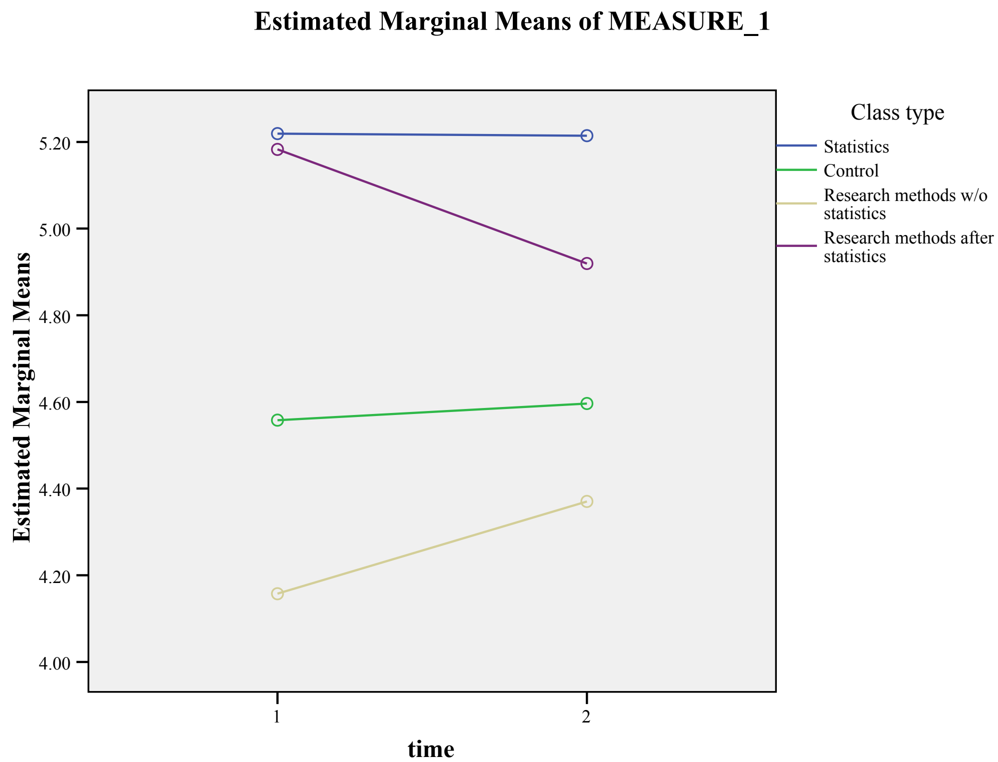
Class type	time	Mean	Std. Error	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
Statistics	1	5.219	.177	4.868	5.570
	2	5.214	.171	4.874	5.554
Control	1	4.558	.205	4.151	4.964
	2	4.596	.199	4.202	4.990
Research methods w/o statistics	1	4.157	.247	3.668	4.646
	2	4.370	.239	3.896	4.844
Research methods after statistics	1	5.183	.188	4.810	5.555
	2	4.919	.182	4.558	5.280

Increases and decreases in pre- to post-test scores among class, for the dependent variable, cognitive competence, are graphically displayed in Chart 6.

Chart 6

Dispositional Element: Attitudes, Cognitive Competence

Pre- to Post-Test Increases and Decreases



Value. For value, there was no significant interaction between class type and time, $p = .134$, and no substantial effect for time, $p = .057$, as shown in Table 62. However, results indicate that the main effect for comparing four types of classes was significant, $p = .02$, as shown in Table 63. Because the mixed between-within subjects analyses of variance found a significant main effect, further analyses were completed post-hoc.

Table 62

Dispositional Element: Attitudes, Value

Interaction and Time Effect

Multivariate Tests

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
time	Wilks' Lambda	.997	.334	1.000	106.000	.565	.003
time * classtyp	Wilks' Lambda	.949	1.899	3.000	106.000	.134	.051

Table 63

Dispositional Element: Value

Main Effect

Tests of Between-Subjects Effects

Measure: MEASURE_1
Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Intercept	4376.227	1	4376.227	2610.533	.000	.961
classtyp	17.456	3	5.819	3.471	.019	.089
Error	177.696	106	1.676			

Post-hoc comparisons from Fisher's LSD test showed two statistically significant outcomes for the dependent variable value, and are summarized in Table 64. First, results indicated a statistically significant difference on pre- and post-test scores between the research methods class with no prior statistics and the control group. This indicates that adult students in the research methods class with no prior statistics developed more negative attitudes about the usefulness, relevance, and worth of statistics in their personal and professional life, while those in the control group had developed more positive attitudes toward the value of statistics at the end of the semester.

Second, there was a statistically significant difference on pre- and post-test scores between the research methods class with prior statistics and the control group. This indicates that adult students in the research methods class with prior statistics had more positive attitudes than those in the control group about the usefulness, relevance, and worth of statistics in their personal and professional life. Estimated marginal means for the dispositional element, value, and class types are illustrated in Table 65.

Table 64

Dispositional Element, Attitude: Value

Post-hoc Comparisons

Multiple Comparisons

Measure: MEASURE_1
LSD

(I) Class type	(J) Class type	Sig.
Statistics	Control	.168
	Research methods w/o statistics	.248
	Research methods after statistics	.078
Control	Statistics	.168
	Research methods w/o statistics	.025
	Research methods after statistics	.003
Research methods w/o statistics	Statistics	.248
	Control	.025
	Research methods after statistics	.730
Research methods after statistics	Statistics	.078
	Control	.003
	Research methods w/o statistics	.730

Table 65

Dispositional Element, Attitude: Value

Estimated Marginal Means

Class type * time

Measure: MEASURE_1

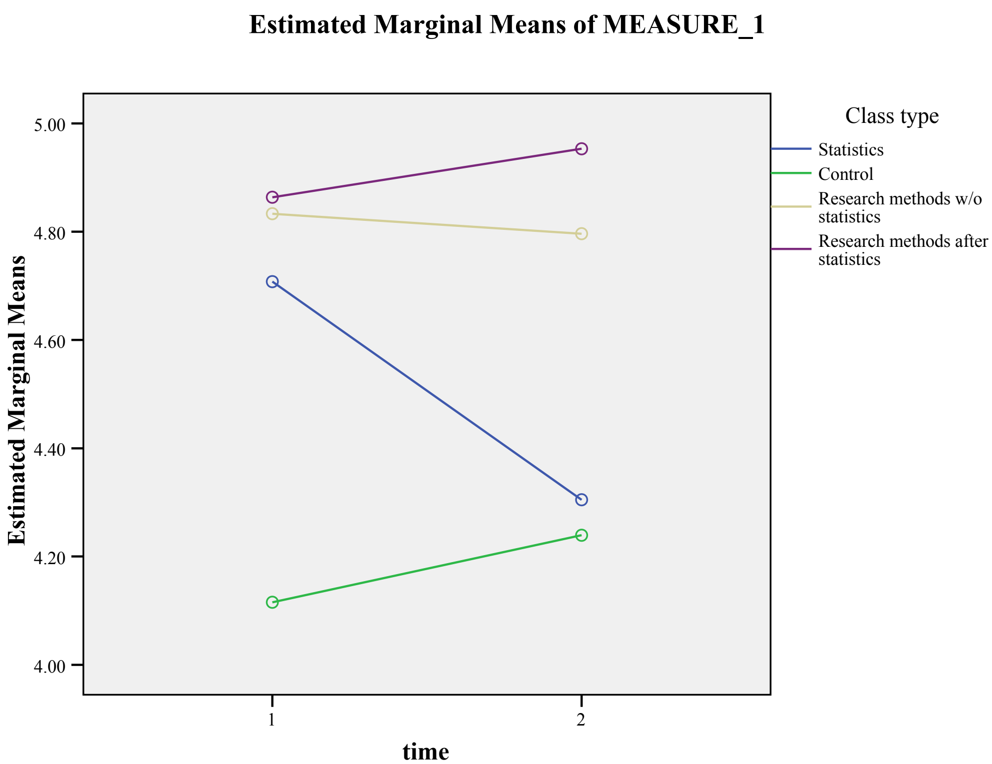
Class type	time	Mean	Std. Error	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
Statistics	1	4.708	.169	4.373	5.043
	2	4.305	.183	3.942	4.668
Control	1	4.115	.196	3.727	4.504
	2	4.239	.213	3.818	4.661
Research methods w/o statistics	1	4.833	.236	4.366	5.300
	2	4.796	.255	4.290	5.303
Research methods after statistics	1	4.864	.180	4.508	5.220
	2	4.953	.195	4.568	5.339

Increases and decreases in pre- to post-test scores among class, for the dependent variable, value, are graphically displayed in Chart 7.

Chart 7

Dispositional Element: Attitudes, Value

Pre- to Post-Test Increases and Decreases



Difficulty. For difficulty, there was no significant interaction between class type and time, p = .88, and no substantial effect for time, p = .83, as shown on Table 66. In addition, as displayed in Table 67, the main effect for comparing four types of classes was not significant, p = 1.91. This indicates that there were no changes in adult students' attitudes about the difficulty of statistics as a subject at the end of the semester. These results are graphically displayed in Chart 8.

Table 66

Dispositional Element: Attitudes, Difficulty

Interaction and Time Effect

Multivariate Tests

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
time	Wilks' Lambda	1.000	.045	1.000	106.000	.832	.000
time * classtyp	Wilks' Lambda	.994	.221	3.000	106.000	.882	.006

Table 67

Dispositional Element: Attitudes, Difficulty

Main Effect

Tests of Between-Subjects Effects

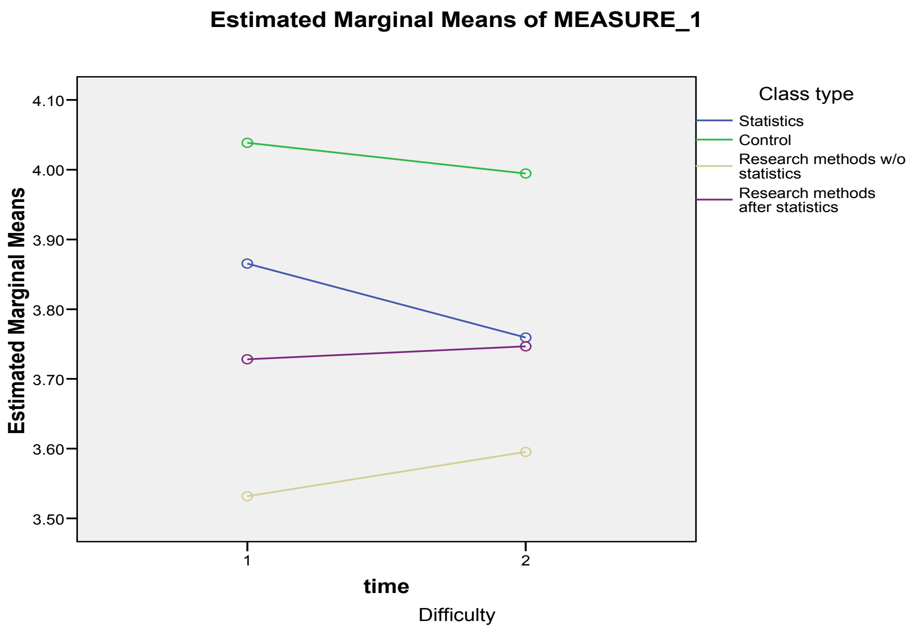
Measure: MEASURE_1
Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Intercept	2956.516	1	2956.516	3039.816	.000	.966
classtyp	4.704	3	1.568	1.612	.191	.044
Error	103.095	106	.973			

Chart 8

Dispositional Element: Attitudes, Difficulty

Pre- to Post-Test Increases and Decreases



Attitudes: *ANCOVA*

Four one-way between-groups of analyses of covariance were conducted to compare the effectiveness of four different types of classes (i.e., statistics, research methods with a prior statistics class, research methods class with no prior statistics and a control group) on scores from each of the elements on the SATS (i.e., affect, cognitive competence, value and difficulty). Participants' scores from each of the elements' pre-test scores of the SATS were used as the covariate in this analysis.

Preliminary checks were conducted to ensure that there were no violations of the assumptions of normality, linearity, homogeneity of variances, homogeneity of regressions slopes, and reliable measurement of the covariate.

After adjusting for pre-test scores from the SATS there were no significant differences among the four types of classes, statistics, the research methods class with prior statistics, research methods class with no prior statistics and the control group on post-test scores for affect, p = .531, cognitive competence, p = .54, value, p = .10 and difficulty, p = .89.

There was a moderate relationship between the pre-and post-test scores on affect as indicated by a partial eta-squared value of .38; a moderate relationship between pre- and post-test scores on cognitive competence as indicated by a partial eta-squared value of .39; a moderate relationship between pre- and post-test scores on value as indicated by a partial eta-squared value of .30 and a moderate relationship between pre- and post-test scores on difficulty as indicated by a partial eta-squared value of .24. Tables 68, 69, 70 and 71 provide details of these analyses.

Table 68

Dispositional Element: Attitudes, Affect

Tests of Between-Subjects Effects

Dependent Variable: Post-test affect

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
PREAFFTOT	55.465	1	55.465	64.003	.000	.379
classtyp	1.920	3	.640	.739	.531	.021

Table 69

Dispositional Element: Attitudes, Cognitive Competence

Tests of Between-Subjects Effects

Dependent Variable: Post-test cognitive competence

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
PRECCTOT	42.242	1	42.242	66.430	.000	.388
classtyp	1.390	3	.463	.729	.537	.020

Table 70

Dispositional Element: Attitudes, Value

Tests of Between-Subjects Effects

Dependent Variable: Post-test value

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
PREVALTOT	36.908	1	36.908	44.258	.000	.297
classtyp	5.416	3	1.805	2.165	.097	.058

Table 71

Dispositional Element: Attitudes, Difficulty

Tests of Between-Subjects Effects

Dependent Variable: Post-test difficulty

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
PREDIFFTOT	17.924	1	17.924	33.802	.000	.244
classtyp	.340	3	.113	.214	.887	.006

In addition the results from the ANCOVA provides the estimated marginal means by class types, according to each element of the SATS, affect, cognitive competence, value and difficulty. Details of these results are shown in Tables 72, 73, 74 and 75.

In sum, the estimated marginal means show us which class scored highest for each of the SATS elements. First, for affect, the estimated marginal means from highest to lowest were the research methods class with prior statistics, the control group, the statistics class, and the research methods class with no prior statistics.

Second, for cognitive competence, the estimated marginal means from highest to lowest were the statistics class, the research methods class with no prior statistics, the control group and the research methods class with prior statistics.

Third, for value, the estimated marginal means from highest to lowest were the research methods class with prior statistics, the research methods class with no prior statistics, the control group, and the statistics class.

And fourth, for difficulty, the estimated marginal means from highest to lowest were the control group, the research methods class with prior statistics, the research methods class with no prior statistics and the statistics class.

To summarize, after controlling for pre-test effects on post-test scores, the highest estimated marginal means for the groups were as follows: for affect, the research methods class with prior statistics; cognitive competence, statistics; value, the research methods class with prior statistics; and difficulty, the control group.

Table 72

Dispositional Element: Attitudes, Affect

Class type

Dependent Variable: Post-test affect

			95% Confidence Interval	
Class type	Mean	Std. Error	Lower Bound	Upper Bound
Statistics	4.005	.160	3.687	4.322
Control	4.244	.183	3.882	4.606
Research methods w/o statistics	3.973	.226	3.526	4.420
Research methods after statistics	4.268	.167	3.936	4.600

Table 73

Dispositional Element: Attitudes, Cognitive Competence

Class type

Dependent Variable: Post-test cognitive competence

			95% Confidence Interval	
Class type	Mean	Std. Error	Lower Bound	Upper Bound
Statistics	5.009	.137	4.737	5.281
Control	4.790	.158	4.476	5.104
Research methods w/o statistics	4.806	.195	4.418	5.193
Research methods after statistics	4.736	.145	4.448	5.023

Table 74

Dispositional Element: Attitudes, Value

Class type

Dependent Variable: Post-test value

			95% Confidence Interval	
Class type	Mean	Std. Error	Lower Bound	Upper Bound
Statistics	4.260	.155	3.954	4.566
Control	4.545	.185	4.178	4.911
Research methods w/o statistics	4.678	.216	4.249	5.106
Research methods after statistics	4.817	.165	4.489	5.144

Table 75

Dispositional Element: Attitudes, Difficulty

Class type

Dependent Variable: Post-test difficulty

Class type	Mean	Std. Error	95% Confidence Interval Lower Bound	Upper Bound
Statistics	3.732	.123	3.487	3.976
Control	3.876	.144	3.590	4.162
Research methods w/o statistics	3.743	.174	3.399	4.087
Research methods after statistics	3.791	.131	3.531	4.051

Gender: *Independent T-Tests.* An independent-samples t-test was conducted to compare the SATS total post-test scores between females and males. This test was found to be statistically significant at an alpha level of .05, p = .046, as shown on Table 76. This indicates that males have more positive overall attitudes toward statistics than females, as shown by total mean scores in Table 77. Additional independent-samples t-tests were used to analyze each element of the SATS, affect, cognitive competence, value, and difficulty with gender.

Table 76

Dispositional Elements: Affect, Cognitive Competence, Value and Difficulty

Independent Samples Test

		Gender t	df	Sig. (2-tailed)
POSTTOT	Equal variances assumed	-2.029	82	.046

Table 77

Dispositional Elements: Affect, Cognitive Competence, Value and Difficulty

Means and Standard Deviations

Group Statistics

	What is your sex?	N	Mean	Std. Deviation	Std. Error Mean
POSTTOT	female	59	4.2506	.73766	.09603
	male	25	4.6057	.72246	.14449

Four independent-samples t-tests compared the mean post-test scores for each element of the SATS, as shown on Table 78. For affect, there was a statistically significant difference between males and females, at an alpha level of .05, $p = .016$, indicating that males have more positive feelings toward statistics than females, as confirmed by mean post-test scores illustrated on Table 79. However, no statistically significant differences were found between males and females for the element, cognitive competence, $p = .056$; value, $p = .47$; and difficulty, $p = .42$.

Table 78

Dispositional Elements: Attitudes, Affect, Cognitive Competence, Value and Difficulty

Gender

Independent Samples Test

		Gender		
		t	df	Sig. (2-tailed)
Post affect	Equal variances assumed	-2.470	82	.016
Post Cognitive Competence	Equal variances assumed	-1.936	82	.056
Post Value	Equal variances assumed	-.726	82	.470
Post Difficulty	Equal variances assumed	-.805	82	.423

Table 79

Dispositional Elements: Attitudes, Affect, Cognitive Competence, Value and Difficulty

Gender: Means and Standard Deviations

Group Statistics

	What is your sex?	N	Mean	Std. Deviation
Post affect	female	59	3.8814	1.23062
	male	25	4.5867	1.10985
Post Cognitive Competence	female	59	4.7797	1.05611
	male	25	5.2667	1.04969
Post Value	female	59	4.5951	1.04797
	male	25	4.7778	1.06911
Post Difficulty	female	59	3.6707	.85684
	male	25	3.8343	.83877

First-Generation Students: *Independent-Samples T-Tests.* An independent-samples t-test was conducted to compare the SATS total post-test scores between adult students who were first-generation college students with those who were not first-generation. No statistically significant difference was found for first-generation status, $p = .32$, as shown in Table 80. Means and standard deviations of the total post-test scores for the SATS are displayed in Table 81. Additional, separate independent-samples t-tests on each element of the SATS were completed next.

Table 80

Dispositional Elements: Attitudes, Affect, Cognitive Competence, Value and Difficulty

First-Generation Status

Independent Samples Test

		First-Generation Status		
		t	df	Sig. (2-tailed)
POSTTOT	Equal variances assumed	1.002	81	.319

Table 81

Dispositional Elements: Attitudes, Affect, Cognitive Competence, Value and Difficulty

First-Generation Status: Means and Standard Deviations

Group Statistics

	First generation college student	N	Mean	Std. Deviation
POSTTOT	yes	42	4.4405	.67665
	no	41	4.2753	.82031

Four independent-samples t-tests were conducted to compare the post-test scores of each element of the SATS, affect, cognitive competence, value and difficulty with first-generation status. There were no statistically significant differences in scores between first-generation adult college students and those who were not first-generation for the element affect, $p = .15$, value, $p = .42$ and difficulty, $p = .43$. However, there was a statistically significant difference in scores between first-generation college students and those who were not first-generation for cognitive competence, $p = .04$. As indicated by the post-test means, as summarized in Table 83, first-generation adult college students scored higher than those who were not first-generation. The independent-samples t-tests results are summarized in Table 82.

Table 82

Dispositional Elements: Attitudes, Affect, Cognitive Competence, Value and Difficulty

Independent Samples Test

		First-Generation Status		
		t	df	Sig. (2-tailed)
Post affect	Equal variances assumed	1.453	81	.150
Post Cognitive Competence	Equal variances assumed	2.088	81	.040
Post Value	Equal variances assumed	-.805	81	.423
Post Difficulty	Equal variances assumed	.786	81	.434

Table 83

Dispositional Elements: Attitudes, Affect, Cognitive Competence, Value and Difficulty

First-Year Status, Means and Standard Deviations

Group Statistics

	First generation college student	N	Mean	Std. Deviation
Post affect	yes	42	4.2738	1.22021
	no	41	3.8821	1.23561
Post Cognitive Competence	yes	42	5.1587	.99044
	no	41	4.6748	1.11862
Post Value	yes	42	4.5767	1.01011
	no	41	4.7615	1.08143
Post Difficulty	yes	42	3.7925	.88974
	no	41	3.6446	.82197

Gender: Paired- Samples T-Tests. A paired-samples t-test was conducted to evaluate the impact of gender on the total pre- and post-test scores from the SATS scale. As displayed in Table 84, no statistically significant difference was found for females between the pre- and post-test scores, $p = .38$, or for males between the pre- and post-test scores, $p = .69$. Means and standard deviations for males and females are summarized in Table 85.

Table 84

Dispositional Elements: Attitudes, Affect, Cognitive Competence, Value and Difficulty

Paired Samples Test

What is your sex?			Paired Differences Mean	Paired Differences Std. Deviation	t	df	Sig. (2-tailed)
female	Pair 1	PRETOTAL - POSTTOT	.06840	.58895	.892	58	.376
male	Pair 1	PRETOTAL - POSTTOT	.04571	.55792	.410	24	.686

Table 85

Dispositional Elements: Attitudes, Affect, Cognitive Competence, Value and Difficulty

Gender: Means and Standard Deviations

Paired Samples Statistics

What is your sex?			Mean	N	Std. Deviation
female	Pair 1	PRETOTAL	4.3190	59	.75897
		POSTTOT	4.2506	59	.73766
male	Pair 1	PRETOTAL	4.6514	25	.62375
		POSTTOT	4.6057	25	.72246

Next, separate paired-samples t-tests examined each element of the SATS, affect, cognitive competence, value and difficulty, with females. No statistically significant differences were found for females between the pre- and post-test scores for affect, p = 1.00, cognitive competence, p = .36, value, p = .27; and difficulty, p = .93. A summary of the paired-samples t-tests is provided in Table 86 and a summary of the means and standard deviations are provided in Table 87.

Table 86

Dispositional Elements: Attitudes, Affect, Cognitive Competence, Value and Difficulty

Gender: Female

Paired Samples Test

What is your sex?			Paired Differences Mean	Std. Deviation	t	df	Sig. (2-tailed)
female	Pair 1	Pre test affect - Post affect	.00000	1.03715	.000	58	1.000
	Pair 2	Pre cognitive competence - Post Cognitive Competence	.10452	.86486	.928	58	.357
	Pair 3	Pre value - Post Value	.13559	.93648	1.112	58	.271
	Pair 4	Pre difficult - Post Difficulty	.00969	.86363	.086	58	.932

Table 87

Dispositional Elements: Attitudes, Affect, Cognitive Competence, Value and Difficulty

Female: Means and Standard Deviations

Paired Samples Statistics

What is your sex?			Mean	N	Std. Deviation
female	Pair 1	Pre test affect	3.8814	59	1.24069
		Post affect	3.8814	59	1.23062
	Pair 2	Pre cognitive competence	4.8842	59	1.09521
		Post Cognitive Competence	4.7797	59	1.05611
	Pair 3	Pre value	4.7307	59	.98891
		Post Value	4.5951	59	1.04797
	Pair 4	Pre difficult	3.6804	59	.84862
		Post Difficulty	3.6707	59	.85684

Likewise, separate paired t-tests examined each element of the SATS, affect, cognitive competence, value and difficulty with males. No statistically significant differences were found for males between the pre- and post-test scores on affect, $p = .89$, cognitive competence, $p = .69$, value, $p = .28$ and difficulty, $p = .72$. A summary of the paired t-tests is provided by Table 88 and the means and standard deviations are provided in Table 89.

Table 88

Dispositional Elements: Attitudes, Affect, Cognitive Competence, Value and Difficulty

Gender: Males

Paired Samples Test

What is your sex?			Paired Differences Mean	Std. Deviation	t	df	Sig. (2-tailed)
male	Pair 1	Pre test affect - Post affect	-.02667	.98446	-.135	24	.893
	Pair 2	Pre cognitive competence - Post Cognitive Competence	-.06667	.84574	-.394	24	.697
	Pair 3	Pre value - Post Value	.16000	.72512	1.103	24	.281
	Pair 4	Pre difficult - Post Difficulty	.05714	.77701	.368	24	.716

Table 89

Dispositional Elements: Attitudes, Affect, Cognitive Competence, Value and Difficulty

Males: Means and Standard Deviations

Paired Samples Statistics

What is your sex?			Mean	N	Std. Deviation
male	Pair 1	Pre test affect	4.5600	25	.90252
		Post affect	4.5867	25	1.10985
	Pair 2	Pre cognitive competence	5.2000	25	.96825
		Post Cognitive Competence	5.2667	25	1.04969
	Pair 3	Pre value	4.9378	25	.84503
		Post Value	4.7778	25	1.06911
	Pair 4	Pre difficult	3.8914	25	.77183
		Post Difficulty	3.8343	25	.83877

First- Generation Status: Paired-Samples T-Tests. A paired-samples t-test was conducted to evaluate the impact of first-generation status on the total pre- and post-test scores from the SATS scale. No statistically significant difference was found for first-generation adult students between the pre- and post-test scores, $p = .71$. Likewise, no statistically significant difference was found for adult students who were not first-generation between the pre- and post-test scores, $p = .22$. A summary of the paired t-tests is provided in Table 90 and the means and standard deviations for total post-test scores are provided in Table 91.

Table 90

Dispositional Elements: Attitudes, Affect, Cognitive Competence, Value and Difficulty

First-Generation Status

Paired Samples Test

First generation college student			Paired Differences Mean	Std. Deviation	t	df	Sig. (2-tailed)
yes	Pair 1	PRETOTAL - POSTTOT	.03656	.62177	.381	41	.705
no	Pair 1	PRETOTAL - POSTTOT	.10366	.52917	1.254	40	.217

Table 91

Dispositional Elements: Attitudes, Affect, Cognitive Competence, Value and Difficulty

First-Generation Status: Means and Standard Deviations

Paired Samples Statistics

First generation college student			Mean	N	Std. Deviation	Std. Error Mean
yes	Pair 1	PRETOTAL	4.4770	42	.70042	.10808
		POSTTOT	4.4405	42	.67665	.10441
no	Pair 1	PRETOTAL	4.3789	41	.76789	.11992
		POSTTOT	4.2753	41	.82031	.12811

Next, separate paired-samples t-tests examined each element of the SATS, affect, cognitive competence, value and difficulty, with first-generation status. No statistically significant differences were found for first-generation adult college students between the pre- and post-test scores on affect, $p = .79$, cognitive competence, $p = .22$, value, $p = .08$ and difficulty, $p = .75$. A summary of the paired-samples t-tests is provided in Table 92 and a summary of the means and standard deviations is provided in Table 93.

Table 92

Dispositional Elements: Attitudes, Affect, Cognitive Competence, Value and Difficulty

First-Generation Adult Students

Paired Samples Test

			Paired Differences				
First generation college student			Mean	Std. Deviation	t	df	Sig. (2-tailed)
yes	Pair 1	Pre test affect - Post affect	.04762	1.14781	.269	41	.789
	Pair 2	Pre cognitive competence - Post Cognitive Competence	-.16270	.84623	-1.246	41	.220
	Pair 3	Pre value - Post Value	.22222	.81057	1.777	41	.083
	Pair 4	Pre difficult - Post Difficulty	-.04082	.81382	-.325	41	.747

Table 93

Dispositional Elements: Attitudes, Affect, Cognitive Competence, Value and Difficulty

First-Generation Adult Students: Means and Standard Deviations

Paired Samples Statistics

First generation college student			Mean	N	Std. Deviation
yes	Pair 1	Pre test affect	4.3214	42	1.13958
		Post affect	4.2738	42	1.22021
	Pair 2	Pre cognitive competence	4.9960	42	1.18395
		Post Cognitive Competence	5.1587	42	.99044
	Pair 3	Pre value	4.7989	42	.84877
		Post Value	4.5767	42	1.01011
	Pair 4	Pre difficult	3.7517	42	.85369
		Post Difficulty	3.7925	42	.88974

Similarly, separate paired-samples t-tests examined each element of the SATS, affect, cognitive competence, value and difficulty, with first-generation status. No statistically significant differences were found for adult students who were not first-generation between the pre- and post-test scores on affect, $p = .97$, on value, $p = .81$, and on difficulty, $p = .41$. But, a statistically significant difference was found between the pre- and post-test scores on cognitive competence, $p = .02$. A summary of the paired-simples t-tests is provided by Table 94 and means and standard deviations for the elements are shown in Table 95.

Table 94

Dispositional Elements: Attitudes, Affect, Cognitive Competence, Value and Difficulty

Not a First-Generation Adult Student

Paired Samples Test

First generation college student			Paired Differences				
			Mean	Std. Deviation	t	df	Sig. (2-tailed)
no	Pair 1	Pre test affect - Post affect	-.00407	.79538	-.033	40	.974
	Pair 2	Pre cognitive competence - Post Cognitive Competence	.30488	.80096	2.437	40	.019
	Pair 3	Pre value - Post Value	.03523	.93076	.242	40	.810
	Pair 4	Pre difficult - Post Difficulty	.11150	.85685	.833	40	.410

Table 95

Dispositional Elements: Attitudes, Affect, Cognitive Competence, Value and Difficulty

Not a First-Generation Adult Student: Means and Standard Deviations

Paired Samples Statistics

First generation college student			Mean	N	Std. Deviation
no	Pair 1	Pre test affect	3.8780	41	1.19433
		Post affect	3.8821	41	1.23561
	Pair 2	Pre cognitive competence	4.9797	41	.94442
		Post Cognitive Competence	4.6748	41	1.11862
	Pair 3	Pre value	4.7967	41	1.05903
		Post Value	4.7615	41	1.08143
	Pair 4	Pre difficult	3.7561	41	.80891
		Post Difficulty	3.6446	41	.82197

Beliefs

A chi-square test was used to examine differences among the independent variable, class types, statistics, research methods class with no prior statistics, research methods class with prior statistics and a control group with students' beliefs on these four dependent variables: *(a) Statistics formulas are easy or not easy to understand, (b) I will or will not make a lot of math errors in statistics, (c) I will feel insecure or secure when doing statistics problems, and (d) statistics is or is not relevant in my life.* Each test result will be listed below. The dependent variables were dichotomous responses for each statement: easy or not easy, will or will not, insecure or secure and is or is not relevant.

A chi-square test for independence indicated no significant association between class types and students' beliefs that statistics formulas are easy or not easy to understand (see Table 96); no significant association for they will or will not make a lot of math errors in statistics; (see Table 97); and no significant association for they will feel insecure or secure when doing statistics problems (see Table 98).

The only belief that showed a significant association was students' belief that statistics is or is not relevant in my life (see Table 99). Results from Table 100 showed 90% of adult students in the research methods class with prior statistics believed that statistics is relevant in their lives, while only 46% of adult students in the control group believed that statistics was relevant in their lives.

Table 96

Statistics Formulas are Easy or Not Easy to Understand

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	4.998	3	.172
N of Valid Cases	107		

Table 97

I will or will not make a lot of math errors in statistics

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	6.254	3	.100
N of Valid Cases	107		

Table 98

I will feel insecure or secure when doing statistics problems

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	7.112	3	.068
N of Valid Cases	107		

Table 99

Statistics is or is not relevant in my life

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	12.873	3	.005
N of Valid Cases	107		

Table 100

Frequencies: Statistics is or is not relevant in my life.

Statistics is or is not relevant in my life because

Class type			Percent	Valid Percent	Cumulative Percent
Statistics	Valid	is relevant	62.9	62.9	62.9
		is not relevant	37.1	37.1	100.0
		Total	100.0	100.0	
Control	Valid	is relevant	42.3	47.8	47.8
		is not relevant	46.2	52.2	100.0
		Total	88.5	100.0	
	Missing	System	11.5		
	Total		100.0		
Research methods w/o statistics	Valid	is relevant	77.8	77.8	77.8
		is not relevant	22.2	22.2	100.0
		Total	100.0	100.0	
Research methods after statistics	Valid	is relevant	90.3	90.3	90.3
		is not relevant	9.7	9.7	100.0
		Total	100.0	100.0	

Open-ended responses. Participants' responses to the pre- and post-test belief statements were extended by adding the word *because* at the end of each statement to gather open-ended responses: *I will feel insecure or secure when doing statistics problems because…I will or will not make a lot of math errors in statistics because…statistics formulas are easy or not easy to understand because…and statistics is or is not relevant in my life because…* Each statement from the open-ended responses was coded according to the initial responses, for example insecure or secure, and were compiled and then separated according to themes. Each of these statements will be examined next with their corresponding themes.

*I will feel insecure or secure when doing statistic problems because…*Three underlying themes emerged from participants' responses to the statement, *I feel insecure* *or secure when doing statistics problems because…*for both types of responses, secure or insecure. Themes centered on students' beliefs in their mathematical abilities, their sense of personal responsibility and a belief in their abilities to master the course content, and confidence or lack of confidence in their instructor's ability to help them understand the material.

Participants who reported feeling secure when doing statistics problems often credited their mathematical abilities, as demonstrated by their responses. Some stated, "secure, because math is one of my favorite subjects…I feel secure, because the math is easy…I enjoy math…and I am good at math." One student clearly linked the two subjects by stating, "I am comfortable with math; therefore, I am comfortable with statistics." One the other hand, participants who reported being insecure often cited their lack of mathematical abilities as the reason. For example, many stated, "I am not good at math…I struggle with math…I do not feel confident with my math skills…I do not like math and it is not one of my favorite subjects…and I had trouble in math since middle school..."

Other participants showed a sense of personal responsibility and a positive belief in their abilities toward completing statistics problems. They reported feeling secure because, "I pay attention in class and complete homework assignments…I practice and study hard…I can learn statistics and with a lot of practice, I will be able to do very well…and the only time I feel insecure doing statistics is when I have not studied the material." Different participants who felt insecure made no comments about their sense of personal responsibility, but did make negative comments that reflected their beliefs in their abilities. Some stated, "I'm not as quick at learning as others…I always make stupid mistakes…I am not always right, I am usually wrong…and I have math anxiety and no matter how hard I study….when I get a math or stats in front of me, I panic."

The third theme focused on the participants' instructors and whether they had confidence in their abilities to help them successfully complete their courses. Some who reported being secure stated, "the teacher explains it very well…I have a good statistics instructor…and I will feel secure when doing statistics problems because if I do not understand, I will ask for help," unlike those who reported being insecure who stated, "the instructor does not clarify…and I am insecure because of the instructor's teaching methods."

*I will or will not make a lot of math errors in statistics because…*Two underlying themes emerged from participants' responses to the statement, *I will or will not make a lot of math errors in statistics because…*For either response, will or will not, both centered on participants' beliefs in their mathematical abilities, and their sense of personal responsibility and perceived self-efficacy.

Reflecting participants' beliefs in their mathematical abilities were those who stated they will not make a lot of math errors in statistics because, "I am good at math…and I like math…I will not because I am math savvy…I am excellent in math…math comes easy to me…I will not because math seems to be one of my best subjects and I normally do well in it." Conversely, participants who reported they will make a lot of math errors viewed their mathematical abilities differently. Participants stated, "I was never really that good at math…math is not my strong subject…I am bad at math…and I will make a lot of math errors in statistics, because I am no good at math and never was."

Participants' sense of personal responsibility and perceived self-efficacy are demonstrated in their statements according to whether they believed they will or will not make a lot of math errors when doing statistics problems. Those who believed they will not make a lot of errors had a positive belief in their abilities. They stated, "I will not make errors because I will be careful…I will take my time…I will not make mistakes if I study hard and learn my definitions…I will study hard…I will go over my answers…I am going to take notes…I will keep practicing the problems until I understand them…and study before the exam." On the other hand, participants who believed they would make a lot of math errors had a negative belief in their abilities. They stated, "I will because it is hard…it is difficult to not to make errors in any math course…I will because I am impatient…I always make a lot of math errors…it is easy to make errors…I will make errors even when I take my time…and I will because I hate math and do not spend a lot of time checking over my work, because I just want to get done with it."

*Statistics formulas are easy or not easy to understand because…*Three underlying themes emerged from participants' responses to the statement, *Statistics formulas are easy or not easy to understand because…*These reflected participants' beliefs in their instructors' abilities, their sense of personal responsibility and their personal beliefs about the complexity of the formulas.

First, participants who believed statistics formulas are easy to understand believed their instructors' abilities to explain the formulas was a key element that contributed to their understanding of them and viewed their instructors in a positive manner. Some stated, "the professor made sure that everyone understood each formula…my teacher helped the class out until everyone was familiar with the formula…they are easy because the professor explains the formulas…easy to understand because the teacher went through them step-by-step…easy when you have a good teacher…I understand the concepts behind the formulas because of the great professor I had…and statistics formulas are easy to understand because they are explained well by the instructor." No comments about instructors were made by participants who believed statistics formulas were not easy to understand.

The second theme concerned participants' personal sense of responsibility when it came to learning and understanding statistics formulas. Those who believed formulas were easy to understand stated, "they are easy because I carefully analyze the problem…easy, if you are willing to study…easy, if you understand what the numbers are for, then you will understand what formulas to use…easy, if a person studies…and easy, because from practice it becomes obvious where numbers in the formulas come from." Opposite beliefs were declared by those who viewed statistics formulas as not easy to understand. Some stated, "statistics formulas are not easy to understand because they are complex…complicated…and require more thinking…they are not easy because they require memorization."

Participants' personal belief about the complexity of the formulas was the third theme from the open-ended responses. Some felt the formulas were easy to understand because all they had to do was "plug the numbers into an equation…they are easy to understand because they are not different than any other math formula, just plug the numbers in…they are simple…easy somewhat because the formulas themselves are part of typical math and algebra." Conversely, those who believed statistics formulas are not easy to understand compared them to learning a foreign language. "Statistics formulas are not easy to understand, because they are almost a whole new language to me…you have to learn the meaning of the letters and where to find those numbers…not easy, because of all the letters, numbers and symbols…it is like you are learning a whole new language, it takes a lot of practice.” Furthermore, some believed statistics formulas were not easy to understand because simply, “they can be confusing…they are complicated …they are too long and similar…there are big formulas…pretty complicated and hard to understand.”

*Statistics is or is not relevant in my life because…*Two underlying themes emerged from participants’ responses to the statement, *Statistics is or is not relevant in my life because…*For either response, is or is not relevant, both centered on participants’ beliefs in their future careers and lives, after graduation.

The most prevalent theme from participants’ statements about the relevance of statistics in their lives was careers. Many participants believed statistics would help them with their careers after graduation. Some stated their beliefs according to their specific career or major; “I will need it to major in business management… It’s because I want to major in accounting…I am an accounting major, stat is important…the criminal justice field uses stats frequently…it is relevant because I am a psychology major…is, I am a sociology major and have to do statistics at some point in my job…it is for my career in the medical field.” Others just referred to statistics as being relevant more simply; “jobs often require knowledge of statistics…I will use it in my job…it is beneficial for my career…and is relevant because I feel that it is good to know and understand statistics when looking for a job.” However, some had opposite views; “Statistics is not relevant; I am working with people, not researching and comparing...it is not relevant, I plan on being a counselor, not a field psychologist…I will not use it in my career.”

Many participants reported that statistics was relevant in their future lives as "it will help me understand certain things in the world…it involves many aspects of our lives… statistics are present in schools, jobs and the media…magazines…voting …newspapers…I am a big sports fan, statistics are used in the major leagues sports all the time…It applies to the weather…It involves real-life situations, such as politics or the economy." A few believed that statistics was not relevant in their future lives. "It is not relevant in my life because I do not use statistics at work or at home…is not, I have no interest in it and no need for it…I am not aware of any way that I will use statistics outside of class…and is not relevant because there are professionals who specialize in statistics and therefore I need not know it."

Critical Stance

Critical stance is the third and final element of the dispositional element and can be defined by an individual's "willingness to invoke action" (Gal, 2004, p. 69), when they encounter statistical messages from the media. It is the idea that individuals do not remain passive when they interpret statistical information, but develop a questioning attitude toward these statistical messages.

Critical Stance: ANOVA. A one-way between-groups analysis of variance was conducted to explore the impact of class types on post-test scores of critical stance. Adult students were in four groups according to the type of class they were enrolled in, statistics, a research methods class with prior statistics, a research methods class with no prior statistics or a control group. No statistically significant differences were found, $p = .26$ among class types as shown in Table 101.

Table 101

Dispositional Element: Critical Stance

Class Type

ANOVA

Total post scores critical stance

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	2.227	3	.742	1.364	.258
Within Groups	57.668	106	.544		
Total	59.894	109			

Critical Stance: Mixed Between-Within Subjects ANOVA. A mixed between-within subjects analysis of variance was conducted to assess the impact of four different types of classes (i.e., statistics, research methods class with prior statistics, research methods class with no prior statistics, and a control group) on participants' scores of critical stance across two time periods, pre- and post-test scores. There was no significant interaction between class types and time, $p = .79$, and no substantial effect for time, $p = .61$, as shown in Table 102. The main effect comparing the types of classes was significant, $p = .024$, as shown in Table 103. Because the main effect was significant, post-hoc analyses followed.

Table 102

Dispositional Element: Critical Stance

Interaction and Time Effect

Multivariate Tests

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
time	Wilks' Lambda	.997	.268	1.000	106.000	.606	.003
time * classtyp	Wilks' Lambda	.990	.347	3.000	106.000	.791	.010

Table 103

Dispositional Element: Critical Stance

Main Effect

Tests of Between-Subjects Effects

Measure: MEASURE_1
Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Intercept	4408.958	1	4408.958	7228.227	.000	.986
classtyp	5.980	3	1.993	3.268	.024	.085
Error	64.656	106	.610			

Post-hoc comparisons were completed using Fisher's LSD test and showed two statistically significant outcomes for the dependent variable, critical stance and are summarized in Table 104. First, the test indicated a statistically significant difference on pre- and post-test scores between the statistics class and the research methods class with prior statistics. Results indicated that both groups scored lower on the post-test, but the research methods class with prior statistics had higher scores on both the pre- and post-test scores than did the statistics class.

Second, there was a statistically significant difference on pre- and post-test scores between the control group and the research methods class with prior statistics, indicating that the control group scores increased from pre- to post-test, while scores for the research methods class with prior statistics decreased from pre- to post-test, albeit both pre-and post-test scores were higher. Estimated marginal means for the dispositional element, critical stance, are summarized in Table 105, and pre- to post-test differences are displayed graphically in Chart 9.

Table 104

Dispositional Element: Critical Stance

Post-hoc: Class Types

Multiple Comparisons

Measure: MEASURE_1
LSD

(I) Class type	(J) Class type	Sig.
Statistics	Control	.673
	Research methods w/o statistics	.343
	Research methods after statistics	.012
Control	Statistics	.673
	Research methods w/o statistics	.211
	Research methods after statistics	.006
Research methods w/o statistics	Statistics	.343
	Control	.211
	Research methods after statistics	.231
Research methods after statistics	Statistics	.012
	Control	.006
	Research methods w/o statistics	.231

.

Table 105

Dispositional Element, Critical Stance

Estimated Marginal Means

Class type * time

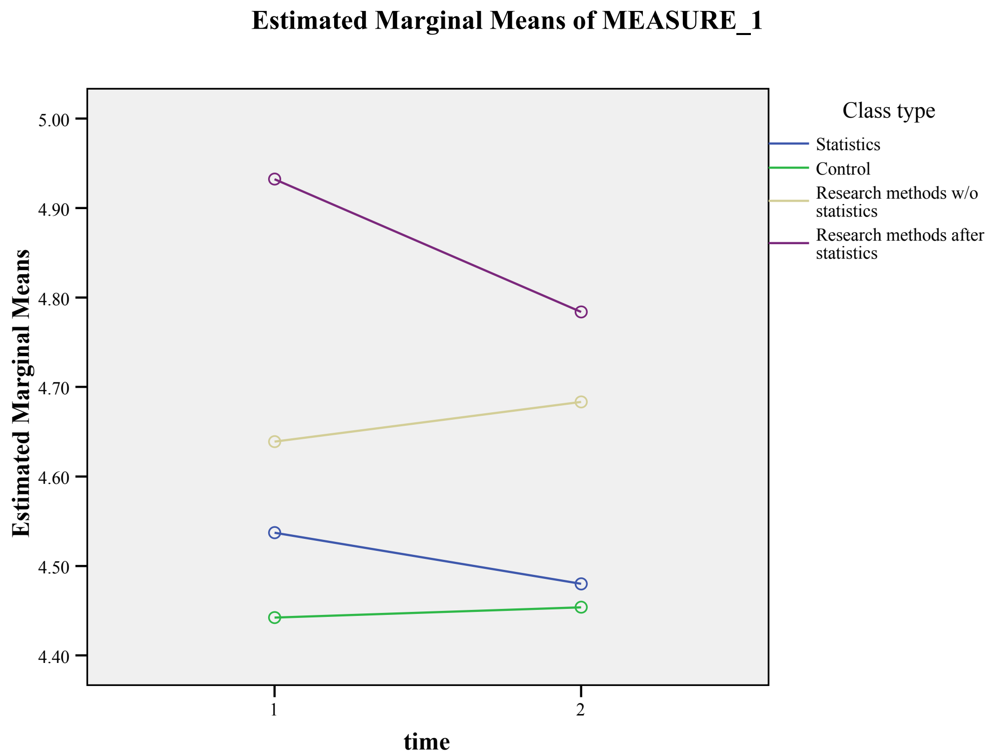
Measure: MEASURE_1

Class type	time	Mean	Std. Error	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
Statistics	1	4.537	.098	4.343	4.731
	2	4.480	.125	4.233	4.727
Control	1	4.442	.114	4.217	4.667
	2	4.454	.145	4.167	4.741
Research methods w/o statistics	1	4.639	.136	4.368	4.909
	2	4.683	.174	4.339	5.028
Research methods after statistics	1	4.932	.104	4.726	5.138
	2	4.784	.132	4.521	5.047

Chart 9

Dispositional Element: Critical Stance

Pre- and Post-Test Increases and Decreases



Critical Stance: ANCOVA

To examine for pre-test effects on post-test scores an ANCOVA was used. The independent variable was the type of class (i.e., statistics, research methods class with prior statistics, research methods class with no prior statistics, and the control) and the dependent variable consisted of the post-test scores from the critical stance scale. Participants' pre–test scores from the critical stance scale were used as the covariate in this analysis.

Preliminary checks were conducted to insure there were no violations of the assumptions of normality, linearity, homogeneity of variances, and homogeneity of regression slopes. After adjusting for pre-test scores on the critical stance scale, there were no significant differences between the four class types on post-test scores on the critical stance scale, p = .86, which is summarized in Table 106. There was a small relationship between the pre- and post-test scores on the critical stance scale, as indicated by a partial eta squared value of .16, which indicates a small effect.

Table 106

Dispositional Element: Critical Stance (ANCOVA)

Tests of Between-Subjects Effects

Dependent Variable: Total post scores critical stance

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
PRECSTOT	9.188	1	9.188	19.901	.000	.159
classtyp	.355	3	.118	.256	.857	.007

Gender: Paired-Samples T-Tests. Two paired-samples t-tests were conducted to examine adult students' pre- and post-test scores of critical stance by gender. No statistically significant difference was found between pre-and post-test scores for females, p = .92. Likewise, no statistically significant difference was found between pre- and post-test scores for males, p = .15. A summary of the paired-samples t-tests results is provided in Table 107 and the means and standard deviations are shown in Table 108.

Table 107

Dispositional Element: Critical Stance

Gender Differences

Paired Samples Test

What is your sex?			Paired Differences Mean	Paired Differences Std. Deviation	t	df	Sig. (2-tailed)
female	Pair 1	PRECSTOT - Total post scores critical stance	.00847	.63634	.102	58	.919
male	Pair 1	PRECSTOT - Total post scores critical stance	.21200	.71491	1.483	24	.151

Table 108

Dispositional Element: Critical Stance

Males and Females: Means and Standard Deviations

Paired Samples Statistics

What is your sex?			Mean	N	Std. Deviation
female	Pair 1	PRECSTOT	4.6271	59	.57710
		Total post scores critical stance	4.6186	59	.68543
male	Pair 1	PRECSTOT	4.8880	25	.56886
		Total post scores critical stance	4.6760	25	.78064

First-Generation Status: Paired-Samples T-Tests. Likewise, two paired-samples t-tests were conducted to evaluate adult students pre- and post-test scores of critical stance by first-generation status. No statistically significant difference was found between pre- and post-test scores for first-generation college students, p = .11. Likewise, no statistically significant difference was found between pre-and post-test scores for not being a first-generation college student, p = .74. A summary of the paired-samples t-test is provided in Table 109 and the means and standard deviations are shown in Table 110.

Table 109

Dispositional Element: Critical Stance

First-Generation Status

Paired Samples Test

First generation college student			Paired Differences Mean	Std. Deviation	t	df	Sig. (2-tailed)
yes	Pair 1	PRECSTOT - Total post scores critical stance	.16429	.66142	1.610	41	.115
no	Pair 1	PRECSTOT - Total post scores critical stance	-.03415	.66468	-.329	40	.744

Table 110

Dispositional Element: Critical Stance

First-Generation Status: Means and Standard Deviations

Paired Samples Statistics

First generation college student			Mean	N	Std. Deviation
yes	Pair 1	PRECSTOT	4.7881	42	.60454
		Total post scores critical stance	4.6238	42	.69205
no	Pair 1	PRECSTOT	4.6293	41	.56136
		Total post scores critical stance	4.6634	41	.73917

Gender: Independent- Samples T-Test. An independent-samples t-test was conducted to compare critical stance post-test scores between males and females. There was no statistically significant difference in scores between males and females, $p = .74$. A summary of the independent-samples t-tests is provided in Table 111 and the means and standard deviations are shown in Table 112.

Table 111

Dispositional Element: Critical Stance

Gender: Males and Females

Independent Samples Test

		Gender		
		t	df	Sig. (2-tailed)
Total post scores critical stance	Equal variances assumed	-.336	82	.737

Table 112

Dispositional Element: Critical Stance

Gender: Males and Females, Means and Standard Deviations

Group Statistics

	What is your sex?	N	Mean	Std. Deviation
Total post scores critical stance	female	59	4.6186	.68543
	male	25	4.6760	.78064

First-Generation Status: Independent-Samples T-Test. An independent-samples t-test was conducted to compare critical stance post-test scores with first-generation status. There was no statistically significant difference in scores between first-generation and not first-generation adult students, $p = .80$. A summary of the independent-samples t-tests is provided in Table 113 and the means and standard deviations are shown in Table 114.

Table 113

Dispositional Element: Critical Stance

First-Generation Status: Yes or No

Independent Samples Test

		First-Generation Status		
		t	df	Sig. (2-tailed)
Total post scores critical stance	Equal variances assumed	-.252	81	.802

Table 114

Dispositional Element: Critical Stance

First-Generation Status: Yes or No, Means and Standard Deviations

Group Statistics

	First generation college student	N	Mean	Std. Deviation
Total post scores critical stance	yes	42	4.6238	.69205
	no	41	4.6634	.73917

Summary

This chapter provided detailed statistical analyses completed with instruments that reflected Gal's Model of Statistical Literacy, which is represented by the knowledge elements, statistical thinking, reasoning, and literacy; and the dispositional elements, attitudes and beliefs, and critical stance. The results of the statistical analyses are summarized briefly within the text and the summary charts.

Knowledge Elements

The knowledge elements were analyzed through a series of statistical tests, which analyzed post-test scores among groups. These were MANOVAs, ANOVAs, independent-samples t-tests and post-hoc comparisons using Fisher's LSD. Mixed between-within subjects ANOVAs were used to examine gains or losses in pre- and post-test scores, along with paired-samples t-tests. And ANCOVAs were used to examine pre-test effects on post-test scores.

A MANOVA showed statistically significant differences among class types on the combined dependent variables of the knowledge elements, statistical thinking, reasoning and literacy. In order to find out where group differences were, further post-hoc comparisons were made using Fisher's LSD on each of the knowledge elements.

Post-hoc comparisons for statistical thinking showed four significant results; the statistics class had higher levels of statistical thinking than the control group; the research methods class with no prior statistics had higher levels of statistical thinking than the control group; the research methods class with no prior statistics had higher level of statistical thinking than the control group; and the research methods class with prior statistics had higher levels of statistical thinking than the statistics class.

Results using post-hoc comparisons for statistical reasoning showed two significant results. The research methods class with prior statistics had higher levels of statistical reasoning than the statistics class; and the research methods class with prior statistics had higher levels of statistical reasoning than the control group.

And post-hoc comparisons for statistical literacy showed three significant results. The research methods class with prior statistics had higher levels of statistical literacy than the statistics class; the research methods class with prior statistics had higher levels of statistical literacy than the control group; and the research methods class with prior statistics had higher levels of statistical literacy than the research methods class with no prior statistics. A summary of these results is provided in Table 115.

Table 115

Summary

Knowledge Elements: Statistical Thinking, Reasoning and Literacy

Class Types: Post-Test Scores

Knowledge Element	Higher Scores	Lower Scores
Statistical Thinking	Statistics Class	Control Group
	RM w/o Statistics	Control Group
	RM after Statistics	Control Group
	RM after Statistics	Statistics Class
Statistical Reasoning	RM after Statistics	Statistics Class
	RM after Statistics	Control Group
Statistical Literacy	RM after Statistics	Statistics Class
	RM after Statistics	Control Group
	RM after Statistics	RM w/o Statistics

To examine learning gains among class types, three mixed between-within subjects ANOVAs were completed on each of the dependent variables, statistical thinking, reasoning and literacy, to examine pre- and post-test scores.

Results from the mixed between-within subjects ANOVA for the dependent variable, statistical thinking, showed the main effect in comparing the types of classes significant. Post-hoc comparisons followed with five significant results. First, there was a statistically significant difference on pre- and post- test scores for the statistics class and the control group. Both classes showed a decrease in pre- to post-test scores, but the statistics class had higher post-test scores than the control group. Second, the research methods class with prior statistics showed an increase in post-test scores, while the control group showed a decrease in post-test scores. Third, the research methods class with no prior statistics showed an increase in post-test scores, while the control group showed a decrease. Fourth, the research methods class with prior statistics showed an increase in post-test scores, while the statistics class showed a decrease. And fifth, the research methods class with prior statistics showed higher post-test scores than the research methods class with no prior statistics. A summary of the results is shown in Table 116.

Table 116

Summary

Knowledge Element: Statistical Thinking: Pre- to Post-Test

Group Differences	Increase	Decrease
Statistics Class		Yes (higher post-test)
Control Group		Yes
RM after Statistics	Yes	
Control Group		Yes
RM w/o Statistics	Yes (higher pre- and post-test scores)	
Control Group		Yes
RM after Statistics	Yes (higher pre- and post-test scores)	
Statistics Class		Yes
RM after Statistics	Yes	
RM w/o Statistics		Yes

Likewise, results for the dependent variable statistical reasoning showed the main effect in comparing the types of classes significant. Post-hoc comparisons followed with three significant results. First, there was a statistically significant difference on pre- and post- test scores between the research methods class with prior statistics and the statistics class, with the research methods class with prior statistics scoring higher on both the pre- and post-test scores than the statistics class. Second, the research methods with prior statistics scored higher on both pre- and post-test scores than those in the control group. Third, the research methods class with prior statistics scored higher on both the pre-and post-test scores than the research methods class with no prior statistics. A summary of the results is shown in Table 117.

Table 117

Summary

Knowledge Element: Statistical Reasoning Pre- to Post-Test

Group Differences	Increase	Decrease
RM after Statistics Statistics Class	Yes (higher pre- and post-test scores) small increase	
RM after Statistics Control Group	Yes (higher pre- and post-test scores) slight increase	
RM after Statistics RM w/o Statistics	Yes (higher pre- and post-test scores) small increase	

The third knowledge element, statistical literacy, also showed the main effect in comparing the types of classes significant. Post-hoc comparisons followed with three significant results. First, there was a statistically significant difference on pre- and post-test scores between the research methods class with prior statistics and the statistics class, with the research methods class with prior statistics scoring higher on both the pre-and post-test scores than the statistics class. Second, there was a statistically significant difference on the pre-and post-test scores for the research methods class with prior statistics and the control group, with the research methods class with prior statistics scoring higher on both the pre-and post-tests, while the control group had lower post-test scores than the pre-test. And third, the research methods with prior statistics had higher pre-and post-test scores, while the research methods class with no prior statistics had decreased post-test scores. A summary of the results is shown in Table 118.

Table 118

Summary

Knowledge Element: Statistical Literacy Pre- to Post-Test

Group Differences	Increase	Decrease
RM after Statistics	Yes (higher pre- and post-test scores)	
Statistics Class	slight increase	
RM after Statistics	Yes (higher pre- and post-test scores)	
Control Group		Yes
RM after Statistics	Yes (higher pre- and post-test scores)	
RM w/o Statistics		Yes

To further explore learning gains between pre- and post-test scores on the knowledge elements, statistical thinking, reasoning and literacy, the independent variables, gender and first-generation students, were analyzed by 12 paired-samples t-tests.

In examining scores for females, no statistically significant differences were found between pre-and post-test scores for statistical thinking and literacy, but significant differences were found for statistical reasoning; females increased scores from pre- to post-tests. For males, no statistically significant differences were found between pre- and post-test scores on any of the variables, statistical thinking, reasoning and literacy. Results for pre- to post-test scores for first-generation adult college students showed no statistically significant differences for the variables statistical thinking and literacy, but did show a statistically significant difference between pre- to post-test scores on statistical reasoning. No statistically significant differences were found between pre- to post-test scores on statistical thinking, reasoning and literacy for adult students who were not first-generation adult college students. Table 119 provides a summary for the paired-samples t-tests and the knowledge elements, statistical thinking, reasoning and literary.

Table 119

Summary

Knowledge Elements: Paired T-Tests

	Male	Female	First-Generation	Not First-Generation
Statistical Thinking	No	No	No	No
Statistical Reasoning	No	Yes	Yes	No
Statistical Literacy	No	No	No	No

To further explore group differences, gender and first-generation status were used as the independent variables to examine differences on post-test scores on each of the dependent variables statistical thinking, reasoning and literacy. Six independent-samples t-tests were used for the analyses.

No statistically significant differences were found between males and females on post-test scores for the dependent variables statistical thinking, reasoning and literacy. Likewise, no statistically significant differences were found between first-generation and not first-generation adult students on the dependent variables statistical thinking, reasoning, and literacy. A summary of the knowledge elements and independent-samples t-tests is shown in Table 120.

Table 120

Summary

Knowledge Elements: Independent T-Tests

	Between Males/Females	Between First-Generation and Not
Statistical Thinking	No	No
Statistical Reasoning	No	No
Statistical Literacy	No	No

One final analysis was completed on the knowledge elements to examine for pre-test influences on post-test scores. ANCOVAs were used for statistical analyses on each of the knowledge elements, statistical thinking, reasoning, and literacy. Results indicated that after adjusting for pre-test scores on each of the knowledge elements, statistically significant differences were found among the four types of classes, statistics, research methods class with prior statistics, research methods class with no prior statistics and the control group. Relationships between pre- and post-test scores ranged from small, statistical thinking, to moderate, statistical literacy and to strong for statistical reasoning.

The fourth knowledge element, critical questions, was analyzed by using a mixed between-within subjects ANOVA. The main effect in comparing the different class types was significant; this resulted in further data analyses by using Fisher's LSD for post-hoc analyses. First, results showed a statistically significant difference on pre- and post-test scores between the research methods class with no prior statistics and the control group; second, results showed a statistically significant difference on pre- and post-test scores between the research methods class with prior statistics and the control group; and third, results showed a statistically significant difference on pre- and post-test scores between the research methods class with prior statistics and the statistics class. A summary of the results are shown in Table 121.

Table 121

Summary

Knowledge Element, Critical Question: Pre- to Post-Test Score

Group Differences	Increase	Decrease
RM w/o Statistics Control Group	Yes (higher pre- and post-test scores) (slight increase)	
RM after Statistics Control Group	Yes (higher pre- and post-test scores) (slight increase)	
RM after Statistics Statistics	Yes (higher pre- and post-test scores) (slight increase)	

To examine the impact of class types on post-test scores on the critical question, a one-way ANOVA was completed. No statistically significant differences were found among class types. And to examine post-test scores differences between gender and between first-generation statuses, two independent-samples t-tests were used. Results showed no statistically significant differences on post-test scores for the critical questions between males and females. Likewise, there were no statistically significant differences between first-generation adult students and those who were not first-generation on post-test scores for the critical questions. These results are summed up in Table 122.

Table 122

Summary

Knowledge Elements: Critical Questions and Class Types

ANOVA	Critical Questions Significant/Yes-No
RM after Statistics	No
RM w/o Statistics	No
Statistics class	No
Control Group	No
Independent T-Test	
Gender	No
First-Generation Status	No

Table 123 provides a summary of the results for the critical questions on pre-and post-test scores examined by using two paired-samples t-tests with gender and first-generation status as the independent variables. For females, there was no statistically significant difference between pre- and post-test scores on the critical questions. Conversely, there was a statistically significant difference between pre- and post-test scores for males on the critical questions. And for students who were either first-generation or not a first-generation adult college student, no statistically significant differences were found between the pre- and post-test scores.

Table 123

Summary

Knowledge Element: Critical Questions, Paired T-Test

Group	Critical Questions
Males	No
Females	Yes
First-Generation	No
Not First-Generation	No

One more test, an ANCOVA, was conducted on the critical questions to examine if the pre-test had affected the post-test scores. After adjusting for the pre-test scores, there was no statistically significant difference among the four class types and post-test scores of the critical questions. A moderate relationship was found between the pre-test and post-test critical questions.

Dispositional Elements

The dispositional elements were analyzed through a series of statistical tests that analyzed post-test scores among groups. These were MANOVAs, ANOVAs, independent-samples t-tests and post-hoc comparisons using Fisher's LSD. Mixed between-within subjects ANOVAs were used to examine gains or losses in pre- to post-test scores along with paired-samples t-tests. And ANCOVAs were used to examine pre-test effects on post-test scores. Categorical data (i.e., beliefs) were analyzed by use a Chi-square and open-ended responses were grouped by themes.

A MANOVA revealed statistically significant differences among class types on the combined dependent variable attitudes—SATS affect, cognitive competence, difficulty and value—but when the results of the dependent variable were considered separately, no statistically significant differences were found when using a Bonferroni adjusted alpha level of .0125.

To examine pre- to post-test gains or losses on the four elements of the SATS scale (i.e., affect, cognitive competence, value and difficulty), four mixed between-within subjects analyses of variance were conducted. Results showed the main effect was significant for affect and a post hoc-test followed revealing three significant outcomes. First, students in the statistics class scored higher on the pre-test than students in the research methods class with no prior statistics, but they scored lower on the post-test, while students in the research methods class with no prior statistics scored higher on the post-test than on the pre-test. Second, students in the research methods class with prior statistics scored higher on the post-test and higher on both the pre- and post-test scores than the research methods class with no prior statistics. Third, students in the control group scored higher on the post-test and higher on both the pre- and post-test scores than the students in the research methods class with no prior statistics.

Likewise, results showed the main effect was significant for cognitive competence. A post-hoc test followed and indicated two significant outcomes. First, students in the statistics class scored higher on both the pre- and post-test than students in the research methods class with no prior statistics, albeit the differences between the pre- and post-test scores were almost non-existent. And students in the research methods class with no prior statistics increased their scores from pre- to post-test. Second, students in the research methods class with prior statistics scored lower on the post-test than on the pre-test, but had higher scores on both the pre- and post-test than students in the research methods class with no prior statistics. And students in the research methods class with no prior statistics showed an increase from pre- to post-test scores.

In the same way, results showed the main effect was significant for value. Post-hoc comparisons followed resulting in two statistically significant outcomes. First, students in the research methods class with no prior statistics had higher pre- and post-test scores than students in the control group, but had lower post-test scores. Students in the control group had higher post-test than pre-test scores. Second, students in both the research methods class with prior statistics and the control group had higher post-test than pre-test scores, but the research methods class with prior statistics had higher pre- and post-test scores than the control group. However, the last element, difficulty, showed no main effect. The dispositional element, affect, is summarized in Table 124, cognitive competence, Table 125, value, Table 126 and difficulty, Table 127.

Table 124

Summary

Dispositional Element, Affect: Pre- to Post-Test Scores

Group Differences for Affect	Increase	Decrease
RM w/o Statistics	Yes (higher pre- and post-test scores)	
Statistics Class	Yes	
RM after Statistics	Yes (higher pre- and post-test scores)	
RM w/o Statistics	Yes	
Control Group	Yes (higher pre- and post-test scores)	
RM w/o Statistics	Yes	

Table 125

Summary

Dispositional Element, Cognitive Competence: Pre- to Post-Test Scores

Group Differences for Cognitive Competence	Increase	Decrease
Statistics Class	almost identical/higher pre- and post-test scores	
RM w/o Statistics	Yes	
RM after Statistics		Yes (higher pre-and post-test)
RM w/o Statistics	Yes	

Table 126

Summary

Dispositional Element, Value: Pre- to Post-Test Scores

Group Differences for Value	Increase	Decrease
RM w/o Statistics	Yes (higher pre- and post-test scores)	
Control Group	Yes	
RM after Statistics	Yes (higher pre- and post-test scores)	
Control Group	Yes	

Table 127

Summary

Dispositional Element, Difficulty, Pre- to Post-Test Scores

Group Differences for Difficulty	Increase	Decrease
No differences among any groups		

Next, gender and first-generation status as the independent variables were analyzed with the SATS's scale components, affect, cognitive competence, value and difficulty through independent-samples t-tests and paired t-tests.

An independent-samples t-test compared total post-test scores between males and females, resulting in males scoring higher than females. When comparing the elements of the scale separately, only affect was found to be statistically significant, with males scoring higher than females. There were no gender differences uncovered for cognitive competence, value and difficulty.

Another independent-samples t-test compared total post-test scores between first-generation adult students and those who were not first-generation. No statistically significant differences were found. When comparing the elements of the scale separately, only cognitive competence was found to be statistically significant. First-generation adult students scored higher than those who were not first-generation students. There were no differences in first-generation status uncovered for cognitive competence, value and difficulty.

Paired-samples t-tests were used to examine pre- and post-test scores on each element of the SATS scale with gender and first-generation status. No statistically significant differences were found for males or females. Likewise, no statistically significant differences were found for adult students who were first-generation and those who were not first-generation on three elements of the SATS, affect, difficulty and value. The only element found to be statistically significant was cognitive competence for adult students who were not first-generation. A summary for the dispositional elements, affect, cognitive competence, value and difficulty by independent t-tests is shown in Table 128 and the paired t-test in Table 129.

Table 128

Summary: Independent-Samples T-Tests

Dispositional Elements: Affect, Cognitive Competence, Value and Difficulty

Independent T-Tests			Groups	
Elements	Males	Females	Fist-Generation	Not First-Generation
Total Post-Test	Yes	No	No	No
Affect	Yes	No	No	No
Cognitive Competence	No	No	No	No
Value	No	No	No	No
Difficulty	No	No	No	No

Table 129

Summary: Paired t-tests

Dispositional Elements: Affect, Cognitive Competence, Value and Difficulty

Paired T-Test				
Elements	Males	Females	Fist-Generation	Not First-Generation
Total Pre-Post	No	No	No	No
Affect	No	No	No	No
Cognitive Competence	No	No	No	Yes
Value	No	No	No	No
Difficulty	No	No	No	No

One final analysis was completed on the elements of the SATS to examine for pre-test influences on post-test scores. ANCOVAs were used for completing statistical analyses on each of the components, affect, cognitive competence, value and difficulty. Results indicated that after adjusting for pre-test scores on each of the elements, no statistically significant differences were found among the four types of classes, statistics, research methods class with prior statistics, research methods with no prior statistics and the control group. Relationships between pre- and post-test scores were moderate for each of the elements of the SATS, affect, cognitive competence, value and difficulty.

The dispositional element, beliefs, was examined by a Chi-square and open-ended statements, which were grouped according to themes. The only belief that indicated a significant association between class types and beliefs was students' belief that statistics is or is not relevant in my life. Students in research methods class with prior statistics believed that statistics is relevant in their lives, while the other groups did not.

Four belief statements were used for participants to respond to in open-ended statements. Each statement will be summarized according to the main themes. For the belief statement, *I will feel insecure or secure when doing statistics problems because*…, four underlying themes emerged for students who felt insecure or secure. These were their beliefs in their mathematical abilities, their sense of personal responsibility, a belief in their abilities to master the course content, and confidence or lack of confidence in their instructor's ability to help them understand the material.

For the second belief statement, *I will or will not make a lot of math errors in statistics because…,* two themes emerged for students who will or will not make a lot of math errors in statistics. These were participants' beliefs in their mathematical abilities, and their sense of personal responsibility and perceived self-efficacy.

The third belief statement*, statistics formulas are easy or not easy to understand because*…, three themes emerged for students who believed statistics formulas are easy or not easy to understand. These reflected participants' beliefs in their instructors' abilities, their sense of personal responsibility and their personal beliefs about the complexity of the formulas.

And the fourth belief statement, *statistics is or is not relevant in my life because*… two themes emerged for students who believed statistics is or is not relevant in their lives.

These themes were centered on participants' beliefs in their future careers and lives, after graduation.

An ANOVA was used to analyze the final dispositional element, critical stance, and class types. Results showed no statistically significant differences among class types and post-test scores. Next, a mixed between-within subjects ANOVA was used to examine differences among class types and pre- to post-test scores. Results showed the main effect to be significant and post-hoc analyses followed.

Post-hoc analyses showed two significant results. First, the research methods class with prior statistics showed higher pre- and post-test scores than the statistics class. However, both groups scored lower on the post-test than on the pre-test.

Second, the research methods class with prior statistics showed higher pre- and post-test scores than the control group, but the research methods class with prior statistics showed lower post-test scores than pre-test scores. The control group increased their scores from pre- to post-test. Table 130 provides a summary for the results of the dispositional element, critical stance.

Table 130

Summary

Dispositional Elements: Critical Stance

Class Types	Increase	Decrease
RM after Statistics	(higher pre-post)	Yes
Statistics Class		Yes
RM after Statistics	(higher pre-post)	Yes
Control Group	Yes	

Paired-samples t-tests were used to examine pre- and post-test scores on critical stance, with gender and first-generation status. No statistically significant differences were found for males or females. Likewise, no statistically significant differences were found for adult students who were first-generation and those who were not first-generation on critical stance.

An independent-samples t-test compared post-test scores on critical stance between males and females, and first-generation status. Results showed no statistically significant results for both gender and first-generation status, as shown in Table 131.

Table 131

Summary

Dispositional Elements: Critical Stance

Independent T-Tests	Group			
	Males	Females	First-Generation	Not First-Generation
Critical Stance	No	No	No	No
Paired Samples T-Test				
Critical Stance	No	No	No	No

One final analysis was completed on the dispositional element, critical stance, to examine for pre-test influences on post-test scores. Results from the ANCOVA indicated that after adjusting for pre-test scores on critical stance, no statistically significant differences were found among the four types of classes, statistics, research methods class with prior statistics, research methods class with no prior statistics and the control group. Relationships between pre- and post-test scores were small. Next, results of the elements of Gal's Model of Statistical literacy are discussed in Chapter 5.

CHAPTER 5

DISCUSSION, IMPLICATIONS, LIMITATIONS, CONCLUSION

The purpose of this final chapter is to discuss the results, implications, and limitations of this study, which examined statistical literacy in adult learners, before and after they have completed a statistics class, a research methods class without prior statistics, and a research methods class with prior statistics. Statistical literacy was examined by using instruments based on Gal's Model of Statistical Literacy. The model is comprised of two main elements; the first is knowledge elements, statistical reasoning, thinking and literacy, and critical questions; the second is dispositional elements, attitudes and beliefs, and critical stance. Using 2 main and 4 sub-hypotheses, both elements were examined with class types, gender, and first-generation status. The results are discussed within the context of the hypotheses that guided this research.

Discussion: Knowledge Elements

The first main hypothesis, that adult learners who have completed a research methods class with prior statistics will be more proficient in their knowledge of statistics than adult learners who have only completed a statistics or research methods class with no prior statistics, was tested and disproved. A discussion of the results for each of the knowledge elements, statistical thinking, reasoning and literacy, and the critical questions follows. These elements are discussed specifically and then more generally.

Statistical Thinking

Results from post-hoc comparisons of statistical thinking among four class types showed significant differences in post-test scores indicating the research methods class with prior statistics had higher scores, than the statistics class and the control group.

However, they did not reach significance with the research methods class with no prior statistics.

To explain the results of post-test comparisons, it is important to revisit the definition of statistical thinking, and the type of questions used to assess it. Questions that embraced statistical thinking were taken from the ARTIST website, and cover the same topics in statistics called context knowledge from Gal's Model of Statistical Literacy. These topics, according to Gal (2002), include how and where the data is collected, the data generation processes, such as the research methodology used, and the processes used to analyze the data. To understand statistics, the data needs to be viewed as numbers within its context, as the context is the source of meaning and the basis for the interpretation of the results (Moore, 1990). These topics are taught in research methods classes.

Accordingly, adult students who were enrolled in a research methods class, with or without a prior statistics class, most likely would have learned the course content—this material was the content of the questions used in statistical reasoning. This could explain why there was no significant difference between the two class types. And adult students who have only taken a statistics class did not have the opportunity to learn the material, which was the basis for the questions in statistical thinking.

From this it is evident that adult students in the statistics class have learning gaps, which will affect their ability to become statistically literate, as some do not have to complete a research methods class for their discipline. In a sense, they learned the mechanics of statistical analysis, but cannot apply this knowledge to the real world without completing a class on research methods to develop an understanding of context knowledge. Context knowledge is “the main determinant of the reader’s familiarity with sources for variation and error” (Gal, 2004, p. 4). Without this knowledge it is impossible to understand why group differences occur, or what alternative explanations exist, or how a study should be completed correctly (Gal, 2002). Further, as Moore (1997) asserts, “a student who emerges from a first statistics class without an appreciation of the distinction between observation and experiments, and of the importance of randomized comparative experiments…has been cheated” (p. 127).

In addition to post-test score comparisons among class types, the data was examined for pre- and post-test gains or losses. Results showed statistically significant increases for the research methods class with prior statistics among the three groups, research methods class with no prior statistics, statistics and the control group. The research methods class with no prior statistics scores increased from pre- to post-test, but the research methods class with prior statistics had significantly higher post-test scores.

This could be representative of their learning, whereas adult students who had taken a statistics class before research methods could better integrate statistical concepts with research methodology, possibly, because they have used real data sets in their prior statistics class and already have developed a deeper understanding of statistical concepts (Morris, 2001). And when completing a research project, previous learned material from their statistics class was easier to re-learn when applying it to a research study.

Two groups, the control group and the statistics class, decreased their scores from pre- to post-test, which could indicate that these adult students did not correctly answer the questions on the pre-test; they may have guessed the multiple choice answers.

Likewise, for their responses on the post-test, because it is evident from the content that makes up the statistical thinking questions, students never had the opportunity to learn this material.

Statistical Reasoning

Results from post-hoc comparisons of statistical reasoning among four class types showed significant differences in post-test scores, indicating that the research methods class with prior statistics had higher scores than the statistics class. No significant difference was seen between the research methods class with prior statistics and the research methods class with no prior statistics.

A possible explanation of these results can be found in the definition of statistical reasoning, and the type of questions used to asses it. Questions that embraced statistical reasoning were also taken from the ARTIST website, and are made up of two components, statistical and mathematical knowledge. Statistical knowledge is comprised of familiarity with the basic terms and ideas related to descriptive statistics, graphical and tabular displays, knowing why the data are needed, how the data can be produced and how statistical conclusions are reached (Gal, 2002). It also includes sampling procedures and research designs, such as experimental and non-experimental methods (Cobb & Moore, 1997). Mathematical knowledge would include, for example, how a mean is computed, or how it is influenced by extreme values in a data set and the use of percentages (Gal, 2002). It also includes a solid understanding of underlying statistical ideas, such as the quantification of variance, repeated sampling, and sampling distributions (Cobb & Moore, 1997). As evident from these descriptions, statistical and mathematical knowledge are not mutually exclusive, but complement one another, and together make up the content of statistical reasoning.

A possible explanation for the differences between the research methods with prior statistics and the statistics class is that statistical and mathematical knowledge are not mutually exclusive. Many topics of statistical reasoning are part of the curriculum in a statistics class; however, some may not be—for example, the topics of experimental and non-experimental designs. These topics, however, are often included in a research methods class. Adult students who have taken a statistics class and a research methods class would have the opportunity to learn and integrate this material from both classes. There is no past research that examined this—the course content of a statistics and a research methods class.

What is interesting from the data is that there were no statistically significant differences between the research methods class with prior statistics and the research methods class with no prior statistics, as many of these topics would not be covered in a research methods class. One possible explanation could be that some students may have completed other mathematical courses where some of the material is similar. It is important to note that the means for the research methods class with prior statistics were higher than the means for the research methods without a statistics class.

In addition to post-test scores comparisons, pre- to post-test learning gains were found in statistical reasoning for the research methods class with prior statistics. This class showed significant differences in higher pre- to post-test scores with the research methods class with no prior statistics, the statistics class and the control group. All three of these classes did increase their score from pre- to post, but the increase was very small.

For the research methods class with prior statistics, where there was a statistically significant difference among the three groups, adult students only correctly answered about 50% of the questions concerning statistical reasoning, which indicates some learning gaps in statistical and mathematical knowledge. These results contribute to the current debate about the amount of mathematics adults need to know to understand concepts of statistics (Cobb & Moore, 1997). Currently, most prerequisites for an introductory statistics class are math courses, such as algebra, but perhaps a prerequisite for statistics should be a course in mathematical probability, as it contains the founding concepts that underlie statistical concepts.

Statistical Literacy

Results from post-hoc comparisons of statistical reasoning among four class types showed significant differences in post-test scores indicating the research methods class with prior statistics had higher scores than the statistics class, the research methods class with no prior statistics and the control group.

To interpret these results it is necessary to revisit the description of statistical literacy. Statistical literacy is the same term used by Gal and by the ARTIST website, and pertains to the understanding of statistical messages of written text, or graphs, which are displayed with a few words. It involves understanding and using the basic language and tools of statistics (i.e., various statistical tests) (ARTIST, 2006; Gal, 2004), and from the descriptions, interpret the results. In order to understand the basic language of statistics, adult students need to have a knowledge of statistical analyses and research methodologies, and be able to apply these to statistical messages in many types of media. To accomplish this, adult students would need to have completed both a statistics and a research methods class, because statistical tools need to be applied within a research context—this is usually accomplished in a research methods class, but not without completing a class in statistics first.

In addition to post-test scores comparisons, learning gains in statistical literacy were statistically significant between pre- to post-test scores between the research methods class with prior statistics and the statistics class, and between the control group and the research methods class with no prior statistics. Pre- to post-test scores increased for the research methods class with prior statistics, increased slightly for the statistics class, and decreased for the research methods class with no prior statistics and the control group.

Learning gaps are evident. For scores to decrease between the pre- to post-test, this could indicate the adult students did not know the correct answer when they completed the pre-test, and later the post-test. And as is evident here, decreasing scores for the research methods class with no prior statistics indicates these adult students do not understand the basic tools and language of statistics, possibly because they never had the opportunity to take a statistics class.

Further examination of the data may give us more insight into the results. The knowledge elements, statistical thinking, reasoning and literacy, were examined for pre-test influences on post-test scores. A small relationship was found between pre- to post-test scores for thinking and literacy, and a moderate one for statistical reasoning. Nevertheless, after controlling for pre-test influences on post-test scores, statistically significant differences were still found for the all of the knowledge elements. An examination of the estimated marginal means for the knowledge elements showed the research methods class with prior statistics had the higher means among the groups, similar to the initial statistical analyses on each of the knowledge elements, statistical thinking, reasoning and literacy. This gives verification that the independent variable, class types, had affected the dependent variable, the knowledge elements, rather than having pre-test scores influence post-test scores.

Critical Questions

In examining the critical questions, results showed there were no significant differences among class types and post-test scores, and the only significant difference between pre- to post-test scores occurred with an increase for the statistics class. Unfortunately, this result means that most adult students in this study are unlikely to challenge reported research results. This may be due to the value they place on statistics or it may occur because they do not have competence in their knowledge of statistics. For the increases in pre- to post-test scores for the statistics class, one explanation could be that real research was brought into the classroom by the instructor for students to analyze. Using real or like real data sets is known to help students develop an understanding of statistical concepts (Proctor, 2002; Cralley & Ruscher, 2001; Morris, 2001), and help students understand the value of statistics while increasing their feeling of competence in their statistical skills. However, it is important to note, the increase from pre- to post-test scores reached statistical significance, but practical significance is questionable, because the increase was very small.

In sum, one possible explanation centers on adult students' learning gaps evident by the type of classes they were enrolled in and the type of knowledge element. A comparison of the courses by class types was compared with the construction of the questions for the knowledge elements. This showed results could be due partly to the material that was covered in each class, statistics and research methods. Each knowledge element reflected particular materials, which were somewhat specific for each class, hence, adult students who were in the classes that covered the particular material related to a specific knowledge element did better on post-test scores. Thus, adult students who were in the statistics class did not have the context knowledge (i.e., statistical thinking and literacy) to combine with their statistical knowledge (i.e., statistical reasoning), and adult students who had completed a research methods class without a statistics class did not have the knowledge of the undying statistical concepts (i.e., statistical reasoning).

However, some questions remain, and there are other possible reasons for these results. For example, this does not explain why the research methods class with prior statistics did not score higher than the research methods class without a prior statistics class, on the knowledge element, statistical thinking, as this is part of the scale that reflected material taught in a research methods class. A more general discussion of the results follows.

Some results, perhaps, can be explained by considering the variety of statistics and research methods classes used in this research. Because this was a multi-campus research endeavor, there were different types of statistics and research methods classes that adult students completed. Not all statistics courses are identical, as some are taught with a focus on mathematics, some on less mathematics and more concepts, some as a general statistics class, and some specific to a discipline, for example psychology. And there are instructional differences also, as some instructors engage students in their learning of statistical concepts by incorporating everyday problems into the course curricula (Lawson, et al., 2003; Vanderstoep & Shaughnessy, 1997). Others incorporate technology into their statistics class and use computer labs, where computer programs allow students to work with data sets in Excel or SPSS (Proctor, 2002; Raymondo & Garrett, 1998; Warner & Meehan, 2001). In fact, technology in the classroom also includes the use of computer simulations to learn specific statistical concepts, for example the central limit theorem (Aberson, et al., 2000, 2002, 2003; Morris et al., 2002; Morris, 2001), and there are hybrid courses (e. g., Bushway & Flowers, 2002; Symanzik & Vukasinovic, 2006). Also, some instructors may not have used any type of technology in their classrooms.

In addition, some statistics courses have added a writing component as part of the course requirement. Past research showed that by adding a writing component to the statistics class, students who participated in the statistics classes with writing assignments scored higher on evaluations of statistical concepts than students who were not in statistics classes that incorporated writing into the curriculum (Rajecki, 2002; Vanderstoep & Shaughnessy, 1997).

Moreover, research methods is taught in many different ways, even though the material may be similar. Some classes require adult students to complete an entire research project, from designing, collecting data, analyzing and writing up the research, while others only require a research proposal, and perhaps some classes do not have a writing component. No past research has examined the various course curricula used in statistics or research methods classes.

And for both class types, statistics and research methods, we do not know what type of teaching method the instructor used, as this could affect adult students' learning.

Some may have used traditional behaviorist teaching methods, where the instructor is the sole information-giver to passive students, and the textbooks contain a fixed world of knowledge (Hanley, 1994). Others may have used a constructivist perspective of pedagogy, where the instructor's role is to "facilitate and negotiate meaning-making with the learner" (Merriam et al., 2007, p. 295), with the purpose of learning being to construct knowledge.

The fourth knowledge element, critical questions, showed no differences among class types for post-test scores, which might indicate that these adult students do not place any value on statistics, and therefore, do not place any value on research printed in the media; or, maybe it indicates they do not have confidence in their statistical abilities. Further examination of this explanation was completed by using the dispositional elements, and will be discussed later in this chapter.

More importantly, pre- to post-test scores were the highest, and increased for the research methods class after statistics on all the knowledge elements, statistical thinking, reasoning and literacy, indicating that this group showed the largest learning gains from the beginning of the semester to the end of the semester.

Accordingly, it would be prudent for future research to examine instructors' teaching methods and what is required in the course curriculum for each class, and then examine statistical literacy again, by using Gal's Model. This is the first time the model has been tested to investigate statistical literacy. There is no prior research that examines statistical literacy in adult students.

Statistical Thinking, Reasoning, Literacy, and Critical Questions: Gender

The first sub-hypothesis, adult learners who are male will be more proficient in their knowledge of statistics than learners who are female, was tested on each of the knowledge elements, statistical thinking, reasoning, literacy and the critical questions, and was disproved. Results showed no significant differences in post-test scores between males and females on any of the knowledge elements, statistical thinking, reasoning, literacy, and the critical questions.

Because there is no research that addresses the knowledge elements, similar research that examines statistics and gender differences will be used for this discussion. These results are consistent with some past research that showed no gender differences in final course grades in statistics classes between males and females (Buck, 1985), and similarly, Ware and Chastain (1991) found no gender difference between males and females when comparing scores for statistical concepts. Other research disagrees. Schram's (1996) research results showed that, in general, females have outperformed males in statistics classes; however, this was when the outcome measured was the final grade; males outperformed females when the outcome measure was tests.

There is no research on gender differences on research methods classes, and only a few studies were found for statistics. Perhaps, because some of the knowledge elements incorporate course materials that include research methods concepts, there were no gender differences, albeit this cannot explain the results for statistical reasoning, which is constructed of both statistical and mathematical concepts.

Further, examination of learning gains for pre- to post-test scores for females resulted in no significant differences for the knowledge elements, statistical thinking, literacy and the critical questions. However, pre- to post-test scores increased for statistical reasoning for females, and for males on the critical questions.

Here is one possible explanation for this. Statistical reasoning combines the elements of statistical and mathematical knowledge, and females are reported to be better at mathematical computation than males, while males performed better than females in problem solving (Hyde et al., 1990; Fennema, & Lamon, 1990). This might also provide us with a partial explanation for why males showed a significant increase from pre- to post-test scores on the critical questions. For males to be competent at problem solving, they must think critically in order to solve the problem. Therefore, males may be more critical of research than females.

However, it is evident that research on gender and statistics is sparse, and there is no past research on gender and statistical literacy, and an important component of statistical literacy, the research methods class. These issues need to be examined further through more research using Gal's Model of Statistical Literacy.

Statistical Thinking, Reasoning, Literacy and Critical Questions: First-Generation Status

The second sub-hypothesis, adult learners who are not first-generation learners will be more proficient in their knowledge of statistics than learners who are first-generation adult learners, was tested on post-test scores on each of the knowledge elements, statistical thinking, reasoning, literacy and the critical questions, and was disproved. Results showed no significant differences in post-test scores between first-generation and not first-generation adult students on any of the knowledge elements, statistical thinking, reasoning, literacy, and the critical questions.

Further, examination of learning gains for pre- to post-test scores for adult students who were not first-generation resulted in no significant differences for all the knowledge elements, statistical thinking, reasoning, literacy and the critical questions. Similarly, for adult students who were first-generation, no learning gains were found in pre- to post-test scores for the knowledge elements, statistical thinking, literacy and the critical questions. However, there was a significant increase for the knowledge element, statistical reasoning. This would indicate that first-generation students had increased their statistical and mathematical abilities from the beginning of the semester to the end.

A possible explanation for the increase on statistical reasoning scores for first-generation adult students could be that they were very low at the beginning of the semester, perhaps due to deficient mathematics skills, as this element is comprised of statistical and mathematical knowledge. These adult students may have sought tutoring to get help with brushing up on their mathematical skills. However, why these results occurred remains unknown, because there is no prior research on first-generation adult students and statistical literacy.

We really do not know what accounts for the difference in statistical reasoning skills and the lack of significance for the other knowledge elements, and no differences in scores between first-generation and not first-generation students. There is no research on statistical literacy and adult students. This is important, and needs to be addressed in research, as approximately one-third of the current college population consists of adults (Schaeffer, 2001).

Discussion: Dispositional Elements

The second main hypothesis, adult learners who have completed research methods class after a prior statistics class will have more of a positive disposition toward statistics than students who have only completed a statistics or research methods class without a prior statistics class, was tested with each of the dispositional elements, affect, cognitive competence, difficulty and value (i.e., elements of the SATS) and was disproved. Post-test scores among groups showed no significant differences on the dispositional elements, affect, cognitive competence, difficulty and value. The critical stance scale had validity issues and will be discussed further in this chapter. A discussion of the results follow, albeit there is a paucity of past research on students' attitudes toward statistics and no previous research on beliefs. Beliefs which underlie attitudes are discussed separately.

Attitudes: Affect

Several differences resulted among class types and the SATS element, affect. First, adult students in the statistics class developed more negative attitudes toward statistics at the end of the semester. This may have occurred because they may have not wanted to take a class in statistics and found it hard to do the computations, especially toward the end of semester. As Gal, et al. (1997) explains, many students do not come into a statistics class ready to learn statistics. Thus, with statistics being a required course in many disciplines, some students may feel they were forced to take it.

Second, adult students in the research methods class with no prior statistics, and those in the research methods class with prior statistics, developed more positive attitudes toward statistics. Conversely, Carnell (2008), Alldredge et al. (2006), and Chadjipadelis and Andeadis (2006), found no differences between their groups for the SATS element, affect.

This change in attitudes for adult students in the research methods class with no prior statistics might have occurred because they did not have to do any statistical computations, and they may have enjoyed the course material in the research methods class. Gal, et al. (1997) informs us that many students hold a commonly held belief that statistics is heavy in mathematics, and many students have developed negative views about their mathematical skills from past experiences. Thus, these students probably did not engage in statistical computations, and therefore did not develop negative attitudes at the end of the class.

And third, adult students who were in the research methods class with prior statistics, may have developed an appreciation for the subject, because they have an understanding of statistics and can integrate this within the context of the topics in research methods. Perhaps a further explanation could be they have mastered their statistics class and are no longer intimidated by the mathematical computations required by statistics. Similarly, Alldredge et al. (2006) found a significant interaction effect between the treatment group and the preliminary algebra test score, for the element, affect. As students' algebra test scores increased, their feeling concerning statistics grew more positive. Another possible explanation is that adult students may have been able to use programs such as Excel or SPSS in their statistics or research methods class, as integration of technology into the classroom has been documented well in the literature (Proctor, 2002; Raymondo & Garrett, 1998; Warner & Meehan, 2001).

Attitudes: Cognitive Competence

Cognitive competence showed several different results among class types. Adult students in the statistics class showed a higher increase in their pre- to post-test scores than those in the research methods class with no prior statistics. This could indicate that adult students from the statistics class, who had some positive learning experiences in statistics, became confident in their intellectual ability to master many different types of statistical computations.

Another possible explanation for these results may be offered by Carnell (2008), who found no differences for cognitive competence in his research, but believed that test performance and past experiences in different types of quantitative classes could impact cognitive competence, negatively or positively. Further, Carnell reports, "students with a more extensive mathematical background might feel differently about statistics than students with a more limited background" (p. 7). Hence, adult students in the statistics class may have had positive past experiences in mathematics classes and/or a more extensive mathematics background than the adult students in the research methods class with no prior statistics.

Further, adult students from the research methods class with no prior statistics may have had to read journal articles, or collect data, or complete data analyses for a research project without having the underlying knowledge of how statistics are used for different types of research. This could indicate low mathematical abilities, as Froelich, Stephenson and Duckworth (2008) found in their research, for cognitive competence, scores decreased for the group with the lowest mathematical abilities.

Surprisingly, adult students from the research methods class with prior statistics showed a decrease in their scores on cognitive competence from pre- to post-test. Lowered attitudes about their intellectual knowledge and skills when applied to statistics may have resulted from their experiences in their research methods class. Most likely, these adult students had to complete a research project in which they collected and analyzed data. They had to use previous learning from statistics, and may only have achieved a superficial understanding of statistical concepts, which led to difficulties in trying to complete their research projects. A possible explanation for these results is similar to the research methods class with no prior statistics, as adult students' cognitive competence is often affected by past experiences. Carnell (2008) found no differences for cognitive competence, but believed that test performance and past experiences in different types of quantitative classes could impact cognitive competence, negatively or positively. These adult students may have had bad experiences in their statistics class. Or as Carnell reports, students with a "more extensive math background might feel differently about statistics than students with a more limited background" (p. 7). Hence, these adult students may have a limited mathematical background.

One other possible explanation can be offered for the adult students in the statistics class, the research methods class with no prior statistics, and the research methods class with prior statistics. Wiberg (2009) found increased scores for the SATS element, cognitive competence. This increase was attributed to the different type of teaching method he used in his revised class, which included data-driven problems and student-centered learning, which is reflective of a constructivist perspective of pedagogy, where the instructor's role is to "facilitate and negotiate meaning-making with the learner" (Merriam et al., 2007, p. 295). Wiberg's other class used a more traditional behaviorist approach, which did not show an increase for cognitive competence scores. Different types of teaching approaches could have affected the increase or decrease in cognitive competence scores.

Attitudes: Value

The SATS element, value, has two interesting research results. Adult students in the research methods class with no prior statistics had developed more negative attitudes about the value of statistics. Past research by Carnell (2008), Alldredge et al. (2006), Chadjipadelis and Andeadis (2006), and Froelich, Stephenson, and Duckworth (2008) showed no significant differences between pre- and post-test scores on the SATS element, value. However, they did not offer any explanation for their results.

One possible reason for these results for this research could be that adult students did not have a statistics class; therefore, they lack an understanding of statistical formulas and the variety of statistical tests that underline the results of research. This lack of knowledge and understanding of statistical computations can devalue the importance of statistics in one's personal and professional life.

Conversely, adult students in the research methods class with prior statistics had increased their attitudes about the value of statistics between pre- and post-test scores. One possible explanation for these results is that, because these adult students did have a statistics class, they developed an understanding of statistical formulas and the variety of statistical tests that underlie the results of research. Understanding and knowledge—knowing what underlies the data—can make it more useful and relevant in their personal and professional lives.

Wiberg (2009) offers another possible explanation. His research showed increased scores for the SATS element, value. Like his interpretation of the results for cognitive competence, he attributed this increase to the different type of teaching method used in the revised course, which included data-driven problems and student-centered learning. This embraced constructivist perspectives on teaching, while his other class, which used a traditional teaching approach (i.e., behaviorist), did not. As constructivism purports, "learning is essentially a process of making sense" (Fox, 2001, p. 30), as an important aspect of learning is about understanding, and in doing so, it takes us beyond the conception of rote learning. By using data-driven problems and student-centered learning, students become engaged in their learning of statistical concepts and can incorporate these ideas into their daily lives, and understand the value of statistics in their personal and professional lives.

Attitudes: Difficulty

Adult students' attitudes concerning statistics as a difficult subject, showed no changes from pre- to post-test scores for any of the class types, statistics, research methods class with a prior statistics, research methods with no prior statistics and the control group. Conclusively, they did not change their attitudes to whether they viewed statistics as being a difficult or not a difficult subject.

Past research showed similar results. Carnell (2008), Alldredge et al. (2006), Chadjipadelis and Andeadis (2006), Froelich et al. (2008), and Wiberg (2009) found no differences among their groups for the SATS element, difficulty. They offer no explanation for these results. However, one possible explanation could be that completing one statistics class, or one research methods class, or one statistics and research methods class is not enough time for students to change their minds about the difficulty of the class(es).

From all this, we can see that there is sparse research on students' attitudes concerning statistics. More research is definitely needed in this area. However, what we do know from this data is that adult students' past experiences, especially in mathematics, can affect their attitudes either negatively or positively toward statistics, and that instructors' teaching methodologies have not only an impact on students' learning, but also on their attitudes as well.

Attitudes, Affect, Cognitive Competence, Value and Difficulty: ANCOVA

Further analysis on the SATS scale was completed to examine for pre-test effects on post-test scores. After adjusting for pre-test scores from the SATS (i.e., affect, cognitive competence, value and difficulty), there were no significant differences across course types. Relationships between the pre- and post-test scores for each of the elements on the SATS were moderate. This effect may be able to explain some of the lack of significance among the class types, but the data did reveal a variety of significant results. And most importantly, results from the ANCOVA did show that research methods with a prior statistics class to have the highest estimated marginal means than any of the other class types, on all elements of the SATS, similar to the data results previously discussed.

Critical Stance

Critical stance examines if adult students are willing to challenge statistical messages they encounter from the media. No differences resulted among the class types, statistics, research methods class with a prior statistics, research methods class with no prior statistics, and the control groups on post-test scores. However, results for pre- to post-test differences showed a decrease for students in the statistics class and the research methods class with prior statistics. This indicates adult students from both classes were less confident in their abilities to challenge statistical messages at the end of the semester than at the beginning. One possible explanation is that they may have believed they understood more about statistics at the beginning of the semester than at the end of the semester, but because of different learning experiences in the classroom that showed the complexity of the subject, they were less confident in their abilities to challenge statistical messages at the end of the semester.

While it is disappointing to see results that indicate adult students are less likely to challenge statistical messages, the data from these analyses might be misleading. In adapting an instrument to reflect Gal's (2004) Model of Statistical Literacy, a separate scale was created to examine critical stance—no such scale has been developed, because the model has never been tested. The scale was tested and did yield small Cronbach alphas; however, this could be due to the scale's construction, as it contained a minimum number of statements. For this reason, it will be excluded from the discussion following.

Attitudes: Affect, Cognitive Competence, Value, Difficulty: Gender

The third sub-hypothesis, adult learners who are male will have more of a positive disposition toward statistics than adult learners who are female, was tested on each of the elements of the SATS scale, affect, cognitive competence, value and difficulty, and was confirmed from the results of total post-test scores. However, when each element was examined separately, the only significant difference between gender was on the element affect, with males scoring higher than females.

This may have occurred because many adult students believe statistics is related to mathematics (Gal, et al., 1997), and a gender difference between males and females has been discussed quite often in the literature on the teaching of mathematics. This phenomenon is described as, doing mathematics is doing masculinity, whereas doing well and doing mathematics is a male-oriented task—females cannot do mathematics as proficiently as males, because they are female (Mendick, 2005). It is interesting to note, although males have more positive feeling concerning statistics than females, there was no significant increase between pre- to post-test scores on any of the knowledge elements, statistical thinking, reasoning and literacy for males. However, for females there was a significant increase between pre- to post-test scores for statistical reasoning. The other SATS elements, cognitive competence, value and difficulty, did not show any differences between males and females.

Total pre- to post-test scores on the combined elements of the SATS showed no differences for males and for females. Likewise, when each of the elements, affect, cognitive competence, value and difficulty, were examined separately for males and females, there were no differences. This indicates that males and females did not change their attitudes toward statistics over the class duration of the semester.

These results are different from past research that examined gender and students' attitudes toward statistics. Carnell (2008), and Chadjipadelis and Andeadis, (2006) found no differences between gender on any of the elements, affect, cognitive competence, difficulty and value on the SATS scale. However, this was the only study that compared gender and the SATS elements. In the other few studies that were found, no gender was reported for Wiberg (2009) and Froelich et al. (2008). Alldredge et al. (2006) reported gender, but did not include it the analyses. Definitely, more research is warranted on gender and adult students' attitudes toward statistics.

Attitudes: Affect, Cognitive Competence, Value, Difficulty: First-Generation

The fourth sub-hypothesis, adult learners who are not first-generation learners will have more of a positive disposition toward statistics than learners who are first-generation adult learners, was tested on each element of the SATS scale, affect, cognitive competence, value and difficulty, and was disproved by post-test scores on the SATS scale.

However, when the elements from the SATS, affect, cognitive competence, value and difficulty were compared separately, cognitive competence was significant. Adult students who were first-generation scored higher than students who were not first-generation. This indicates that adult students who were first-generation have stronger beliefs about their intellectual knowledge and skills when applied to statistics. One possible explanation could be that first-generation adult students may have stronger beliefs about their abilities, because they are the first ones in their family to be a college student.

Interestingly, the only pre- to post-test difference was on the same element, cognitive competence, and no differences were found for affect, value and difficulty. The difference was found for adult students who were not first-generation students. No differences were found for first-generation students on any of the elements on the SATS.

Accordingly, scores for the adult students who were not first-generation college students increased significantly for cognitive competence, and hence, their beliefs about their confidence in their intellectual knowledge and skills when applied to statistics, but those who were first-generation did have higher post-test scores. One possible explanation could be that adult students who were not first-generation did well on exams in their classes, leading them to feel more competent in their abilities over the course of the semester.

However, we really do not know why these differences were found. There is no past research on first-generation or not first-generation adult students and statistical literacy, or their attitudes toward statistics. In the studies that were found examining students' attitudes toward statistics, Wiberg (2009), Froelich et al. (2008), Alldredge et al. (2006), Carnell (2008), and Chadjipadelis and Andeadis (2006), no adult student status was noted. Clearly, this is a new area of research that needs more attention, because as more and more college students are first-generation, their needs for learning may be different than students who are not first-generation.

Beliefs and Class Types

Unfortunately, there was no research found on students' beliefs in relation to their attitudes toward statistics. This is an area where research is needed, since students' beliefs underlie their attitudes. Even though there is no research on students' beliefs in relation to their attitudes, it is recommended in the literature that this needs examination. Gal and Ginsburg (1994) strongly suggest when examining students' attitudes toward statistics, which is accomplished by using a Likert-type scale, this scale should be used in conjunction with open-ended questions that reflect the elements from the scale. This will allow students to "describe the intensity and frequency of specific emotional responses, and elaborate on their source" (Gal & Ginsburg, n.p). This will give insight into the beliefs that underlie attitudes toward statistics. Hence, this part of the research is cutting-edge and needs to be followed up with more research that includes students' beliefs, along with their attitudes.

Results concerning adult students' beliefs were examined by class types (i.e., statistics, research methods class with no prior statistics, research methods class with prior statistics and a control group). No significant differences were found for the dependent variables *statistics formulas are easy or not easy to understand*, *I will or will not make a lot of math errors in statistics*, and *I will feel insecure or secure when doing statistics problems*. The only belief to reach significance was, *statistics is or is not relevant in my life*, and was significant for the research methods class with prior statistics. This is interesting, because results from the SATS scale on the element value for the research methods class with prior statistics showed high pre- and post-test scores for this group, and hence confirms the results. And the only group to reach significance with the research methods class with prior statistics was the control group, as their scores decreased from pre- to post-tests.

One possible explanation for these results could be that adult students enrolled in the research methods class with a prior statistics class experienced the culmination of both classes—understanding the statistics that underlie the resulting data and the concepts learned in a research methods—therefore, they can understand why statistics is relevant in their lives professionally and personally. But we really do not know until more research is completed.

Beliefs: Open-Ended Statements

As beliefs undermine our attitudes, data from four belief statements were further extrapolated by using open-ended responses for each of the statements. Each one will be discussed next.

Open-ended responses from the statement, *I will feel insecure or secure when* *doing statistics problems because* ….give us some insight into participants' beliefs about their academic level of comfort when completing statistics problems. Three themes emerged from the data, for both types of responses, secure or insecure: (a) adult students' beliefs in their mathematical abilities, (b) their sense of personal responsibility, and (c) confidence or a lack of confidence in their instructor's ability to help them master the material.

The first theme centered on adult students' beliefs in their mathematical abilities, probably because they hold the belief that statistics is the same discipline as mathematics—it is related, but different, as mathematics is about the certainty of numbers; statistics is concerned with the uncertainty of numbers. They believe mathematics often requires performing massive hand computations with large equations; therefore, so does statistics (Gal, et al., 1997). Albeit, in an introductory statistics class, some hand computations are used, but more often, computer programs compute the data (Proctor, 2002; Raymondo & Garrett, 1998, Warner & Meehan, 2001). At this level, statistics in general is concerned more on the interpretation of what the data results are either from by-hand calculations or computer-generated mathematical computations.

Adult students also relate their previous learning experiences in mathematics to their performance in statistics, whether past experiences were negative or positive. In other words, if I succeeded in mathematics, I will succeed in statistics, or if I failed in mathematics, I will fail in statistics. These results are somewhat supported. As Tobias (1994) suggests, adults' early memories of learning mathematics, which are often negative, are triggered when students become confused now in mathematics courses, which results from failing to understand some mathematical concept. This leads to losing a sense of confidence in their abilities and a loss of control over their comprehension, and it is similar in the statistics classroom, because adult students link statistics to mathematics.

The second theme focused on adult students' sense of personal responsibility. Many felt by paying attention in class, studying, practicing statistical problems and completing homework assignments, they would do very well in the class. On the other hand, some adult students simply had negative views about their abilities, which reflected a lack of personal responsibility—for example, it did not matter if they studied or not, they were doomed to fail, so consequently, they just gave up without trying. One possible explanation for this result could be adult students' past learning experiences, when they tried and failed before.

The third theme represents adult students' views about the role their instructors played in helping them master the class material. Those who were secure in their abilities credited this belief to their instructor's availability and willingness to help them, if the need arose. Insecure adult students often had negative views about their instructors, and held their instructors responsible for their lack of mastering the class material. Maybe some students were enrolled in classes where behaviorist methods of teaching were employed and others, in classes where constructivist perspectives were used by the instructors.

Importantly, there is no past research on these topics. New research is needed to understand the relationship between adult students' past class experiences and how this impacts their learning of statistics now, particularly in areas of mathematical achievement. In addition, adult students' sense of personal responsibility, and the relationship between a instructor and student and how it plays a role in his or her learning, needs further examination.

The second open-ended statement centered on these adult students' beliefs that they will or will not make a lot of math errors in statistics. And similar to the first statement, beliefs centered on (a) their mathematical abilities, (b) their sense of personal responsibility and (c) their perceived self-efficacy.

Again, many adult students' responses were based on their previous experiences in mathematics, and whether it was negative or positive was reflected in their responses. Those who believed in their mathematical abilities felt they would do well with the math computations that are required for the statistics class. On the other hand, those who did not believe they had any mathematical abilities felt they would make a lot of errors in mathematical calculations required for the statistics class. As explained by Gal, et al. (1997), "other than the commonly held belief that statistics is heavily mathematical and that statistics is a somewhat difficult discipline, students' beliefs about statistics as a domain remain mostly unexplored" (p. 4).

Adult students' sense of personal responsibility is apparent in those who believed they could correctly complete the mathematics computations required, because they would study, practice, and take notes in class. In other words, *if I work at it, I know I can master the class material*. Conversely, those who believed they would make a lot of errors had negative beliefs in their abilities, because some admittedly would not spend time working and practicing the mathematical computations. It did not matter to them if they tried, because in past experiences they tried and failed before. Past experiences in relation to adult students' beliefs is documented in the literature, as often they carry baggage from past experiences that can include negative beliefs about themselves in relation to mathematical issues (McLeod, 1992). However, what are the other past learning experiences that may affect adult students' beliefs about personal responsibility? Also, how do these beliefs affect their perceived self-efficacy? This is a new area for research to explore and investigate.

The third open-ended statement centered on the adult students' beliefs on whether or not statistics formulas are easy or not easy to understand. Three themes were reflected in their beliefs: (a) their confidence in their instructor's abilities to help them understand the formulas, (b) their sense of personal responsibility, and (c) their personal beliefs about the complexity of the formulas.

Adult students who believed they were able to master statistics formulas credited the ability of their instructor, who explained the formulas in precise details. No comments about instructors were made by adult students who felt statistics formulas were hard to understand. Possibly, those who believed they could master statistics formulas were enrolled in a class with instructors who had constructivist perspectives, and they participated in some type of active learning. Or perhaps there was a good relationship between the instructor and adult students, which made them comfortable in asking questions about formulas in class.

Second, a sense of personal responsibility emerged again. Adult students who felt they had mastered statistics formulas believed they achieved this goal through studying and practicing. Others, who viewed statistics formulas as not easy to understand, showed no sense of personal responsibility, and credited their failure to understand them to the fact that they are just hard, complex and require a lot of memorization. Again, these beliefs could relate to adult students' past experiences, especially in relation to mathematics. Tobias (1994) suggests adults' early memories of learning mathematics, which are often negative, affect their beliefs in their ability to do well in a statistics class. However, it could be more than just their past experiences in mathematics; it could be other learning experiences or life experiences. It is well known that adults bring with them knowledge from their own experiences, and new knowledge is constructed internally by transforming, organizing, and reorganizing previous knowledge (Cobb, 1994), as well as externally through the environment, and social factors, which are influenced by culture, language and social interactions.

The complexity of the statistics formulas was the third theme. Some adult students who felt formulas were easy to understand believed all you had to do was plug in the numbers. In a sense this is troublesome, because while it represents some learning, it is representative of superficial learning of statistics formulas. An understanding of why the numbers belong to a particular part of a formula enables a deeper understanding of statistical concepts. Others, who felt statistics formulas are too complex to understand, compared learning the formulas to learning a foreign language, because each letter in the formula represented a corresponding number from a statistics problem. Perhaps these adult students who struggled to understand statistics formulas were trying to understand the formulas at a deeper learning level than those who believed the formulas were easy to understand. In this sense, maybe the adult students who believed they were easy never really understood what the formulas represented.

One explanation could be the type of teaching methodologies the instructors used in the classroom. A behaviorist orientation could result in superficial learning of statistics formulas where the teacher was the sole information-giver to passive students (Hanley, 1994). In this setting, students simply understood where to put the numbers into the formulas; they did not understand why the numbers belonged to certain parts of the formulas.

Differently, students who felt the formulas were hard to understand, could have been part of a constructivist classroom, where the instructor tried to "facilitate and negotiate meaning-making with the learner" (Merriam, et al., 2007, p. 295). In this sense, the formulas had a particular function for each letter in relation to a statistics problem, and with a deeper of understanding of why and when one uses the formula, students would be able to construct their knowledge about formulas. Some in this research seemed to struggle to understand them.

The fourth open-ended statement centered on the adult students' beliefs on whether statistics was relevant or not relevant in their lives. Emerging themes were their beliefs about statistics in their future careers and lives after graduation. Many adult students believed statistics was relevant, because they would encounter it in their future careers. Conversely, some responses were troublesome, as some stated they would be working with people and not analyzing data.

The second important theme that emerged was adult students' beliefs that statistics would be relevant in their personal lives, as it applied to the weather, politics, sports, voting and economics. But some believed statistics would not be relevant in their personal lives and it should be left for the professionals. They believed learning statistics was a waste of time, and would never use it outside the classroom.

For students to have beliefs about the relevance or lack of relevance of statistics in their lives, these beliefs may have come from social influences, for example, their families. In the previous discussion about attitudes and first-generation status, there was no difference between adult students who were first-generation and those who were not first-generation on the attitudinal element, value. This result could be due to familial influences on adult students, as maybe their families did not place any value on statistics, and hence they did not either. And what might have influenced family members may be the lack of discussion on statistical literacy in mainstream society, as the adult literacy scale does not include statistical literacy. It does not include basic statistical concepts that may be relevant to issues in the media or work contexts (Gal, 2002). Interestingly, what is missing from the discourse in adult literacy may have impacted numerous adults negatively, as they may place no value on statistics.

Again, students' beliefs in statistics is a new research frontier—an area that has not been researched. However, from this discussion, we are beginning to understand some of the adult students' beliefs that may impact their ability to learn. These are: (a) the instructor's abilities to help them understand the formulas, (b) their sense of personal responsibility and (c) their personal beliefs about the complexity of the formulas. Research needs to expand and investigate further in this area, as there are numerous factors that can influence adults about the importance of statistics in their personal and professional lives.

This discussion was to provide insight into the research results for this study on statistical literacy and adult students. Five out of six hypotheses were disproved; however, this could be due to the combination of the variables, which are made up of the knowledge and dispositional elements of Gal's Model of Statistical Literacy. It might have been better to make the hypotheses more parsimonious—that is, breaking down the knowledge and dispositional elements into their simplest components and state hypotheses that would represent each one. Nevertheless, these results provide us with a wealth of information on the significance, and inherent possibilities, in statistical literacy to the field of adult education and implications for the teaching of statistical literacy.

Implications for Policy

From the discussion of the research results, three important implications for policy emerged for adult students to become statistically literate: (a) adult students need to complete both a statistics and a research methods course, (b) prerequisites for statistics should be re-examined, and (c) adult literacy needs to include components of statistical literacy on their measurement of adult literacy. A discussion of each implication follows.

First, from the research it is evident that adult students engaged in becoming college educated need to take both statistics and research methods classes in order to become statistically literate. In some disciplines, both classes are not required in the curriculum; only a statistics or a research methods class is required. Results from the data on the knowledge elements suggest learning gaps in these adult students. Those who completed only a statistics class scored lower on the knowledge elements questions that incorporate statistical concepts with research methodology—they understood how the statistics are computed, but lacked knowledge to apply it to a variety of research designs.

Likewise, adult students who had completed only a research methods class scored lower on the knowledge elements questions that were constructed of statistical and mathematical concepts. In a sense, it is like they are getting half an education toward becoming statistically literate. For society to move toward educating adults to become statistically literate, both classes need to be a requirement for the curriculum. And further, results showed that students who took both classes had increases in pre- and post-test scores on the knowledge elements of statistical thinking, reasoning and literacy.

Second, another policy issue that surfaced when examining students' beliefs about the complexity of statistics formulas was the type of mathematics that is required as a prerequisite for many statistics classes. In examining adult students' beliefs about statistics, some reported that statistics formulas are not easy to understand. To believe a formula is complex is another way of saying, I do not understand it. Perhaps, this could be due to the lack of mathematical education students become engaged in at the university level, or the type of mathematics courses they had taken. One course that is usually missing from the curriculum is a mathematics course on probability. This course is the backbone of statistics, as without it, statistics courses lose the meaning of the data. Hotelling et al. (1948) explains, "the whole foundation of descriptive statistical methods, of inductive inference, and the design of experiments rest upon probability theory" (p.105). However, classes on probability theory as a prerequisite for statistics are not very common in the social sciences, and are more likely to be a course in algebra. Maybe a course on probability would be more purposeful for students who are going to take an introductory statistics class.

And third, it became apparent from the research that there was no difference between first-generation and those who were not first-generation adult students when examining the value of statistics from the SATS scale. This is an important result if one considers one possible reason why there was no difference between these adult students. Many times social influences can affect the way we think about particular things in life, and the most important social influence an adult student can have is their family. Both types of families might not have placed any value on learning statistics, because it has been a missing element in the discourse on adult literacy in mainstream society. Adult literacy does not include basic statistical concepts that may be relevant to issues in the media or work contexts (Gal, 2002). Adult literacy includes, “where the numbers to be used have to be located in different types of forms or texts; where mathematical operations to be performed have to be inferred; or where quantitative information has to be gleaned from graphs or tables” (Gal, p. 23).

Accordingly, the discussion about statistical literacy does not end in the college curriculum, but rather starts at this point. One duty of a responsible government is to provide statistical information about the welfare of its citizens, and should be studied by all who aspire to improve the state of the nation (Schaeffer, 2001); hence, education toward statistical literacy has much broader implications through adult literacy. Adult literacy should be expanded to include basic statistical concepts that are used in everyday life. Because being illiterate in statistics alienates individuals to “the culture of silence, the masses are mute; that is they are prohibited from creatively taking part in the transformation of their society…” (Freire, 1970; 1998, p. 486). It is possible that parts of Gal’s Model of Statistical Literacy could be integrated to examine adults’ statistical literacy skills on adults who do not attend college.

In sum, policy implications are: first, adult college students need to take both a statistics and a research methods class; second, curriculum development needs to re-examine statistics courses prerequisites; is an algebra course sufficient, or would a course in probability be more purposeful? And third, statistical literacy needs to become inclusive when examining adult literacy.

Implications for Teaching

Adult students' attitudes can affect outcomes of learning, and their underlying beliefs form these attitudes. From this research it is evident that some of their beliefs may be affected negatively or positively within the classroom. Albeit, as previously stated, there is no previous research that examined adult students' beliefs toward statistics, but from this research it is evident that classroom techniques that provide a means for positive beliefs in adult students were focused on constructivist perspectives occurring within the classroom settings.

Albeit, a constructivist perspective focuses on individual subjective meanings, von Glasersfeld (1992), as previously stated, informs us that just because we can communicate and can agree on certain things, does not mean that we experience objective reality (i.e., a truth about the world), but consensual domains. Hence, "all our experience is subjective, but we manage in communication with those around us, to render our subjective meanings intersubjective and to create consensual domains" (Maturana as cited Schoenfeld, 1992, p. 291). In other words, these consensual domains are constructed out of our "in-context experience of each others' speeches and actions" (Goldin, 1990, p. 35). To communicate, individuals do not necessarily need to have identically shared meaning of things; only compatible meanings are necessary. It is these shared beliefs that become important in communication between the instructor and the learner (von Glasersfeld). It is through these consensual domains that constructivist ideas emerged from this research to inform us of the implications of teaching.

Accordingly, within these constructivist perspectives are the implications for teaching, which complement the needs of the adult students and offer solutions to help them succeed. Adult students' beliefs centered on a variety of topics, which are (a) personal responsibility, (b) past experiences, (c) perceived self-efficacy, (d) different experiences, and (e) the value of statistics. These will be discussed next within a constructivist paradigm.

Constructivist Perspective

Coinciding with our data, the teaching methodologies of mathematics instructors became a main topic in educational reform in the teaching of mathematics through the efforts of the National Council of Teachers of Mathematics (NCTM) in 1989, and soon expanded to include the teaching of statistics. Before the reform, most mathematics was taught by behaviorist teaching methods, where the instructor was the sole information-giver to passive students—instructors lecture as they transfer their thoughts and meanings to the students (Hanley, 1994). Different from this method, the NCTM stressed constructivist perspectives on learning. In a constructivist classroom, the instructor's role is to "facilitate and negotiate meaning-making with the learner" (Merriam et al., 2007, p. 295), with the purpose of learning to construct knowledge. Learning then becomes an active process rather than a passive one, as "knowledge is constructed rather than innate or passively absorbed" (Fox, 2001, pp. 24-25). To create this type of learning environment, instructors need to create a supportive atmosphere where students feel safe to explore statistical concepts (Aberson, et al., 2000).

Personal Responsibility. In addition, negotiating meaning-making with adult students requires them to change their roles in the classroom, which requires a certain amount of personal responsibility. They need to become actively engaged in their own learning through willing participation in hands-on activities (Mvududu, 2005), asking questions and making their own analogies and coming to their own conclusions (Merriam et al., 2007). And important to constructivist perspectives, many adult students responded that their sense of personal responsibility was their underlying belief in whether they felt secure or insecure when doing statistics problems, will or will not make a lot of math errors in statistics, and whether statistics formulas are or are not easy to understand. As instructors change their teaching methodologies to a constructivist pedagogy, and students are willing to become personally responsible for their learning, this match, between instructors and adult students, will have a positive effect on their beliefs and in their learning of statistics.

Past Learning Experiences. Especially valuable to this discussion is adult students' past learning experiences, as this was their responses to whether they will feel secure or insecure when doing statistics problems or whether they will or will not make a lot of math errors in statistics. Past learning experiences in mathematics for adult students were either negative or positive, and these experiences were related to their abilities to complete mathematical computations in their statistics class. To understand how past learning experiences can influence current learning experiences in the classroom, it is important to reexamine Piaget's ideas of cognitive constructivism.

Drawing on Piaget, it is understood that adults construct their own meaning of the world from their experiences and bring these experiences into the statistics classroom. These include their ideas, opinions, values and beliefs, which then relate to their abilities as to whether they will succeed or fail the class. Accordingly, if adults had bad experiences learning mathematics during their earlier days of school, it can leave them with negative views of their mathematical skills, which can affect their abilities to complete statistics problems correctly (Gal & Ginsburg, 1994). Instructors need to be sensitive to these issues for adult students and offer possible solutions, which can include recommending them to seek tutoring, or offering to meet with them outside of class time to help them learn and change their past beliefs in their abilities.

Perceived Self-Efficacy. Adult students' sense of perceived self-efficacy was also found to affect their beliefs on whether they could learn statistics. This was a belief stated on whether they would feel secure or insecure when doing statistics problems and whether they will or will not make a lot of math errors in statistics. As previously stated, adult students' beliefs in their abilities often emerge from prior learning experiences in which they failed or succeeded. For adult students who believe they will make a lot of math errors in statistics, it is important for them to know that statistics is not the same as mathematics, but uses some statistical equations, which enable researchers to come to a conclusion about phenomena. Mathematics is different from statistics in that mathematics is the science of "numbers and their operations, interrelations, combinations, generalizations, and abstractions, space configurations and their structure, measurement, transformations and generalizations…"(Manaster, 2001, pp. 67-68). On the other hand, statistics uses some mathematical calculations for data to allow interpretations from the results. Interpretations often combine ideas about chance and data that lead to inferences and interpreting statistical results. Important conceptual ideas, for example distributions, center, spreads, association, and sampling are necessary to understand, in order to be able to reason in statistics (Garfield & Chance, 2000). Hopefully, when adult students can understand the difference between mathematics and statistics, they may realize if they failed at mathematics before, it does not mean that they will fail at statistics.

In addition to adult students' understanding that mathematics is different than statistics, instructors need to be sensitive to adult students' attitudinal dispositions, because these factors can have an impact on the learning and teaching process in statistics (Mvududu, 2005). Additionally, because adult students bring into the classroom knowledge from their own experiences, new knowledge is constructed by transforming, organizing, and reorganizing previous knowledge (Cobb, 1994). Through these processes, according to constructivism ideas, "learning is an active process and knowledge is constructed rather than innate or passively absorbed" (Fox, 2001, pp. 24-25). And importantly, one of the best ways to help adult students develop a positive sense of self-efficacy is to have them actively engage in the learning process. This can be accomplished with the use of technology in a statistics class in a number of ways.
One way is to incorporate technology through the use of web-based learning tools, allowing adult students to learn specific topics (i.e., Central Limit Theorem, statistical power, or hypothesis testing) in statistics through computer-based simulations which include analysis, charts and graphs (Aberson et al., 2000; 2002; 2003). Another way is to incorporate computer programs, such as SPSS or Excel into the class curriculum in order for students to work with data sets. This gives adult students hands-on experience in data analysis as they enter and analyze data (Proctor, 2002; Raymondo & Garrett, 1998; Warner & Meehan, 2001).

Different Experiences. It is important for instructors to remember, when they present the same lesson to a variety of adult students, learning experiences may result differently for each, as "all knowledge is idiosyncratic and personal" (Fox, 2001, p. 29).

Hence, some adult students may need more instructional support from instructors or skilled peers who can help them bridge the gap between their current skill level and the desired one (Mvududu, 2005). Learners construct their own sets of meanings or understandings as, “knowledge is not a mere copy of the external world; nor is knowledge acquired by passive absorption or by simple transference from one person to another…Knowledge is made, not acquired” (Phillips, 2000, p.7). Hence, there are multiple ways in which adult students can construct their own knowledge. Constructing knowledge can be accomplished in the classroom through the use of small in-class group projects, or it can be accomplished by an instructor who includes time for interaction and discussion with adult students (Mills, 2003). Group work can also consist of adult students working together to develop a research question, design a data collection strategy, analyze data, or give an oral presentation of the results to the class (Chance, 2003). These provide sensory experiences, which can help adult students construct their knowledge about statistical concepts. Ideas about sensory experiences go back to the early ideas of constructivism. As explained by Bredo (2000) “John Locke viewed knowledge as synthesized from elementary sensory experience” (p. 128) and Kant believed that both “mental organization and sensory input are involved in knowing” (p. 129).

Moreover, classroom techniques that include adults working together in the classroom or on course projects reflect Vygotsky’s zone of proximal development, his most popular constructivist idea, in which importance is placed on social interactions with more knowledgeable others (Fosnot, 1996). Through these interactions, students can learn things they could not learn on their own (Wadsworth, 1996).

Value of Statistics. The final belief statement, statistics is or is not relevant in my life, examined how valuable learning statistics is to adult students, because if little value is placed on learning statistics, the learning may become superficial and only relevant to their grade. Adult students need to know statistics has value—for their future careers and everyday lives—and there are a couple of ways this can be accomplished in the class room.

One way is to use real data sets in the classroom for analyses and interpretation (Raymondo & Garret, 1998). Another is to bring regularly reported research from the media containing health and medicine issues into the classroom to enhance adult students' statistical reasoning abilities (Lawson, et al., 2003). The use of real-world research can bridge the gap between reality and numbers, as the context makes it meaningful (Garfield et al., 2002) and can help them to understand the value of learning statistics.

As adult students' beliefs are the underlying forces of their attitudes, it is important for them to have a positive belief in their abilities toward statistics and believe in the value of learning them. Hence, from this research we can see many important teaching implications for the statistics classroom. In summary, instructors need to have constructivist perspectives on teaching methods, as the old behaviorist teaching methods are obsolete; learning now consists of active student engagement in the classroom. This can be achieved in a number of ways; one is to include the use of technology by using computers for simulation and data analysis, another includes the use of real data sets. Or instructors can bring real research into the classroom from the media for discussion and critique, to enable a sense of value for statistics, and adult students can learn from one another through active engagement in small group projects completed in the classroom.

Implications for Future Research

Future research should continue to test Gal’s Model of Statistical Literacy; however, there were a few problems that arose during this research, which may indicate some modifications to the model are needed. An important issue that arose during the data analysis was the low validity scores on the critical stance scale. Because Gal’s Model of Statistical Literacy was never tested before, it emerged that no scale was available to examine critical stance; therefore, a scale was constructed by using the definition of critical stance. Perhaps, a new scale should be constructed for critical stance or be replaced by a scale to measure perceived self-efficacy. Based on the results from the open-ended statements, adult students’ beliefs in their self-efficacy were found to be important. Also, a scale examining self-efficacy could be used in tandem with the SATS scale—one for attitudes, and one to help examine the beliefs that underlie their attitudes. Maybe adult students’ critical stance cannot be operationalized by a scale; however, there are tested scales that measure perceived self-efficacy.

Because adult students’ self-efficacy beliefs were strongly stated by them on the open-ended responses, it would be practical for future research to measure perceived self-efficacy in students who are in a statistics or a research methods class and the teaching methodologies of the instructors. Instructors’ teaching style was strongly voiced by the adult students in relation to whether or not they would be successful in the class. This could provide us with more insight to how students’ perceptions of themselves and others play a role in their learning. Also, prior to engaging in research, it would be helpful to examine the course curriculum of the statistics and research methods classes, as there is variability in the course content (i.e., topics, etc.).

Aside from the cognitive factors related to statistical literacy, a discussion on the knowledge elements, statistical thinking, reasoning and literacy, needs to be addressed. Albeit the elements are broken down in to basic categories related to statistical literacy, such as (a) statistical thinking encompasses understanding of statistics related to the context, or more simply stated—to the concepts of research methods, (b) statistical reasoning encompasses statistical and mathematical knowledge, and (c), statistical literacy encompasses the reading of non-prose text, there is no differentiation between descriptive and inferential statistics.

Most statistics courses, especially introductory courses, are broken down into two distinct categories, descriptive and inferential statistics. "Descriptive statistics is the branch that deals with describing raw data in the form of graphics and sample statistics," while "inferential statistics is the branch that deals with inferring, or estimating population parameters from sample data" (Bennett, Briggs & Triola, 2003, p. 7). Albeit statistical literacy is defined as reading of non-prose text, it also includes "understanding and using the basic language and tools of statistics, knowing what statistical terms mean, understanding the use of statistical symbols and recognizing and being able to interpret representations of data" (ARTIST, 2006, n.p.). This definition is broad and can embody both categories, descriptive and inferential statistics.

Further, in examining the element, statistical reasoning, which is statistical and mathematical knowledge, this element also combined descriptive and inferential statistics. As previously discussed, statistical and mathematical knowledge embraces the ideas, such as (a) knowing why data are needed and how data can be produced, (b) familiarity with basic terms and ideas related to descriptive statistics, (c) familiarity with basic terms and ideas related to graphical and tabular display, (d) understanding the basic notions of probability, and (d) knowing how statistical conclusions or inferences are reached (Gal, 2004).

Unlike the other two elements, statistical thinking is a combination of descriptive and inferential statistics used in combination with the concepts learned in a research methods course. Hence, statistical thinking would be accomplished after completion of a statistics and a research methods course.

Separating the two branches of statistics, descriptive and inferential, is crucial to consider when examining learning gaps in adult students' learning of statistics. It would give us a more precise indicator where their learning deficits lay, either in descriptive or inferential statistics. In both types, descriptive and inferential, statistical and mathematical knowledge are embedded; similarly, but different. Because of the importance of prior mathematics experience, as demonstrated by adult students in this research, it would prudent to understand in which branch of statistics they are having difficulty; it is one branch or two? This would allow a better understanding of their prior skills in order for an instructor to focus on helping them to develop knowledge in both branches, descriptive and inferential. As it has already been stated, the reasoning involved in data-based statistical inference "is harder for students to grasp and explain than the comparable symbol-based problems and proofs in a typical calculus course" (Steen, p. 62).

Hence, when examining statistical literacy holistically, all of the elements from

Gal's (2004) Model of Statistical Literacy embrace all the necessary elements that form the concept of statistical literacy. But as Gal stated, "the elements in the proposed model should not be viewed as fixed and separate entities, but as a context-dependent dynamic set of knowledge and dispositions that together enable statistically literate behavior" (p.51), which may be a good way to examine statistical literacy; however, this may not be the best way to examine deficits in adult student's statistical literacy.

Accordingly, it may be fruitful to modify the knowledge elements, statistical literacy and reasoning, into two new categories, which embrace descriptive and inferential statistics. These categories would include statistical and mathematical knowledge, respectively, as statistical and mathematical knowledge is embedded within each. Separation of these categories would give more insight into the specific deficits in adult students' learning of statistics, and concepts in research methods, necessary to become statistically literate. It is indubitably apparent that more research is warranted on statistical literacy.

Other demographic characteristics should be considered in future research also. These would include adult students' ethnicity and prior mathematics courses. In this study, ethnicity was collected, but most students were Caucasian, hence, ethnicity differences would not be found.

And importantly, due to the paucity of research on students' attitudes toward statistics and notably their beliefs, more research needs to be undertaken in this area. This is a new frontier that needs to expand and investigate gender differences, and definitely needs to include adult students, as they are one group of students who have been missed in research that examines the teaching of statistics. Beliefs are ultimately very important, as beliefs underlie attitudes, and attitudes affect learning. And if negative beliefs systems can change to positive, then adult students can become statistically literate.

Limitations

It is important to briefly discuss the limitations of this dissertation research. It would have been impossible to conduct this research as a true experimental design with random sampling. Hence, this research used a quasi-experimental design. Therefore, it is limited in its ability to generalize the results of statistical literacy to all adults who are enrolled in college. Adult students who participated in this research were from small rural colleges on the east coast of the United States. This further limits the generalizabiity due to demographic variables.

In addition, because this research was a multi-campus effort, there were many different types of statistics and research methods classes, which were taught in a variety of ways, as there were multiple instructors teaching these classes. And because this research spanned many campuses, there were a variety of times that these classes met All these factors could have affected the research results and are acknowledged.

Conclusion

In conclusion, using Gal's Model of Statistical Literacy to guide this research allowed us to examine a variety of variables, which embrace statistical literacy. From these results, many important implications for policy and teaching were found. This dissertation should be the starting point for many more research endeavors into this unexplored frontier, as there is no previous research that has examined statistical literacy in adult students.

Because numbers underlie everyday decisions, from quantitatively based proposals that shape public policy in education and health (Steen, 2003) to decisions regarding political candidates (Moreno, 2001), there is a need to be statistically literate in order to sort social facts from social fallacies. Hence, adults need to be statistically literate, which should be the outcome of their learning experiences in college. In a sense, are we educating our adult students for statistical literacy, or are we just educating?

References

Aberson, C. L., Berger, D. E., Healy, M. R., & Romero, V. L. (2003). Evaluation of an interactive tutorial for teaching Hypothesis Testing Concepts. *Teaching of Psychology, 30*(1), 76-79.

Aberson, C. L., Berger, D. E., Healy, M. R., & Romero, V. L. (2002). An interactive tutorial for teaching statistical power. *Journal of Statistics Education, 10*(3) Retrieved March 6, 2007 from: http://www.amstat.org/publications/jse/

Aberson, C. L., Berger, D. E., Healy, M. R., Kyle, D, J., & Romero, V. L. (2000). Evaluation of an interactive tutorial for teaching the Central Limit Theorem. *Teaching of Psychology, 27*(4), 289-291.

Abercrombie, N., Hill, S., & Turner, B. (2000). *The penguin dictionary of sociology* (4th ed.). New York: Penguin Books.

Alldredge, J. R., Johnson, H.D., & Sanchez, J.J. (2006). Does viewing video of statistics in action affect students' attitudes? *Proceeding of the International Conference on the Teaching of Statistics, Salvador, Brazil.*

Andrews, R., Biggs, M., & Seidel, M. (1996*). The Columbia world of quotations*. Columbia University Press. Retrieved September 9, 2005 from http://www.bartleby.com/66/99/16799.html

ARTIST (2006). Defining and distinguishing statistical literacy, statistical reasoning and statistical thinking. Retrieved January 1, 2007 from: http://app.gen.umn.edu/artist/glossary.html

Babbie, E. (2004). *The practice of social research* (10th ed.). Belmont, CA: Wadsworth Publishing.

Bailey, R.P. (1941). On the treatment of certain problems of elementary probability. *The American Mathematical Monthly, 4*(84), 254-256.

Baten, W. D. (1950). Another opinion concerning statistical training below the college level. *The American Statistician, 4*(4), 24.

Baten, W. D. (1934). Combining two probabilities functions. *The Annuals of Mathematical Statistics, 5*(1), 13-20.

Baten, W. D. (1932). Comparison of certain probabilities. *The American Mathematical Monthly,* 39(5), 256-261.

Bechtel, W. (1988). *Philosophy of science.* Lawrence Erlbaum: Hillsdale, New Jersey.

Belcastro, A. & Purslow, V. T. (2006). *An integrative framework: Meeting the needs of the new-traditional student*. Paper presented at the Faculty Work and the New Academy meeting of the Association of American Colleges and Universities, Chicago, IL.

Bennett, J. O., Briggs, W. L., & Triola, M. F. (2003). *Statistical reasoning for everyday life.* New York: Addison Wesley.

Benton, T., & Craib, I. (2001). *Philosophy of social science: The philosophical foundations of social thought.* New York: Palgrave.

Borowski, E. J., & Borwein, J. M. (1991*). The Harper Collins dictionary: Mathematics.* New York: Harper Perennial.

Brendo, E. (2000). Reconsidering social constructivism: The relevance of George Herbert Mead's Interactionism. In D. C. Phillips (Ed.). *Constructivism in education: Opinions and second opinions on controversial issues* (pp. 127-158). Chicago: IL: University of Chicago Press.

Brooks, C. (1987). Superiority of women in statistics achievement. *Teaching of Psychology, 14*(1), 45.

Brown, M., Askew, Baker, D., Demuir, H., & Millett, A. (1998). Is the national numeracy strategy research based? *British Journal of Educational Studies, 46*(4), 362-385.

Bryant, C. (1985). *Positivism in social theory and research.* New York: St. Martin's Press.

Buck, J. (1985). A failure to find gender differences in statistics achievement. *Teaching of Psychology, 12*(2), 100.

Bushkovitch, A. V. (1970). Philosophy of science as a model for all philosophy. *Philosophy of Science, 37*(2), 307-311.

Bushway, S. D., & Flower, S. M. (2002). Helping criminal justice students learn statistics: A quasi-experimental evaluation of learning assistance. *Journal of Criminal Justice Education, 13*(1), 35-56.

Camp, A.H. (1932). Definitions of probability. *The American Mathematical Monthly, 39*(5), 285-288.

Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research.* Chicago, IL: Rand McNally and Company.

Carlson, R. (1978). Review of special issue on statistical education. *The American Statistician,32*(4), 139-140.

Carnell, L.J. (2008). The effect of a student-designed data collection project on attitudes toward statistics. *Journal of Statistics Education, 16*(1): Retrieved January 30, 2009 from: http://www.amstat.org/publications/jse/

Carnevale, A. P., & Desrochers, D. M. (2003). The democratization of mathematics. In B.L.Madison & L. A. Steen (Eds.), *Quantitative literacy: Why numeracy matters for schools and colleges* (pp. 7-20). Princeton, NJ: The National Council on Education and the Disciplines.

Carpenter, C. (1963). *History of American schoolbooks.* Philadelphia, PA: University of Pennsylvania Press.

Carter (1827). Improvement in arithmetic instruction*. American Journal of Education, 2*(1), 30-36.

Cerrito, P. (1999). Teaching statistical literacy. *College Teaching, 47*(1), 9-13.

Chance, B. L. (2005). Integrating pedagogies to teach statistics. In J. B.Garfield (Ed.) *Innovations in Teaching Statistics.* (pp. 101-110). Washington DC: The Mathematical Association of America.

Chadjipadelis, T., & Andreadis, I. (2006). Use of projects for teaching social statistics: case study. *Proceeding of the International Conference on the Teaching of Statistics, Salvador, Brazil.*

Cherulnik, P. D. (2001). *Methods for behavioral research: A systematic approach.* Thousand Oaks, CA: Sage Publications.

Clemm, R., & Gregory, R. (2000). Preparing adult students to be better decision makers. In I. Gal (Ed.) *Adult numeracy development: Theory, research, practice.* Cresskill, NJ: Hampton Press.

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

Cobb, P. (1994). Where is the mind? Constructivist and sociocultural perspectives on mathematical development. *Educational Researcher, 23*, 13-20.

Cohen, P.C. (2003). Democracy and the numerate citizen: Quantitative literacy in historical perspective. In B.L. Madison & L. A. Steen (Eds.), *Quantitative literacy: Why numeracy matters for schools and colleges* (pp. 7-20). Princeton, NJ: The National Council on Education and the Disciplines.

Cohen, P. C. (2001). The emergence of numeracy. In L. A. Steen (Ed), *Mathematics and democracy: The case for quantitative literacy.* Princeton, NJ: The National Council on Education and the Disciplines.

Copeland, A.H. (1937). Expansion of certain logical functions. *The American Mathematical Monthly, 44*(4), 213-218.

Cozby, P.C. (2004). *Methods in behavioral research* (8^{th} ed.). New York: McGraw Hill.

Cralley, E. L., & Ruscher, J.B. (2001). To share or not to share: Unique data sets facilitate performance in a psychology statistics course. *College Student Journal, 35*(4), 498-504.

Cremin, L. A. (1970). *American education: The colonial experience.* New York: Harper & Row Publishers.

Dacey, M. F. (1971). Mathematics for the undergraduate in the social sciences. *The American Mathematical Monthly, 78*(7), 784-788.

Dauphinee, T. L., Schau, C., & Stevens, J. (1997). Survey of attitudes toward statistics. *Structural Equation Modeling. 4,*129-141.

delMas, R., Garfield, J., & Chance, B. (2006). The Web-based ARTIST: An online resource for the assessment of instructional outcomes. Retrieved January 1, 2007 from: https://app.gen.umn.edu/artist/index.html

delMas, R.,Garfield, J., Ooms, A., & Chance, B. (2006). *Assessing students conceptual understanding after a first course in statistics*. Paper presented at the annual meetings of the American Educational Research Association, San Francisco, CA.

Derry, S., Levin, J. R., & Schauble, L. (1995), Stimulating statistical thinking through situated simulations. *Teaching of Psychology*, *22*, 51-57.

Dutka, S., & Kafka, F. (1950). Statistical training below the college level. *The American Statistician, 4*(1), 6-7.

Federer, W. T. (1978). Some remarks on statistical education. *The American Statistician, 32*(4), 117-121.

Field, A. (2005). *Discovering statistics using SPSS* (2nd ed.). Thousand Oaks, CA: Sage Publications, LTD.

Fitzpatrick, P. J. (1955). The early teaching of statistics in American colleges and universities. *The American Statistician, 9*(5), 12-18.

Fosnot, C. T. (1996). *Constructivism: Theory, perspectives and practice.* New York: Teachers College Press.

Fox, R. (2001). Constructivism examined. *Oxford Review of Education, 27*(1), 23-35.

Freire, P. (1970:1998). The adult literacy process as cultural action for freedom*. Harvard Educational Review, 69*(4), 480-499.

Froelich, A. G., Stephenson, W. R., & Duckworth, W. M. (2008).Assessment of materials for engaging students in statistical discovery. *Journal of Statistics Education, 16*(2): Retrieved January 30, 2009 from: http://www.amstat.org/publications/jse/

Gal, I. (2004). Statistical literacy: Meanings, components, responsibilities. In D. Ben-Zvi & J. Garfield (Eds.) *The challenge of developing statistical literacy, reasoning and thinking.* (pp. 47-78). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Gal, I. (2002). Systemic needs in adult numeracy education. *Adult Basic Education, 12*(1), 20-33.

Gal, I. (2000). Statistical literary: Conceptual and instructional issues. In D. Coben, J. O'Donoghue & G. E. Fitzsimons (Eds.) *Perspectives on adults learning mathematics* (pp.135-140). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Gal, I. (1997). Numeracy: Becoming literate with numbers. *Adult Learning, 9*(2).

Gal, I. (1995). Statistical tools and statistical literacy: The case of the average. *Teaching Statistics, 24*(1), 1-18.

Gal, I., & Ginsburg, L. (1994). The role of beliefs and attitudes in learning statistics: Towards an assessment framework. *Journal of Statistics Education, 2*(2) Retrieved February 9, 2007 from: http://www.amstat.org/publications/jse/

Gal, I., Ginsburg, L., & Schau, C. (1997). Monitoring attitudes and beliefs in statistics education. In I. Gal & Garfield, J. (Eds.) *The assessment challenge in statistics education.* (pp. 37-51). Netherlands: IOS Press.

Gane, M. (2006). *Auguste Comte.* New York: Routledge.

Garfield, J.B. (2003). Assessing statistical reasoning. *Statistics Education Research Journal, 2*(1), 22-38.

Garfield, J., & Chance, B. (2000). Assessment in statistics education: Issues and challenges. *Mathematical Thinking and Learning, 21*(1&2), 99-125.

Garfield, J., & Hogg, B., Schau, C., & Whittinghill, D. (2002). First courses in statistical science: The status of educational reform efforts. *Journal of Statistics Education, 10*(2) Retrieved May 26, 2007: from http://www.amstat.org/publications/jse/

Gaudard, M., & Hahn, G. J. (1991). An undergraduate concentration in applied statistics for mathematics majors. *The American Statistician, 45*(2), 115-120.

Goldin, G. A. (1990). Chapter 3: Epistemology, constructivism and discovery learning in mathematics. *Journal for Research in Mathematics Education,* 31-47.

Goodall. G. (2005). News and notes. *Teaching Statistics, 27*(3), 96.

Hanley, S. (1994). *On Constructivism* The University of Maryland at College Park. (NSF No. DUE 9255745).

Hart, D. K. (1964). Saint-Simon and the role of the elite. *The Western Political Quarterly, 17*(3), 423-431.

Hausfather, S. J. (1996). Vygotsky and schooling: Creating a social context for learning. *Action in Teacher Education, 18,* 1-10.

Hogg, R. V. (1991). Statistical literacy: Improvements are badly needed. *The American Statistician, 45*(4), 342-343.

Hotelling, H., Bartky, W., Deming, W. E., Milton, F., & Hoel, P. (1948). The teaching of statistics. *The Annals of Mathematical Statistics, 19*(1), 95-115.

Hotelling, H., (1940). Resolutions of the teaching of statistics. *The Annals of Mathematical Statistics, 11*(4), 472.

Howe, K. R., & Berv, J. (2000). Constructing constructivism, epistemological and pedagogical. In D. C. Phillips (Ed.). *Constructivism in education: Opinions and second opinions on controversial issues* (pp.19-40). Chicago, IL: University of Chicago Press.

Ho Yu, C. (2005). *Philosophical foundations of quantitative research methodology.* New York: University Press of America.

Huff, D. (1954). *How to lie with statistics.* New York: Norton.

Hyde, J.S., Fennema, E., & Lamon, S.J. (1990). Gender differences in mathematics Performance. *Psychological Bulletin, 107*, 139-155.

Jackson, C. S. (1917). A problem in probability. *The American Mathematical Monthly, 24*(2), 73-76.

James, T. (2005). Making sense of positivism. *Sociology Review, 15*(1), 22-24.

Johnson, H. D. & Dasgupta, N. (2005). Traditional versus non-traditional teaching: Perspectives of students in introductory statistics classes. *Journal of Statistics Education, 13*(2). Retrieved June 4, 2007 from: http://www.amstat.org/publications/jse/

Jones, P. J., & Coxford, A. F. (1970). Mathematics in the evolving schools. In the National Council of Teachers of Mathematics, *A history of mathematics education in the United States and Canada.* (pp.11-22). Washington DC: National Council of Mathematics Teachers.

Keeler, C. M., & Steinhorst, K. M. (1995). Using small groups to promote active learning in the introductory statistics course: A report from the field. *Journal of Statistics Education,3*(2). Retrieved May 23, 2007 from http://www.amstat.org/publications/jse/

Kiefer, A. J. (1948). *American children through their book.* Philadelphia, PA: University of Pennsylvania Press.

Kirk, R. E. (2002). *Teaching introductory statistics: Some things I have learned*. ERIC Document Reproduction Service No.ED 473 611

Kivinen, O., & Ristela, P. (2003). From constructivism to a pragmatist conception of learning. *Oxford Review of Education, 29*(3), 363-375.

Kline, T. (2005). *Psychological testing: A practical approach to design and evaluation.* Thousand Oaks, CA: Sage Publications Inc.

Kraus, J. W. (1961). The development of a curriculum in the early American colleges. *History of Education Quarterly, 1*(2), 64-76.

Kvam, P. H. (2000). The effect of active learning methods on student retention in engineering statistics. *The American Statistician, 34*(2), 136-140.

Lawson, T.J., Schwiers, M. Doellman, M., Grady, G. & Keinhofer, R. (2003). Enhancing students' ability to use statistical reasoning with everyday problems. *Teaching of Psychology, 30*(2), 107-110.

Lebanon Valley College (2007). *Programs of Study.* Retrieved September 30, 2007 from: http://www.lvc.edu/sociology/programs.aspx?pageType=sociology&bhiw=1143

Lee, J., Sax, L., Kim, K., & Hagedom, L. (2004). Understanding students' parental education beyond first-generation status. *Community College Review, 32*(1), 1-20.

Lee, M. W. (1960). Statistics and education for business. *The American Statistician, 14*(1), 16-19.

Manaster, A. B. (2001). Mathematics and numeracy: Mutual reinforcement. In L. A. Steen (Ed), *Mathematics and democracy: The case for quantitative literacy.* Princeton, NJ: The National Council on Education and the Disciplines.

Mancuso, S, (2001). Adult centered practices: Benchmarking study in higher education. *Innovative Higher Education, 25*(3), 165-181.

Mathematics: Review of the Cambridge course of mathematics (1822). *American Journal of Science and Arts, 5*(2), 304- 326.

Matthers, M. R. (2000). Appraising Constructivism in science and mathematics education. In D. C. Phillips (Ed.). *Constructivism in education: Opinions and second opinions on controversial issues* (pp.161-192). Chicago, IL: University of Chicago Press.

McCarty, L.P., & Schwandt, T. A. (2000). Seductive Illusions: Von Glasersfeld and Gergen on epistemology and education. In D. C. Phillips (Ed.). *Constructivism in education: Opinions and second opinions on controversial issues* (pp. 41-85). Chicago, IL: University of Chicago Press

McKeachie, W. J. (2002). *McKeachie's teaching tips.* Boston, MA: Houghton Mifflin Company.

Mendick, H. (2005). Mathematical stories: Why do more boys than girls choose to study mathematics at AS-level in England? *British Journal of Sociology of Education, 26*, (2), 235-251.

Merriam, S. B., & Caffarella, R. S. & Baumgartner, L. M (2007). *Learning in adulthood* (3rd ed). San Francisco, CA: John Wiley & Sons, Inc.

Mills, J.D. (2003). A theoretical framework for teaching statistics. *Teaching Statistics, 25*(2), 56-58.

Monroe, W.S. (1912). A chapter in the development of arithmetic teaching in the United States. *The Elementary School Teacher, 13*(1), 17-24.

Moore, D. S. (1997). New pedagogy and new content: The case of statistics. *International Statistical Review, 65*(2), 123-165.

Moore. D. S. (1990). Uncertainty. In L.A. Steen (Ed). *On the shoulders of giants: New approaches to numeracy.* pp-95-137. Washington D.C.: National Academy Press.

Moore, T. L., & Roberts, R. A. (1989). Statistics at Liberal arts colleges. *The American Statistician, 43*(2), 80-85.

Moore, T. L., & Witmer, J.A. (1991). Statistics within departments of mathematics at Liberal Arts Colleges. *The American Mathematical Monthly, 87*(5), 431-436.

Moreno, J.L. (2002, July). *Toward a statistically literate citizenry: What statistics everyone should know.* Proceedings of the Sixth International Conference on Teaching Statistics, Cape Town, South Africa.

Morris, E. (2001). The design and evaluation of Link: A computer-based learning system for correlation. *British Journal of Educational Technology, 32*(1), 39-52.

Morris, E., Joiner, R., & Scanlon, E. (2002). The contribution of computer-based activities to understanding statistics. *Journal of Computer Assisted Learning, 18,* 114-124.

Mvududu, N. (2005). Constructivism in the statistics classroom: From theory to practice. *Teaching Statistics, 27*(2), 49-54.

National Centers for Educational Statistics, United States Department of Education (NCES) (2006). *Key concepts and features of the 2003 national assessment of adult literacy.* Retrieved November 11, 2006 from http://nces.ed.gov/

National Council of Teachers of Mathematics (NCTM). (2007). http://www.nctm.org

Nelson, J. (1763). *An essay on the government of children, under three general heads, viz. health, manners and education.* London, England: R. and J. Dodsley.

Neuman, W. L. (2000). *Social research methods: Qualitative and quantitative approaches* (4^{th} ed.). Boston: Allyn and Bacon.

Nietz, J.A. (1961). *Old Textbooks.* Pittsburg, PA: University of Pittsburg Press.

Orcutt, J.D., & Turner, J. B. (1993). Shocking numbers and graphic accounts: Quantified images of drug problems in the print media. *Social Problems, 40* (2), 190-206.

Orrill, R. (2001). Preface, mathematics, numeracy and democracy. In L.A. Steen (Ed.) *Mathematics and democracy: The case for quantitative literacy* (pp. i-xx). Washington, DC: The Woodrow Wilson National Fellowship Foundation.

Pallant, J. (2007). *SPSS survival manual* (3rd ed,). Chicago, IL: Open University Press.

Pallant, J. (2006). *SPSS survival manual* (2nd ed.). Chicago, IL: Open University Press.

Parker, M., & Leinhardt, G. (1995). Percent: A privileged proportion. *Review of Educational Research, 63*(4), 421-481.

Passow, I.A. (1962), After Crowther. What? *Comparative Education Review. 5*(3), 175-181.

Paulos, J.A. (1989). *Innumeracy: Mathematical illiteracy and its consequences.* New York: Hill and Wang.

Pennsylvania State University (2007). *University bulletin.* Retrieved October 25 from: http://bulletins.psu.edu/bulletins/bluebook/university_course_descriptions.cfm?letter=P&dept=PSYCH

Pennsylvania State University Harrisburg (2007). *University bulletin.* Retrieved September 30, 2007 from: http://bulletins.psu.edu/bulletins/bluebook/baccalaureate_ degree_programs.cfm?letter=P&program=psyc.htm

Pennsylvania State University Harrisburg (2007a). *University bulletin.* Retrieved September 30, 2007 from: http://bulletins.psu.edu/bulletins/bluebook/baccalaureate_degree_programs.cfm?letter=C&program=crimj.htm

Pennsylvania State University Harrisburg (2007b). *University bulletin.* Retrieved September 30, 2007 from: http://bulletins.psu.edu/bulletins/bluebook/baccalaureate_degree_programs.cfm?letter=S&program=socio.htm

Pennsylvania State University, Schuylkill (2007)/ *University bulletin.* Retrieved September 30, 2007 from: http://bulletins.psu.edu/bulletins/bluebook/baccalaureate_degree_programs.cfm?letter=A&progr am=aojbs.htm

Pennsylvania State University, Schuylkill (2007a). *University bulletin.* Retrieved September 30, 2007 from: http://bulletins.psu.edu/bulletins/bluebook/baccalaureate_degree_programs.cfm?letter=A&progr am=aojbs.htm

Phillips, D.C. (1997). How, why, what, when and where: Perspectives on constructivism in psychology and education. *Issues in Education, 3*(2), 1-37.

Proctor, J. L. (2006). Academic achievement and statistical knowledge: A comparison of criminal justice and noncriminal justice majors. *Journal of Criminal Justice Education, 17*(1), 143-161.

Proctor, J. L. (2002). SPSS vs. Excel: Computing software, criminal justice students and statistics. *Journal of Criminal Justice Education, 13*(2), 433-442.

Pugalee, D.K. (1999). Constructing a model of mathematical literacy. *The Clearing House, 73*(1), 19-22.

Rajecki, D.W. (2002). Personal ad content analysis teaches statistical applications. *Teaching of Psychology, 29*(2), 119-122.

Raymondo, J. C., & Garrett, J. R. (1998). Assessing the introduction of a computer laboratory experience into a behavioral science statistics course. *Teaching Sociology, 26*(1), 29-37.

Robson, D. W. (1983). College founding in the new republic (1776-1800). *History of Education Quarterly, 23*(3), 323-341.

Rose, A. D. (1997). Expanding the scope of mathematics instruction. *Adult Learning, 9*(2).

Rose, N. (1991). Governing by numbers. Figuring out democracy. *Accounting, Organization, and Society, 16* (7), 673-692.

Rosen, L.P., Weil, L., & Von Zastrow, C. (2003). In B.L. Madison & L. A. Steen (Eds.), *Quantitative literacy: Why numeracy matters for schools and colleges* (pp. 7-20). Princeton, NJ: The National Council on Education and the Disciplines.

Rowbotham, E. J. (1919). *Story lives of great scientists.* London: Wells, Gardner, Darton & Co.

Rumsey, D. J. (2002). Discussion: Statistical literacy: Implications for teaching, research, and practice. *International Statistical Review, 70*(1), 32-36.

Rumsey, D. J. (2002a). Statistical literacy as a goal for introductory statistics courses. *Journal of Statistics Education, 10,* (3).) Retrieved August 16, 2005 from: http://www.amstat.org/publications/jse/

Russell, B. (1963). Logical Atomism*.* In A. J. Ayer (Ed.) *Logical Positivism.* (pp. 31-52). United States: The Free Press of Glencoe.

Schaeffer, R. L. (2001). Statistics and Quantitative Literacy. In B.L. Madison & L. A. Steen (Eds.), *Quantitative literacy: Why numeracy matters for schools and colleges* (pp. 7-20). Princeton, NJ: The National Council on Education and the Disciplines.

Schau, C., Dauphinee, T., Del Vecchio, A., & Stevens, J. (1991). Survey of attitudes toward statistics: (SATS) brief report. University of New Mexico.

Schau, C., Dauphinee, T., Del Vecchio, A., & Stevens, J. (1995). The development and validation of the survey of attitudes toward statistics. *Educational and Psychological Measurement, 55*(5), 868-875.

Schield, M. (1999). *Statistical literacy: Thinking critically about statistics*. Paper presented at the meeting of the Association of Public Data Users (ADPU) Retrieved from: http://web.augsburg.edu/~schield/MiloPapers/984StatisticalLiteracy6.pdf

Schoenfeld, A. H. (1992). Radical constructivism and the pragmatics of instruction. [Review of the book *Radical constructivism in mathematics education*]. *Journal for Research in Mathematics Education, 23*(3), 290-295.

Schram. C. (1996). A meta-analysis of gender differences in applied statistics achievement. *Journal of Educational Research and Behavioral Statistics, 21*(1), 55-70.

Schultz, D. P., & Schultz, S. E. (2000). *A history of modern psychology* (7th ed.). New York: Harcourt Brace College Publishers.

Shaddish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference.* New York: Houghton Mifflin Company.

Shaughnessy, J.J., Zechmeister, E. B., & Zechmeister. (2003). *Research methods in psychology* (6th ed.) New York: McGraw Hill.

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

Simon, W. (1956). History for Utopia: Saint-Simon and the idea of progress. *Journal of the History of Ideas, 17*(3), 311-331.

Sleight, E.R. (1943). Arithmetic according to Cocker. *National Mathematics Magazine, 17*(6), 248-257.

Smith, D. E., & Ginsburg, J. (1934). *A history of Mathematics in America before 1900.* Chicago, IL: The Open Court Publishing Company.

Smith, R. A., & Davis, S. F. (2007). *The psychologist as detective: An introduction to conducting research in psychology.* Upper Saddle River, NJ: Pearson.

Statistical Literacy (2007). *Home page*. Retrieved January 16, 2007 from http://www.statlit.org/

Steen, L.A. (2004). *Achieving quantitative literacy: An urgent challenge for higher education.* The Mathematical Association of America: Washington DC.

Steen, L.A. (2003). Data, shapes, symbols: Achieving balance in school mathematics. In B.L.Madison & L. A. Steen (Eds.), *Quantitative literacy: Why numeracy matters for schools and colleges* (pp. 7-20). Princeton, NJ: The National Council on Education and the Disciplines.

Steen, L.A. (2001). The case for quantitative literacy. In L. A. Steen (Ed), *Mathematics and democracy: The case for quantitative literacy.* Princeton, NJ: The National Council on Education and the Disciplines.

Steen, L. A. (1997). *Why numbers count: Quantitative literacy for tomorrow's America.* New York: College Entrance Examination Board.

Steen, L.A. (1990). Numeracy. *Daedalus, 119*(2), 211-231.

Symanzik, J. & Vukasinovic, N. (2006). Teaching an introductory statistics course with Cyber Stats, an electronic textbook. *Journal of Statistics Education, 14*(1). Retrieved February 23, 2007 from: http://www.amstat.org/publications/jse/

Tabachnick, B., & Fidell, L. S. (2007) *Using multivariate statistics.* (5th ed.). Boston, MA: Allyn and Bacon.

Thompson, K. (1976). *Auguste Comte: The foundation of sociology.* London: Cox & Wyman Ltd.

Tobias, S. (1994). *Overcoming math anxiety.* New York: W.W.Norton Co.

Tobin, K., & Tippins, D.J. (1993). Constructivism as a referent for teaching and learning. In K. Tobin (Ed.), *The practice of constructivism in science education* (pp. 3-22). Hillsdale, NJ: Erlbaum.

Utts, J. (2003). What educated citizens should know about statistics and probability. *The American Statistician, 57*(2), 74-79.

Utts, J., Sommer, B., Acredolo, C. Maher, M. W., & Matthews, H. R. (2003). A study comparing traditional and hybrid internet based instruction in introductory statistics class. *Journal of Statistics Education, 11*(2). Retrieved June 4, 2007 from: http://www.amstat.org/publications/jse/

Vanderstoep, S. W., & Shaughnessy, J. J. (1997). Teaching a course in research methods improves reasoning about real-life events. *Teaching of Psychology, 24*(2), 122-124. von Glasersfeld, E. (1992). Constructivism reconstructed: A reply to Suchting. *Science and Education, 1*, 379-384.

University of Illinois, (2009). *Understanding the data.* Retrieved February 4, 2009 from: www.turf.uiuc.edu/research/summaries/1994/understanding.pdf

University of Pittsburgh, Greensburg (2007). *Course Requirements* Retrieved September 30, 2007 from: http://www.upg.pitt.edu/prospective/P_AdministrationOfJustice.htm

Wade, B. (2003). *Statistical anxiety in undergraduate students and anxiety reduction techniques.* Unpublished master's thesis. Pennsylvania State University, Harrisburg, PA.

Wadsworth, B. J. (1996). *Piaget's theory of cognitive and affective development.* New York: Longman Publishers.

Wallman, K.K. (1993). Enhancing statistical literacy: Enriching our society. *Journal of the American Statistical Association, 88*(421), 1-8.

Ward, B. (2004). The best of both worlds: A hybrid statistics course. *Journal of Statistics Education, 12*(3). Retrieved June 4, 2007 from: http://www.amstat.org/publications/jse/

Ware, M. & Chastain, J. (1991). Developing selection skills in introductory statistics. *Teaching of Psychology, 18*(4), 219-222.

Warner, B. C., & Meehan, A.M. (2001). Microsoft Excel as a tool for teaching basic statistics. *Teaching of Psychology. 28*(4), 295-298.

Watson, J. M. (2002). Discussion: Statistical literacy before adulthood. *International Statistical Review, 70*(1), 26-30.

Weist, L. R., Higgins, H. J. & Frost, J. H. (2007). Quantitative literacy for social justice. *Equity & Education, 40,* 47-55.

Wiberg, M. (2009). Teaching statistics in integration with psychology. *Journal of Statistics Education, 17*(1). Retrieved January 25, 2009 from: http://www.amstat.org/publications/jse/

Wilks, S. S. (1947). Personnel and training problems in statistics. *The American Mathematical Monthly, 54*(9), 525-528.

Wright, C. D. (1888). The study of statistics in colleges. *American Economic Association Publications, 3*(1), 5-28.

Appendix A

Instruments to Measure Statistical Literacy

Gal's Model		Instrumentation
Knowledge Elements		
Literacy Skills	Statistical Literacy	ARTIST
Statistical & Mathematical Knowledge	Statistical Reasoning	ARTIST
Context Knowledge	Statistical Thinking	ARTIST
Critical Questions	Asking the Worry Questions on reported research in the news	Gal's Worry Questions
Dispositional Elements		
Beliefs & Attitudes	Attitude scale and opened-ended questions	Schau's SATS
	Extension of open-ended responses	
Critical Stance	Scale	Wade's SCS

Appendix B

Pre-Test Survey

1. Please enter you PSU ID (abc180)

2. What is your marital status?

____single
____married
____divorced
____widowed
____living with a partner

3. How many credits are you taking this semester?

4. What is your mother's education level?

____no high school dipolma
____High school or GED
____Some college or vocational training but no degree
____2 year college degree
____4 year college degree
____Master's degree
____PhD

5. What is your father's education level?

____No high school
____High school or GED
____Some college or vocational training but no degree
____2 year college degree
____4 year college degree
____PhD

6. Are you presently taking a statistics course this semester?

yes
no

7. Have you previously completed a statistics course?

yes
no

8. If you have took a prior statistics course, in what college department was the course taken? For example, was the course listed in the catalogue as a psychology course (Psych 200), or a Math course (Stat 200)?

____psychology
____sociology
____math
____other
____do not recall

9. Have you previously completed a research method course?

yes
no

10. Are you presently taking a research methods course this semester?

yes
no

11. If you are taking a research methods course, did you take a statistics course prior to taking Research Methods?

yes
no
not applicable

12. Did you start college the same year you graduated from high school?

yes
no

13. Are you dependent on your parents financially?

yes
no
partially dependent

14. Are you employed?

yes
no

15. Are you responsible for the care of another person (e.g., a child or adult)?

yes
no

16. What is your class ranking?

____freshman
____sophomore
____junior
____Senior

17. What is your age?

18. What is your sex?

____female
____male

19. What is your race/ethnicity?

____white non-hispanic
____black non-hispanic
____asian/pacific islander
____hispanic
____other

20. What is your grade point average (GPA)?

____0-1.0
____1.1 - 2.0
____2.1-3
____3.1-4.0

Strongly Disagree Neither Agree nor Disagree Strongly Agree

21. I will like statistics.

1 2 3 4 5 6 7

22. I will feel insecure when I have to do statistics problems.

1 2 3 4 5 6 7

23. I will have trouble understanding statistics because of how I think.

1 2 3 4 5 6 7

24. Statistics formulas are easy to understand.

1 2 3 4 5 6 7

25. Statistics is worthless.

1 2 3 4 5 6 7

26. Statistics is complicated subject.
1 2 3 4 5 6 7

27. Statistics should be a required part of my professional training.
1 2 3 4 5 6 7

28. Statistical skills will make me more employable.
1 2 3 4 5 6 7

29. I will have no idea of what's going on in statistics.
1 2 3 4 5 6 7

30. Statistics is not useful to the typical professional.
1 2 3 4 5 6 7

31. I will get frustrated going over statistics tests in class.
1 2 3 4 5 6 7

32. Statistical thinking is not applicable in my life outside my job.
1 2 3 4 5 6 7

33. I will use statistics in my everyday life.
1 2 3 4 5 6 7

34. I will be under stress during statistics class.
1 2 3 4 5 6 7

35. I will enjoy taking statistics courses.
1 2 3 4 5 6 7

36. Statistical conclusions are rarely presented in everyday life.
1 2 3 4 5 6 7

37. Statistics is a subject quickly learned by most people.
1 2 3 4 5 6 7

38. Learning statistics requires a great deal of discipline.
1 2 3 4 5 6 7

39. I will have no application for statistics in my profession.
1 2 3 4 5 6 7

40. I will make a lot of math errors in statistics.
1 2 3 4 5 6 7

41. I will be scared by statistics.
1 2 3 4 5 6 7

42. Statistics involves massive computations.
1 2 3 4 5 6 7

43. I can learn statistics.
1 2 3 4 5 6 7

44. I will understand statistics equations.
1 2 3 4 5 6 7

45. Statistics is irrelevant in my life.
1 2 3 4 5 6 7

46. Statistics is highly technical.
1 2 3 4 5 6 7

47. I will find it difficult to understand statistics concepts.
1 2 3 4 5 6 7

48. Most people have to learn a new way of thinking to do statistics.
1 2 3 4 5 6 7

Circle how you feel and complete the sentence.

49. I will feel INSECURE OR SECURE when doing statistics problems because

50. Statistics IS or IS NOT relevant in my life because

51. I WILL or WILL NOT make a lot of math errors in statistics because

52. Statistics formulas are EASY or NOT EASY to understand because

53. I never challenge newspaper articles that report research when it is written by a governmental official.
1 2 3 4 5 6 7

54. I am confident in my ability to challenge data that is written by government officials that may lead to the creation of social programs.
1 2 3 4 5 6 7

55. I tend to be skeptical about advertisements for medications when they are advertised by well know public figures.
1 2 3 4 5 6 7

56. I can interpret most charts representing research that is reported in the media.
1 2 3 4 5 6 7

57. I am uncomfortable discussing research reported in the media with family and fiends.
1 2 3 4 5 6 7

58. I feel confident in my ability to challenge research presented by the media.
1 2 3 4 5 6 7

59. Research reported in newspapers always contains enough information to make an informed decision about the validity of the study.
1 2 3 4 5 6 7

60. Internet news sources reporting current research, often lack complete information.
1 2 3 4 5 6 7

61. Research that is reported in peer-reviewed journals always uses correct statistical analysis.
1 2 3 4 5 6 7

62. I have no idea what I am reading when I read articles about research.
1 2 3 4 5 6 7

63. A college student is gathering data on the type of transmissions students prefer to drive. The variable is "type of transmission" 1 = manual transmission, 2= automatic transmission. What type of variable is this?

____nominal
____ordinal
____ratio
____interval

64. In a study to assess whether aspirin reduces the risk of pregnant women developing hypertension, 34 pregnant women were randomly assigned to receive a low dosage of aspirin every day, while 31 pregnant women received a placebo every day. Of those in the aspirin group 4 developed hypertension during their pregnancy, compared to 11 of those in the placebo group. This study would be considered an

____experiment
____observational study

65. A set of data are put in numerical order and a statistic is calculated that divides the data into two equal parts. Which of the following statistics was computed?

____mean
____interquartile range
____standard deviation
____median

66. A psychologist is studying the effects of drug and electroshock therapy on a subject's ability to solve simple tasks. The number of tasks completed in a 10 minute period is recorded for subjects who have been randomly assigned to compare four treatment groups drug with electroshock, drug with out electroshock, no drug with electroshock, and no drug/no electroshock. Sixty-four participants were randomly selected students form a large university. Identify the population and the sample.

67. A recent article in an educational research journal reports a correlation of +.80 between math achievement and overall math aptitude for a large sample of students. It also reports a correlation of -.80 between math achievement and math anxiety for the same group of students. Only students with scores on all three measures were included in the study. Which of the following interpretations is the most correct?

____The correlation of +.80 indicates a stronger relationship than the correlation of -80.
____The correlation of +.80 is just as strong as the correlation of -.80.
____It is impossible to tell which correlation is stronger.

68. Suppose we want to know the opinions of college students about a proposed tuition increase and we decided to ask 500 students to participate in a survey. Give an example of how to take a random sample of 500 college students at this college.

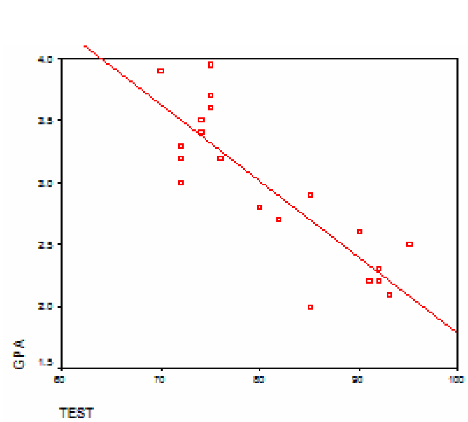
69. Suppose students want to determine whether the caffeine in a cup of regular coffee consumed each morning before class can improve the performance of a typical student on a statistics exam. Suppose all students in the 8 am section are given the treatment (1 cup of coffee) and all students in the 9am section are not permitted to have any caffeine before class. It turned out that the students in the 9am section did better on the exam. This is not a well-designed study. Explain why "confounding" makes it difficult to draw conclusions about the effect of caffeine on students' exam performance from this study. (Identify a potential confounding variable).

70. A group of 30 introductory statistics students took a 25 item test. The mean and standard deviation were compared and the standard deviation was 0. You know that

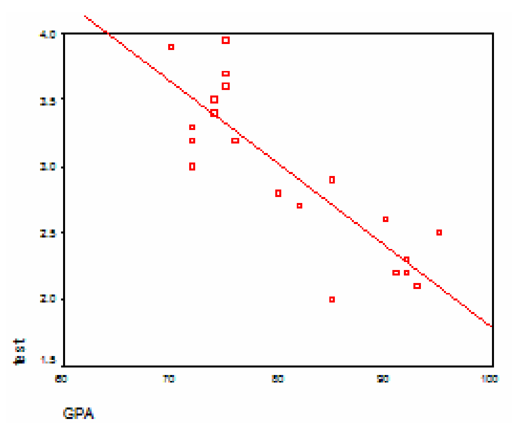
_____About half of the scores were above the mean
_____The test was so hard that everyone missed all items.
_____A calculation error must have been made in determining the standard deviation.
_____Everyone correctly answered the same number of items

71. The director of admissions in a small college administered a newly designed entrance test to 20 freshman students selected at random. The purpose of the study was to determine whether students' grade point average (GPA) at the end of their freshman year can be predicted form the entrance test score.

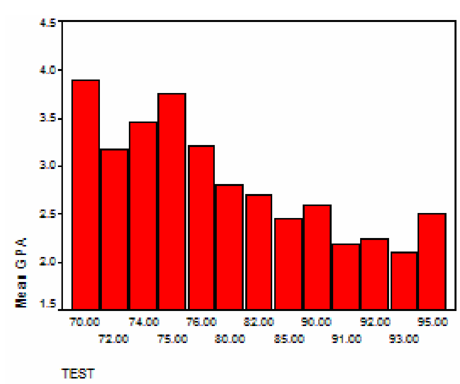
Select the chart that is most appropriate for this research study.

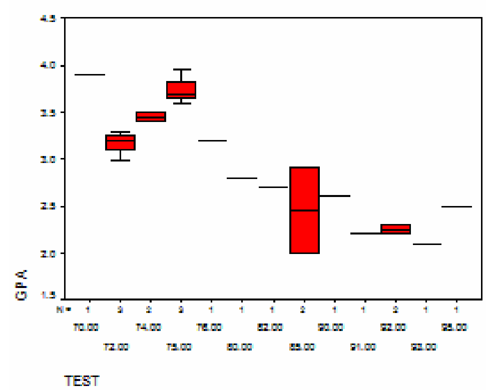


A



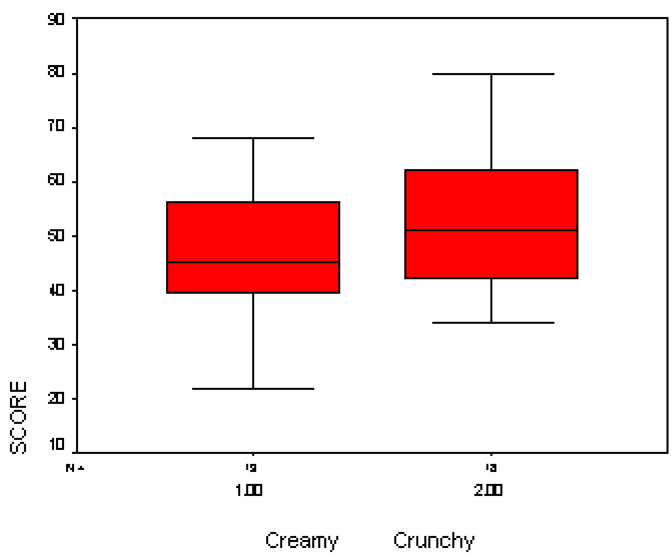
B





C **D**

A
B
C
D



72. Comment on the side by side box plots--which group has a higher median?
Reasoning:

73. The following statement is statistically incorrect. Explain what is wrong. Student ratings of professors' teaching and colleagues' rating of their research have a correlation of ran =+1.21. This shows that professors who are good teachers are also good researchers.

74. A researcher uses a chi-square test to determine if there is a relationship between categorical variables. Which of the following p values indicate the strongest evidence of such a relationship?

A. .02
B. .001
C. .05
D. .10

75. People who eat lots of fruits and vegetables have lower rates of colon cancer than those who eat little of these foods. Fruits and vegetables are rich in antioxidants such as vitamins A, C, and E. Will taking antioxidants help prevent colon cancer? A clinical trial studied this question with 864 people who were at risk of colon cancer. The participants were divided into four groups; 1) daily beta carotene, 2) daily vitamins C and E, 3) all three vitamins every day 4) daily placebo. After four years the researchers were surprised to find not significant differences in colon cancer among the groups. What type of statistical analyses would be used to test for difference between four groups on evidence of colon cancer?

A. MANOVA
B. Independent T Test
C. ANOVA
D. Chi-Square

76. (New questions same example) People who eat lots of fruits and vegetables have lower rates of colon cancer than those who eat little of these foods. Fruits and vegetables are rich in antioxidants such as vitamins A, C, and E. Will taking antioxidants help prevent colon cancer? A clinical trial studied this question with 864 people who were at risk of colon cancer. The participants were divided into four groups; 1) daily beta carotene, 2) daily vitamins C and E, 3) all three vitamins every day 4) daily placebo. After four years the researchers were surprised to find not significant differences in colon cancer among the groups. Researchers concluded that there were no statistically significant differences after analyzing the data. What does no significant difference mean in describing the results of this study?

77. Suggest one possible lurking variable that could explain why people who eat lots of fruit and vegetables have lower rates of colon cancer.

78. Does knowing a college student's SAT score tell us anything about his or her first year college grade point average? What type of statistical analysis should be used in trying to determine the answer to this?

A. Test one mean against a hypothesized constant.
B. Test the difference between two means (independent samples)
C. Test the difference in means between two pairs of dependent samples
D. Test that a correlation coefficient is not equal to 0 correlation analysis.
E. Use a chi-square test of association.

79. Bob and Bill each bought one ticket 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?

A. Bill
B. Bob
C. They have an equal chance of winning.
D. Not enough information to tell.

80. A newspaper article on drunk driving cited data on traffic deaths in Rhode Island. Forty-two percent of all fatalities occurred on Friday, Saturday and Sunday, apparently because of increased drinking on the weekends. Are you surprised that 42% of accidents occur on those days? Why or Why not?

Read the following short article; would you use any of the questions following the article to challenge the results?

81. Severe Weight Loss Can Be Caused By Chewing Gum, Doctors Report Science Daily (Jan. 11, 2008) —

Doctors warn of excess sorbitol intake, a widely used sweetener in "sugar-free" products such as chewing gum and sweets in an article in the British Medical Journal. http://www.sciencedaily.com/ Sorbitol has laxative properties and is poorly absorbed by the small intestine.

Their advice follows the cases of two patients with chronic diarrhea, abdominal pain and severe weight loss. Although extensive investigations were carried out, final diagnosis was only established after detailed analysis of eating habits. On questioning, both patients admitted consuming substantial amounts of sugar-free gum and sweets. The first patient (a 21 year old woman) chewed large amounts of sugar-free gum, accounting for a total daily dose of 18-20g sorbitol (one stick of chewing gum contains about 1.25g sorbitol). The second patient (a 46 year-old man) reported chewing 20 sticks of sugar-free gum and eating up to 200g of sweets each day, which together contained around 30g sorbitol.

After both patients started a sorbitol free diet, diarrhea subsided, normal bowel movements resumed and weight gain was achieved. As possible side effects are usually found only within the small print on foods containing sorbitol, consumers may be unaware of its laxative effects and fail to recognize a link with their gastrointestinal problems, write the authors. In conclusion, they say, our cases demonstrate that sorbitol consumption can cause not only chronic diarrhea and functional bowel complaints but also considerable unintended weight loss (about 20% of usual body weight).

Worry Questions

1. Where did the data come from? What kind of study was it? Is this kind of study reasonable in this context?

2. Was a sample used? How was it sampled? How many people did actually participate? Is the sample large enough? Did the sample include people/units which are representative of the population? Is the sample biased is some way? Overall, could this sample reasonably lead to valid inferences about the target population?

3. How reliable or accurate were the instruments or measures (tests, questionnaires, interviews) used to generate the reported data?

4. What is the shape of the underlying distribution of raw data? Does it matter how it is shaped?

5. Are the reported statistics appropriate for this kind of data? Was an average to summarize ordinal data; is a mode a reasonable summary? Could outliers cause a summary statistic to misrepresent the true picture?

6. Is a given graph drawn appropriately, or does it distort the data?

7. How was this probabilistic statement derived? Are there enough credible data to justify the estimate of likelihood given?

8. Overall, are the claims made here sensible and supported by the data? Is correlation confused with causation, or a small difference made to loom large?

9. Should additional information or procedures be made available to enable me to evaluate the sensibility of these arguments? Is something missing? Did the writer conveniently forget to specify the base of a reported percent-of-change, or the actual sample size?

10. Are there alternative interpretations for the meaning of the findings or different explanations for what caused them e.g., an intervening or a moderator variable affected the results? Are there additional or different implications that are not mentioned?

APPENDIX C

Post-Test Survey

1. Please enter you PSU ID (abc180)

2. What is your marital status?

____single
____married
____divorced
____widowed
____living with a partner

3. How many credits are you taking this semester?

4. What is your mother's education level?

____no high school dipolma
____High school or GED
____Some college or vocational training but no degree
____2 year college degree
____4 year college degree
____Master's degree
____PhD

5. What is your father's education level?

____No high school
____High school or GED
____Some college or vocational training but no degree
____2 year college degree
____4 year college degree
____PhD

6. Are you presently taking a statistics course this semester?

yes
no

7. Have you previously completed a statistics course?

yes
no

8. If you have took a prior statistics course, in what college department was the course taken? For example, was the course listed in the catalogue as a psychology course (Psych 200), or a Math course (Stat 200)?

____psychology
____sociology
____math
____other
____do not recall

9. Have you previously completed a research method course?

yes
no

10. Are you presently taking a research methods course this semester?

yes
no

11. If you are taking a research methods course, did you take a statistics course prior to taking Research Methods?

yes
no
not applicable

12. Did you start college the same year you graduated from high school?

yes
no

13. Are you dependent on your parents financially?

yes
no
partially dependent

14. Are you employed?

yes
no

15. Are you responsible for the care of another person (e.g., a child or adult)?

yes
no

16. What is your class ranking?

____freshman
____sophomore
____junior
____senior

17. What is your age?

18. What is your sex?

____female
____male

19. What is your race/ethnicity?

____white non-hispanic
____black non-hispanic
____asian/pacific islander
____hispanic
____other

20. What is your grade point average (GPA)?

____0-1.0
____1.1 - 2.0
____2.1-3
____3.1-4.0

Strongly Disagree Neither Agree nor Disagree Strongly Agree

21. I like statistics.

1 2 3 4 5 6 7

22. I feel insecure when I have to do statistics problems.

1 2 3 4 5 6 7

23. I have trouble understanding statistics because of how I think.

1 2 3 4 5 6 7

24. Statistics formulas are easy to understand.

1 2 3 4 5 6 7

25. Statistics is worthless.

1 2 3 4 5 6 7

26. Statistics is complicated subject.

1 2 3 4 5 6 7

27. Statistics should be a required part of my professional training.

1 2 3 4 5 6 7

28. Statistical skills will make me more employable.

1 2 3 4 5 6 7

29. I have no idea of what's going on in statistics.

1 2 3 4 5 6 7

30. Statistics is not useful to the typical professional.

1 2 3 4 5 6 7

31. I get frustrated going over statistics tests in class.

1 2 3 4 5 6 7

32. Statistical thinking is not applicable in my life outside my job.

1 2 3 4 5 6 7

33. I use statistics in my everyday life.

1 2 3 4 5 6 7

34. I am under stress during statistics class.

1 2 3 4 5 6 7

35. I enjoy taking statistics courses.

1 2 3 4 5 6 7

36. Statistical conclusions are rarely presented in everyday life.

1 2 3 4 5 6 7

37. Statistics is a subject quickly learned by most people.

1 2 3 4 5 6 7

38. Learning statistics requires a great deal of discipline.

1 2 3 4 5 6 7

39. I will have no application for statistics in my profession.
1 2 3 4 5 6 7

40. I make a lot of math errors in statistics.
1 2 3 4 5 6 7

41. I am scared by statistics.
1 2 3 4 5 6 7

42. Statistics involves massive computations.
1 2 3 4 5 6 7

43. I can learn statistics.
1 2 3 4 5 6 7

44. I understand statistics equations.
1 2 3 4 5 6 7

45. Statistics is irrelevant in my life.
1 2 3 4 5 6 7

46. Statistics is highly technical.
1 2 3 4 5 6 7

47. I found it difficult to understand statistics concepts.
1 2 3 4 5 6 7

48. Most people have to learn a new way of thinking to do statistics.
1 2 3 4 5 6 7

49. I feel INSECURE OR SECURE when doing statistics problems because

50. Statistics IS or IS NOT relevant in my life because

51. I WILL or WILL NOT make a lot of math errors in statistics because

52. Statistics formulas are EASY or NOT EASY to understand because

53. I never challenge newspaper articles that report research when it is written by a governmental official.

1 2 3 4 5 6 7

54. I am confident in my ability to challenge data that is written by government officials that may lead to the creation of social programs.

1 2 3 4 5 6 7

55. I tend to be skeptical about advertisements for medications when they are advertised by well know public figures.

1 2 3 4 5 6 7

56. I can interpret most charts representing research that is reported in the media.

1 2 3 4 5 6 7

57. I am uncomfortable discussing research reported in the media with family and fiends.

1 2 3 4 5 6 7

58. I feel confident in my ability to challenge research presented by the media.

1 2 3 4 5 6 7

59. Research reported in newspapers always contains enough information to make an informed decision about the validity of the study.

1 2 3 4 5 6 7

60. Internet news sources reporting current research, often lack complete information.

1 2 3 4 5 6 7

61. Research that is reported in peer-reviewed journals always uses correct statistical analysis.

1 2 3 4 5 6 7

62. I have no idea what I am reading when I read articles about research.

1 2 3 4 5 6 7

63. A college student is gathering data on the type of transmissions students prefer to drive. The variable is "type of transmission" 1 = manual transmission, 2= automatic transmission. What type of variable is this?

____nominal
____ordinal
____ratio
____interval

64. In a study to assess whether aspirin reduces the risk of pregnant women developing hypertension, 34 pregnant women were randomly assigned to receive a low dosage of aspirin every day, while 31 pregnant women received a placebo every day. Of those in the aspirin group, 4 developed hypertension during their pregnancy, compared to 11 of those in the placebo group. This study would be considered an

____experiment
____observational study

65. A set of data are put in numerical order and a statistic is calculated that divides the data into two equal parts. Which of the following statistics was computed?

____mean
____interquartile range
____standard deviation
____median

66. A psychologist is studying the effects of drug and electroshock therapy on a subject's ability to solve simple tasks. The number of tasks completed in a 10 minute period is recorded for subjects who have been randomly assigned to compare four treatment groups drug with electroshock, drug with out electroshock, no drug with electroshock, and no drug/no electroshock. Sixty-four participants were randomly selected students form a large university. Identify the population and the sample.

67. A recent article in an educational research journal reports a correlation of +.80 between math achievement and overall math aptitude for a large sample of students. It also reports a correlation of -.80 between math achievement and math anxiety for the same group of students. Only students with scores on all three measures were included in the study. Which of the following interpretations is the most correct?

____The correlation of +.80 indicates a stronger relationship than the correlation of -.80.
____The correlation of +.80 is just as strong as the correlation of -.80.
____It is impossible to tell which correlation is stronger.

68. Suppose we want to know the opinions of college students about a proposed tuition increase and we decided to ask 500 students to participate in a survey. Give an example of how to take a random sample of 500 college students at this college.

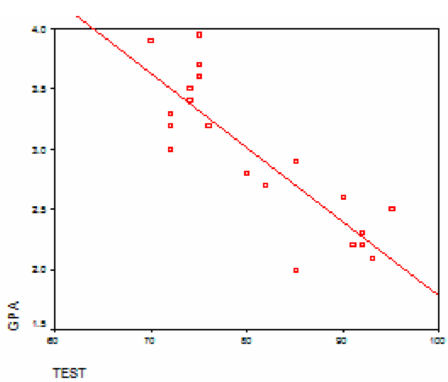
69. Suppose students want to determine whether the caffeine in a cup of regular coffee consumed each morning before class can improve the performance of a typical student on a statistics exam. Suppose all students in the 8 am section are given the treatment (1 cup of coffee) and all students in the 9am section are not permitted to have any caffeine before class. It turned out that the students in the 9am section did better on the exam. This is not a well-designed study. Explain why "confounding" makes it difficult to draw conclusions about the effect of caffeine on students' exam performance from this study. (Identify a potential confounding variable).

70. A group of 30 introductory statistics students took a 25 item test. The mean and standard deviation were compared and the standard deviation was 0. You know that

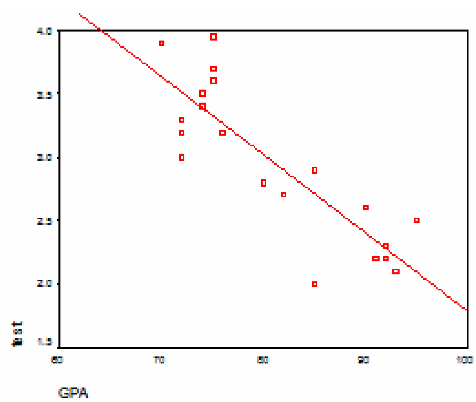
_____About half of the scores were above the mean
_____The test was so hard that everyone missed all items.
_____A calculation error must have been made in determining the standard deviation.
_____Everyone correctly answered the same number of items

71. The director of admissions in a small college administered a newly designed entrance test to 20 freshman students selected at random. The purpose of the study was to determine whether students' grade point average (GPA) at the end of their freshman year can be predicted form the entrance test score.

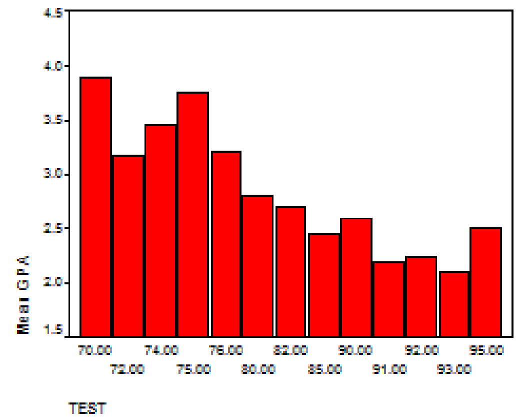
Select the chart that is most appropriate for this research study.



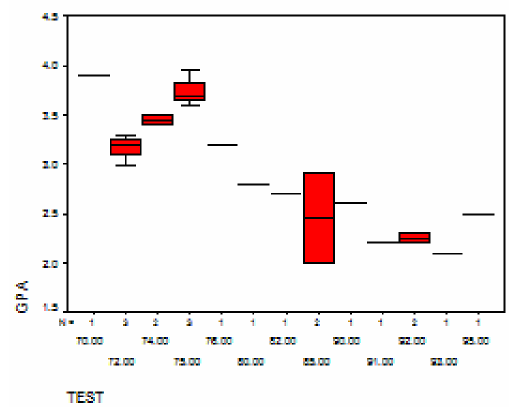
A



B

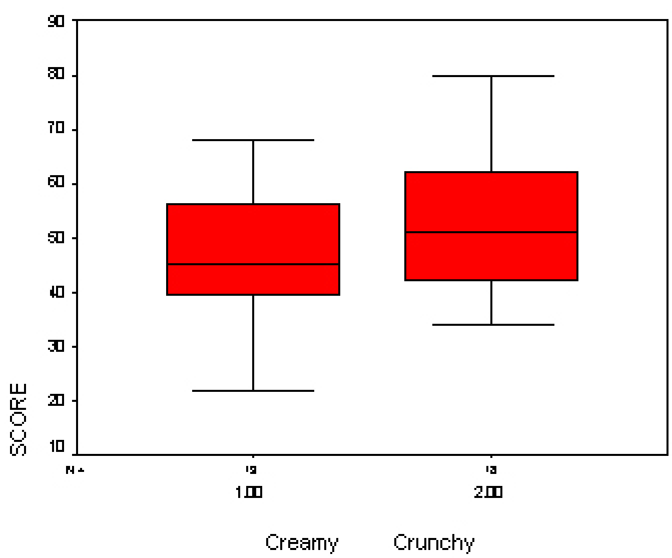


C



D

A
B
C
D



72. Comment on the side by side box plots--which group has a higher median?
Reasoning:

73. The following statement is statistically incorrect. Explain what is wrong. Student ratings of professors' teaching and colleagues' rating of their research have a correlation of ran =+1.21. This shows that professors who are good teachers are also good researchers.

74. A researcher uses a chi-square test to determine if there is a relationship between categorical variables. Which of the following p values indicate the strongest evidence of such a relationship?

A. .02
B. .001
C. .05
D. .10

75. People who eat lots of fruits and vegetables have lower rates of colon cancer than those who eat little of these foods. Fruits and vegetables are rich in antioxidants such as vitamins A, C, and E. Will taking antioxidants help prevent colon cancer? A clinical trial studied this question with 864 people who were at risk of colon cancer. The participants were divided into four groups; 1) daily beta carotene, 2) daily vitamins C and E, 3) all three vitamins every day 4) daily placebo. After four years the researchers were surprised to find not significant differences in colon cancer among the groups. What type of statistical analyses would be used to test for difference between four groups on evidence of colon cancer?

A. MANOVA
B. Independent T Test
C. ANOVA
D. Chi-Square

76. (New questions same example) People who eat lots of fruits and vegetables have lower rates of colon cancer than those who eat little of these foods. Fruits and vegetables are rich in antioxidants such as vitamins A, C, and E. Will taking antioxidants help prevent colon cancer? A clinical trial studied this question with 864 people who were at risk of colon cancer. The participants were divided into four groups; 1) daily beta carotene, 2) daily vitamins C and E, 3) all three vitamins every day 4) daily placebo. After four years the researchers were surprised to find not significant differences in colon cancer among the groups. Researchers concluded that there were no statistically significant differences after analyzing the data. What does no significant difference mean in describing the results of this study?

77. Suggest one possible lurking variable that could explain why people who eat lots of fruit and vegetables have lower rates of colon cancer.

78. Does knowing a college student's SAT score tell us anything about his or her first year college grade point average? What type of statistical analysis should be used in trying to determine the answer to this?

A. Test one mean against a hypothesized constant.
B. Test the difference between two means (independent samples)
C. Test the difference in means between two pairs of dependent samples
D. Test that a correlation coefficient is not equal to 0 correlation analysis.
E. Use a chi-square test of association.

79. Bob and Bill each bought one ticket 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?

A. Bill
B. Bob
C. They have an equal chance of winning.
D. Not enough information to tell.

80. A newspaper article on drunk driving cited data on traffic deaths in Rhode Island. Forty-two percent of all fatalities occurred on Friday, Saturday and Sunday, apparently because of increased drinking on the weekends. Are you surprised that 42% of accidents occur on those days? Why or Why not?

Read the following short article; would you use any of the questions following the article to challenge the results?

81. Severe Weight Loss Can Be Caused By Chewing Gum, Doctors Report Science Daily (Jan. 11, 2008) —

Doctors warn of excess sorbitol intake, a widely used sweetener in "sugar-free" products such as chewing gum and sweets in an article in the British Medical Journal. http://www.sciencedaily.com/ Sorbitol has laxative properties and is poorly absorbed by the small intestine.

Their advice follows the cases of two patients with chronic diarrhea, abdominal pain and severe weight loss. Although extensive investigations were carried out, final diagnosis was only established after detailed analysis of eating habits. On questioning, both patients admitted consuming substantial amounts of sugar-free gum and sweets. The first patient (a 21 year old woman) chewed large amounts of sugar-free gum, accounting for a total daily dose of 18-20g sorbitol (one stick of chewing gum contains about 1.25g sorbitol). The second patient (a 46 year old man) reported chewing 20 sticks of sugar-free gum and eating up to 200g of sweets each day, which together contained around 30g sorbitol.

After both patients started a sorbitol free diet, diarrhea subsided, normal bowel movements resumed and weight gain was achieved. As possible side effects are usually found only within the small print on foods containing sorbitol, consumers may be unaware of its laxative effects and fail to recognize a link with their gastrointestinal problems, write the authors. In conclusion, they say, our cases demonstrate that sorbitol consumption can cause not only chronic diarrhea and functional bowel complaints but also considerable unintended weight loss (about 20% of usual body weight).

Thus, the investigation of unexplained weight loss should include detailed dietary history with regard to foods containing sorbitol.

Worry Questions

1. Where did the data come from? What kind of study was it? Is this kind of study reasonable in this context?

2. Was a sample used? How was it sampled? How many people did actually participate? Is the sample large enough? Did the sample include people/units which are representative of the population? Is the sample biased is some way? Overall, could this sample reasonably lead to valid inferences about the target population?

3. How reliable or accurate were the instruments or measures (tests, questionnaires, interviews) used to generate the reported data?

4. What is the shape of the underlying distribution of raw data? Does it matter how it is shaped?

5. Are the reported statistics appropriate for this kind of data? Was an average to summarize ordinal data; is a mode a reasonable summary? Could outliers cause a summary statistic to misrepresent the true picture?

6. Is a given graph drawn appropriately, or does it distort the data?

7. How was this probabilistic statement derived? Are there enough credible data to justify the estimate of likelihood given?

8. Overall, are the claims made here sensible and supported by the data? Is correlation confused with causation, or a small difference made to loom large?

9. Should additional information or procedures be made available to enable me to evaluate the sensibility of these arguments? Is something missing? Did the writer conveniently forget to specify the base of a reported percent-of-change, or the actual sample size?

10. Are there alternative interpretations for the meaning of the findings or different explanations for what caused them e.g., an intervening or a moderator variable affected the results? Are there additional or different implications that are not mentioned?

Curriculum Vitae: Barbara Wade

Education

Pennsylvania State University, Harrisburg: M.A., Community Psychology and Social Change 2003
Pennsylvania State University, Harrisburg: D. Ed., Adult Education, 2009
Pennsylvania State University, Schuylkill: Bachelor of Science, Psychology; and Bachelor of Science, Criminal Justice

Publications

Wade, B. (2007). An evaluation of adult education programs within correctional institutions. *Adult Basic Education, 1*(1), 27-31.

Goodfellow, M, & Wade, B. (2007). The digital divide and first-year students. *Journal of Student Retention: Research, Theory and Practice, 8*(4), 425-436.

Couch, S., & Wade, B. (2003). I want to barbecue bin Laden: Humor after 9/11. *International Journal of Mass Emergencies and Disasters, 21*(2), 67-86.

Recent Conference Presentations

Wade, B., and Goodfellow, M. (2008). *Barriers to Statistical Literacy among Students Enrolled in Research Methods Classes: Implications for Instruction and Curriculum*. Presented at Annual Meeting of the Pennsylvania Sociological Society, Harrisburg, PA October, 31-November 1.

Couch, S., Wade, B. (2007, August). *Claimsmaking, Scared Ground, and the Demise of the International Freedom Center.* Paper presented at the Society for the Study of Social Problems, 57th Annual Meeting, New York, NY.

Couch, S., Wade, B., Kindler, J. (2007, February). *Victims' Groups Following the 9/11 Terrorist Attacks: Collective Identity, Collective Memory and Collective Witness.* Paper presented at the Eastern Sociological Society, 77th Annual Meeting, Philadelphia, PA.

Currently, Barbara Wade teaches for psychology at Penn State, Schuylkill Campus.