FACTORS RELATED TO UNDERGRADUATE PSYCHOLOGY MAJORS
LEARNING STATISTICS

Dissertation Proposal
Submitted to
Temple University Graduate Board

In Partial Fulfillment
Of the Requirements for the Degree
DOCTOR OF PHILOSOPHY

---

by
Tamarah Faye Smith
May 2013

Examining Committee Members:

Frank Farley, Advisory Chair, Educational Psychology
Joseph DuCette, Educational Psychology
William Fullard, Educational Psychology
Catherine Fiorello, School Psychology
Jerry Stahler, Geography and Urban Studies

ABSTRACT

Factors Related to Undergraduate Psychology Majors Learning Statistics

Tamarah Faye Smith

Doctor of Philosophy: Educational Psychology

Major Advisor: Dr. Frank Farley

The American Psychological Association (APA) has outlined goals for psychology undergraduates. These goals are aimed at several objectives including the need to build skills for interpreting and conducting psychological research (APA, 2007). These skills allow psychologists to conduct research that is covered in the media (Farley et al. 2009) and influences policy and law (Fischer, Stein & Heikkinen, 2009; Steinberg, Cauffman, Woolard, Graham & Banich, 2009a; Steinberg, Cauffman, Woolard, Graham & Banich, 2009b). One of the fundamental courses required for building these skills is statistics, a course that begins at the undergraduate level. Research has suggested that performance after completing statistics courses is weak for many students (Garfield, 2003; Hirsch & O'Donnell, 2001; Konold et al. 1993; Mulhern &Wylie, 2005; Schau & Mattern, 1997). The current study examined factors that may be related to performance on a statistical test. A sample of 231 students enrolled in or having already completed a statistics course for psychology majors completed a statistical skill questionnaire, built by the author, to measure performance with four APA outlined goals. To measure student attitudes the Survey of Attitudes Toward Statistics (SATS-36; Schau, 2003) was completed with adapted questions to measure perceived attitudes of peers and faculty toward statistics. Finally, questions pertaining to classroom techniques and content areas covered were assessed.

Building off of social cognitive theory (SCT; Bandura, 1986) and expectancy-value theory (Eccles & Wigfield, 2002), it was expected that lower attitudes, such as low value and low interest, among the students and those perceived to be held by faculty and peers would be related to lower performance on the statistical test. A series of linear regressions were conducted and revealed no significant relationship between perceived faculty attitudes and performance. Students' own liking and positive affect ratings were positive predictors of performance indicating a gain of 3-4% on the statistical test. However, an interesting negative relationship emerged with respect to students' value of statistics and peer interest scores where performance on the statistical test decreased as value and peer interest increased. This may be demonstrating issues pertaining to the SATS-36 validity when measuring students' value as well as issues with the items created to measure perceived peer interest. The results of a factor analysis on perceived attitude measures for peers and faculty suggest that the need for more items is necessary, particularly for faculty attitudes. Finally, this study provides a first look at the performance of a sample of psychology students with APA goals for quantitative reasoning. Results showed that students performed best at reading basic descriptive statistics (M=74.5%), and worst when choosing statistical tests for a given research hypothesis (M=30%). Performance on questions pertaining to confidence intervals (M=38%) and discriminating between statistical and practical significance (M=39%) was also low.

Future research can address limitations of this study by expanding the sample to include a broader range of psychology undergraduates and including additional items for measuring perceived attitudes. Other methodological approaches, such as experimental design and directly measuring faculty attitudes, should also be considered. Finally, further research and replication are necessary to determine if scores on the statistical test will continue to be low with other samples and varying question formats. These results can then be used to generate conversation about why and how students are, or are not, learning the appropriate quantitative skills.

ACKNOWLEDGEMENTS

The path that has led me to completing this dissertation has been filled with numerous amazing people without whom I would not be where I am today. Firstly, I owe the deepest appreciation to my children, Tiffany and Jonathan. Tiffany, without the sacrifices you have made to share your mother's time with her ever-continuing schoolwork, I would never have reached this level. Beyond this, you have given me more love and hope than I can thank you enough for. Together, these gifts from you have made me this dream a reality of me. From the depths of my heart, thank you. And, dearest Jonathan, you have taught me to be patient and work steady and strong. Both of these have been essential for getting through this final step to earn my degree. Your smile gave me the strength I needed to keep going and has kept *me* smiling during the process. My heart overflows with gratitude and love for all of this. Thank you.

Secondly, to my mother and father, sisters and brother and Pat, I owe you all deep thanks. Our life as a family has been tumultuous, and I have contributed my fair share to that. You always loved me and supported me, and my children, all in an effort to help me along my path, even when it seemed unreasonable. Thank you. Thirdly, to my dear Andrew, these final months of my project have challenged me in ways I could not have imagined. You were there unfailingly every step of the way. Your wisdom has been priceless and your love needed. For all of this, thank you.

Fourthly, there are several individuals in my academic and professional careers that have contributed to my journey in ways that I will forever be grateful. To my advisor, Dr. Frank Farley, and my committee members, who gave me the support and autonomy every doctoral students dreams of, thank you. To Dr. Anthony Tomasco and

Dr. Kathleen McKinley, for always believing in me, and pushing me along, thank you. To Dr. Edna Barenabum, for pointing me in the right direction, thank you. Dr. Randell Rieger, Dr. Stephanie Ives, Julie Nettleton and Kate Ward-Gaus, thank you for creating the rock solid foundation on which my profession now stands. Finally, to all those individuals who contributed to allowing a young woman like myself to be mother, student and professional, thank you.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

The role of psychological research is vast as highlighted by its common presence in the media (Farley et al., 2009) and influence on public policy and law (Fischer, Stein & Heikkinen, 2009; Steinberg, Cauffman, Woolard, Graham & Banich, 2009a; Steinberg, Cauffman, Woolard, Graham & Banich, 2009b). The skills needed to conduct and interpret this type of research include research design and statistics. Courses to teach these skills are embedded in many psychology programs starting at the undergraduate level, and the American Psychological Association has outlined guidelines for ensuring that undergraduate psychology majors are learning the appropriate quantitative skills in these courses (APA, 2007).

Research has examined the performance of students after completing courses that teach these skills, and this research has been largely focused on statistics courses. Data have illustrated that regardless of completing statistics courses, students' level of understanding statistical concepts is weak (Garfield, 2003; Hirsch & O'Donnell, 2001; Konold, et al. 1993; Mulhern &Wylie, 2006; Schau & Mattern, 1997). The investigation of factors that may be related to poor or strong performance in statistics has been well established. Typically, these studies examine students' attitudes toward statistics or particular pedagogical approaches and how these attitudes and approaches impact learning (see Chapter 2). When narrowing in on results from samples including psychology majors, research has shown that these students tend to have low value for statistics and rate it high in difficulty (Carmona, 2005; Coetzee, 2010; Finney & Schraw, 2003; Schau, 2003). It is suspected that this partially explains the corresponding low levels of achievement in statistics among these students with low value and high difficulty ratings of the subject. However, the attitudes these students perceive others to have (i.e., their peers and faculty), have not been investigated even though multiple theories, such as social cognitive theory and expectancy-value theory, might suggest these perceptions would influence levels of learning in similar ways that their own attitudes would impact learning. In addition to not examining the roles of others' attitudes, there appears to be no research that systematically investigates if evidence based approaches are being used to teach statistics to psychology majors. Rather, studies typically focus on the ability of one technique to increase performance. While this is important information, it does not answer whether these empirically supported techniques are being used. Failure to use them in exchange for using an ineffective technique may explain poor performance.

To build on this body of literature, the current study examined not only the relationship of a student's own attitudes on statistical skills but also included the role of their perceptions of peers' and faculty's attitudes on statistical skills. In addition, whether they are being taught using techniques that have been shown to be effective was analyzed for its potential relationship to statistical skills. The role of each of these individually and together was considered for the first time as previous research has not appeared to include these variables. To accomplish this, social cognitive theory (SCT; Bandura, 1986) and expectancy-value theory

(Eccles & Wigfield, 2002) were used to define a theoretical model of how both environment (i.e., peers, faculty and classroom experiences) and individual student attitudes influence statistical skills. SCT would suggest that a poor learning environment and negative attitudes towards statistics would both individually and interactively lead to poor statistical skills. In turn, these poor skills would reinforce poor attitudes and justify limited effort on the part of faculty's instruction and students' course participation. Expectancy-value theory provides detailed components within each of these levels (learning environment, attitudes and skills) that help provide measurable outcomes to test this model.

As mentioned, research that includes these multiple factors has been scarce. However, investigations have been conducted that provide some insight into the outcome of the current study. For example, research has provided evidence of poor quantitative reasoning (Garfield, 2003; Hirsch & O'Donnell, 2001; Konold, Pollstsek, Well, Lohmeier & Lipson, 1993; Mulhern &Wylie, 2006; Schau & Mattern, 1997) and low value of statistics among undergraduates (Carmona, 2005; Coetzee, 2010; Finney, 2003; Mills, 2004; Schau, 2003; Tempelarr, Van deer Loeff & Gijselaers, 2007). As will be reviewed below, other literature has provided guidelines for best approaches when teaching statistics to undergraduates, but it is not known if these methods are being used and how they may be related to student attitudes and achievement.

Goals and Objectives

The current study examined student, peer and faculty attitudes, teaching techniques used in statistics class, as well as statistical skills to achieve the following goals:

- Highlight potential weaknesses in the statistical skills of students with particular focus on the extent to which the APA guidelines for quantitative literacy in the undergraduate major are met.
- Examine both student attitudes and their perceptions of peer and faculty attitudes toward statistics.
- Explore the role of various factors in the learning environment on poor statistical skills; specifically, the use of empirically supported teaching techniques and covering APA recommended topics.
- The relationship of these environmental variables, attitudes and skills were studied to attempt to explain the potential weaknesses in psychology students' statistical skills.

CHAPTER 2

REVIEW OF LITERATURE

Background

Statistical skills may be often viewed as an advanced outcome of academic study; however, such skills are essential for many people. In fact, some have argued they are necessary for ample participation in a democratic society (Cobb & Moore, 1997). These skills can be applied in many areas including activities to determine policy, interpreting research to guide practice, and in interpreting medical, commercial and education related information. An inability to reason with statistics may lead to a decreased ability to contribute in all of these areas. Practice in psychology is certainly not exempt from this need for statistical reasoning. Indeed, two individuals who developed several statistical techniques still used today, Karl Pearson and Charles Spearman, did so in an effort to study psychological phenomenon with greater accuracy. This importance has persisted throughout the years. A study reported in the American Psychologist in 1958 presents an interesting train of thought at that time. Here, Holder, Leavitt and McKenna surveyed chairpersons of psychology departments to explore what courses would be essential for the psychology major. An overwhelming percentage of the chairpersons (96%) selected statistics as an ideal course for students in a psychology major. While there was variability in terms of what course might be ideal for *all* psychology majors regardless of specialization, only two courses were unanimously selected: statistics and experimental psychology. It may not, then, be surprising that today the American Psychological Association has provided guidelines for the undergraduate major that include statistical reasoning (2007). Regardless of the history and guidelines in the discipline, statistical reasoning continues to be limited among psychology students (Mulhern & Wylie, 2005). There are many factors that could be examined as potential reasons for this disconnect between psychology's guidelines and the quantitative reasoning skills of psychology students. Social cognitive theory and expectancy-value theory can be used to understand what these issues are and how they are at work.

Social Cognitive Theory

Social cognitive theory (SCT; Bandura, 1986) suggests that there is an interactive model of reciprocal causation that facilitates learning between the environment and personal characteristics including cognition and behavior. This model is different from social learning theory as it extends to include the involvement of thought, as reflected in the individual not the social environment, in influencing behavior. In addition, the reciprocal nature of three factors is considered rather than a unidirectional relationship of the environment on individual behavior.

The influence of thought is a complicated issue with many potential directions in which to influence the environment and behavior. Bandura provides an important description of the potential positive and negative outcome of thought on both these factors.

> To say that people base many of their actions on thought does not necessarily mean they are always objectively rational. Rationality depends on reasoning skills which are not always well developed or used effectively. Even if people know how to reason logically, they make faulty judgments when they base their inferences on inadequate information or fail to consider the full consequences of different choices. Moreover, they often missample and misread events in ways that give rise to erroneous conceptions about themselves and the world around them. When they act on their misconceptions, which appear subjectively rational given their errant basis, such persons are viewed by others as behaving in an unreasoning, if not downright foolish, manner. Thought can thus be a source of human failing and distress as well as human accomplishment. (Bandura, 1986, pp. 19)

Here we see thought allowing an individual to evaluate the environment (misreading events) and behavior (anticipating outcomes). Such evaluations then allow us to determine what behavior we deem appropriate for ourselves. The resulting action can then shape the reaction of others, a piece of our environment, and our interpretation (via thought) of this reaction can then influence our future behavior.

This final piece highlights the importance of thought in the role of self-reflection. Self-reflection can also be faulty and create erroneous beliefs. The actions that result from these beliefs may often confirm the inaccurate belief. Our beliefs can help shape our level of efficacy which then influences our later behavior. The role of thought, in this case in the form of self-reflection, is an important addition to the SCT.

The reciprocal nature of the environment, cognition and behavior is also an important contribution of SCT. The model is neither of one-sided determinism nor one-sided interactionism. In the view of one-sided determinism, the person and situation are separate and involved in a unidirectional relationship such that the environment influences the person (i.e., behaviorism) or the person influences the environment (i.e., personal determinism). SCT purposes that the person influences the environment, which is then acting back upon the person. This is in contrast to the "behaviorless" person described in one-sided interactionism. Bandura points out that "…behavior is an interacting determinant, not a detached byproduct that plays no role in the production process" (Bandura, 1986, pp. 23). It is clear, then, that SCT is suggesting three sets of interacting factors: 1) the relationship between cognition and behavior; 2) the relationship between cognition and the environment; and, 3) the relationship between the environment and behavior.

The reciprocal nature of the environment, cognition and behavior described by SCT provides a useful model for understanding the attitude of students and faculty in undergraduate psychology departments and the statistical skills of these individuals. In a fairly simple scenario, the environment in which students study statistics can impact their attitude of statistics. These attitudes may then impact their behavior. The behavior and attitude then reinforces the environmental factors. For example, when students experience an environment where optimal learning is not an option whether it be a result of faculty who do not have the appropriate expertise to instruct the course, use of pedagogical tools that are not effective or observing an attitude from faculty and other students that insinuates statistics are not relevant outside of the course and are very difficult or impossible to learn, students' self-reflection on this environment may lead to their own negative attitudes. Students may begin to believe that learning statistics is too difficult for them to achieve. In addition, the limited resources put toward teaching the course and attitudes of faculty and students may impress upon the student that success, even if attainable, is not highly valued. Together, these may lead to decreased effort, class performance and ultimately poor statistical skills on the part of the student. The reciprocal nature of this model would then suggest that these attitudes reinforce the limited effort of faculty when teaching the course and the negative attitudes of other students. It is interesting to note the interaction between student and faculty in this example. Not only does the faculty's teaching impact the students' learning, but the students' lack of success in the course, a function of environmental and cognitive factors, also potentially influences faculty's teaching of the course (the environment).

While SCT provides a framework for understanding this issue, the above example omits important pieces. For example, the role of students' and faculty's expectancies is also important. Here, Bandura discusses two types of expectancy: outcome expectation and efficacy expectation. A person's understanding of how to achieve a certain task describes his or her outcome expectation. That is, they understand what measures are necessary to obtain the desired outcome. In contrast, efficacy expectation defines one's belief that they are capable of these measures so that they may be successful. These cognitive expectancies are essential in understanding how psychology students learn statistics. First, knowledge of what is necessary to achieve in a statistics course may be difficult for students who spend much time in courses quite different. For example, psychology students may be well versed at reading qualitative descriptions of human behavior, writing down key facts from lectures and synthesizing this information for test taking and paper writing; however, when entering a statistics course, they may not have the knowledge of how best to read through quantitative texts, what notes are important to make during a lecture and even how to proceed with assignments and projects that involve data. Therefore, their outcome expectancy may be challenged by not knowing what steps are necessary for success. It follows that their efficacy expectation will then suffer. Certainly, it would be difficult for students to expect themself to be successful in a course when they are not clear what steps need to be taken to reach that success. In addition, the environment in which the learning is taking place may also impact the students' expectations of the success. For example, when the attitude of others reflects that it is difficult to achieve in the course, this may influence the extent to which other students feel they are able to achieve in the course.

An additional important consideration in the theoretical model is the role of one's value of a given task. While expectancies define whether a person feels they are capable ("Can I?"), value defines whether a person desires to do a task ("Why should I?"). In the event that students' outcome and efficacy expectations are high for a given task, they may still have no vested reason to attempt the task. That is, their value for completing the task may be low. Expectancy-value theory adds the idea of value to the self-efficacy model of expectancies created by Bandura. In this model, value is often defined in terms of interest. Interest may be intrinsic by reflecting an individual's desire to complete a task or it may be extrinsic reflecting situational motivation to complete a task. Intrinsic interest is theorized to derive from a need to maintain a certain level of stimulation that may lead people to choose particular tasks. In contrast, extrinsic interest is often the result of an environment being set up in such a way as to motivate one to act.

Expectancy-Value Theory

When combining the expectancy and value components together, Eccles and Wigfield (2002) define a wide range of factors that influence both expectations and values. Among these are behavioral related factors (performance, persistence and task choice), environmental factors (behaviors and beliefs of socializer, culture, and history), and several cognitive factors. These cognitive components can be divided into task specific beliefs and individual perceptions. Task specific beliefs include competence, difficulty and goals/self-schema while individual perceptions include others' attitudes and expectations, affect memories and interpretations of prior achievement. Figure 2.1 illustrates how these factors, outlined by expectancy-value theory, affect each other and might be categorized into the three main components of SCT: behavior, cognition and environment.

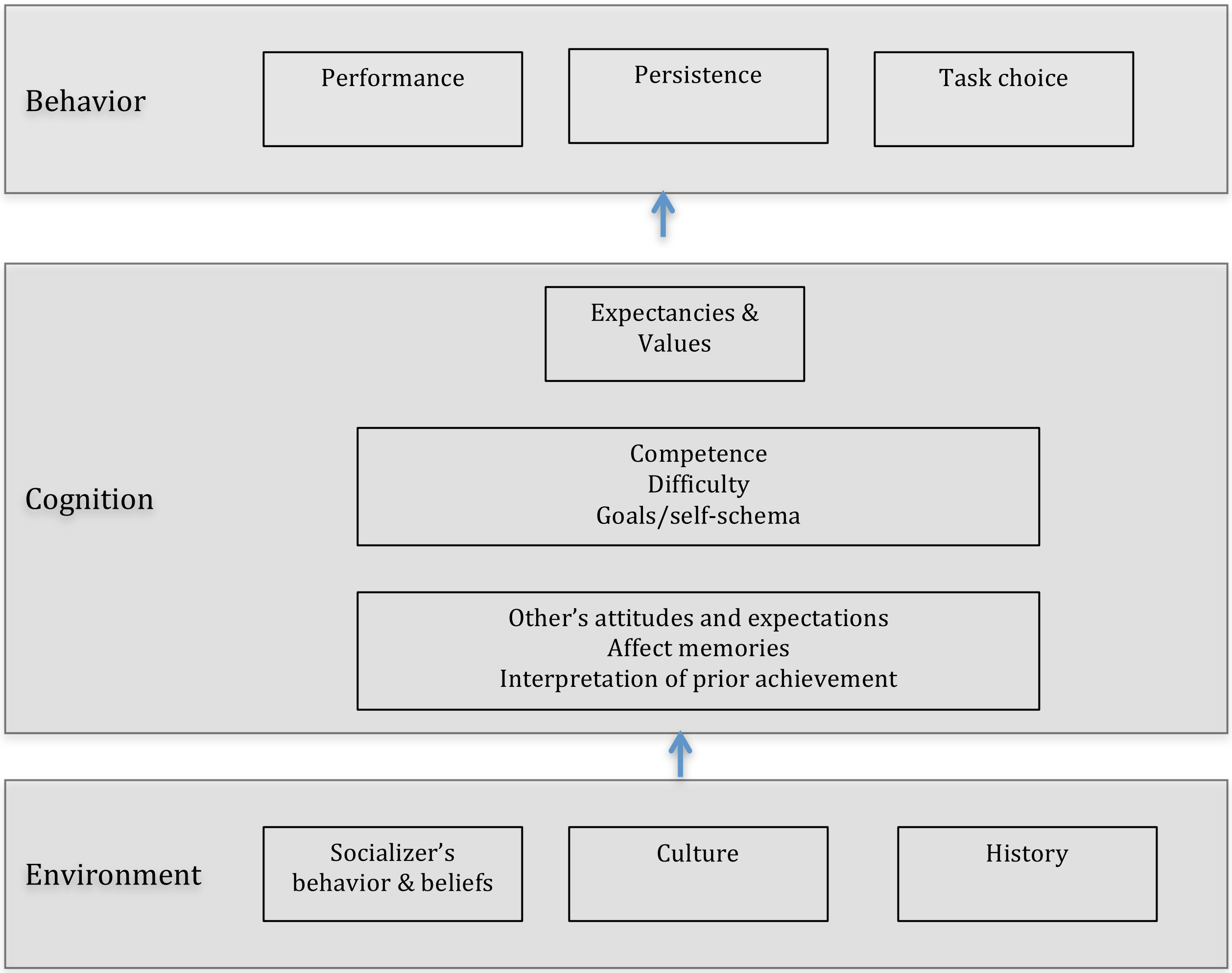

Figure 2.1. Combined theories. Expectancy-Value theory overlapped with Social Cognitive Theory.

While this suggests overlap between the two theories, they remain distinct in several ways. One of these distinctions which is important here is the direction of the relationship(s). As stated earlier, SCT posits a reciprocal relationship between behavior, cognition and environment. In contrast, expectancy-value theory outlines a relationship that is more focused on a unidirectional relationship. Specifically, environmental factors (behaviors and beliefs of socializer, culture, and history) influence individual perceptions (others' attitudes and expectations, affect memories and interpretations of prior achievement), which influence task beliefs (competence, difficulty and goals/self-schema) and ultimately define the expectancies and values that predict behaviors (performance, persistence and task choice). Nonetheless, it has become clear that both theoretical models, alone and together, provide a framework for understanding the current state of quantitative knowledge in psychology.

Taking these two theories together, the above example can now be expanded to include the previously missing important considerations. The behaviors and beliefs in the culture of undergraduate psychology studies may be such that little adequate resources are provided for students to learn such as poor instruction or ineffectual faculty. The attitude of faculty and other students may be perceived by the learner in such a way that it creates a general negative experience. This negative affect, when recalled by the learner, may then influence the learner's beliefs about a given task. Specifically, a negative culture and affect may lead a student to have limited goals with respect to achieving (expectancy) and their confidence in being able to achieve may be hindered by the limited resources for learning. Students' interest or value in learning may then suffer as a result. Together, this would predict poor performance, persistence and/or task choice.

This example includes several assumptions. First, it is assumed that a poor environment exists within undergraduate psychology departments specifically with respect to the expertise and attitudes of faculty and resources given for courses. Secondly, it assumes poor attitudes exist among students, which leads to poor performance. Finally, this example has assumed that statistical skills are not adequate.

Learning Environment

There is support in the literature for some of these assumptions. Several factors help define the environment in which undergraduate psychology majors learn including teacher expertise, pedagogy used and attitudes of faculty. A number of studies exist that have examined the ways in which undergraduate psychology majors best learn statistics. Approaches have included using computers, calculators and spreadsheets or using context such as life examples to illustrate statistical concepts. Results of using these approaches have been mixed. In some instances, no significant improvements in students' performance on statistical measures have been observed when using calculators (Collins & Mittag, 2005) or when using computers (Morris, 2001) to better understand statistical concepts. Some qualitative evidence has suggested computer simulations may help to reduce misconceptions about statistical concepts such as correlations (Morris, 2001). Other research has found possible confounded results (Morris, Joiner & Scanlon, 2002). Here, significant increases in understanding central tendency and borderline increases in understanding correlation were observed for an experimental group. However, a control group also had borderline increases in understanding central tendency suggesting that confounding factors or poor reliability and validity of measures may be at play.

Positive effects of using computers resulted when adding additional computer lab time for an experimental group of students (n=22) and compared achievement to a control group (n=22) that had no extra class time added (Smith, 2003). The experimental group had a significantly higher percentage of correct responses compared to the control group on four open-ended questions. Basturk (2005) examined the benefit of using computer assisted instruction (CAI) in an introductory statistics course with graduate psychology students by adding a one hour weekly computer lab using SPSS to demonstrate and practice content covered in the lecture. Midterm and final exam grades were compared between an experimental and control group and showed that students attending the CAI lab had significantly higher grades on both exams. In addition, this group's exam scores increased on average from the midterm to final while the lecture only group's exam scores decreased from the midterm to the final. These findings suggest that the use of lab time and CAI, specifically using SPSS to practice techniques covered in class, can help increase students' performance when combined with lecture. The results here may conflict with the insignificant results of the previous studies which used computers only to illustrate concepts because of these later studies including broader based use of computers for practicing work and possibly due to the nature of this practice adding to class time.

When examining the role of context to illustrate statistical techniques and concepts, Schoenfelder, Olson, Bell and Tom (2007) found students' knowledge test scores increased significantly after a data collection activity counting roses on bushes and calculating and graphing sample means and confidence intervals. Wiberg (2009) reviewed the outcomes from a student-centered curriculum compared to a more traditional instructionalist curriculum when statistics were incorporated into an existing psychology course. Results revealed significantly higher exam scores and more positive attitudes toward statistics for the students receiving student centered curriculum.

Lawson, Schwiers, Doellman, Grady and Kelnhofer (2003) used written examples of real-life activities and measured whether the activity increased psychology students' ability to make this type of application on their own.   As expected, the authors found significantly higher gains in correct test responses than control groups on the statistical reasoning test. In another study, Christopher and Marek (2009) mixed cookie dough in class to demonstrate the concepts of main effects and interactions. Students across two semesters (n=50) had significant increases on post-test knowledge scores after the cookie demonstration was presented in their class; however, there was no control group to help differentiation the source of this increase. Overall, these studies suggest that using content and application may be helpful in students learning statistics.

These studies on effective ways of learning statistics are limited as they often employed small samples, have author derived tests, no direct replications and use inferior statistical tests when analyzing data. In addition, few studies have explored the frequency with which these approaches are being used, what specific topics are being covered with these approaches and they have not elaborated on the expertise of the faculty who teach them. This is critical information for understanding the environment in which students are learning. A good pedagogy may be ineffective if the faculty overseeing the course is not familiar enough with the material to begin with. In addition, certain subject areas have been identified as most important for psychology undergraduates to study (APA, 2006). Failure to cover these areas may leave students unprepared for future work in the field.

One study has examined a limited number of these additional environmental factors. Friedrich, Buday and Kerr (2000) conducted a national survey of undergraduate psychology departments using similar methodology to that which was previously used in surveying graduate programs (see Aiken, West & Millsap, 2008). Results indicated that 93% of departments require a statistics course as part of the psychology curriculum.

However, who teaches these courses and the methods being used in teaching the courses are not known with the exception of using computerized statistical packages (87% of sample). Whether these packages are being used to help demonstrate and learn concepts supported by research or simply to analyze data is not known. Most concerning is the fact that the study found that critical areas of content outlined by the APA are lacking in the statistics classroom for undergraduate psychology majors. One hour or less of a semester was reported as being given to address each of the following APA guideline topics: confidence interval estimation, power analysis, effect size estimation, graphical analysis of data and APA style for reporting statistics.

Together, this literature provides some evidence that the environment students are learning statistics in is not adequate. According to Friedrich et al. (2000), APA guidelines for content are not being taught in the classroom. Covering these topics may allow faculty to draw connections between the course content and its relevance in the field, a connection that may have been influential in the increased performance of students in the previously discussed content and activity based studies. While there is some literature to suggest that these types of classroom techniques are best, results are mixed suggesting more investigation is needed. In addition, there is a lack of review to determine whether those methods, which are shown to be effective, are being used. With regard to faculty, it is not known who is instructing the courses with respect to skill level and expertise with statistics. Given previous research that examines the training of doctoral students in psychology, it should not be assumed that all faculty have adequate statistical training (Aiken et al., 2008).

An additional factor that may be critical in the learning environment is that of the attitudes of those around the learner including their peers and faculty. There appears to be no research addressing the rate at which students' sharing their negative attitudes about statistics occurs and what effect it may have. In addition, the role of faculty's attitudes has also been understudied. It is interesting that the attitudes toward statistics of *students* tends to be the primary focus when examining the relationship between attitudes and achievement, particularly considering those conducting this research may be the same individuals teaching the course. Faculty attitudes and effort warrant investigation, as they are a critical piece of the learning process. Certainly, the relationship between teacher attitude and student learning has been established in K-12 settings. One related study evaluated faculty's role in this issue and found that, in addition to students' attitudes, teachers' labor (i.e., effort put into preparing and teaching a course) was a significant predictor of student achievement (Sabiote, Pérez & Llorente, 2010). In another study examining students and teachers from 1,000 public and private schools, early education teachers' attitudes were related to children's reading achievement. More positive attitudes indicated a positive increase in reading scores while scores decreased with a more negative teacher attitude (Palardy & Rumberger, 2008). This is similar to results that were previously uncovered by Goddard, Hoy and Hoy (2000).

Marlin (1991) showed an impressive relationship of teacher attitude on student achievement. Using a sample of high school juniors and seniors (*n*=602) and their economics teachers (*n*=50), teacher attitude was the second highest predictor of students' achievement on the standardized Test of Economic Literacy (TEL). In fact, a one-point increase in teachers' attitude was equal to a predicted 8% increase in students TEL score.

Additional research has shown that collective responsibility, that is the level of responsibility a teacher takes for student learning, was significantly predictive of achievement gains in math and science for a large sample of high school students (Lee, Smith & Croninger, 1997).

The role of teacher attitude may not only be related to student performance but also have a strong relationship with students' attitudes. Limited evidence has shown a strong correlation between teacher support, including student perceived teacher interest and university science students' enjoyment in the course (Martin-Dunlap & Fraser, 2007).

Students' Attitudes Toward Statistics

While the attitudes of faculty and peers may impact the environment a student is in, the attitudes of an individual student are also important as they represent the student's cognition. Research has revealed a consistent pattern of students' attitudes towards statistics prior to completing a statistics course. Two primary measures have been used to determine these attitudes. When using the Survey of Attitudes Toward Statistics (Schau, Stevens, Dauphinee, & Del Vecchio, 1995), students score their attitude on four subscales: affect, cognitive competence, difficulty and value. Overall, ratings tend to be neutral in terms of affect toward the subject and their cognitive competence for learning statistics. Typically, students rate the statistics as somewhat difficult. The value students place on the subject varies slightly. Business, mathematics and economics majors tend to rate statistics as slightly higher in value compared to social science students. These results have been established with samples of American university students of mixed majors (Schau, 2003), business majors (Mills, 2004) and psychology majors (Finney, 2003). Samples of social science students from Spain (Carmona, Martinez, & Sanchez, 2005), psychology students from South African (Coetzee & Van der Merwe, 2010) and Dutch business and economic majors (Tempelarr, 2007) have also revealed similar patterns of attitudes among students.

When using an alternative measure, the Attitudes Toward Statistics questionnaire (ATS; Wise, 1985), the subscales course and field are measured. Course ratings report students' attitudes toward their statistics course. Field ratings report students' attitudes toward statistics as they are applicable to the student's given field of study. Results of studies using the ATS have shown the majority of students have fairly neutral ratings of statistics in their course and high ratings of statistics in their field. Again, a variety of samples have been used including undergraduate students of mixed majors (Waters, Martelli, Zakrajsek, & Popovich, 1988; see also Vanhoof et al., 2006), business, accounting and economic majors (Mvududu, 2003) and psychology majors (Shultz & Koshino, 1998). These results are consistent with both American samples (Elmore et al., 1993; Mvududu, 2003; Shultz et al., 1998; Waters et al., 1988) and international samples (Mvududu, 2003; Mji, 2009).

Additional research that has examined the attitudes of *psychology* students in depth indicates mixed attitudes, at best. Finney and Schraw (2003) sampled students taking an introductory statistics course through an Educational Psychology department (n=103) and found somewhat positive ratings of cognitive competence, high ratings of difficulty but neutral ratings in terms of value. This indicated that students felt somewhat confident in their ability to comprehend statistics, a course they rated as very difficult; however, it is discouraging that they rated statistics as not having high value for their future in psychology as reflected by their low value ratings. These findings were duplicated with a sample of industrial psychology majors (*n*=235) in South Africa (Coetzee & Van der Merve, 2010). Here, the slightly older undergraduate students (mean age, *M*=31) resulted in reasonable estimates of competence, high levels of difficulty, neutral ratings in terms of value with high levels of interest and planned effort for a statistics course.

In an alternative qualitative study of students' attitudes, Ruggeri, Dempster, Hanna and Cleary (2008) conducted focus groups with a sample of British psychology majors (*n*=196). Results revealed similar attitudes in terms of valuing statistics; that is, participants had little value for statistics. Comments illustrated that students did not identify statistics as a relevant part of psychology, felt it was not important to be a part of psychology as it had decreased in its value over time and reported they would "never" use statistical skills in practice.

Together, these studies illustrate that while some students may have favorable attitudes towards statistics, psychology majors appear to not have these same positive attitudes. Specifically, they see statistics as very difficult and having little value. Understanding these attitudes is important because these cognitions help to create the beliefs and schemas that guide decisions on behavior such as course engagement which in turn impacts learning and the level statistical skills one may obtain.

Statistical Skills

Literature on statistical skills highlights that students have problematic views of statistical concepts. There are high rates of misconceptions. This may be a function of learning experiences. For example, effort-based learning approaches (e.g., repetitious practice of calculations) in statistics classes tend to lead to lower skills and higher adherence to misconceptions (Tempelaar, Gijselaeaers & Schim van der Loeff, 2006). This would be consistent with research that shows students develop strong computational skills in statistics likely as a result of the high levels of repetitious problem solving found in effort-based learning courses (Pollatsek, Lima & Well, 1981). These findings highlight not only that students have high rates of misconceptions accompanying their computation skills, but also deeper knowledge appears to be absent for many students.

When focusing on psychology undergraduate majors specifically, similar results have been found. Using a broad sample of incoming psychology undergraduate students in the United Kingdom (*n*=890), Mulhern and Wylie (2006) found quantitative skills were poor as illustrated by the average score on a reasoning test (43%). It is more concerning that among the different areas of understanding assessed, probability and sampling received some of the poorest scores. While these skills were measured prior to taking a statistics course, the results of other studies suggest that even after formal training, statistical skills are poor (Garfield, 2003; Hirsch & O'Donnell, 2001; Konold, Pollstsek, Well, Lohmeier & Lipson, 1993; Schau & Mattern, 1997). Qualitative studies have echoed this finding when examining students' knowledge of statistical concepts such as samples, means, standard deviations and the Central Limit Theorem (Groth & Bergner, 2005; Mathews & Clark, 2003).

While these reasoning outcomes are important, they are not inclusive of the APA outlined specific goals for psychology undergraduate majors to achieve in terms of quantitative skills. These goals include the ability to distinguish between statistical significance and practical significance, describing effect size and confidence intervals and interpreting conclusions in research reports are listed as important skills. Studies have failed to address these skills in psychology undergraduate students. This leaves the field with an unclear understanding of what the statistical skills of undergraduates are when leaving college. While research has shown that broader conceptual knowledge is problematic, it is not known if the specific skills outlined as necessary by the APA have been obtained.

Hypothesis

As noted above, psychology students tend to rate statistics as a difficult course, have little value for the subject and they are somewhat neutral to high when rating their cognitive competence. Research has illustrated that these attitudes are related to learning in statistics (Tempelarr et al., 2007). In addition, the studies summarized above have shown that learning is also a function of the environment, specifically, content and application focused pedagogy and faculty attitudes. As such, it may be that negative attitudes together with problematic learning experiences might predict poor statistical skills in statistics. The purpose of this study is to investigate how these factors relate to undergraduate psychology students' statistical skills. Specific hypotheses include:

1) Similar to previous findings (Coetzee & Van der Merve, 2010; Finney & Schraw, 2003; Ruggeri et al., 2008), it is expected that attitudes of psychology undergraduates will illustrate high ratings of difficulty, neutral to high ratings of competence and low levels of value for statistics. While there is no literature that has examined the perceptions of peers and psychology faculty, it may be reasonable to expect attitudes will be similar among peers and the faculty teaching the course given their belonging in the same field of study.

2) Friedrich et al. (2000) reported that psychology faculty have limited experience with statistics, feel somewhat unprepared for teaching the course, spent less time preparing for statistics course compared to other courses, use a limited number of empirically supported classroom techniques when teaching and spent a limited amount of time on APA outlined goals. It is expected that students' perceptions of faculty will reflect these trends. Specifically, students will rate faculty as having limited competence with the subject, putting low effort into teaching and report few empirically supported teaching techniques and APA content areas being present in statistics class.
3) The statistical skills of students are expected to be poor when measuring understanding of the APA outlined areas of importance. Given the low scores obtained by Lawson et al. (2003) when measuring statistical skills (30% or less), it is expected that the average scores in this study will be similar.
4) Finally, based on SCT and expectancy-value theory outlined above, measures of student attitude and environmental factors (teaching techniques and faculty and peer attitudes) will significantly predict statistical skills among students. The direction of the relationship is expected to be such that lower attitude scores (reflecting more negative attitudes) and lower environmental scores (reflecting poor environment) will predict lower statistical skill scores.

CHAPTER 3

METHODS

Participants

Undergraduate psychology majors were sampled from a small liberal arts college and large private university in southeastern Pennsylvania. The total sample size was *N*=231 which consisted primarily of psychology majors. The students were surveyed during their statistics course that satisfied the requirements of their psychology degree, or, for students who had already completed the statistics course, during other class time.

Measures

Statistical Skills

The main dependent variable in this study was student performance on a measure of statistical skills. To measure statistical skills, a series of questions were created that reflect the five APA outlined goals for quantitative literacy in the undergraduate major (see Appendix A). While these goals are the product of a selected committee and subject to debate and compromise among that committee, they are not without limitation. However, building off previous research that has used these goals (Friedrich et al., 2003) they were chosen as a point of reference for building a measure of statistical skills. The specific goals are provided in the APA Guidelines for the Undergraduate Major in psychology (APA, 2007) and include: the ability to interpret basic statistical results, distinguish between statistical significance and practical significance, describe effect size and confidence intervals, interpret conclusions in research reports and choose appropriate statistical analyses to be used for evaluating a given hypothesis. To assess the ability of students to interpret basic results, the questionnaire begins by asking students to interpret basic statistical results from a table of central tendency and variability as well as determining which z-scores in a given list are statistically significant. To measure students' ability to distinguish statistical and practical significance, a scenario is given that provides the results of a t-test including the t-value, p-value and effect size. Students are asked to interpret whether the results of the t-test are statistically significant and whether the effect is small, medium or large. Knowledge of confidence intervals is assessed using two true/false questions adapted from the Assessment Item Bank for Statistics Tools Project (Garfield, delMas & Chance, n.d.). Finally, to assess the ability of students to choose appropriate tests for a given research hypothesis, a series of six research questions are presented and students are asked to select which test from a list of statistical analyses is most appropriate given the research question.

Attitudes Toward Statistics

A variety of measures exist to determine students' attitudes towards statistics. Perhaps the most commonly used measure is the Survey of Attitudes Toward Statistics (SATS; Schau, Stevens, Dauphinnee & Del Vecchio, 1995). The SATS-36 was used to measure the attitudes of students in this sample. The SATS-36 includes 36 positively and negatively worded statements that are rated on a 7-point Likert scale anchored at each end (strongly disagree/agree) and in the middle (neither disagree nor agree). Sixteen questions on knowledge, mathematics experience and demographics follow. Again, Likert scales are used where appropriate and use respective scales (i.e. very poorly/very well; not at all/great deal; not at all likely/very likely).

The SATS-36 includes six subscales: 1) *Affect* evaluates the student's general attitude toward statistics; 2) *Cognitive competence* evaluates students' perceptions of their ability to achieve in a statistics course; 3) *Value* evaluates the attitude of how valued statistics are in the world; 4) *Difficulty* evaluates the student's attitude regarding how difficult statistics is in practice; 5) *Interest* evaluates student's interest in the statistics; and, 6) *Effort* evaluates student's effort put towards learning statistics. Several studies have demonstrated high reliability and validity of both the original version of the survey, the SATS-28 (Dauphne, Schau & Stevens, 1997; Hilton, Schau & Olsen, 2004; Schau, et al., 1995; Schau, 2003) and one study has shown similar results with the expanded version, the SATS-36 (Tempelaar et al., 2007).

Learning Environment

Variables that can contribute to the learning environment are broad in scope and in this study include: perceptions students hold of their faculty's and peers' attitudes toward statistics, the instructional approaches used in statistics class and content areas that are covered in the class. Using the SATS-36 as a guide, author derived questions were created to assess students' perceptions of peer attitudes. Items were measured on a 7-point Likert scale anchored at each end (strongly disagree/agree) and in the middle (neither disagree nor agree). Questions included items that assess perceived value ("My peers value statistics") perceived effort ("My peers complete assignments on time") perceived interest ("My peers are interested in statistics") and perceived competence ("My peers are capable of achieving in statistics class").

While the SATS was developed to measure student attitudes, faculty attitudes are also an important piece of the environment being measured here. Again, the SATS served as a guide to create items that measured students' perceptions of faculty attitudes. Items were similar to those created for students with appropriate adjustments made where necessary. As before, questions included items that assess perceived value ("My faculty value statistics") perceived effort ("My faculty prepares well for each class") perceived interest ("My faculty are interested in statistics") and perceived competence ("My faculty are able to communicate statistical information well").

The SATS-36 also includes questions relating to demographics and on effort and experience with past mathematics and statistics courses. Two additional lists were included for students who have taken statistics in college to assess what, if any, empirically supported classroom approaches the students remember being used and what subject areas were addressed from a list of areas highlighted by the APA as goals for the undergraduate major.

Procedure

Permission was obtained from the Institutional Review Board at each institution as well as the authors' institution, Temple University. All participation was voluntary and no incentive was given. For the main portion of the study, all student data were collected during class time after receiving permission from faculty. Due to the small size of the psychology department at the liberal arts school, all faculty members in that department were approached and asked to allow the researcher to survey students during class. This allowed access to all students in the department who had taken statistics or were currently enrolled, spread across a total of 11 classes. At the larger University, faculty teaching the statistics course and higher level classes that require statistics as a pre-requisite were approached for permission to allow the research to survey students during classes. This resulted in nine classes being surveyed. While it is unknown how many different faculty members taught the statistics courses to students who had already completed the course, five separate faculty members taught the statistics courses currently in progress.

During the first or last 15 minutes of each class, students were invited to complete the SATS-36 with added questions pertaining to peer attitudes, faculty attitudes, instructional methods used and areas covered in statistics class. In addition, the statistical skills questionnaire was given. All materials were counterbalanced to account for potential survey fatigue. All participants were required to complete a consent form. Consent forms were kept separate from the questionnaires to ensure that no name was attached to the data.

Prior to collecting data for the main portion of the study, a pilot study was conducted with a small (N=9) group of students who were completing a research methods course with the author. All student surveys were given during the first week of class in the Fall 2012 semester. After completing the surveys, students were asked for feedback on the surveys including but not limited to clarity of directions and questions and ease of completing. Results from the pilot study were used to fine-tune the surveys as deemed necessary and are summarized below.

Data Analysis

Pilot Study

Survey results from the pilot study were analyzed primarily in a descriptive manner. Central tendency and variability for each survey item, composite scores and subscale scores were computed to explore the possibility of ceiling or floor effects. Questions were edited to best address potential issues presented in both the data and students' verbal feedback regarding the instruments used. Of the nine students, seven had completed a statistics course and two were currently enrolled in a statistics course. Students were given a packet to complete during the first fifteen minutes of class time that included a consent form, SATS-36 with an additional 21 items to measure perceived peer and faculty attitudes toward statistics, a list of classroom techniques and content areas from which students were to select those that were used in their statistics course and the statistical skills questionnaire. The amount of time it took to complete the surveys was 15-20 minutes with half of the class completing it in 15 minutes and half completing it between 15 and 20 minutes. Four students were observed behaving in ways that could indicate distress or fatigue during the process. For example, one student was sighing deeply throughout the process while another was stretching with her eyes closed for a short period of time. After collecting all the surveys, students were asked for their feedback regarding the length of the procedure. The students unanimously agreed that the length was not excessive. When probed for more information by providing the signs observed by the researcher, students agreed that this was a result of having just returned to classes from the weekend.

Additional feedback was requested for clarity of the directions and questions by reviewing the surveys using an overhead projector. One section/survey was addressed at a time. Students were comfortable with all the directions and questions with one exception. Confusion existed regarding the interpretation of the question, “Number of credit hours earned toward the degree you are currently seeking (don’t count this semester)” listed on the last page of the SATS-36. It was unclear to the students how to report this information if they had more than one major; the interpretations of the students included reporting the number for what they would consider their first major or reporting the number for their overall progress toward graduation. No other substantial comments were given that would suggest necessary changes to the surveys.

Descriptive statistics for the SATS-36, perception questions and statistical skill questionnaire showed no indication of floor or ceiling effects. Subscale scores for the SATS-36 are shown in Table 4.1. The pattern of responses is typical to that reported in the literature. That is, this sample of students reported higher ratings of difficulty for statistics than students reported by Tempelarr et al. (2007) studying economics, business and mathematics, (M=4.41 vs. M=3.59, $t(8)$=3.086, p=.015). This sample of students also had lower ratings of the value statistics hold for them than the comparison group (M=3.12 vs. M=5.05, $t(8)$=-9.144, p<.001).

Table 4.1. Descriptive Statistics for SATS-36 and Attitude Perceptions Using Pilot Data

	N	*Min*	*Max*	*Mean*	*SD*
<u>SATS-36</u>					
Competence					
Student	9	3.00	5.00	4.09	.61
Peer	9	4.67	5.67	5.33	.28
Faculty	9	2.50	7.00	6.83	.35
Statistics faculty	9	2.00	7.00	5.67	1.93

Table 4.1. (continued)

	N	*Min*	*Max*	*Mean*	*SD*
Value					
Student	9	2.44	4.44	3.12	.63
Peer	9	3.50	7.00	5.50	1.17
Faculty	9	6.00	7.00	6.83	.35
Effort					
Student	9	4.00	7.00	5.94	1.05
Peer	9	3.00	6.50	5.42	1.22
Statistics faculty	9	2.00	7.00	4.92	1.60
Interest					
Student	9	1.00	5.50	4.44	1.43
Peer	9	3.00	5.00	4.22	.67
Affect					
Student	9	2.33	5.33	3.61	1.06
Difficulty					
Student	9	3.29	5.71	4.41	.80

Note. Missing statistics for groups above indicates that these items were not assessed for these missing groups.

When assessing perceptions of peer and faculty attitudes, subscale scores were created to mimic those of the SATS-36 (i.e., value, affect, competence, etc.). The number of peer and faculty items where limited to include only the necessary items that would be relevant to the overall hypothesis. This allowed the length of the survey to be controlled. However, while analyzing the pilot data, it was noted that some items needed to be added in order to examine the relationship between students' attitudes with those of their peers and faculty. Specifically, the average value score for students could not be compared with the value they perceive the faculty who teach statistics hold because no value items were include for these faculty. As such, value items were added to the statistics faculty list of questions. These items include the questions, "The faculty who teach in my psychology department: believe statistics are worthless" and "…use statistics in their profession."

When comparing subscale scores between students and their peers, results showed that students had similar ratings for their interest in statistics compared to their peers (M=4.22 vs. M=4.44, p=.613). However, an interesting pattern emerged when students rated their peers' and faculty's value of statistics and cognitive competence with statistics. There were significant differences when comparing students' ratings of their own value of statistics (M=3.12) to their perception of their peers' value of statistics (M=5.50, $t(8)$=-4.280, p=.003), and when compared to their faculty's value of statistics (M=6.83, $t(8)$=-14.498, p<.001). When comparing levels of cognitive competence, students again rated themselves (M=4.09) significantly lower than their peers (M=5.33, $t(8)$=-5.260, p=.001) and their department faculty (M=6.83, $t(8)$=-10.012, p<.001) and approached a significant difference when compared to the faculty who teach statistics (M=5.67, $t(8)$=-1.906, p=.093). With respect to the effort put into statistics class, students rated themselves and their peers as putting high levels of effort in (M=5.94 and

M=5.4) and rated faculty teaching statistics as putting in lower levels of effort (*M*=4.9). These patterns may have important implications in interpreting the hypothesis. For example, does a larger difference between a student's self-rating of competence and their perception of their peer's competence help to predict the student's performance in statistics? Such questions will be examined with the larger dataset.

With respect to skill scores, students' performance here was similar to previous findings that have found low scores on statistical skill measures (Lawson et al., 2003). For example, Lawson et al. found the typical student to obtain only 30% of questions correct after engaging with a technique to help students learn statistics. Here, students had an average score of 53% on the statistical skills questionnaire (see Table 4.2). However, when breaking the questionnaire down into sections based on the APA guidelines for quantitative reasoning, average scores ranged from a low of 33% when needing to choose statistical tests to a high of 82% when interpreting basic statistical results.

Table 4.2. Percent Correct for Statistical Skills Questionnaire Using Pilot Data

	N	*Min*	*Max*	*Mean*	*SD*
Total	9	39	82	53	14
Interpretation	9	62	100	82	11
Significance	9	00	75	36	28
Confidence intervals	9	20	80	51	20
Choosing tests	9	00	83	33	25

While reviewing the results for the statistical skill questionnaire, it became apparent that the response choices for section two needed to be adapted. In the original version, the questions asked students to place a check next to each *z*-score in a given list that is significant at the p<.05 level. This method does not allow the researcher to distinguish between a *correct* response when a z-score is not significant (i.e., no check mark made) with a *blank* response when a student is unsure of the answer (i.e., no check mark made). To alleviate this, the final version required students to circle either “significant” or “non-significant”.

In addition to collecting pilot data, and in an effort to assess validity, expert feedback on the statistical skills questionnaires’ ability to measure the APA guidelines was requested from individuals who have taught statistics each for more than thirty years in higher education. One professor has taught statistics to undergraduate psychology majors at a small liberal arts school while one has taught graduate level psychology students at a large University. Feedback included the need to write questions in a “clearer context,” add more questions regarding statistical and practical significance and decrease the number of questions regarding confidence intervals. To address these concerns, each question was given a clearer context by stating a clear research design at the beginning, an additional question regarding z-score interpretations was added, three of the five confidence interval question were removed and the questions regarding statistical and practical significance were rewritten (see Appendix A). The resulting questionnaire of 20 questions was given to two students for review. Their scores were both 70% showing a higher performance than the overall pilot sample average of 51%. This was taken as an indication of the questionnaire being clearer. When marking these two questionnaires, it was apparent that there was some ambiguity in determining the responses to the question regarding effect size. It was decided to make the effect size in question equal to .85 rather than .76 to provide an example that was more definite in terms of the size of the effect.

The final data that were examined measured the frequency of techniques being used during statistics class and the frequency of APA recommended areas being covered in class. Table 4.3 shows that there is some variability with respect to techniques that are used. Some techniques were reported as being used by all students (i.e., mathematical examples, computer simulations, real life examples) and two techniques had very low frequencies (i.e., mnemonics and manipulatives). All other methods were reported by 44.4% to 77.8% of students.

All but one student reported covering all APA recommended areas in their statistics course. This is particularly interesting given the low overall statistical skill scores students received (53%). The main study utilized a larger sample to investigate the relationship between these two factors in greater detail.

Table 4.3. Frequency of Using Research Supported Techniques and Covering APA Goal Content Using Pilot Data

Techniques/Content	Frequency	Percent Used/Covered
Math examples	9	100
Computer simulations	9	100
Calculators/Spreadsheets	7	77.8
Real life examples	9	100
Activities	6	66.7
Fun/Humor	4	44.4
Stories	6	66.7
Mnemonics	0	0
Manipulatives	2	22.2
Interpreting results	8	88.9
Statistical/Practical Sig.	8	88.9
Effect size	8	88.9
Confidence intervals	8	88.9
Interpreting conclusion	8	88.9
Choosing tests	8	88.9

Validity Analysis

SATS-36

Chronbach α reliability estimates were calculated for each subscale of the SATS-36. Resulting coefficients ranged from .36 to .92. For the interest and effort subscales, values were similar to that obtained during the development of the SATS-36 (Schau, 2003) and other published research (Tempelaar et al., 2007). Values were lower than previously obtained values for affect (α=.50), cognitive competence (α=.36), value (α=.67) and difficulty (α=.58; see Table 4.4). Basic exploratory factor analysis was conducted using principal components, varimax rotation followed by a reanalysis using maximum likelihood with varimax rotation. In each case eigenvalues were used to determine factors and both models produced the identical seven factors. These factors had eigenvalues above 1.0, ranging from 1.03 to 8.09. The scree plots indicated strong jumps for the first six factors and a smaller jump for factors seven. Given that the seventh factor included only one item, had a eigenvalue just above 1.0 (1.03) and illustrated a small jump on the scree plot, the factor analysis was re-run fixing the number of components to six to examine the possibility of using the six-factor model, as proposed by the authors of the SATS-36 (Schau, 2003). The six factors that were extracted were identical to both seven-factor models with the exception of merging the item composing the seventh factor into factor four. Together, all three sets of these results are similar reproductions of the findings in the literature using exploratory factor analysis and confirmatory factor analysis (Schau et al., 1995; Dauphne, Schau & Stevens, 1997; Schau, 2003; Hilton, Schau & Olsen, 2004; Tempelaar et al., 2007).

Table 4.4 Chronbach Alpha Values and Descriptive Statistics for SATS-36 Subscales

	N	*M (SD)*	*95% CI*	Chronbach α	Schau (2003)
Affect	231	3.90 (.98)	[3.77, 4.03]	.50	.80-.89
Competence	231	4.33 (.81)	[4.23, 4.44]	.36	.77-.88
Value	231	3.78 (.92)	[3.67, 3.91]	.68	.74-.90
Difficulty	231	4.11 (.81)	[4.01, 4.22]	.58	.64-.81
Interest	231	3.95 (1.56)	[3.77, 4.18]	.92	n/a
Effort	231	6.03 (.94)	[5.93, 6.16]	.73	n/a

Note. CI = confidence interval.

However, there were some interesting deviations in the factors found here from the SATS-36 author's original intention for grouping the items. Specifically, items on the competence, affect and difficulty subscales grouped in a way that is not as intended. Nonetheless, the groupings appear to make appropriate theoretical sense. Factors one through three are comprised of items that were intended to measure similar factors including value, difficulty and effort. However, factor four includes a mix of difficulty and cognitive competence items; factor five includes a mix of cognitive competence and affect items; and, factor six includes a mix of affect, value and interest items. Careful inspection of these items reveals an interesting pattern. The difficulty and competence items comprising factor four appear to be related, not only in their wording, but also in the sense that they all measure cognitive competence despite two items originally being labeled as difficulty items. These two difficulty items differ from the remaining difficulty items that loaded on factor three. The factor three difficulty items do not assume anything about an individual's competence while the two items loading with competence items do. For example, "statistics is a complicated subject" discusses the subject rather than the individual, and "most people have to learn a new way of thinking to do statistics," although focusing on the individual, assumes a *lack* of competence in the subject prior to the course. In contrast, both items that loaded with competence items (factor three), "statistics formulas are easy to understand," and "statistics is a subject quickly learned by most people," assume that one would have a level of competence. For example, one must be cognitively competent to "understand" and "quickly learn" a subject. It is not then surprising that these two "difficulty" items loaded with cognitive competence items creating factor four.

The remaining competence items grouped on factor five. These items, "I have trouble understanding statistics because of how I think," "I will have no idea what is going on in this statistics course," and "I make a lot of math errors in statistics," are worded negatively (which contrast to the competence items that loaded on factor four) and loaded with affect items that reflect negative emotions (i.e., "I feel insecure...," "I get frustrated…," "I am under stress…," and "I am scared by statistics."). This provides another interesting pattern, that is, these three negatively worded competence items are grouped with items that measure negative affect. This pattern appears sensible such that these difficulties (i.e., difficulty making sense of what goes on it class and making a lot of errors) and affect (feeling insecure, frustrated and stressed) would likely be related.

Factor six also included affect items; however, these affect items were positive (i.e., "I like statistics" and "I enjoy taking statistics courses") and they are grouped with the two positively worded value items. Negatively worded value items loaded separately to create factor two. This is consistent with the pattern of competence and affect items discussed above in that items that are positively worded group together and items that are negatively worded group together. While one might expect negative and positive value items to be serving as a counterbalance and group together, the data suggest that the positive value items are measuring a more similar construct to the positive affect items. For example, students who like statistics may be more likely to positively value the subject and this creates an overarching factor that is being measured by the positively worded affect and value items. Finally the four interest items were loaded with these items. Again, this seems to be a reasonable grouping as one may be more interested in something they feel positive about and value, and vice versa.

Based on this item analysis, it seemed appropriate to conduct further analysis using the six subscale groupings that resulted from this item analysis in addition to the original subscale groupings of the SATS-36 and compare the results. The labels given to the six factors from this analysis were: effort (new), value (new), difficulty (new), ease, negative affect and positive affect.

Faculty Perceptions

The peer and faculty perception items were created for this study and have no established reliability or validity measures. Items were created to mimic the subscales used in the SATS-36 (e.g., affect, value, effort, etc.). Exploratory factor analysis and reliability coefficients were used to determine the extent to which these items fit these subscales.

The faculty attitude questions were analyzed using an exploratory factor analysis using principal components, varimax rotation followed by a reanalysis using maximum likelihood with varimax rotation. In each case, eigenvalues were used to determine factors and the models produced the identical three components accounting for 66.9% of the variance. These factors had eigenvalues above 1.0, ranging from 1.05 to 3.90. The scree plots indicated strong jumps for all three factors.

The items making up each factor are presented in Table 4.5. As can be seen there, the four items intended to measure affect and interest grouped together along with one value item creating a similar "positive affect" construct as that found when measuring student attitudes measured on the SATS-36. The three items intended to measure cognitive competence grouped together with two items that measure effort. This is a different pattern than what was observed with the SATS-36. There, competence items loaded with difficulty and affect items, but not effort items. This difference may be a result of the both the faculty competence and effort items being derived from teaching performance. For example, effort items included "works hard to teach the [statistics] course well," and "completes grading in a timely fashion," while competence items include items such as "when students ask a question in statistics class, my statistics professor gives a clear answer." It seems fitting, then, to suggest that these items might be measuring a similar attitude, particularly one that is reflecting teaching competence. The final factor included two items, "never cancels statistics class," an item intended to measure effort and, "statistics is worthless," an item intended to measure value. The theoretical relevance of grouping these two items is not clear and the items are weakly negatively correlated. As such, it was determined that the items may not be appropriate measures of an underlying construct measuring a similar faculty attitude, and, as such, should be kept separate.

Table 4.5 Faculty Items with Corresponding Factor Analysis Components

Factor	Item (original subscale)
Factor 1: Teaching Competence	
	Is able to communicate statistical information well to others (C)
	When students ask a question in statistics class, my statistics professor gives a clear answer (C)
	Works hard to teach the [statistics] course well (D)
	Is capable of using statistics well (C)
	Completes grading in a timely fashion (E)
Factor 2: Positive affect	
	Enjoy using statistics (A)
	Are interested in statistics (I)
	Like statistics (A)
	Would choose to teach a statistics course (I)
	Use statistics in their profession (V)

Table 4.5 (continued)

Factor	Item (original subscale)
Factor 3: Other	
	Never cancels statistics class (E)
	Statistics is worthless, (V)

Note. C = cognitive competence; D = difficulty; E = effort; A = affect; I = interest; V = value.

Table 4.6 shows that subscales for the faculty survey had high reliability coefficients for the first factor, teaching competence (α=.87), as well as the second factor of positive affect (α=.85).

Table 4.6 Chronbach Alpha Values and Descriptive Statistics for Faculty Subscales

Faculty Attitudes	*N*	*M (SD)*	*95% CI*	Chronbach α
F1: Teaching Competence	226	6.13 (1.07)	[5.98, 6.27]	.87
F2: Positive Affect	226	5.69 (.99)	[5.56, 5.82]	.85
Never cancels class	226	6.30 (1.19)	[6.14, 6.46]	---
Statistics is worthless	226	1.61 (1.23)	[1.45, 1.77]	---

Note. CI = confidence interval; F1 = Factor one; F2 = Factor two.

Peer Perceptions

Peer perception items were analyzed using exploratory factor analysis with principal components, varimax rotation followed by a reanalysis using maximum likelihood with varimax rotation. In each case, eigenvalues were used to determine factors and the models produced the identical four components accounting for 62.6% of the variance. These factors had eigenvalues above 1.0, ranging from 1.06 to 3.83. The scree plots indicated reasonable jumps for all four factors. A list of the items and their grouping is presented in Table 4.7. Some of the factors did result in similar natured items grouping together. For example, items such as, "my peer would choose to take a statistics class," and "my peers will use statistics in their profession," grouped together indicating a level of interest and value with statistics such that the peer may be interested in a career that uses statistics and will seek out the appropriate academic courses to prepare for that. Another interesting grouping is that of "my peers are often frustrated in statistics class," and "statistics takes a lot of discipline." These items appear to measure the focus and effort that a student may need to do well in a statistics course. In contrast to these groupings, factor one illustrates several items that were grouped that did not match theoretical expectations. This included the items such as "Often receive high marks in statistics class" and "Are capable of doing well in statistics class" grouping with items such as "Like statistics" and "Are interested in statistics." Overall, items did not load according to the subscale groupings that were intended when the survey was created and it was decided to use the individual items in analyses rather than the factor analyzed groupings.

Table 4.7 Peer Items with Corresponding Factor Analysis Components

Factor	Item (original subscale)
Factor 1	Often receive high marks in statistics class (C)
	Are capable of doing well in statistics class (C)
	Will learn statistics quickly (D)
	Study well for statistics class (E)
	Like statistics (A)
	Are interested in statistics (I)
Factor 2	Would choose to take statistics as an elective (I)
	Will use statistics in their profession (V)
Factor 3	Often skip statistics class (E)
	Believe statistics is worthless (V)
Factor 4	Need a great deal of discipline to learn statistics (D)
	Are frustrated while in statistics class (A)

Note. C = cognitive competence; D = difficulty; E = effort; A = affect; I = interest; V = value.

Descriptives

Survey results from the main study were analyzed in two parts. Firstly, descriptive statistics were computed and used to describe generalizability of the sample and examine the first three research hypotheses. Sample characteristics were also summarized. For students, frequencies of gender and year in school were tabulated.

Central tendency and variability for age, GPA and number of mathematics and statistics classes completed were also tabulated. For all items measuring student attitudes and perceived attitudes, central tendency and variability for subscales were tabulated. Central tendency and variability for the statistical skill questions were calculated for overall scores and by sections.

Regression Analysis

To address the fourth research hypothesis that attitude and environment will significantly predict statistical skills among students, the second part of the analysis included the use of regression models. It was expected that lower student attitude scores (reflecting more negative attitudes) and lower peer and faculty attitude scores would predict lower statistical skill scores. As such, attitude measures (student SATS-36 subscale scores) and scores for perceptions of peers and faculty attitudes and number of empirically based teaching techniques used were regressed on students' statistical skills scores. Other demographic variables and number of previous statistics course were also examined in the regression model.

CHAPTER 4

RESULTS

Main Study

Descriptives

Demographics

A total of 231 students were surveyed for the study. Demographic data are provided in Table 4.8 and Table 4.9. The majority of students were female (81%) with an average age of *M*=22.03 (4.12). The average number of credits completed for their degree was *M*=67.77 (31.89) indicating that a large number of students in the sample were likely in their sophomore or junior years. Students reported having completed an average of *M*=3.92 (1.21) high school mathematic courses and *M*=2.84 (1.32) college statistics courses. The sample included mostly students who had already completed a college statistics course (*N*=169) although a number of students (*N*=61) were still in progress of completing their first college statistics course. All data were collected toward the end of the semester.

While the majority of students were majoring in psychology (78%), the next most frequent major for students in the sample was in the sociology/criminology department at one of the institutions (14%). These students complete the same statistics courses as the psychology majors that are cross listed between psychology and sociology/criminology at their institution and were therefore included in the sample. Potential differences between psychology majors and non-psychology majors were explored and found to be insignificant.

Table 4.8. Descriptives for Age, Credits Completed and Math and Statistics Courses Taken

	N	*Min*	*Max*	*Mean*	*SD*
Age	231	19.00	53.00	22.03	4.12
Credits completed	222	3.00	160.00	67.77	31.89
High school math	230	0	8.00	3.92	1.21
College statistics	231	0	8.00	2.84	1.32

Table 4.9. Frequencies for Sample Participants' Gender and Major

	N	*%*
Gender		
Male	43	18.6
Female	187	81.0
Major		
Psychology	180	80.4
Sociology	10	4.5
Criminology	23	10.8
Other	11	4.3

Attitudes Toward Statistics

Subscale scores for the SATS-36 were computed on the six dimensions: affect, value, cognitive competence, interest, difficulty and effort as well as the six grouping that resulted from the factor analysis. Descriptive data are presented in Table 4.10. Overall, students reported their affect (M=3.89), value (M=3.78) and interest (M=3.95) toward statistics just above the middle score. Their sense of competence with the subject M=4.33 and difficulty mastering the subject M=4.11 were slightly higher than this. The amount of effort they reported they would put into a statistics course was high (M=6.03). Although the original and new subscales consist of different groupings of the items, an attempt was made to examine potential similarities and differences. When comparing value scores, the new grouping yielded a slightly lower average (M=3.52 vs. M=3.78) but this was not significantly different when comparing confidence intervals. In contrast, the new difficulty rating was significantly higher (M=4.53 vs. M=4.11) likely because "ease" questions in the original difficulty rating were excluded from the new grouping. These "ease" questions had an average of M=4.14. If this was combined with the remaining difficulty items, the average would be similar to the original difficulty average. When comparing affect scores to the new negative affect and positive affect the pattern is as expected. Both the negative affect and positive affect averages were higher than the original affect average. The high scoring competence and interest items that were included in these new groupings are likely increasing these averages. Overall, the patterns are as expected.

Table 4.10 Descriptive Data for SATS-36 Subscales

		Grouping for Subscales			
		Original SATS-36		Factor Analysis	
	n	*M* *SD*	*95% CI*	*M* *SD*	*95% CI*
Value	231	3.78 (.92)	[3.67, 3.91]	3.52 (1.43)	[3.34, 3.71]
Difficulty	231	4.11 (.81)	[4.01, 4.22]	4.52 (1.14)	[4.38, 4.67]
Ease	231	--	--	4.14 (1.05)	[4.00, 4.28]
Effort	231	6.03 (.94)	[5.93, 6.16]	6.03 (.94)	[5.93, 6.16]
Affect	231	3.90 (.98)	[3.77, 4.03]	--	--
Negative affect	231	--	--	4.03 (1.35)	[3.85, 4.20]
Positive affect	231	--	--	3.93 (1.31)	[3.77, 4.10]
Interest	231	3.95 (1.56)	[3.77, 4.18]	--	--
Competence	231	4.33 (.81)	[4.23, 4.44]	--	--

Descriptive data for the faculty attitude factor analysis derived subscale scores are provided in Table 4.11. These values indicate that students feel their faculty are highly competent to teach and use statistics (*M*=6.13) and that they feel their faculty have fairly positive affect for the subject (*M*=5.69). Note that students report much higher levels for their faculty's competence than themselves (*M*=4.33). In terms of the positive affect grouping, the students rate their faculty as having much higher affect than themselves (*M*=3.93).

Table 4.11 Descriptive Data for Faculty Attitude Subscales

	n	*M* (*SD*)	*95% CI*
Teaching Competence	226	6.13 (1.08)	[5.99, 6.27]
Positive affect	226	5.69 (.99)	[5.56, 5.82]
Other			
Never cancel class	226	6.30 (1.19)	[6.15, 6.46]
Statistics is worthless	226	1.61 (1.24)	[1.45, 1.78]

Because the factor groupings created for peer attitudes had no clear theoretical construct, peer attitudes are reported by individual items in Table 4.12. Ratings for items were fairly neutral ranging between 3.08 and 4.58. Exceptions to this included items rating peers' frustration during class (*M*=4.94), expectation of peers to use statistics in their profession (*M*=5.12) and peers being capable of doing well in their statistics course (*M*=5.32). These reflect that students perceive their peers as fairly capable with the subject regardless of somewhat high levels of frustration during class. Further, they agree that their peers will likely use statistics in their profession. It is interesting that this would be the case given that the rating of this item for the students themselves is much lower than this (*M*=3.50).

Table 4.12 Descriptive Data for Peer Attitudes Items

Peer Attitudes	*n*	*M* *SD*
Receives high marks	225	4.22 (1.11)
Capable of doing well	225	5.32 (1.02)
Learns statistics quickly	225	4.32 (1.05)
Studies for statistics well	225	4.58 (1.30)
Interested in statistics	225	3.60 (1.10)
Would choose to take statistics	225	3.08 (1.38)
Will use in profession	225	5.12 (1.84)
Skips class often	225	3.67 (1.55)
Believes statistics is worthless	225	3.64 (1.51)
Needs discipline for course	225	3.29 (1.24)
Gets frustrated in course	225	4.94 (1.34)

Teaching Techniques and Content Covered

Analysis of the teaching techniques used in the class showed that students report an average of M=5.53 (1.94) techniques reviewed in the published literature being utilized in their statistics course. This is approximately 62% of the 14 possible techniques listed on the survey. The most common techniques reported were mathematic examples (100%), use of calculators and spreadsheets (85%), real life example (85%), computer simulation (80%) and activities (65%). All other techniques were reported being used by less than 55% of the respondents.

In terms of the content areas that students reported faculty had covered in statistics course, the average was high. A total of M=5.41 areas were reported as being covered representing 90% of the total areas listed on the survey. Table 4.13 provides the frequency for each of these. Each area corresponded with an APA goal and 80% or more of the respondents reported that all areas were covered in their statistics course.

Table 4.13. Frequency of Content Areas Covered in Statistics Class

Content Area	*N*	*%*
Interpreting results	138	97.83
Significance	138	80.43
Effect size	138	89.13
Confidence intervals	138	88.41
Interpreting conclusion	138	92.75
Appropriate tests	138	92.03

Statistical Skills

The average number of items students had correct on the statistical skills test was M=10.8 or 51.5%. However, when breaking down the test into the four components representing APA goals, the percentage correct varied as follows: 74.5% for interpreting results; 39% for interpreting statistical vs. practical significance; 38% for applying confidence intervals; and, 30% for choosing the correct statistical test for a given scenario. These results are in direct contrast with students' reports of having covered these techniques in class. In addition, when breaking down the data between students who have completed statistics and those who have not, no significant differences were evident. Figure 3.2 illustrates the spread of scores on the tests. The data appear relatively normal on this graph with a skewness of *M*=-.166 (*SE*=.16) and kurtosis of *M*=-.11 (*SE*=.63). The average score was 51.5% and a majority of students falling below a score 75%. This suggests very poor statistical skills among the students in the sample.

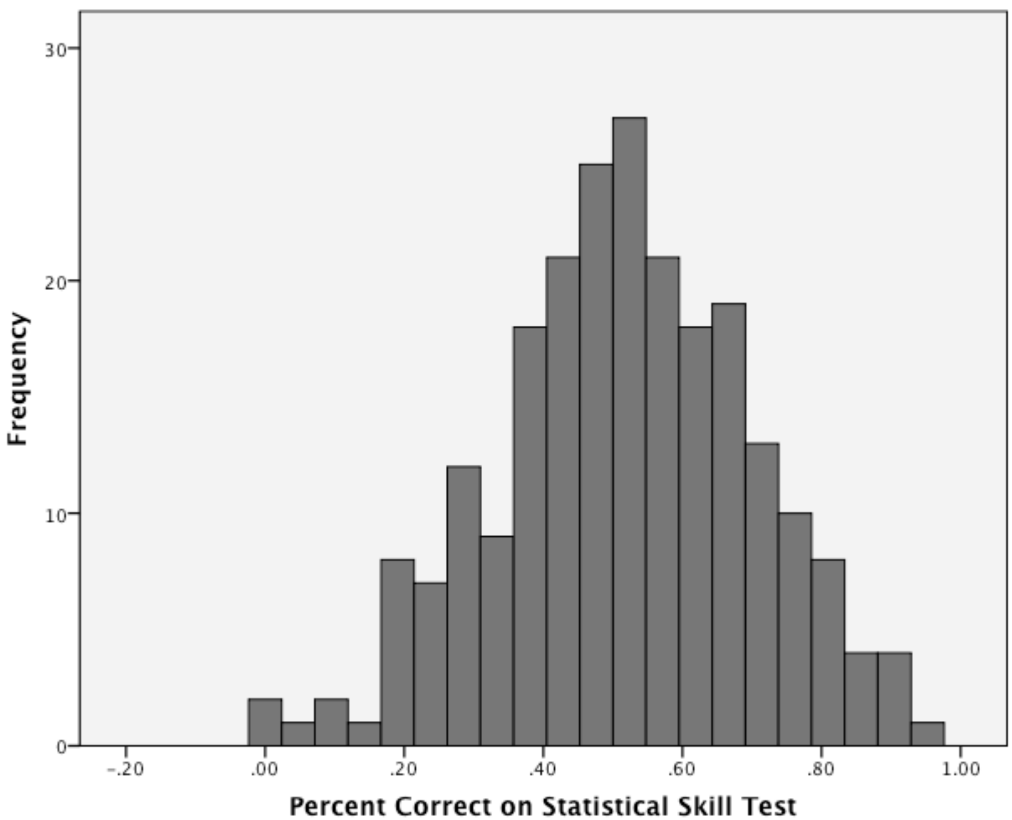

Figure 4.2. Histogram. Distribution of statistical skill test scores.

Regression Analysis

The final hypothesis for the study was that measures of attitude and environment would significantly predict statistical skills among students. It was expected that students who achieve higher scores on the skill questionnaire would have higher (i.e., more positive) ratings of attitudes for themselves, peers and faculty. To examine this, the second part of the analysis included the use of regression models. Attitude measures (student SATS-36 subscale scores), subscale scores for perceptions of faculty attitudes, peer attitude items, and the number of empirically based teaching techniques used in statistics class were regressed on students' statistical skills scores.

Correlation coefficients between the dependent variable, performance, and each independent variable are provided in Appendix C. Scatterplots between the dependent variable and predictor variables highlighted three potential outliers (see Figure D.1 and Figure D.2 in the Appendix). Each case was reviewed for potential typos in the data entry. As no such errors were found, it was determined that the first round of models would be examined with and without each outlier to determine potential influence on the data. Criteria for exclusion were set to be more than a 10% change in R^2 upon removing any outlier value. The removal of these outlier values influenced the R^2 less than 1% in all cases. Homoscedasticity and linearity of the outcome variables used were assessed by graphing the studentized residuals against predicted values for the final models. These graphs are presented in Appendix F. The graphs illustrate values equally spread above and below the reference line of zero indicating that the assumptions of homoscedasticity and linearity appear to be met. Normality of the outcome variable was graphed using a histogram (see Figure 4.2) and appeared to be normality distributed. Given the similar nature of the predictors used in the regression models, collinearity was assessed during the modeling process. Any VIF score over 10 was to be considered problematic; however, no VIF scores reached this level.

Regression analyses were conducted using SPSS tools for forward and backward modeling. In the backward model, all chosen variables were entered and then removed one at a time as a variable failed to meet the predetermined alpha level of .30. In contrast, the forward model included one variable at a time starting with the variable that is most highly correlated with statistical skill scores. The process stopped when no remaining variables that met the predetermined alpha level of .20 for inclusion remained in the model. The modeling process followed two paths. First, the original SATS-36 subscale scores were entered. The significant predictors that resulted from this process were then entered into a second round of modeling that added faculty subscale scores defined by the factor analysis (e.g., teaching competence and positive affect) as well as the two individual item composing the third factor (e.g., never cancel class and believing statistics is worthless). Finally, the combination of the significant SATS-36 subscale scores and faculty attitude measures were entered into a third round of modeling with the eleven peer attitude items. The resulting models using forward and backward techniques were then compared to choose a final model. This process was then repeated using the same steps with the exception that the subscale scores for the original SATS-36 student attitude items were replaced with subscale scores on the subscales defined by the factor analysis. The two final models, one using the original SATS-36 subscales and one using the factor analysis defined subscales, were then compared.

When examining the ability to predict overall average statistical skill scores using the original SATS-36 subscale scores, both the forward and backward models yielded two significant predictors. They included students' value of statistics (b_1=-.06) and students' interest in statistics (b_2=.029). This final model was significant, $F(2, 228)=16.96$, $p<.001$, and the amount of variance in the outcome accounted by this model was relatively small, R^2=.13. Subscale faculty scores of teaching competence and positive affect and two individual items (never cancel class and believing statistics is worthless) were added into the model. No faculty attitude measures made significant predictions of performance. Finally, the peer attitude items were added into the model. Both the backward and forward methods resulted in one significant peer attitude predictor, peer liking of statistics (b_3=-.024). Adding this variable resulted in a significant model, $F(3, 221)=13.08$, $p<.001$, but the amount of variance explained remained low, $R^2 = .15$.

The second modeling procedure replaced the original SATS-36 subscale scores with those defined by the factor analysis. Both the forward and backward methods using these scores resulted in two significant predictors, student value (b_1 = -.046) and positive affect (b_2 = .025). The model was significant, $F(2, 228)=25.65$, $p<.001$, and the amount of variance explained was low, $R^2 = .18$. Faculty attitude measures were added to the model; however, none were significant. Finally, the peer attitude items were added and both the forward and backward methods yielded two significant predictors, peer liking of statistics (b_3 = -.02) and expectation of peers using statistics in their profession (b_4 = -.02). The model was significant, $F(4, 220)=15.40$, $p<.001$, and the amount of variance explained increased slightly, $R^2 = .22$.

Table 4.14 provides a summary of the two final regression models. Using both the original SATS-36 subscales and those defined by the factor analysis yielded very similar results. In both models, student value and interest are significant predictors. In model one, using the original SATS-36 subscales, these are labeled value and interest and included items intended to measure these constructs. For model two, using the factor analysis subscales, value is also significant; however, the construct labeled positive affect is the second predictor. Note that this second predictor includes the original interest items. It also includes two value items. Therefore, for student attitudes, it appears that the difference between model one and model two is the way in which the value and interest items are grouped. It should be noted, however, that the positive affect construct in model two also includes two affect items that are similar to the interest items. They are "I like statistics," and "I enjoy statistics."

Similar results between model one and two were also obtained when examining faculty and peer attitude measures. There were no significant faculty measures in either model, and both models included peer liking of statistics as a significant predictor. One distinction is that model two included the expected use of statistics professionally by their peer as a significant predictor. Together, this suggests that both models have similar significant variables and would result in somewhat similar predictions of statistical skill performance. The second model was chosen here as the best model due to the increase in R^2 from .15 to .22.

Table 4.14 Final Regression Model Parameters, Significance and R^2 Values

	β	*F*	R^2	ΔR^2
<u>Model 1</u>				
Intercept	.669	13.08***	.15	--
SATS-36 Subscales				
Value	-.054*			
Interest	.033**			
Peer Attitudes				
Liking	-.024**			
<u>Model 2</u>				
Intercept	.695	15.40***	.22	.07
Factor Analysis Subscales				
Value	-.044***			
Positive affect	.037***			
Peer Attitudes				
Liking	-.02*			
Use in profession	-.02*			

*p<.05, **p<.01, ***p<.001

To examine the ability of other factors, beside attitudes, to predict statistical skill performance, the number of teaching techniques used in the statistic course along with demographic variables (e.g., gender, age, psychology major vs. non-psychology major, credits completed) were added one by one to the final model. Results showed that neither teaching techniques nor demographic variables were significant predictors (*p*s>.05)

In the regression models, value scores were consistently related to statistical skills in a negative way such that higher values resulted in lower performance. This outcome is counterintuitive not only at the theoretical level but also with respect to previous literature. Exploratory examination of the sample was conducted to try to identify patterns that may help to explain this outcome. Students were separated into two groups, high value and low value, by grouping above and below the average possible score of 4.5. This resulted in N=187 in the low value group and N=44 in the high value group. An independent samples t-test, adjusting for unequal variances determined by Levene's test (p=.029), showed significant differences in the average value scores between these groups, $t(55.35)=-18.70$, $p<.001$. The average for the high value group was M=5.30 and for the low value group M=3.40.

These groups were then compared on each of the remaining SATS-36 subscale scores. There was a significant difference between the high and low value groups for affect ($F(1, 229)=9.56$, p=.002), competence ($F(1, 229)=34.69$, $p<.001$), difficulty ($F(1, 229)=45.71$, $p<.001$) but not significantly different on effort (p=.801). Students with high value were much closer to neutral ratings for affect (M=4.3) and interest (M=4.9) and these ratings were significantly higher than the ratings provided by students with low value, M=3.8 and M=4.2, respectively. This indicates that students with low value also have negative affect toward and interest in statistics and this is significantly lower than ratings provided by students with high value for the subject. These groups also differed significantly with respect to their perceived cognitive competence with the subjects with those in the high value group rating their competence .4 points above the average score and the low value group rating their competence .4 points below the average score.

In terms of difficulty and effort ratings between groups, an interesting pattern emerges. Both the low group (*M*=6.0) and the high group (*M*=6.1) report a high level of effort. However, the group with low value reports statistics to be significantly higher in terms of difficulty (*M*=4.3) than the high value group (*M*=3.4). These results provide a pattern that is expected: students with high value also have positive affect and interest in the subject, rate it is less difficult than others and put a high level of effort into their work. In contrast, students with low value also have negative affect and interest in the subject, feel that it is more difficult than others and put a high level of effort into the course. However, why is it that the students with low value, affect, interest and competence, perform *better* than students with high value, affect, interest and competence? An initial explanation might be that there is a difference in terms of the effort of students in each of these groups. However, the effort scores between these two groups is not significantly different. One might expect that students with lower competence ratings as well as interest, value and affect would require more effort. Certainly, one would expect their effort is higher if they are outperforming the high value group. An explanation for this phenomenon may lie in the way in which the effort construct is measured. For instance, previous studies have found that higher value leads to higher performance. However, this has been discovered largely with samples of students in business and economic statistics courses that may be more mathematically based than the statistics courses completed by the psychology students in this sample. As such, these previous samples may have rated effort as being measured by “time” spent applying the mathematical equations to data in an effort to practice the steps they will need to follow in potential exam problems, an example of more effort based learning. The sample of psychology students would likely be spending more time producing and interpreting the results of statistical tests as produced by statistical software and these types of interpretations will be somewhat unique depending on the data that are producing them. This implies that the type of effort needed to do well on these types of interpretation problems is not measured only by "time" or amount of practice but rather is better measured by the quality of time (i.e., what the students are doing with their time) spent to gain understanding of the underlying principles being applied. More mathematically based and/or effort-based statistics course work can be mastered more easily by shear repetition of such problems but would not result in higher performance for a psychology based statistics course that emphasizes running and interpreting computer analysis. This leads to an explanation of why this smaller subset of high value students in this sample actually performed lower than the subset of low value students in this sample. Perhaps the quality of their effort was to try to complete problems by applying a general pattern of steps to the problem rather than spending time understanding the underlying principles and how to apply these to varying scenarios. This would impact things such as their ability to choose appropriate statistical tests for different scenarios and interpret outcomes of statistical tests. In fact, this is exactly where they deviated significantly in their performance from low value students in this study. Students in the high value group had an average of 22% when choosing tests compared to 32% for the low group ($t(228)=2.00$, $p=.047$, $d=.34$), and, when interpreting the results of a statistical test, the high group had an average of 22% compared to 43% for the low group (t(83.42)=4.61, p<.001, d=.77).

An alternative explanation for this phenomenon might be that the construct of value may not include a key component for these students, which is directly related to performance. The items on the SATS-36 that measure value are examining a student's value of the subject in general. Value is comprised by items such as "statistics is worthless" and "statistics is irrelevant in my life." What the value construct does not measure is the students' value of the success in the subject as it pertains to their academic achievement. That is, how much does a student value doing well so that they can maintain the desired GPA, continue in their major and even maintain their self-pride? This is an important piece of the expectancy-value theory that appears to be missing from the SATS-36 measurement. Expectancy-value suggests that one must have some level of value for completing a task in order to engage in it. Perhaps for some students, particularly those that have performed better than others (i.e., the low value group), have this level of value with respect to wanting to do well in the course and as such have performed better. Unfortunately, this type of value is not measured by the SATS-36 and as such cannot be modeled here.

CHAPTER 5

CONCLUSION

Summary of Results

The purpose of this study was to investigate the attitudes of psychology students toward statistics as well as their perceived attitudes of their peers and faculty toward statistics and the relationship these attitudes might have on a student's performance in statistics. Other environmental variables, which may have an impact on performance, such as teaching techniques used in the statistics course and content areas covered, were also assessed. It was expected that the students' attitude ratings would result in high ratings of difficulty and low levels of value for statistics and that their peers and faculty would be perceived to have similar attitudes. Results showed that, on average, students did have fairly low levels of value for statistics (M=3.78). Their affect and level of interest in the course were also low (M=3.90 and M=3.95, respectively). In terms of their perceived difficulty of the course, their ratings were not as high as expected (M=4.11). Their perceived cognitive competence with the subject was rated similarly with an average of M=4.33. However, they reported putting in high levels of effort for the course (M=6.03).

It is interesting to compare these ratings with that of their peers and faculty. Overall, students appear to feel that they are similar to their peers in that they do not have a lot of interest or positive affect toward statistics but they perceive that the faculty are very interested and feel positively toward the subject. Regardless of this level of interest and positive affect, they still believe that all parties- themselves, peers and faculty- have low levels of value for the subject. When considering the difficulty of the subject and level of effort put into achieving in the class, students believe the subject to not be overwhelming difficult for themselves or their peers though they do believe that it is very difficult for their faculty to teach the course. The level of effort reported by students was interesting. They appear to believe that both they and the faculty work very hard at the subject with average scores at M=6.03 and M=6.29, respectively; however, they do not rate their peers as putting as much effort into the course, M=4.10. When considering the difference between students and their peers, it at first may seem that this is reflecting a belief that they are not as competent as their peers and will need to work harder. When looking at competence scores, it appears this is not the case, as they rate themselves similarly to their peers (M=4.33 and M=4.74, respectively). Perhaps, there is a tendency to overemphasize the amount of work they are doing because they are not present to observe their peers when performing work for the course. This may also explain the relationship to faculty ratings. Given that the survey items asked about faculty effort primarily during the course instruction, they see the work that faculty are doing, and therefore rate their effort accordingly.

Together, these attitude ratings paint a picture that is not very positive. Particularly concerning is the idea that students do not value statistics. For those who have spent time around undergraduates taking statistics, these results may not be surprising along with the students' perception that their peers do not value statistics; however, it is very concerning that they do not believe their faculty value statistics, a subject that is critical to the pursuit of objective evaluation which allows our discipline to be considered a science.

When assessing the performance of students on the statistical skills questionnaire, it was expected that performance would be low. Lawson et al. (2003) reported that students achieved an average of 30% on a statistical test given during a research study. Findings in this study were similar. Overall, the average score on the skills questionnaire was 51.5%. The questionnaire consisted of four main content areas, all which were reported as covered in statistics class by more than 80% of the students. Nonetheless, when breaking down performance in each of these covered areas, average scores were between 30% and 39% for three sections and 74.5% for one section. The later of these is perhaps the most basic where students were asked to interpret results from a table (e.g., *N, M, SD*) and to indicate which of four z-scores would be significant at an alpha level of .05. The remaining three sections covered content that, although more complicated, is fundamental to understanding and applying statistics. For example, the ability to interpret when a test is significant (i.e., $p<.05$) or when the effect size is large (i.e., $d>.80$) is a basic and necessary skill for both conducting and reading research. The APA guidelines clearly articulate this as a necessary skill and students report having covered this content in statistics class. While perfect performance may not be expected when testing these skills after completing introductory statistics course at the undergraduate level, it seems that the average score of 39% obtained in this section is well below proficiency.

The final hypothesis for this study was that factors such as attitudes and teaching techniques used in the class would predict students' performance on the skills questionnaire such that lower attitude scores (reflecting more negative attitudes) and fewer teaching techniques would predict lower statistical skill scores. Results found that there was no relationship between performance on the skills questionnaire and teaching techniques or perceptions of faculty attitudes. Across both models, results did show student interest in, value of and positive affect toward statistics as well as perceived peer interest and liking were significant predictors of statistical skill performance. Interest predicted a 3.3% gain in performances. This suggests that students with low interest (*M*=2) would increase approximately 6.6% in their overall performance score. In contrast, students with high interest (*M*=6) would increase 19.8% increase in their performance. Positive affect toward the subject also predicted significant gains of 3.7%. A student with low scores on positive affect (*M*=2) could expect, on average, a 7.4% increase in performance while a student with a higher score (*M*=6) could expect a 22.2% average increase. These predicted high gains in performance highlight the strong impact that student interest and positive affect appear to be playing in performance for this sample.

In contrast to these findings, value and perceived use of statistics in their peers' profession predicted decreases in performance. As discussed earlier, these results conflict with prior evidence as well as theoretical models as you would not expect higher value or peer attitude scores to result in lower performance. The possible explanations of systematic differences in effort and interpretations of value being the cause for this negative relationship must be evaluated in future research.

Limitations and Implications for Research

The sample used in this study was taken from two institutions. While the institutions were different in many regards, they do not include the broad characteristics of psychology students, departments or universities overall. Future studies should be performed in an attempt to be duplicated the results of this study with varying samples.

This study examined very specific factors to assess their relationship to performance in statistics class. While they were developed from a theoretical model and supported by prior research, they are not exhaustive. Other variables that might be related to psychology majors learning statistics that can be included in future studies include characteristics related to different psychology departments, experience of those teaching the statistics course and future goals of students.

Further investigation is needed to assess the appropriateness of the statistical skills questionnaire used as well as the attitude measures used for peers and faculty. The skills questionnaire was designed to measure four of the quantitative goals outlined by the APA. Upon searching, there appeared to be no previously published tool to measure these goals. The author derived questions and experts who have taught statistics to psychology majors reviewed them. It would be useful to create alternative questions and compare performance between two versions of the questionnaire to examine criterion validity. Also, content validity might be examined by comparing results on this questionnaire to other performance measures such as course exams and assignments.

There were several limitations of the current study including the small number of items used to assess peer and faculty attitudes. While this was necessary for this study so as to not over burden participants with a large number of items on the surveys, it is problematic for analyzing reliability and validity of the measurement. This study was not able to create subscale scores using the constructs from the SATS-36. While the items were framed so that they would measure similar constructs, the inferred interpretation of these items by students resulted in different groupings than expected and several of the items were negatively correlated. For example, two items were given to measure a peer's value of statistics. These were, "My peers will use statistics in their profession" and "My peers believe statistics is worthless." One might believe that anticipated use of statistics in a profession would be correlated with the belief that statistics is valued (i.e., not worthless); however, an alternative explanation might be that students believe statistics are valuable but they do not anticipate using them in their profession. The later would result in a negative correlation such as that found here. Had additional items been included in the survey to measure value of statistics, such as desired and/or expected grade in the course, a composite score may have been able to be derived from positively correlated items.

In this study, peer and faculty attitudes were not directly measured but instead rated as a perception of the student. It would be interesting to compare these results to those directly rating peer and faculty ratings. This would allow further investigation into why the measures of peer attitudes and faculty attitudes did not yield more significant predictors as well allow for further validity testing on the measure of peer and faculty attitudes. The type of research design necessary for doing this has its own limitations. First, it might not be feasible in terms of the number of institutions that would be necessary to obtain a reasonable sample size of faculty willing to participate. Second, to best compare an individual's performance with their peers' ratings it might be necessary to conduct a study where a few selected individuals with differing peer groups provide measures of performance and these results are compared to their averaged peer ratings to avoid a constant peer score rating that would exist if using all the peers in the group to provide measures of performance.

Another research design that may be helpful in examining the attitudes of faculty is to employ an experimental design where two groups are presented the same lecture on a given statistics topic; however, the attitude of the faculty could be systematically varied during the presentation. Results could examine the change in a pre-post test given on the content for effect of the attitude presented by the faculty to examine a more causal relationship. This design could be used in an actual classroom as well as with computer modules to further assess the potential confounds that can cloud the results of classroom comparisons (e.g., classroom dynamic, day/time of class, location, etc.).

Implication for the Field

The performance on the statistical skills questionnaire by students in this sample was very poor. Their ability to interpret basic statistical conclusions produced an average score (*M*=74.5%), but their performance in other areas such as interpreting the results of statistical tests (*M*=39%) and choosing appropriate statistical tests (*M*=30%) was well below a satisfactory score. It may be important to consider the practicality of the APA guidelines for quantitative reasoning. This poor performance leaves several questions with regard to its cause, particularly after indicating covering the content area in class and the use of empirically tested methods being used. A quick explanation already mentioned may be that the measurement tools are not valid. Another possible explanation might be the impracticality of meeting the guidelines. It may not be uncommon for a statistics course to be scheduled for one semester. Accomplishing the goals of the APA guidelines such as “choosing appropriate tests given particular hypotheses” is challenging to meet in one semester given the large number of tests available. As such, faculty may attempt to meet the goal by covering each test in brief in an effort to include all tests in the one semester course. If this is the case, the result can be seen here in the form of poor performance when applying statistical skills. This level of performance has heavy bearing on the field in the event that such students will pursue higher levels of education in psychology that may lead them to conducting or interpreting research. Based on research examining the extent of statistical education in psychology graduate programs (Aiken et al., 2008), it should not be assumed that this deficit in knowledge would be made up for in graduate studies. This possibility might be playing out in the literature. For example, Bakker and Wicherts (2011) recently reported that psychological research may often include errors when applying statistical techniques. They found that 15% of a sample of psychological journal articles had inaccurate statistical tests published that changed the direction of the results. Of most concern is the finding that the majority of these inappropriate conclusions had been made to support the authors' hypothesis; however, by correcting the statistical analysis, the results would no longer support these findings. The implication of this is that psychologists who are looking to scientific research to guide their practices or recommendations may be misguided.

Attention to this issue has not been limited to the professional sector. Begley (2010) published a critical Newsweek article claiming psychologists are using poor research to guide their practice. Interestingly, the reference Begley used as a basis for this critique was problematic with respect to research design (see Farley, 2010). More recently, news of a psychologist publishing the results of several fraudulent psychological studies has been aired. This invited comment from several psychologists who feel that the field has gone astray in its ability to properly perform research ("Fraud Scandal", 2011) and these concerns were presented and discussed at the 2011 APA Annual Convention (Farley, Jennings & Smith-Dyer, 2011). Currently, an investigation is underway to re-examine psychological research studies published in three highly cited journals to determine the extent to which research in the field is implementing these problematic research methods ("Is Psychology Undone", 2012). It should be pointed out that it is not only within the published research that these shortcomings may exist. Even the study within this dissertation has limitations that can lead to problematic conclusions.

The importance of psychological research being done correctly should not be underestimated. The role of psychologists in society has great importance. In a recent review of popular news outlets it was reported that close to half of news stories covered a psychological topic (Farley et al., 2009). Typically, these reports cited psychological research (44%) or included an interview from experts in the area (66%). While the topics that were covered in that study example may not be controversial, the expertise provided by psychology, as a science, has been called upon to aid in policy and law surrounding controversial issues. For example, recently there was a debate in the literature regarding the role of psychologists' influence over laws concerning abortion rights for females under 18 and the death penalty for youth under 18 (Fischer, Stein & Heikkinen, 2009; Steinberg, Cauffman, Woolard, Graham & Banich, 2009a; Steinberg, Cauffman, Woolard, Graham & Banich, 2009b). When considering the appropriateness for females under 18 to seek parental consent before seeking an abortion, the American Psychological Association (APA) provided in a Supreme Court amicus brief that these young women have the necessary developmental maturity to make decisions around abortion without parental consent. In contrast to this decision, the APA argued that the death penalty is not appropriate for youth under 18 because they may lack the appropriate developmental maturity to understand the consequences of their behavior and control impulses in the same way that an individual over the age of 18 may. This argument ultimately influenced the abolishment of the juvenile death penalty. For some, these decisions presented as a contradictory stance for psychology to take. As mentioned above, this led to a great deal of conversation in the literature regarding the appropriateness of the decisions, and importantly, the ability for psychology to make such recommendations (Fischer et al., 2009; Steinberg et al., 2009a; Steinberg et al., 2009b). Certainly, this event highlights the serious issues psychologists may influence. As a scientific discipline, psychology aims to be grounded in research and this research then shapes our understanding in the field and application of psychology to real world issues such as those presented here. Without the appropriate tools to effectively interpret scientific research on these issues, psychologists may make inappropriate decisions. In addition, the skills of those who actually conduct this research are also of most importance to ensure that the field is grounded as a scientific discipline.

References Cited

Aiken, L. S., West, S. G., Millsap, R. E. (2008). Doctoral training in statistics, measurement, and methodology in psychology. *American Psychologist, 63*(1), 32-50.

American Psychological Association, Task Force on Psychology Major Competencies. (2007). *APA guidelines for the undergraduate psychology major.* Washington, D C: Author.

Auster, C. J. (2000). Probability sampling and inferential statistics: An interactive exercise using M&M's. *Teaching Sociology, 28*(4), 379-85. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=EJ648144&site=ehost-live&scope=site

Bakker, M. & Wicherts, J. M. (2011). The (mis)reporting of statistical results in psychology journals. *Behavioral Research, 43,* 666-678.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory.* Princeton, New Jersey: Prentice-Hall.

Basturk, R. (2005). The effectiveness of computer-assisted instruction in teaching introductory statistics. *Educational Technology & Society, 8*(2), 170-178. www.tojet.net/articles/ 918.pdf.

Begley, S. (2009, October). Ignoring the evidence. Why do psychologists reject science? *Newsweek.* Retrieved from http://www.thedailybeast.com/newsweek/2009/10/01/ignoring-the-evidence.html.

Blanco, A. (2011). Applying social cognitive career theory to predict interests and choice goals in statistics among Spanish psychology students. *Journal of Vocational*

Behavior, 78(1), 49-58. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=EJ907879&site=ehost-live&scope=site; http://dx.doi.org/10.1016/j.jvb.2010.07.003

Blanco, M., & Ginovart, M. (2010). How to introduce historically the normal distribution in engineering education: A classroom experiment. *International Journal of Mathematical Education in Science and Technology, 41*(1), 19-30. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=EJ882145&site=ehostlive&scope=site;http://www.informaworld.com/openurl?genre=article&id=doi:10.1080/00207390903060628

Brooks, G. P., & Raffle, H. (2005). Fish: A new computer program for friendly introductory statistics help. *Teaching Statistics: An International Journal for Teachers, 27*(3), 81-88. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=EJ694418&site=ehost-live&scope=site; http://dx.doi.org/10.1111/j.1467-9639.2005.00221.x

Carmona, J., Martinez, R. J., Sanchez, M. (2005). Mathematical background and attitudes toward statistics in a sample of Spanish college students. *Psychological Reports, 97,* 53-62.

Christopher, A. N., & Marek, P. (2009). A palatable introduction to and demonstration of statistical main effects and interactions. *Teaching of Psychology, 36,* 130-133.

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

Cobb, P., Wood, T., Yackel, E., Nicholls, J., Wheatley, G., Trigatti, B., et al. (1991). Assessment of a problem centered second grade mathematics project. *Journal for Research in Mathematics Education, 22*(2), 3-29.

Collins, L. B., & Mittag, K. C. (2005). Effect of calculator technology on student achievement in an introductory statistics course. *Statistics Education Research Journal, 4*(1), 7-15.

Coetzee, S., & and Van der Merwe, P. (2010). Industrial psychology students' attitudes towards statistics. *SA Journal of Industrial Psychology, 36*, 1-8. Retrieved from www.sajip.co.za.

Dauphinee, T. L., Schau, C., & Stevens, J. J. (1997). Survey of attitudes toward Statistics: Factor structure and factorial invariance for women and men. *Structural Equation Modeling, 4*(2), 129-141.

delMas, R. C., & Garfield, J. (1999). A model of classroom research on action: Developing simulation activities to improve students' statistical reasoning. *Journal of Statistical Education, 7*(3).

Dinov, I. D., Sanchez, J., & Christou, N. (2008). Pedagogical utilization and assessment of the statistic online computational resource in introductory probability and statistics courses. *Computers & Education, 50*(1), 284-300. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=EJ777404&site=ehost-live&scope=site; http://dx.doi.org/10.1016/j.compedu.2006.06.003

Dunn, P. K. (2005). We can still learn about probability by rolling dice and tossing coins. *Teaching Statistics: An International Journal for Teachers, 27*(2), 37-41. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=EJ687829&site=ehost-live&scope=site; http://dx.doi.org/10.1111/j.1467-9639.2005.00205.x

Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology, 53,* 109-132.

Elmore, P. V., Levis, E. L., & Bay, M. L. G. (1993). Statistics achievement: A function of attitudes and related experience. Paper presented at the annual meeting of the American Educational Research Association, Chicago IL.

Farley, F., Jennings, H., & Smith-Dyer, T. (2011, August). *Examining psychological science-Some inconvenient truths in the new monastic order.* Paper session presented at the American Psychological Association Annual Convention, Washington, DC.

Farley, F. (2009, Winter). Pass the salt. *Amplifier,* 10-11.

Farley, F., McMenamin Bakley, A., Dai, T., Funk, K., Litchfield, B., Smith- Dyer, T., & Walsh, J. (2009, Winter). Public psychology: Demand, supply- an exploration, *Amplifier,* 10-11.

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

Fischer, K. W., Stein, Z., & Heikkien, K. (2009). Narrow assessments misrepresent development and misguide policy. *American Psychologist, 64*(7), 595-600.

Fraud scandal fuels debate over practices of social psychology. (2011, November). *The Chronicle.* Retrieved from http://chronicle.com/article/As-Dutch-Research-Scandal/129746/

Friedrich, J., Buday, E., & Kerr, D. (2000). Statistical training in psychology: A national survey and community on undergraduate programs. *Teaching of Psychology,* *27*(4) 248-257.

Garfield, J. B. (2003). "Assessing statistical reasoning," *Statistics Education Research Journal*, *2*(1), 22-38. http://www.stat.auckland.ac.nz/~iase/serj/SERJ2(1).pdf

Garfield, J., delMas, B., & Chance, B. (n.d.) *Tools for Teaching and Assessing Statistical Inference*. http://www.tc.umn.edu/~delma001/stat_tools/

Goddard, R. D., Hoy, W. K., & Hoy, A. W. (2000). Collective teacher efficacy: Its meaning, measure and impact on student achievement. *American Educational Research Journal, 37*(2), 479-507.

Groth, R., and Bergner, J. (2005). "Pre-Service elementary school teachers' metaphors for the concept of statistical sample," *Statistics Education Research Journal*, *4*(2), 27-42. http://www.stat.auckland.ac.nz/~iase/serj/SERJ4(2)_groth_bergner.pdf

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences, 33,* 61-135.

Hilton, S. C., Schau, C., & Olsen, J. A. (2004). Survey of Attitudes Toward Statistics: Factor structure invariance by gender and by administration time. *Structural Equation Modeling, 11*(1)*,* 92-109.

Hirsch, L. S., and O'Donnell, A. M. (2001). "Representativeness in statistical reasoning: Identifying and assessing misconceptions," *Journal of Statistics Education*, *9*(2). http://www.amstat.org/publications/jse/v9n2/hirsch.html

Holder, W. B., Leavitt, G. S., & McKenna, F. S. (1958). Undergraduate training for psychologists. *The American Psychologist, 13,* 585-588.

Holland, B. K. (2007). A classroom demonstration of hypothesis testing. *Teaching Statistics, 29*(3), 71-73.

Is psychology about to come undone? (2012, April). *The Chronicle.* Retrieved from http://chronicle.com/blogs/percolator/is-psychology-about-to-come-undone/29045

Jones, G. T., Hagtvedt, R., & Jones, K. (2004). A VBA-based simulation for teaching simple linear regression. *Teaching Statistics: An International Journal for Teachers, 26*(2), 36-41. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=EJ687786&site=ehost-live&scope=site; http://dx.doi.org/10.1111/j.1467-9639.2004.00149.x

Kamin, L. F. (2010). Using a five-step procedure for inferential statistical analyses. *American Biology Teacher, 72*(3), 186-188. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=EJ884176&site=ehost-live&scope=site; http://caliber.ucpress.net/doi/abs/10.1525/abt.2010.72.3.11

Knypstra, S. (2009). Teaching statistics in an activity encouraging format. *Journal of Statistics Education, 17*(2). www.amstat.org/publications/jse/v17n2/knypstra.html

Konold, C., Pollatsek, A., Well, A. D., Lohmeier, J., & Lipson, A. (1993). "Inconsistencies in students' reasoning about probability," *Journal for Research in Mathematics Education, 24*, 392-414.

Kraus, S. (2010). Monstrous methods for teaching central tendency concepts. *Teaching Statistics: An International Journal for Teachers, 32*(1), 21-23. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=EJ869632&site=ehost-live&scope=site; http://www.blackwell-synergy.com/doi/abs/10.1111/j.1467-9639.2009.00403.x

Lawson, T., Schwiers, M., Doellman, M., Grady, G., & Kelnhofer, R. (2003). Enhancing students' ability to use statistical reasoning with everyday problems. *Teaching of Psychology, 30*(2), 107-110. doi:10.1207/S15328023TOP3002.

Lee, V. E., Smith, J. B., & Croninger, R. G. (1997). How high school organization influences the equitable distribution of learning in mathematics and science. *Sociology of Education, 70*, 128-150.

Lesser, L. M., & Pearl, D. K. (2009). Functional fun in statistics teaching: Resources, research and recommendations. *Journal of Statistics Education, 16*(3) Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=EJ841775&site=ehost-live&scope=site; http://www.amstat.org/publications/jse/v16n3/lesser.pdf

Libman, Z. (2010). Integrating real-life data analysis in teaching descriptive statistics: A constructivist approach. *Journal of Statistics Education, 18*(1) Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=EJ892395&site=ehost-live&scope=site; http://www.amstat.org/publications/jse/v18n1/libman.pdf

Marlin, J. W., (1991). State-Mandated economic education, teacher attitudes, and student learning. *Journal of Economic Education, 22,* 5-14.

Martin-Dunlap, C., & Fraser, B. J. (2007). Learning environment and attitudes associated with an innovative science course designed for prospective elementary teachers. *International Journal of Science and Mathematics Education, 6,* 163-190.

Mathews, D., & Clark, J. (2003). "Successful students' conceptions of mean, standard deviation and the central limit theorem," unpublished manuscript. http://www1.hollins.edu/faculty/clarkjm/papers.htm

Mills, J. A. (2004). Students' attitudes toward Statistics: Implications for the future. *College Student Journal, 38*, 349-361.

Mji, A. (2009). Differences in university students' attitudes and anxiety about statistics. *Psychological Reports, 104,* 737-744.

Morris, E., Joiner, R., & Scanlon, E. (2002). The contribution of computer-based activities to understanding statistics. *Journal of Computer Assisted Learning, 18*(2), 114-124. 10.1046/j.0266-4909.2002.00219.x .

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. doi:10.1111/1467-8535.00175.

Mulhern, G., & Wylie, J. (2005). Mathematical prerequisites for learning statistics in psychology: Assessing core skills of numeracy and mathematical reasoning among undergraduates. *Psychology Learning and Teaching, 5*(2). 119-132.

Mvududu, N. (2003). A cross-cultural study of the connection between students' attitudes toward statistics and the use of constructivist strategies in the course. *Journal of Statistics Education*, *11*(3). Retrieved from www.amstat.org/publications/jse/v11n3/mvududu.html

Neumann, D., Hood, M., & Neumann, M. (2009). Statistics? You must be joking: The application and evaluation of humor when teaching statistics. *Journal of Statistics Education, 17*(2), Retrieved from www.amstat.org/publications/jse/ v17n2/neumann.htmllivepage.apple.com

Newmann, F. M., & Wehlage, G. G. (1995). Successful school restructuring: A report to the public and educators. Washington, D. C.: Office of Education Research and Improvement.

Palardy, G. J., & Rumberger, R. W. (2008). Teacher effectiveness in first grade: The importance of background qualifications, attitudes, and instructional practices for student learning. *Educational Evaluation and Policy Analysis, 30*(2), 111-140.

Pollatsek, A., Lima, S. & Well, A. D. (1981) Concept or computation: Students' understanding of the mean. *Educational Studies in Mathematics, 12*, 191–204.

Ruggeri, K., Dempster, M., Hanna, D., & Cleary, C. (2008). Experiences and expectations: The real reason nobody likes stats. *Psychology Teaching Review, 14*(2), 75-83.

Sabiote, C. R., Pérez, J. G., & Llorente, T. P. (2010). An empirical study of presage variables on the teaching-learning of statistics in the light of contemporary research on competencies. *Electronic Journal of Research in Educational Psychology, 8*(1), 235-262.

Schau, C. (2003, August). *Students' attitudes: The "other" important outcome in statistics education.* Paper presented at the Joint Statistical Meetings, San Francisco.

Schau, C., & Mattern, N. (1997). "Use of map techniques in teaching statistics courses." *The American Statistician, 51*(2), 171-175.

Schau, C., Stevens, J., Dauphinee, T. L., & Del Vecchio, A. (1995). The development and validation of the survey of attitudes toward statistics. *Educational and Psychological Measurement, 55,* 868-875.

Schoenfelder, E., Olson, R., Bell, M., & Tom, K. (2007). Stop and smell the roses: An activity for teaching the central limit theorem. *Psychology Learning & Teaching, 6*(2), 80-84. Retrieved from PsycINFO database.

Shultz, K. S., & Koshino, H. (1998). Evidence of reliability and validity for Wise's Attitude Toward Statistics scale. *Psychological Reports*, *82*, 27-31.

Smith, B. (2003). Using and evaluating resampling simulations in SPSS and excel. *Teaching Sociology, 31*(3), 276-87. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=EJ679815&site=ehost-live&scope=site

Steinberg, L., Cauffman, E., Woolard, J., Graham, S., & Banich, M. (2009a). Are adolescents less mature than adults? Minors' access to abortion, the juvenile death penalty, and the alleged APA "flip-flop". *American Psychologist, 64*(7), 583-594.

Steinberg, L., Cauffman, E., Woolard, J., Graham, S., & Banich, M. (2009b). Reconciling the complexity of human development with the reality of legal policy. *American Psychologist 64,* (7), 601-604.

Tempelaar, D. T., Gijselaers, W. J., & Schim van der Loeff, S. (2006). "Puzzles in statistical reasoning," *Journal of Statistics Education, 14*(1), 1-26.

Tempelarr, D. T., Van deer Loeff, S. S., & Gijselaers, W. H. (2007). A structural equation model analyzing the relationship of students' attitudes toward statistics, prior reasoning abilities and course performance. *Statistics Education Research Journal, 6*(2), 78-102.

Vanhoof, S., Castro Sotos, A. E., Onghene, P., Verschaffel, l., Van Dooren, W., Van den Noortgate, W. (2006). Attitudes toward statistics and their relationship with short and long term exam results. *Journal of Statistics Education, 14*(3).

Waters, L. K., Martelli, T. A., Zakrajsek, T., & Popovich, P. M. (1988). Attitudes toward statistics: An evaluation of multiple measures. *Educational and Psychological Measurement*, *48,* 513-516.

Waters, L. K., Martelli, T. A., Zakrajsek, T., & Popovich, P. M. (1988). Attitudes toward statistics: An evaluation of multiple measures. *Educational and Psychological Measurement*, *48,* 513-516.

Wiberg, M. (2009). Teaching statistics in integration with psychology. *Journal of Statistics Education*, *17(1)*, Retrieved from www.amstat.org/publications/jse/v17n1/wiberg.html

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

Zieffler, A., Garfield, J., Alt, S., Dupuis, D., Holleque, K., & Chang, B. (2008). What does research suggest about the teaching and learning of introductory statistics at the college level? A review of the literature. *Journal of Statistics Education, 16*(2). Retrieved from www.amstat.org/publications/jse/v16n2/zieffler.html

APPENDICES

Appendix A

Survey of Attitudes Toward Statistics

© Schau, 1996, 2003

DIRECTIONS: The statements below are designed to identify your attitudes about statistics. Each item has 7 possible responses. The responses range from 1 (strongly disagree) through 4 (neither disagree nor agree) to 7 (strongly agree). If you have no opinion, choose response 4. Please read each statement. Mark the one response that most clearly represents your degree of agreement or disagreement with that statement. Try not to think too deeply about each response. Record your answer and move quickly to the next item. Please respond to all of the statements.

	Strongly disagree			Neither disagree nor agree			Strongly agree
I tried to complete all of my statistics assignments.	1	2	3	4	5	6	7
I worked hard in my statistics course.	1	2	3	4	5	6	7
I like statistics.	1	2	3	4	5	6	7
I feel insecure when I have to do statistics problems.	1	2	3	4	5	6	7
I have trouble understanding statistics because of how I think..	1	2	3	4	5	6	7
Statistics formulas are easy to understand.	1	2	3	4	5	6	7
Statistics is worthless.	1	2	3	4	5	6	7
Statistics is a complicated subject.	1	2	3	4	5	6	7
Statistics should be a required part of my professional training.	1	2	3	4	5	6	7
Statistical skills will make me more employable.	1	2	3	4	5	6	7
I will have no idea of what's going on in this statistics course.	1	2	3	4	5	6	7
I am interested in being able to communicate statistical information to others.	1	2	3	4	5	6	7

	Strongly disagree			Neither disagree nor agree			Strongly agree
Statistics is not useful to the typical professional.	1	2	3	4	5	6	7
I tried to study hard for every statistics test.	1	2	3	4	5	6	7
I get frustrated going over statistics tests in class.	1	2	3	4	5	6	7
Statistical thinking is not applicable in my life outside my job.	1	2	3	4	5	6	7
I use statistics in my everyday life	1	2	3	4	5	6	7
I am under stress during statistics class.	1	2	3	4	5	6	7
I enjoy taking statistics courses.	1	2	3	4	5	6	7
I am interested in using statistics.	1	2	3	4	5	6	7
Statistics conclusions are rarely presented in everyday life.	1	2	3	4	5	6	7
Statistics is a subject quickly learned by most people.	1	2	3	4	5	6	7
I am interested in understanding statistical information.	1	2	3	4	5	6	7
Learning statistics requires a great deal of discipline.	1	2	3	4	5	6	7
I will have no application for statistics in my profession.	1	2	3	4	5	6	7
I make a lot of math errors in statistics.	1	2	3	4	5	6	7
I tried to attend every statistics class session.	1	2	3	4	5	6	7

	Strongly disagree			Neither disagree nor agree			Strongly agree
I am scared by statistics.	1	2	3	4	5	6	7
I am interested in learning statistics.	1	2	3	4	5	6	7
Statistics involves massive computations.	1	2	3	4	5	6	7
I can learn statistics.	1	2	3	4	5	6	7
I understand statistics equations.	1	2	3	4	5	6	7
Statistics is irrelevant in my life.	1	2	3	4	5	6	7
Statistics is highly technical.	1	2	3	4	5	6	7
I find it difficult to understand statistical concepts.	1	2	3	4	5	6	7
Most people have to learn a new way of thinking to do statistics.	1	2	3	4	5	6	7

Please notice that the labels for each scale on the rest of this page change from item to item.

How well did you do in mathematics courses you have taken in the past?	Very poorly 1	2	3	4	5	6	Very well 7
How good at mathematics are you?	Very poor 1	2	3	4	5	6	Very good 7
In the field in which you hope to be employed when you finish school, how much will you use statistics?	Not at all 1	2	3	4	5	6	Great deal 7
How confident are you that you can master introductory statistics material?	Not at all confident 1	2	3	4	5	6	Very confide nt 7

Are you required to take this statistics course (or one like it) to complete your degree program?	Yes 1			No 2			Don't know 3
If the choice had been yours, how likely is it that you would have chosen to take any course in statistics?	Not at all likely 1	2	3	4	5	6	Very likely 7

DIRECTIONS: For each of the following statements mark the one best response. Notice that the response scale changes on each item.

What is your major? If you have a double major, pick the one that bests represents your interests.

1. Arts/Humanities
2. Biology
3. Business
4. Chemistry
5. Economics
6. Education
7. Engineering
8. Mathematics
9. Medicine/Pre-Medicine
10. Psychology
11. Sociology
12. Criminology
13. Other

Current grade point average (please estimate if you don't know; give only one single numeric response: e.g., 3.52). If you do not yet have a grade point average, please enter 99: ____________

For each of the following three items, give one single numeric response (e.g., 26). Please estimate if you don't know exactly.

Number of credit hours earned toward the degree you are currently seeking (don't count this semester): ____________

Number of high school mathematics and/or statistics courses completed: ____________

Number of college mathematics and/or statistics courses completed (don't count this semester): ____________

In order to describe the characteristics of your class as a whole, we need your responses to the following items.

Your sex: 1. Male 2. Female

Your age (in years): _______________

The following questions ask about your peers who are majoring in psychology and the psychology faculty at this college.

Use the following scale to rate your degree of agreement or disagreement when thinking about your peers in the psychology major.

Most of **my peers who are majoring in psychology**:	Strongly disagree			Neither disagree nor agree			Strongly agree
will use statistics in their profession	1	2	3	4	5	6	7
would choose to take statistics as an elective	1	2	3	4	5	6	7
often skip statistics class	1	2	3	4	5	6	7
are capable of doing well in statistics class	1	2	3	4	5	6	7
like statistics	1	2	3	4	5	6	7
believe statistics is worthless	1	2	3	4	5	6	7
are frustrated while in statistics class	1	2	3	4	5	6	7
are interested in statistics	1	2	3	4	5	6	7
study well for statistics class	1	2	3	4	5	6	7
will learn statistics quickly	1	2	3	4	5	6	7
need a great deal of discipline to learn statistics	1	2	3	4	5	6	7
often receive high marks in statistics class	1	2	3	4	5	6	7

Use the following scale to rate your degree of agreement or disagreement when thinking about the faculty who teach in this psychology department.

The faculty **who teach in the psychology department**:							
like statistics	1	2	3	4	5	6	7
believe statistics are worthless	1	2	3	4	5	6	7
are interested in statistics	1	2	3	4	5	6	7
enjoy using statistics	1	2	3	4	5	6	7
use statistics in their profession	1	2	3	4	5	6	7
would choose to teach a statistics course	1	2	3	4	5	6	7

For the following, please rate your degree of agreement or disagreement for each statement as it pertains to your statistics professor you had (or currently have) at this college:

My **statistics professor**:	Strongly disagree			Neither disagree nor agree			Strongly agree
completes grading in a timely fashion	1	2	3	4	5	6	7
works hard to teach the course well	1	2	3	4	5	6	7
is able to communicate statistical information well to others	1	2	3	4	5	6	7
never cancels statistics class	1	2	3	4	5	6	7
when students ask a question in statistics class, my statistics professor gives a clear answer	1	2	3	4	5	6	7
is capable of using statistics well	1	2	3	4	5	6	7

Which of the following techniques were used in your statistics course at this college (please answer only if you have COMPLETED your statistics course):

____Mathematical examples

____Computers to illustrate mathematical concepts

____Calculators or spreadsheets to illustrate mathematical concepts

____Real life examples

____Activities

____Fun/Humor

____Stories

____Mnemonics

____Manipulatives

Which of the following topics were covered in your statistics course at this college (please answer only if you have COMPLETED your statistics course):

____Interpreting basic statistical results

____Distinguishing statistical significance and practical significance

____Effect size

____Confidence intervals

____Interpreting statistical results as validating conclusions made in research (i.e., test statistics supporting claims of significance or non-significance)

____using appropriate tests for different levels of measurement

Appendix B

Assessment of Statistical Skills

DIRECTIONS: The following questions ask varying questions pertain to statistical concepts. Please read each question carefully and provide your best answer. Please remember that your answers are anonymous, have no impact on your course grades and your faculty will not have access to your results.

Questions 1 – 6 are based on the data in Table 1 which represents the results of a hypothetical administration of the SAT Quantitative Test.

Table 1: Results for SAT Quantitative Test

Statistic	Value
n	100
Mean	700
Median	500
Mode	500
Standard Deviation	300

Q1: How many people were sampled?

Q2: What was the mathematical average for the SAT score?

Q3: Does it appear that the data are normally distributed?

Q4: What measure listed in the table provides information about the spread of the data?

A series of tests were run on the SAT data presented in Table 1. First, z-scores were calculated for each student to determine any outliers. An outlier was defined as having a score more than two standard deviations from the mean.

Q5: Use the following information to determine which students are outliers. Circle the correct response on the right.

Student # 1 has a z-score of 1.64	Outlier	Not an outlier
Student # 2 has a z-score of 2.35	Outlier	Not an outlier
Student # 3 has a z-score of 0	Outlier	Not an outlier
Student # 4 has a z-score of -2.21	Outlier	Not an outlier

Q6: What score on the SAT Quantitative test did Student # 3 obtain?

The next four questions are based on the following hypothetical example: A clinical psychologist was interested in testing the effects of a new treatment for anxiety. He randomly assigned 30 subjects to two groups: Group A received the treatment, which lasted four weeks; Group B was assigned to a waiting list control. A standardized test of anxiety was given to all subjects at the end of the four weeks. This test has a maximum score of 30 where a higher score indicates a greater amount of anxiety. The psychologist obtained the following data:

Table 2: Results of the Experiment on Anxiety

	Mean	Standard Deviation	Value of t-test	Value of p	Value of d
GROUP A	17.80	4.23	2.24	.033	.85
GROUP B	20.93	3.39			

Q7: Did the treatment significantly affect anxiety?

Q8: What statistic did you use to determine if the treatment affected anxiety?

Q8: Is this a meaningful difference?

Q10: What statistic did you use to determine if this is a meaningful difference?

Questions 11 and 12 are based on the following:

A 95% confidence interval is calculated for a set of weights and the resulting confidence interval is 42 to 48 pounds. Indicate whether the following two statements are true or false.

Q11:	A total of 95% of the individual weights are between 42 and 48 pounds.	True	False
Q12:	If 200 confidence intervals were generated using the same process, about 10 of the confidence intervals would not include the population mean (μ).	True	False

Questions 13 through 18 are based on the following:

Researchers at the National Institute of Health have developed a new depression scale. The test is scored on a scale of 0-50 with higher scores indicating higher levels of depression. The scale was given to a large national sample and it was determined that the mean of the test is 25 with a standard deviation of 5 (these values, therefore, are considered to be the population mean and standard deviation).

Please match the appropriate statistical test from the list below that would used to answer each research question related to the scenario above.

a. One-way between subjects ANOVA
b. One-sample t-test
c. Spearman correlation
d. Repeated measures ANOVA
e. Pearson correlation
f. Chi-square tests

Q13: A professor gives the test to his class of students and finds that the mean for this group of students is 35. Which test would he use to determine if his students are significantly more depressed than the population on which the test was normed?

____

Q14: The test was given to a sample of 15 women and 10 men. The mean for women was 24 and the mean for men was 21. Which test would he use to determine if the two means were significantly different from each other? __________________________

Q15: A teacher of statistics gives the test before and after the midterm exam in her class. Which statistical test would be used to decide if there is a significant difference between these two means?

Q16: Which test can be used to determine if there is a relationship between income (in dollars) and scores on the depression test? ___________________________

Q17: What test can be used to determine if there is a relationship between ethnicity (African American, Caucasian, Hispanic) and scores on the depression test?

____

Q18: In reviewing the scoring protocols for the test, it was discovered that some of the test takers did not complete all of the items. To analyze this, the tests were coded as "completed" or "not completed". Which test would be used to determine if a higher percentage of males completed the test as compared to females?

____

Appendix C

Table C.1 Correlation Coefficients for Regression Variables

Measure	Statistical Skill
Affect	-.04
Competence	-.17**
Value	-.26***
Difficulty	.15*
Interest	.20**
Effort	.17**
Faculty: Competence	.04
Faculty: Positive affect	-.05
Faculty: Cancel class	-.03
Faculty: Statistics worthless	-.01
Peer: Use in profession	-.01
Peer: Choose to take	-.10
Peer: Skips class	-.02
Peer: Capable of doing well	.01
Peer: Likes statistics	-.13*
Peer: Statistics worthless	.05
Peer: Frustrated in class	.06

Table C.1 (continued)

Measure	Statistical Skill
Peer: Interested	-.05
Peer: Study well	-.02
Peer: Learn quick	.08
Peer: Discipline to learn	-.02
Peer: High marks	-.02

**p*<.05, ***p*<.01, ****p*<.001

Appendix D

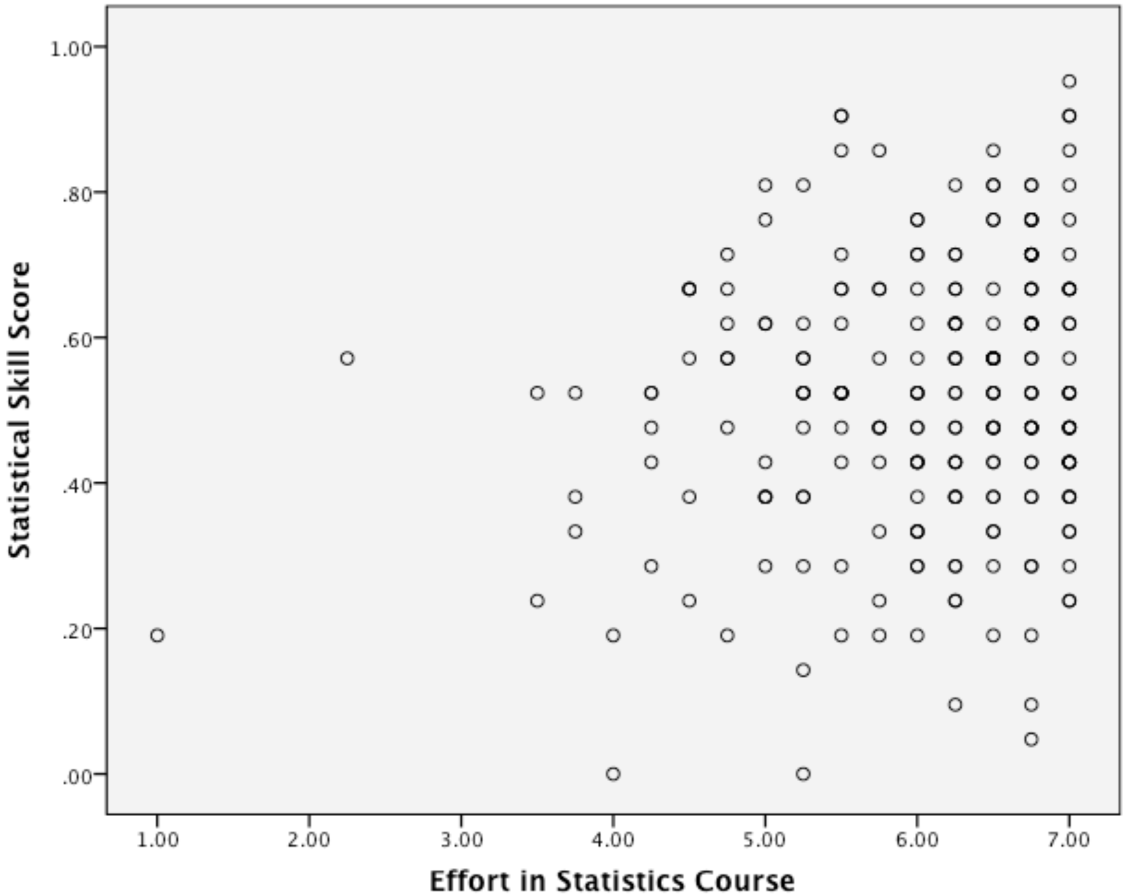

Figure D.1. Scatterplot. Relationship between statistical skill test scores and rating of effort in statistics course.

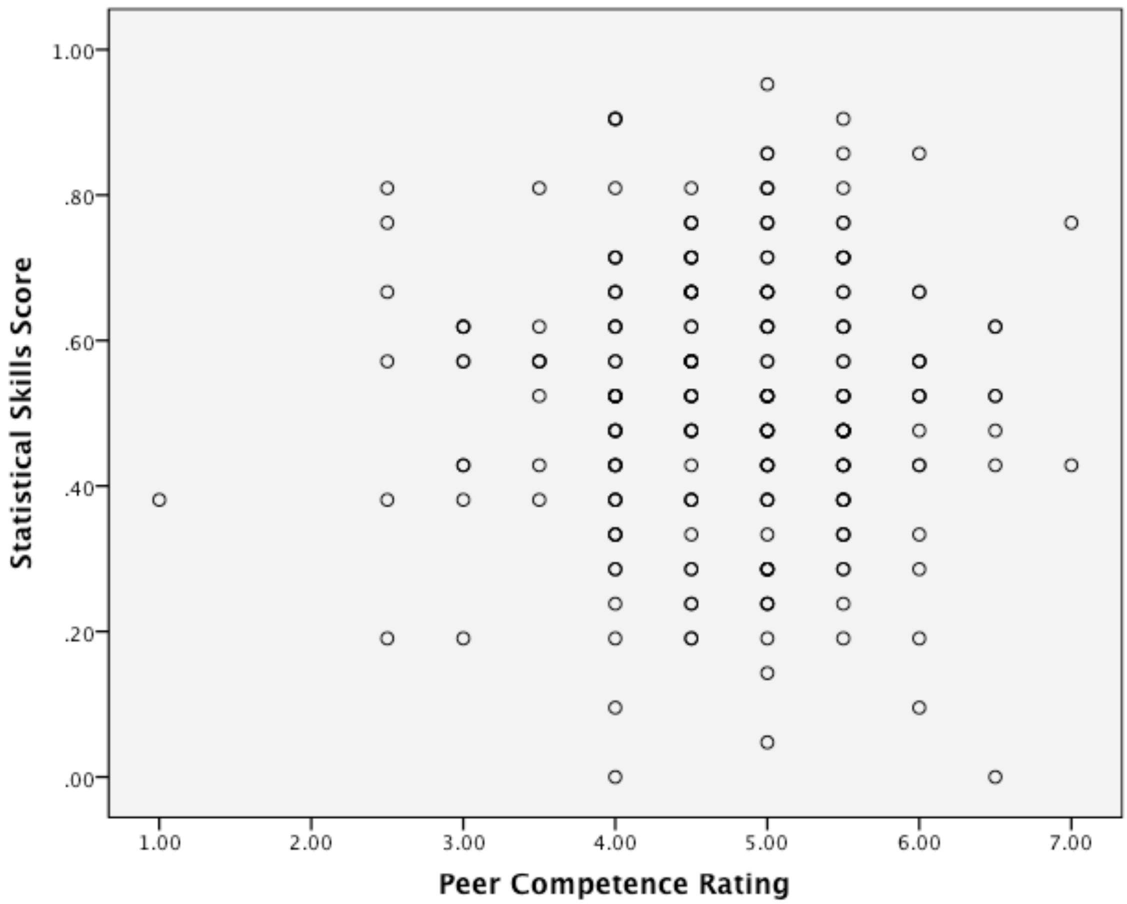

Figure D.2. Scatterplot. Relationship between statistical skills test scores and rating of peer competence with statistics.

Appendix E

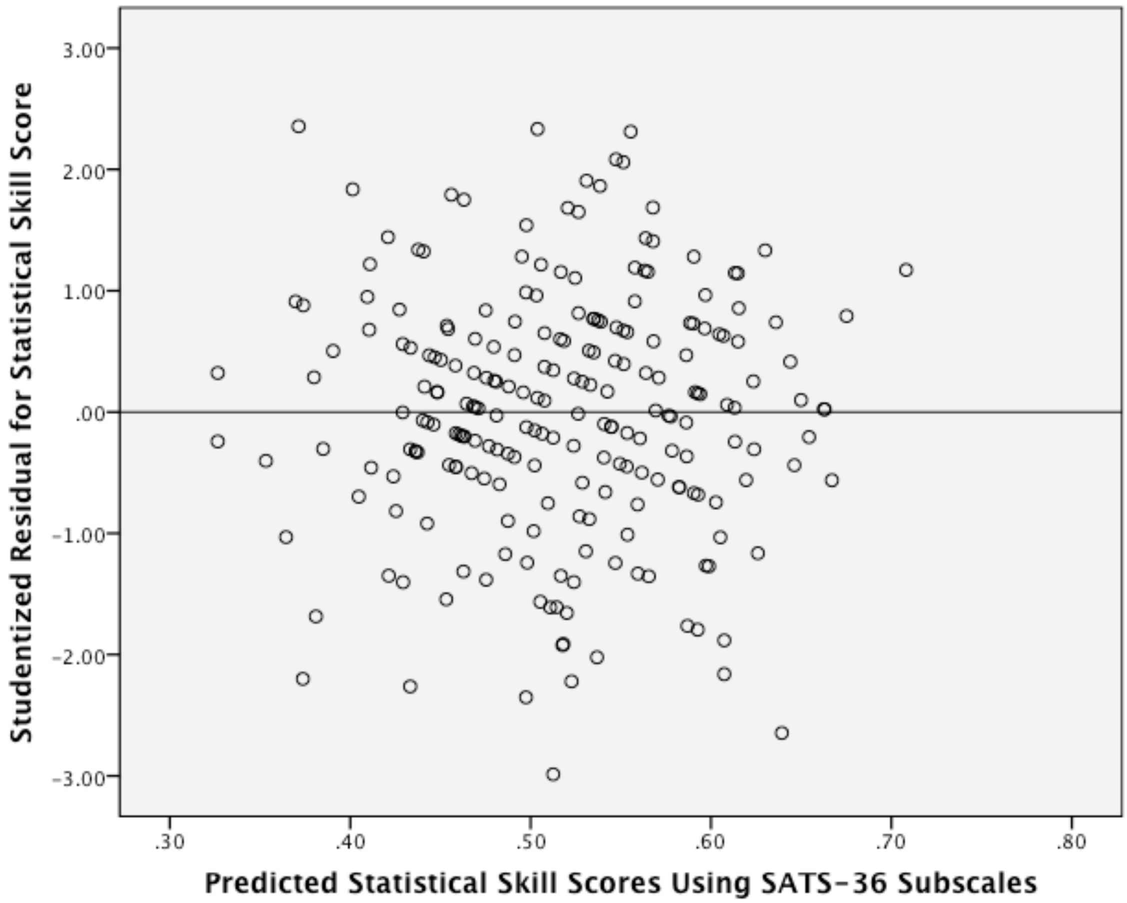

Figure E.1. Scatterplot. Relationship between predicted statistical skills test scores using original SATS-36 subscales and residuals between predicted and observed statistical skills scores.

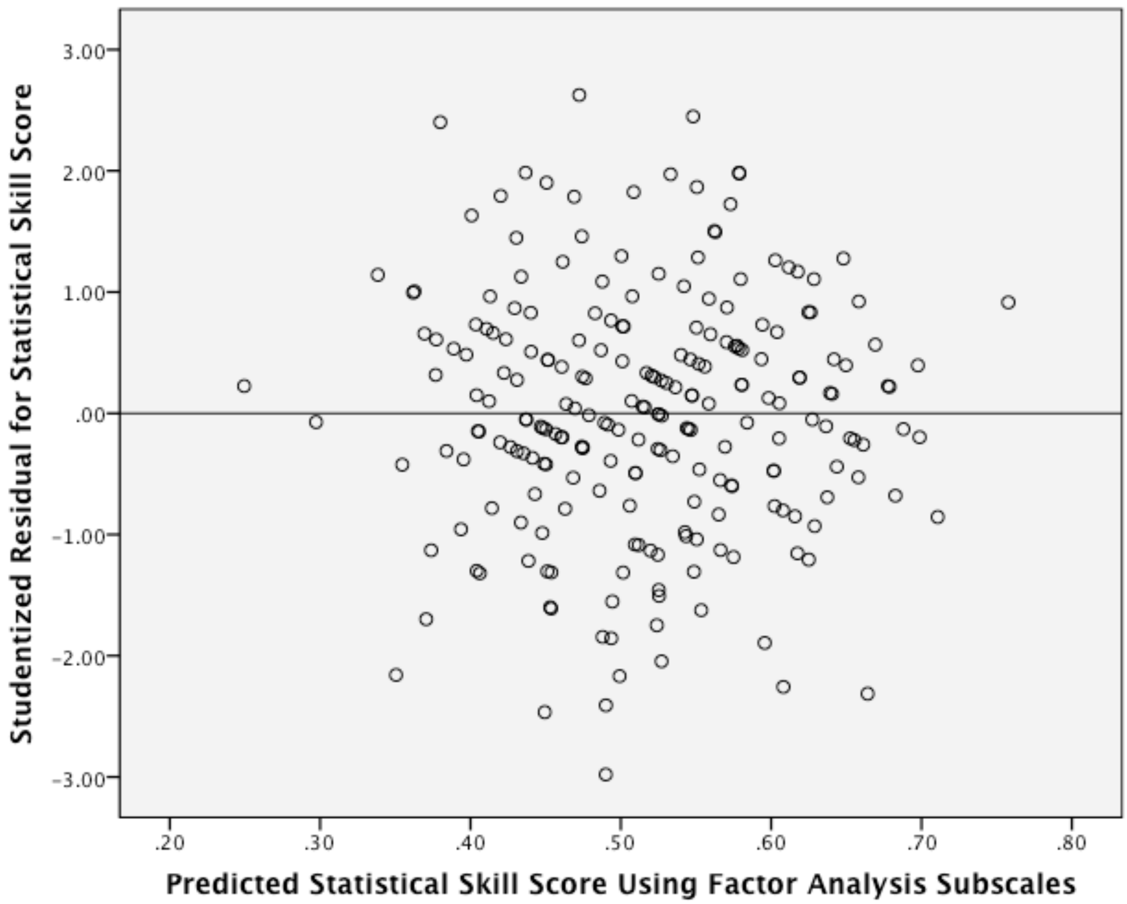

Figure E.2. Scatterplot. Relationship between predicted statistical skills test scores using factor analysis subscales and residuals between these predicted and observed scores.

Appendix F

TEMPLE UNIVERSITY

Office for Human Subjects Protections
Institutional Review Board
Medical Intervention Committees A1 & A2
Social and Behavioral Committee B
Unanticipated Problems Committee

Student Faculty Conference Center
3340 N Broad Street - Suite 304
Philadelphia, Pennsylvania 19140
Phone: (215) 707-3390
Fax: (215) 707-9100
e-mail: irb@temple.edu

Certification of Approval for a Project Involving Human Subjects

Protocol Number:	**20813**
PI:	**FARLEY, FRANK**
Review Type:	EXPEDITED
Approved On:	10-Sep-2012
Approved From:	10-Sep-2012
Approved To:	09-Sep-2013
Committee:	B BEHAVIORAL AND SOCIAL SCIENCES
School/College:	EDUCATION (1900)
Department:	EDUCATION:PSYCHOGICAL STUDIES (19040)
Project Title:	Factors related to psychology majors learning statistics

---

The IRB approved the protocol **20813**.

If the study was approved under expedited or full board review, the approval period can be found above. Otherwise, the study was deemed exempt and does not have an IRB approval period.

Before an approval period ends, you must submit a "Continuing Review Progress Report" to request continuing approval. Please submit the form **at least 60 days before the approval end date** to ensure that the renewal is reviewed and approved and the study can continue.

Finally, in conducting this research, you are required to follow the Policies and Procedures, the Investigator Manual, and other requirements found on the Temple University IRB website: http://www.temple.edu/research/regaffairs/irb/index.html

Please contact the IRB at (215) 707-3390 if you have any questions.

Millersville University
SEIZE THE OPPORTUNITY

P.O. Box 1002
Millersville, PA 17551-0302
www.millersville.edu

Department of Psychology
Phone: 717-872-3093
Fax: 717-871-2480

10/11/12

Tamarah Smith
610 King of Prussia Road
Radnor, PA 19087

Dear Ms. Smith:

Members of the Millersville University Institutional Review Board (IRB) have reviewed your proposed research, "Factors Related to Undergraduate Psychology Students learning Statistics." Dr. Kate Green and I agree that this research qualifies as minimal risk. Your proposal has been approved by an expedited review process. However, we do ask that you ensure student confidentiality by indicating that professors teaching the statistic courses will not have access to the data while teaching the course and will never have access to identifying student information (please send an email confirming this).

Approval for use of human subjects in this research is given for a period of one year from this date. If your study extends beyond 10/11/13 you must again contact the IRB for re-approval six weeks before the expiration date.

By accepting this decision, you agree to notify the Chair of (1) any additions or changes in procedures for your study that modify subjects' risk and (2) any events that affect the safety and well being of subjects.

Thank you for cooperating with our efforts to maintain compliance with federal regulations for the protection of human subjects.

Sincerely,

Karena Rush, Ph.D.
Chairperson
Millersville University Institutional Review Board

Cc: Members of the Millersville University IRB
Rita Smith-Wade-El

CABRINI COLLEGE

September 14, 2012

Ms. Tamarah Smith-Dyer
Department of Sociology
Cabrini College

Re: IRB Review

Dear Tamarah,

Thank you for forwarding a copy of your research proposal, titled *'Factors Related to Psychology Majors Learning Statistics'*, to the IRB. The submission was received on 9/12/12. Since the study has been reviewed and approved by the Temple University Institutional Review Board the review does not have to be duplicated by the IRB at Cabrini College. A copy of your study will be kept on file. Thanks for informing us that the data will be collected within the Cabrini Community. Feel free to begin collecting your data at any time.

Best of luck with your research endeavors!

Tony J. Verde

Dr. Tony J. Verde
Chair of the IRB
Cabrini College
tverde@cabrini.edu
610-902-8530

610 King of Prussia Road / Radnor, Pennsylvania 19087-3698 / 610-902-8100 / www.cabrini.edu