Summer 2014

Investigating the discrepancies between student perceptions and faculty expectations of graduate-level statistics preparation

Devon Hopkins Whetstone
James Madison University

Investigating the Discrepancies Between Student Perceptions and Faculty Expectations of Graduate-Level Statistics Preparation

Devon Hopkins Whetstone

A thesis submitted to the Graduate Faculty of

JAMES MADISON UNIVERSITY

In

Partial Fulfillment of the Requirements

for the degree of

Master of Arts

Department of Graduate Psychology

August 2014

Acknowledgements

I want to thank all the people without whom this project would never had come to fruition. First, I want to sincerely thank my advisor Kethera Fogler, who took a chance taking me on as her student at such a crucial time in my graduate school career. Without her guidance, support, and especially her patience I would not have been able to complete this thesis project. I also want to express my gratitude to my committee members Kenn Barron and Tracy Zinn, for their support and flexibility in helping me put together the best thesis possible. Additionally, I want to acknowledge Michael Hall and Robin Anderson, who were instrumental in helping me to make the best decision and transition that allowed for this project to become a reality. I am so grateful to all of you.

I am thankful for my parents and family and their continued support for my academic and career goals. I am especially grateful to have been blessed with the most patient and supportive husband I could have asked for. Marc, I love you and thank you for holding me up when I needed it most.

Table of Contents

Abstract

This study investigated the discrepancies between students' perceived preparation for graduate-level study in statistics and faculty expected levels of preparation for incoming graduate students. Two 25-item surveys on a 6 point Likert scale were developed and administered to a faculty sample and a student sample for comparison. Overall results showed that students' perceived level of preparation were equal to faculty expected levels of preparation with respect to statistical knowledge. That is, both groups endorsed the middle of the scale indicating that students generally felt somewhat prepared for graduate-level statistics and faculty generally expected students to feel somewhat prepared for their statistics course. Limitations of the study and suggestions for future research are discussed.

CHAPTER 1

Introduction

A master's or doctoral degree in Psychology is one of the most common graduate degrees awarded each year. In 2012, 13% of the doctoral degrees and 7.4% of the master's degrees received were in the Social and Behavioral Sciences (Gonzales, Allum, & Sowell, 2013). With such a large percentage of graduate degrees awarded in the social science field, it is imperative that undergraduate institutions ensure that students are adequately prepared for graduate study with respect to graduate school prerequisites.

Before discussing student learning objectives designed to adequately prepare students for graduate school, it is necessary to know what course prerequisites are required by graduate programs. McGregor (1937) published results from a survey of 41 psychologists from a variety of psychological disciplines as to what should be the necessary prerequisites for graduate study in psychology. The goal of the study was to provide some groundwork for establishing standard prerequisites, because at the time none existed. The results showed a 90% agreement across the board that training in statistics was a *necessary* prerequisite for graduate study in Psychology, and 100% of the respondents felt that statistics was at least a *desirable* prerequisite. Statistics was one of only two subjects (the other being Biology) out of nine in which there was 100% agreement in the *desirable* category (McGregor, 1937).

In the nearly 80 years since McGregor (1937) was published, undergraduate institutions have established course requirements for Psychology majors that are designed to provide students with the knowledge and skills needed to be successful in graduate school, and more recent survey reports show that the desired prerequisites have not changed much since then. Norcross, Hanych, and Terranova (1996) reported several statistics related to characteristics of graduate programs, including demographic information, application requirements, and course prerequisites. Out of 1554 graduate institutions, 56.5% *required* students to have taken a statistics course, and 28.7% *preferred* students to have taken a statistics course, totaling a combined percentage of 85.2%. Additionally, 40% *required* students to have taken a research methods course, and 26% *preferred* students to have taken a research methods course, totaling a combined percentage of 66%. Statistics and Research Methods were the two most desired prerequisite courses out of all the options listed by a large margin, as developmental psychology was the distant third most required/preferred course with a combined percentage of 35.9%. It is clear that training in statistics is paramount to graduate level training in psychological sciences. Therefore, it is essential that undergraduates with the intentions of continuing into graduate-level study be prepared in statistical knowledge and skills.

In order to set standard expectations for optimal psychology programs, a task force was created by the Board of Educational Affairs (BEA) of the American Psychological Association (APA). The purpose of the task force was to articulate the performance expectations of psychology majors at the end of undergraduate study in a document known as *The APA Guidelines for the Undergraduate Psychology Major*. This task required a rigorous assessment protocol with respect to establishing learning outcomes that were reasonable to a variety of types of programs, such as curriculums delivered online versus campus-based institutions. In the end, the task force was successful in creating a set of guidelines that could be applied across educational contexts. These guidelines are to be considered as a "living" document, that is, one that requires constant revision over time. Although the document was created in 2002, it was updated in 2013 to account for improvements regarding improving learning outcomes, and is referred to as *The Guidelines 2.0.*

The Guidelines 2.0 include five overarching goals for a psychology major, four are skills-based and one is content-based. The goals are: (1) knowledge base in psychology, (2) scientific inquiry and critical thinking, (3) ethical and social responsibility, (4) communication, (5) professional development. For the purposes of this paper, the focus will be on goals 2 and 4, as these are the goals most relevant to discussing the importance of statistical training (APA, 2013).

Goal 2 contains Outcome 2.4, which states that students should be able to "interpret, design, and conduct basic psychological research" as a result of the psychology program (APA, 2013, p. 22). In order to meet this objective, students need to be able to: "evaluate the effectiveness of a quantitative and qualitative research method in addressing a research question", "use quantitative and/or qualitative analyses to argue for or against a particular hypothesis", and "design and adopt high-quality measurement strategies that enhance reliability" (p.22). Additionally, Goal 4 contains Outcome 4.1, which states that students should be able to "demonstrate effective writing for different purposes as a result of the psychology program" (p. 30). In order to meet this outcome, students need to be able to "communicate quantitative data in statistics, graphs, and tables" (pp. 30-31). All of the previous outcomes and skills have one thing in common, and that is they are all necessary for a student to be successful in a statistics course.

It is clear that there is a consensus that statistics courses are necessary for students to take in college in order to adequately prepare for graduate-level study, and that there is an established standard for an optimal psychology program via the *APA Guidelines for the Undergraduate Psychology Major*. The next important issue is to determine if students are actually being prepared for graduate-level statistics in the psychological sciences. Also, it is necessary to determine if the statistical knowledge and skills that students come away with after graduating college is adequate with respect to what is *expected* that they know upon entering a graduate-level statistics course. That is, are students actually prepared with the same knowledge and skills that the graduate instructor expects students to possess upon entering the graduate program?

There has been little research on the adequacy of the preparation of students for graduate-level statistics. However, Jannarone (1986) described how at the University of South Carolina, the Psychology program accepted capable students who often had a weak background in statistics. The program ran into several problems upon accepting these types of students. The lack of statistics preparation led to an increase in anxiety in the students, an increase in dropout rate, and problems with getting through the material in a reasonable amount of time arose when teaching a single group of students with varying levels of statistical knowledge.

The program attempted to fix the issue by preventing the problems in the first place with the incoming students, and came up with three interventions (Jannarone, 1986). First, the program sent a letter to all incoming students at the beginning of the summer prior to their first semester. The letter outlined the statistical concepts they were expected to know coming in, and some recommended readings to help them prepare. Second, the program administered a diagnostic exam the first week of classes in order to determine which students were adequately prepared and which students were unprepared. Finally, the program offered students with knowledge deficits to take an introductory statistics course to compensate for their deficiencies prior to taking the graduate-level course required for the program. As a result of these interventions, the program observed significant improvements on the diagnostic exam, and incoming students were better prepared than previous cohorts, as demonstrated by how quickly they were able to move through the curriculum in the graduate-level course. Additionally, course grades improved; in previous years many students received a C or lower, and after the intervention no student received below a B. Jannarone (1986) is the only acknowledgment of an issue with incoming graduate students being unprepared for graduate-level statistics, as well as a suggested method to solve the issue.

Although there is extensive research and literature describing how statistics is difficult to teach and difficult to learn, there is little research that links undergraduate-level statistics deficits with graduate-level statistics expectations, or even investigates if such a discrepancy exists. It is necessary to establish if a discrepancy exists because it would be detrimental to the field of Psychology if what was observed at the University of

South Carolina is actually a widespread problem. In other words, it is imperative to know if a lack of statistical knowledge and skills creates problems for first-year graduate students in psychology that lead to undesirable outcomes, like what was observed by Jannarone (1986). Once it is established that there is a problem, then possible solutions can be implemented.

The purpose of the current study is to try and determine if there is a disconnect between what students in graduate psychology programs perceived that they learned in statistics as undergraduates, and what they were expected to know coming in as first-year graduate students. This will be accomplished by administering surveys to both graduate faculty who teach statistics and first-year psychology graduate students. The surveys attempt to gather data regarding the level of preparation first-year graduate students felt they had prior to their first graduate-level statistics course, as well as the level of preparation for graduate-level statistics that faculty expect their students to have prior to beginning their graduate studies.

It is hypothesized that there is a discrepancy between students perceptions of what they knew and what they were expected to know by their instructors. If results show that students did not feel prepared for graduate-level statistics, it will open the door to many possible additional research questions. The next steps would be to establish why students feel unprepared, as it could be because of an inadequate undergraduate curriculum, or possibly too high of an expectation of statistical knowledge from the graduate program. Another possibility is a lack of communication between undergraduate and graduate programs with respect to student learning objectives. Lack of communication could result in unrealistic expectations from graduate faculty. For example, since there are established learning objectives regarding what undergraduate psychology majors should know upon completion of their program, it is logical to assume that graduate faculty expect that the students entering their class have met those objectives. If undergraduate faculty and graduate faculty do not communicate about students meeting objectives, graduate faculty expectations may be too high. Regardless of the outcome, the results of the study would call for further research into the cause of the discrepancies.

If results show that students felt adequately prepared for graduate-level study, then it would allow for further research into what factors contributed to their preparation. For example, teaching techniques could be investigated, study habits, or even textbook use. The results will be useful and informative regardless of what they show, and therefore this study is a necessary start to expand upon the current research in this domain.

CHAPTER 2

Literature Review

Over the past few decades, statistics courses have evolved substantially with respect to content. As a result of once being taught as part of math curricula with a heavy emphasis on theory, statistics courses were taught to very few students (Chervany, Collier, Fienberg, Johnson, & Neter, 1977; Parke, 2008). Over time, statistics has expanded to other academic disciplines and now reaches a much larger and more academically diverse group of students. Due to this evolution, there has been a call for change in the way statistics courses are taught regarding concepts and methodology (Cobb, 1992; Moore, 1997; Parke, 2008). With developments in technology making statistical software accessible to students, statistical analysis has become as easy as pointing and clicking a computer mouse. Therefore, the theoretical mathematical understanding underlying statistical computation is no longer required for conducting statistical analyses. Statistics courses are now offered across disciplines and are taught at a more conceptual level (Garfield, 1995; Garfield & Ben-Zvi, 2007; Moore, 1997). As a result, there is a broader conceptualization of the construct of statistical literacy, that is, what it means for a student to know and understand statistical concepts.

Graduate programs in psychology often require incoming students to have taken at least one statistics course with the assumption that students enter the program with a certain level of statistical knowledge prior to taking a graduate-level statistics course. This assumption raises questions about the concept of statistical literacy, that is, what it means to know and understand statistical concepts. Additionally, it raises questions about how statistics are taught to students and about the assessment practices of undergraduate statistics courses. The following literature review examines the construct of statistical literacy and recommendations for assessment of students in statistics courses. Establishing how statistical literacy is conceptualized and discussing assessment with respect to student learning objectives provides a theoretical background to the present study's research question. The current study was designed to be a first step in understanding a gap in the literature with respect to discrepancies between student perceived preparation and faculty expectations for graduate-level study in statistics. Therefore, the following does not contain very much discussion directly related to the research question, as the current literature is so scarce in this area.

Statistical Literacy, Reasoning, and Thinking

There are several definitions for statistical reasoning, statistical literacy, and statistical thinking in the literature (ASA, 2005; Ben-Zvi & Garfield, 2004; Chance, 2002; Chervany et al., 1977; Delmas, 2002; Delmas et al., 2007; Gal, 2002; Garfield et al., 2002; Rumsey, 2002; Wade & Goodfellow, 2009; Wallman, 1993). Chervany et al. (1977) defined statistical reasoning as what a student is able to do with statistical content (recalling recognizing, and discriminating among statistical concepts) and the skills students show in a step by step problem solving process. The three step process the authors outlined consists of: (1) comprehension, (2) planning and execution, and (3) evaluation and interpretation. Comprehension refers to seeing a problem and classifying it as a specific type of problem. Planning and execution refers to applying the appropriate method to solve the problem. Evaluation and interpretation refers to interpreting the outcome of the problem (Chervany et al, 1977).

Garfield (2002) presents a five-level model of statistical reasoning: (1) idiosyncratic reasoning, (2) verbal reasoning, (3) transitional reasoning, (4) procedural reasoning, and (5) integrated process reasoning. Idiosyncratic reasoning refers to when the student knows some terminology related to a concept, but does not fully understand them and may even use terms incorrectly. Verbal reasoning refers to when a student has a verbal understanding of a concept, but cannot apply the knowledge to actual behavior (e.g., a student might be able to define a term, but may not understand how it integrates with other statistical concepts). Transitional reasoning refers to when a student is able to correctly identify one or two dimensions of a statistical process, but cannot fully integrate these dimensions (e.g., a student understands a relationship between two concepts, but cannot conduct the steps of a statistical process that utilizes the concepts). Procedural reasoning refers to when a student is able to correctly identify all of the dimensions of a statistical concept or process, but does not understand the process (e.g., a student can follow a step by step procedure but cannot explain the process and does not have confidence in their predictions). Integrated process reasoning refers to when a student has complete understanding of a statistical concept or process, can coordinate the rules and behavior, and can explain it in their own words confidently (Garfield et al., 2002).

Although there is no agreed upon definition of statistical literacy, there are common themes in the literature that allow for a general understanding of how these topics are conceptualized. Literacy, reasoning, and thinking are closely related cognitive processes, and the terms are sometimes used interchangeably. However, all three relate to the general understanding and use of statistics. When outlining goals for students in statistics courses, it is essential to clarify what is meant by these terms. The following describes some of the current perspectives on statistical literacy, reasoning and thinking.

Generally, *statistical literacy* refers to skills used in understanding statistical information and research results, including: data organization, tables, data management, and an understanding of statistical concepts (ASA, 2005; Ben-Zvi & Garfield, 2004; Gal, 2002; Rumsey, 2002; Wade & Goodfellow, 2009; Wallman, 1993). *Statistical reasoning* refers to the way people make sense of statistical information, including: making connections, interpretation, and explaining statistical processes (Ben-Zvi & Garfield, 2004; Garfield et al., 2002). *Statistical thinking* refers to understanding why and how statistical investigations are conducted, including understanding sampling, how inferences are made from samples to population, and how to critique and evaluate study results (ASA, 2005; Ben-Zvi & Garfield, 2004; DelMas, 2002; Chance, 2002; Garfield et al., 2002; Rumsey, 2002).

Over time, the construct of statistical literacy and the method in which statistics courses are taught have changed. These changes can be traced back to the 1970s, when a shift occurred in the material that was taught in introductory statistics courses. It was well established that a primarily mathematical and probability-based curriculum was no longer feasible or necessary for introductory statistics courses with the evolution of technology and the need for general statistics understanding across disciplines (Chervany et al., 1977; Parke, 2008). Therefore, a need to expand upon the construct of statistical literacy came about, as well as the need for a broader statistics curriculum and to develop and assess student-learning objectives that aligned with this new type of statistics course.

Mathematical Statistics and Limited Assessment

Early on in statistics education, statistical topics were taught within a mathematics course, or as a separate course but still based within a mathematical framework. That is, classes were taught with a heavy emphasis on the equations underlying statistical theory and also relied heavily on probability theory and inference (Chervany et al., 1997; Moore, 1997). Over time, with the developments in technology and statistical computing, a change occurred with respect to the way statistics was taught. In the 1970s, with technology advances in statistical software, the theoretical mathematical background of statistics became less emphasized. Statistical concepts and applied skills became more emphasized, as software was used to do most of the mathematical computations. With this change came a need to stress conceptual understanding, linking concepts, using data, interpreting results, and drawing conclusions, rather than memorization, computational skills, and procedural rules (Parke, 2008). Also, with this evolution in the way statistics was taught came the need for a new framework to assess student learning in introductory statistics courses.

Chervaney, et al., (1977) developed a framework to be used in creating instruments to measure conceptual learning and problem-solving skills related to statistical reasoning. At this time, introductory statistics was taught typically without calculus as a required prerequisite, and was often the only statistics course a student would ever take. Additionally, many fields require some ability to use and understand statistics, so an introductory course was often required across disciplines. As a result, enrollment in these courses was increased, and many resources were devoted to improving them, such as developing textbooks designed to engage students, and using software as a teaching tool (Chervany et al., 1977; Parke, 2008).

Even though a large amount of resources were devoted to the new statistics courses, there was dissatisfaction with the curriculum in the professional community (Parke, 2008). Despite the recognition of the necessary shift from a purely mathematical framework to a conceptual framework, many instructors were still teaching introductory courses based on probability and mathematics. Additionally, even in concepts-based courses, much of the focus was on the details of techniques and computations, and not on application. In other words, students were memorizing formulas and could *calculate* a standard deviation, but could not *explain* what a standard deviation is and how it is used

(Garfield, 1995; Parke, 2008). Also, students were dissatisfied with introductory statistics because they found it boring, difficult, and unexciting. Although a conceptually based construct of statistical reasoning was proposed by Chervany et al., (1977), that emphasized the importance of assessment in introductory courses, the authors did not provide any indication as to how instructors were to implement these ideas into their courses (Parke, 2008). The dissatisfaction with these courses was most likely the primary reason for devoting resources to developing materials for them. Since so much time and so many resources were dedicated to improving introductory statistics courses, one would expect to find research evaluating them. However, at that time no studies existed despite the fact that there was dissatisfaction both in the professional field, and in the classroom.

A Call for Reform in Statistics Education

In the seminal publication, *Heeding the Call for Change*, George Cobb describes the changes in the field of statistics, and also reflects upon the separation of statistics from mathematics. This separation between the two fields was recognized 15 years prior, yet little progress had been made to adjust how statistics was taught in the classroom to reflect that separation. Cobb (1992) then proceeds to provide recommendations for teaching to reflect the changes. First, echoing Chervany et al., (1977), it is recommended that teachers emphasize statistical thinking. A solid foundation in basic statistics should be a priority so that students perform better in advanced courses. He also recommends emphasizing data and concepts, and focusing less on theory and "recipes", or formulas. This recommendation aligns with utilizing a more applied approach to teaching basic statistics, in which students get immersed in data and learn concepts by working with real data sets. Also, with statistical software, students benefit from the automatic calculations and graphics, which allow them to learn concepts without being bogged down with learning the mechanics. This point is emphasized by Cobb when describing how lectures do not work as well as most instructors might think, meaning that deep level understanding of concepts does not occur as much as expected in a lecture-based class format. This leads to the final recommendation of fostering active learning. In an active learning environment, students engage in activities like group discussion, exercises, and demonstrations. The goal of utilizing an active learning environment is to engage students in learning statistics, and to prevent the feelings of dissatisfaction and boredom that had been reported by students since the 1970s (Cobb, 1992; Garfield, 1995).

In response to Cobb (1992), the 1990s was a decade of research on pedagogical change regarding statistics course content, how students learn statistics, and what assessment methods should be used to really get at student learning objectives that deal with conceptual understanding of topics, interpreting data, and evaluating quantitative information. Garfield (1995) discussed the importance of proper assessment in statistics courses. Goals for students that aligned with previous research and recommendations related to a concepts-based curriculum were outlined:

> "I believe that we really want students to gain an understanding of ideas such as the following:

a) The idea of variability of data and summary statistics.

b) Normal distributions are useful models though they are seldom perfect fits.

c) The usefulness of sample characteristics (and inference made using these measures) depends critically on how sampling is conducted.

d) A correlation between two variables does not imply cause and effect.

e) Statistics can prove very little conclusively although they may suggest things, and therefore statistical conclusions should not be blindly accepted." (p.26)

Although these goals are not exhaustive, it is important to note that none of them are rooted in complicated mathematical theory, and actually reference the knowledge of more basic foundational statistical concepts. Additionally, none of the goals reference the importance of the memorization of a formula.

Other publications echoed the points made in Garfield (1995), but also emphasized the use of technology and how it could be used to facilitate deeper level understanding of concepts. For example, Moore (1997) discussed technologies and how software, multimedia, and graphing calculators serve content and pedagogy and that new forms should be utilized to improve instruction and to facilitate learning. What is important to note about the literature published in the 1990s and early 2000s is how closely it aligns with the recommendations from previous research. In other words, nothing new was being said regarding the evolution of statistics courses, and how instruction needed to be changed to reflect student-learning objectives that reflected foundational conceptual understanding. However, all of the articles referred to the need for a change in pedagogy (Chance, 2002; Garfield, 1995; Garfield et al., 2002; Garfield, 2003; Rumsey 2002). Even though the shift from a mathematically based format to a concepts-based format was recognized over a decade earlier, there was still no observable practical change in how introductory courses were being taught.

Recent Developments in Statistics Education

The past few years have also produced a large amount of research on statistics education for introductory courses. Many new topics were covered, most notably the research on misconceptions about statistics and how they affect learning, how students learn statistics, technology and online instruction, and the role of affect in success in statistics (Garfield, 1995; Garfield & Ben-Zvi, 2007). Also, many of the previous recommendations to changes in pedagogy were revisited, such as: providing an active learning environment, using technological advances to teach concepts, and how learning outcomes should reflect foundational concepts (Garfield & Ben-Zvi, 2007; Garfield, Zieffler, Kaplan, Cobb, Chance, & Holcomb, 2012).

Additional research on including applied research projects in undergraduate courses provided evidence that active learning environments were beneficial to conceptual understanding (Forster & MacGillivray, 2010; Halvorsen, 2010; Kuiper, 2010). Additionally, using visualization techniques in the classroom were shown to be beneficial to understanding data analysis, and statistical projects integrated into introductory statistics courses increased knowledge of how statistics can be applied to a wide range of disciplines (Bowman, 2010, Wickham, 2010). These studies provided even more evidence that lecture-based formats and computation-based pedagogy is not conducive to students gaining a deep level understanding of basic statistics. Therefore, even today there is still a need for instructors to not only recognize the change in course content, but to use teaching methods that are more aligned with learning objectives regarding conceptual understanding of statistical concepts.

Aligning Goals with Methods

Teaching introductory statistics courses today using a mathematical framework when many of the students are from non-mathematical backgrounds facilitates an environment where it is difficult to learn statistical concepts. Additionally, when students are consistently lectured without any active engagement, it creates a learning environment that is boring (Chervany et al., 1977). It in turn encourages students to "get by" by choosing to memorize equations for the sake of passing exams, and does not allow for them to engage with the material and utilize applied tasks and projects to facilitate deeper level understanding of the concepts. For example, say there is a student who is consistently presented with a powerpoint slide in lectures with the equation for calculating a standard deviation. The student is consistently instructed that standard deviation is an essential concept to know, and memorizes the equation. As a result, the student is then able to plug and chug a standard deviation value. This memorization of an equation would most likely *not* align with a student-learning objective determined by the instructor (Garfield, 1995). To make matters worse, when asked on a test to calculate a standard deviation, the student would most likely answer the item correctly, and receive a higher grade on the test. There are several validity concerns at work here. First, what the instructor is teaching is *not* consistent with the concepts that students should be able to know after taking the course such as how to make appropriate use of statistical inference, and how to communicate the results of statistical analysis (ASA, 2005). With respect to pedagogy and assessment, instructors' goals are most likely to teach students how to use statistics in applied settings, interpret concepts and data, and evaluate information. However, they are actually teaching students how to plug and chug numbers into equations, and then testing them on how well they compute numbers to obtain a statistical value. Students then go on to other advanced statistics courses in which it is assumed they know say, the definition of a standard deviation and how it is used, when really all they know is how to calculate one using a formula they memorized (Garfield, 1995; Joliffe, 1991).

In order to address the serious misalignment of pedagogy and student-learning objectives, the American Statistical Association published the *Guidelines for Assessment and Instruction in Statistics Education* (GAISE) in 2005. The purpose of these guidelines was to develop a set of goals that are conducive to the increased use of technology, active learning environments, and the wider academic audience of the modern statistics course (ASA, 2005).

ASA (2005) outlines what it means to be statistically educated. It is important to note that the authors do not suggest specific topics to be covered, but general skills related to statistical literacy and the ability to think statistically. For example, it is recommended that students should: "Believe and understand why data beat anecdotes and that correlation does not imply causation, recognize sources of bias in experiments, understand how to communicate the results of an analysis and statistical inference, and how to critique media reports of statistics" (pp. 11-12). ASA (2005) also discusses actual recommendations to teachers as to how implement methods to achieve the goals in classrooms, expanding upon Cobb (1992). It provides instructors with clear and specific descriptions of issues with traditional methods by presenting examples and analogies. For example, the carpentry analogy:

> In week 1 of the carpentry (statistics) course, we learned to use various kinds of planes (summary statistics). In week 2, we learned to use different kinds of saws (graphs). Then, we learned about using hammers (confidence intervals). Later, we learned about the characteristics of different types of wood (tests). By the end of the course, we had covered many aspects of carpentry (statistics). But I wanted to learn how to build a table (collect and analyze data to answer a question) and I never learned how to do that. (p. 15)

Each recommendation is supplemented by an example, analogy, and several suggestions for instructors with the primary goal of clearly articulating and encouraging instruction that is best suited for the modern statistics course (ASA, 2005).

Beginning in the 1970s and continuing over 30 years later was the notion that statistics courses were evolving out of a traditional mathematics framework into a more concepts-based format (Chervany et al., 1977; Parke, 2008). The reasons for this evolution were due to the improvement of technology and statistical software, which allowed the more involved computations to be carried out by computers rather than by hand, as well as the demand for a more applied approach so that the material would be accessible to students across many academic domains (Moore, 1997). However, students are often still taught statistics in a computation heavy manner that is still mostly lecture-based (Garfield & Ben-Zvi, 2007). This begs the question as to what students are actually learning in introductory statistics when an active learning environment is generally not being utilized. Additionally, the consequences of failing to adhere to goals and objectives of the modern statistics course with respect to graduate-level study are unclear. In other words, if instructors are not adequately mapping their course curriculum to concepts-based student-learning objectives and instead are teaching students memorization and regurgitation, how does that affect students who go on to graduate-level study in psychology? This thesis project seeks to take the first step to answer these questions by first investigating graduate students' perceived level of preparation for graduate-level statistics against faculty expectations of graduate student knowledge in statistics. The results will begin to pave the way for future studies to investigate teaching methods that lead to the desired outcome of consistency between student preparation for graduate level statistics courses and faculty expectations of student preparedness for those courses. Additionally, results will serve as a springboard that may lead to further communication between undergraduate and graduate programs regarding psychology major student learning objectives.

CHAPTER 3

Method

Hypotheses

It is expected that there will be a discrepancy between the students' perceived preparation and faculty's expected preparation scores, as the literature does not offer any evidence to support undergraduate and graduate programs taking action to prevent a gap. Therefore, it is predicted that faculty participants will endorse the upper end of the scale more than student participants. In other words, faculty will expect their students to be more prepared than students perceive themselves to be. There are also predicted group differences within the two samples. First, it is expected that faculty with a strong quantitative background (e.g., Quantitative Psychology) will expect their students to be more prepared than faculty with a less quantitative background (e.g., Counseling Psychology), as those trained in programs with more quantitative rigor may expect the same out of those they teach. Likewise with the student survey, it is expected that students enrolled in programs with more quantitative focus will endorse higher levels of perceived preparation than those enrolled in programs with less of a quantitative focus, as students committed to a field with more quantitative rigor may have prepared themselves better for study in that field.

Descriptive analyses may reveal a possible pattern among demographic variables, but there are no expected differences between groups based on race, gender, or age. For example, results may indicate a pattern of gender differences on levels of perceived preparation amongst the student respondents. If a pattern emerges, statistical analyses will be conducted to examine the significance of those differences.

Participants

There were two samples from which data will be collected for this study. The first sample consisted of graduate students in Psychology graduate programs, and the second sample consisted of faculty who teach statistics in Psychology graduate programs. The participant pool was determined by a search of 350 Psychology programs with concentrations primarily in the foundational Psychology disciplines (e.g., Cognitive, Quantitative, Social, etc.). The search was also limited to programs offering master's and/or doctoral degrees (i.e., not solely certificate programs) and limited to programs within the United States. Because graduate study in Psychology is one of the most popular degrees sought, it was necessary to set these limitations in order to adhere to a realistic and reasonable sampling procedure.

Materials

The faculty and student surveys were created by researching statistics topics that were likely to be taught in an undergraduate-level introductory course. The statistics topics addressed in the surveys were chosen from existing course undergraduate syllabi (ToP, 2014) and introductory psychology statistics textbooks. A six-point Likert response scale ranging from 1-Very Unprepared, to 6-Very Prepared was chosen to allow respondents to gauge their level of preparation or their expected level of preparation on a spectrum that was large enough to differentiate individuals, but still reasonable in that each response option was meaningful (see Appendix A). In order to allow respondents to reflect on either their classroom experience or their expectations, a mid-point response option was not provided. Both surveys were identical with respect to item wording. However, the student survey asks respondents to reflect upon their *perceived* level of preparation for graduate-level statistics, and the faculty survey asks respondents to reflect upon their *expected* level of their students' preparation for their graduate-level statistics course. The surveys are comprised of 35 items designed to assess statistical concepts and skills, with an additional nine demographic items on the student survey and seven on the faculty survey.

Content validity and instrument development. Content validity addresses the issues as to whether items on a survey measure the desired content (Lynn, 1986). There are three steps in the developmental stage of content validation: domain identification, item generation, and instrument construct (Carmines & Zeller, 1979; DeVellis, 1991; Nunnally & Bernstein, 1994). Domain identification was achieved by a conducting literature review of introductory statistics textbooks and existing introductory statistics course syllabi. Common themes were identified and were used to create a pool of items. The next step in the developmental stage of item generation consisted of sending a short survey regarding the relevance of the topics covered to a small focus group of faculty experts for review. The sample of faculty experts consisted of instructors who were instructors of undergraduate statistics courses, and will not be included in the study sample pool. Their feedback allowed for revision of the original items for the final step in the developmental stage of content validation, instrument construction (Grant & Davis, 1997; Lynn, 1986).

The 35 item stems were designed to target reflections on higher levels of the Cognitive Dimension of the Bloom's Taxonomy hierarchy. Items primarily focused on the "*Understand*", "*Apply*", and "*Analyze*" levels. Because the "*Remember*" level focuses more on surface level learning such as memorization and recognition, items regarding recalling definitions are not included in the surveys (Krathwohl, 2002). For example, item three on the student survey reads, "*Explain* when to use the following analyses…" and then provides a list of statistical tests. Item three was designed to assess the level of preparation regarding procedural knowledge at the second tier of the Bloom's Cognitive Dimension (see Appendix B). It required students to reflect on their level of preparation with respect to going beyond regurgitating a textbook definition, and asks if they were prepared to determine when a specific analysis was appropriate to use. Therefore, the overall goal of the survey is to assess preparation for graduate-level study with respect to *deeper-level* learning (i.e., understanding, applying, and analyzing) that goes beyond memorization and regurgitation.

Procedure

The surveys were sent out electronically via Qualtrics survey software. The overall procedure involved e-mailing the faculty survey to each institution's department head asking to forward the survey to the faculty member who teaches the first graduate statistics course that students take in their program. If the faculty accepted, they were asked at the end of their survey if they were willing to send their students a similar survey. Faculty members who agreed to pass along a survey to their students received an e-mail containing the student survey to forward to them. The precise steps of the procedure are outlined below:

1. E-mail department heads the faculty survey
2. Request the faculty survey be passed along to the appropriate faculty member
3. Faculty are presented with a consent form and upon providing an electronic signature are presented with the survey
4. After completing the survey, faculty are asked if they are willing to pass along the student survey to their former students
5. Upon agreement, the faculty are sent the student survey within 48 hours
6. Faculty forward student survey to previous students

7. Students are presented with a consent form and upon providing an electronic signature are presented with the survey
8. Data are collected and stored via Qualtrics for later analysis

One reminder e-mail was sent to the department heads of each institution in the sample pool to increase participation four weeks after the initial e-mail was sent. The reminder e-mail contained the survey links and reminded department heads of the previous e-mail and request that if they have not already done so, to forward the surveys along to their faculty and students.

Data Analysis

An exploratory factor analysis (EFA) was conducted on the survey items in order to determine the underlying structure. Results of the EFA dictated how the surveys were scored, that is, they will identify any subscales that may exist or if the surveys should be scored with an overall mean. Means for each subscale were calculated and compared to determine if any discrepancies exist and to what extent they exist.

Means were also compared at the item level, as each item on the faculty survey has a corresponding item on the student survey. The differences in the means indicated the extent of the discrepancies between students and faculty responses. It provided insight as to the statistical topics and software skills that faculty expect their students to understand upon enrolling in a graduate-level course, and how prepared students actually felt with respect to those topics. Both of the data sets were examined to ensure that distributions were not bimodal, that is to determine if there were subgroups within each sample. This step in analysis is necessary so as to not calculate an artificial mean that is not a true representation of the how participants responded.

Statistical analyses were conducted to determine if there were relationships between students' graduate major and their scores, and faculty academic background and their scores for each subscale. Demographic data was also examined to determine any patterns in responses. If a pattern emerged it was analyzed statistically.

CHAPTER 4
Results

Sample Demographics

The student sample was comprised of 92 students who completed the survey in its entirety. Of the 92 respondents: 73.9% were male, 80.4% were white, 6.5% were Asian/Pacific Islander, 3.3% were African-American, 4.3% were multiracial, 3.3% were other races, and 1.1% declined to answer. Ages of the respondents ranged from 22 to 51 years. Graduate majors included: 31.5% Clinical Psychology, 19.6% Experimental Psychology, 17.4% General Psychology, 13.0% Cognitive Psychology, 6.5% School Psychology, 4.3% Social Psychology, 3.3% Counseling Psychology, 2.2% Quantitative Psychology, and 2.2% Developmental Psychology.

The faculty sample was comprised of 37 statistics instructors who completed the survey in its entirety. Of the 37 respondents: 54.1% were male, 81.1% were white, 5.4% were Asian/Pacific Islander, 5.4% were African-American, 2.7% were multiracial, and 5.4% declined to answer. Ages ranged from 30 to 69 years. The fields in which the respondents primarily teach were: 37.8% Quantitative Psychology, 13.5% Cognitive Psychology, 5.4% Developmental Psychology, 5.4% Social Psychology, 5.4% Clinical Psychology, 8.1% Experimental Psychology, 8.1% School Psychology, and 16.2 Other Psychology disciplines. The levels of statistics taught by the respondents were 37.8% Introductory, 43.2% Intermediate, and 16.2% Other levels. Those who responded "Other" were asked to describe the level of graduate statistics they taught. Some examples of answers were: “Introductory *and* Intermediate”, “Advanced”, and “I do not understand this question”.

Exploratory Factor Analysis

It is important to note that the small sample sizes of both the student and faculty samples resulted in data analysis and interpretation difficulties. It is generally accepted that an exploratory factor analysis (EFA) is a proper technique when the sample size is adequate, that is, approximately ten subjects per item (Nunally, 1978). Using this recommendation to conduct an EFA would require 350 respondents per survey, which far exceeds the 92 student respondents and 37 faculty respondents. However, since the purpose of this study is to investigate a new area of research, it is completely exploratory in nature. Therefore, an EFA was conducted on the student data and the resulting model was used to score both the student data and the faculty data in order to investigate what the factor structure of the surveys might be.

An EFA was conducted in SPSS v.20 on the student data in order to reduce data dimensions for comparison with the faculty sample. An examination of the Kaiser-Meyer Olkin measure of sampling adequacy suggested that the sample was factorable (KMO = .91) (Tabachnik & Fiddell, 2001).

Extraction and rotation method. A principal axis factoring (PAF) extraction method was used in order to both reduce the large number of variables to a smaller number of factors and to describe the relationships among the variables. An oblique rotation was used in favor of an orthogonal rotation because the student questionnaire was not designed to assess distinct factors, and therefore it is reasonable to assume that the factors are correlated.

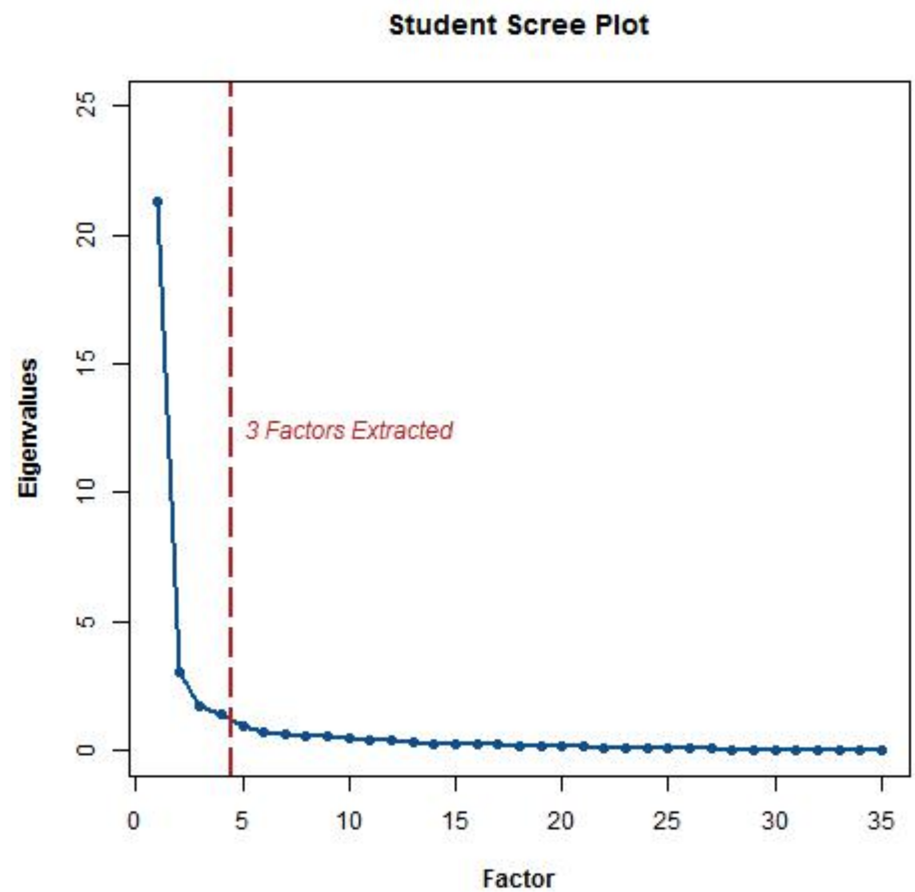
Number of factors and items removed. When SPSS was allowed to extract factors based on eigenvalues greater than 1, four factors were extracted. However, the analysis resulted in several item loadings less than .30, multiple crossloadings, and one factor with three items. A second EFA was conducted in which items were forced to load on three factors. The second analysis resulted in much cleaner results with far fewer item loadings less than .30, fewer crossloadings, and each factor has at least 5 items. Additionally, the pattern in which items that loaded on to each factor made logical sense and were able to be named. Therefore, it was determined that a three-factor structure best fit the student data. Results did not indicate that any item needed to be removed from analysis.

Factor labels. The first factor is comprised of 13 items. Using Krathwohl (2002) to classify the items, it is clear that the items relate to the "*Procedural Knowledge*" category on the Knowledge Dimension (p. 214), and the "*Understanding*" and "*Apply*" categories of the Cognitive Dimension (p. 215) of the revised Bloom's Taxonomy. All items reflect some sort of procedural step (e.g., executing analyses, interpreting analysis results) of a statistical technique. Therefore, the first factor was named "Application of Procedures". The second factor is comprised of 17 items. The items relate to the "*Conceptual Knowledge*" category of the Knowledge Dimension (p. 214), and the "*Understanding*" category of the Cognitive Dimension (p. 215). All items reflect some sort of conceptual understanding of statistics (e.g., explaining standard deviation, differentiating between descriptive and inferential statistics). Therefore, the second factor was named “Understanding of Statistical Concepts”. The third factor is comprised of 5 items. The items relate to the “*Procedural Knowledge*” category of the Knowledge Dimension (p. 214) and the “*Understanding*” category of the Cognitive Dimension (p. 215). All items reflect interpretation of figures and plots (e.g., interpret histograms). Therefore, the third factor was named “Figure Interpretation”. Table 1 provides the EFA results. Factor loadings less than |.32| were not included, as this is the recommended minimum loading for an item (Tabachnik & Fiddell, 2001). Additionally, a scree plot showing the number of factors extracted is provided in Figure 1.

Table 1. Obliquely Rotated Factor Loadings for 35 Survey Items

Item	Factor loading 1	2	3	Communality
1		0.81		0.83
2		0.88		0.88
3	-0.32	0.92		0.88
4		0.68		0.86
5		0.60		0.78
6		0.65		0.83
7		0.79		0.88
8		0.83		0.87
9		0.80		0.86
10		0.69		0.94
11		0.70		0.95
12		0.68		0.91
13		0.74		0.92
14		0.58		0.76
15		0.81		0.87
16		0.86		0.87
17	0.36	0.51		0.76
18	0.60		0.32	0.89
19	0.77			0.95
20	0.69		0.34	0.92
21	0.89			0.98
22	0.89			0.96
23	0.68			0.91
24	0.80			0.93
25	0.84			0.95
26	0.68			0.97
27	0.74			0.95
28	0.49			0.90
29	0.64			0.91
30	0.60			0.92
31			0.69	0.87
32			0.59	0.94
33			0.65	0.94
34			0.59	0.75
35	0.38		0.66	0.89
Eigenvalue	21.29	3.06	1.69	
% of Variance	60.82	8.74	4.82	

**Loadings > .32*



Descriptive Statistics

Survey scores were compared by subscale and item by item for both samples. Table 2 reports means and standard deviations, and Table 3 provides effect sizes on both the raw scale (i.e., the difference in points on the 6-point Likert scale), and Cohen's *d* estimates. Item-level comparisons were calculated on the raw scale, as seen in Table 4. The range of responses for all items on both surveys was 1 to 6, with the exception of Item 1 on the faculty survey ("How prepared do you expect your students to be with respect to explaining frequency distributions") which ranged from 2 to 6. However, the majority of both student and faculty respondents endorsed the middle responses options (i.e., 3 "Somewhat Unprepared" and 4 "Somewhat Prepared"). Additionally, responses were normally distributed for all items on both surveys.

Table 2. Descriptive Statistics for Student and Faculty Samples by Subscale

	Perceived Prep Students (*N* = 92)		Expected Prep Faculty (*N* = 37)	
Subscale	*M*	*SD*	*M*	*SD*
Application of Procedures	3.39	1.49	3.37	1.31
Understanding of Statistical Concepts	3.81	1.26	3.91	1.21
Figure Interpretation	3.83	1.47	3.42	1.24

Table 3. Effect Sizes Between Perceived and Expected Preparation by Subscale

	Raw Scale	Cohen's *d*
Application of Procedures	0.02	.01
Understanding of Statistical Concepts	0.11	.08
Figure Interpretation	0.41	.30

Table 4. Descriptive Statistics and Mean Differences by Item

	Perceived Prep Student		Expected Prep Faculty		Perceived - Expected
	M	*SD*	*M*	*SD*	Mean Difference
Item1	3.38	1.72	3.54	1.37	-0.16
Item2	3.83	1.84	3.78	1.48	0.05
Item3	3.68	1.72	3.51	1.37	0.17
Item4	3.63	1.82	3.57	1.50	0.06
Item5	3.37	1.81	3.32	1.47	0.05
Item6	2.84	1.69	3.00	1.18	-0.16
Item7	2.76	1.59	3.00	1.31	-0.24
Item8	3.70	1.82	3.72	1.49	-0.02
Item9	3.86	1.76	3.51	1.57	0.35
Item10	3.44	1.64	3.33	1.60	0.11
Item11	2.92	1.63	2.86	1.38	0.06
Item12	2.86	1.57	3.08	1.44	-0.22
Item13	3.99	1.77	3.73	1.52	0.26
Item14	4.21	1.41	4.68	1.18	-0.47
Item15	3.90	1.45	4.35	1.23	-0.45
Item16	4.46	1.32	4.49	1.27	-0.03
Item17	4.60	1.29	4.33	1.35	0.27
Item18	3.15	1.37	3.27	1.41	-0.12
Item19	3.62	1.54	3.33	1.39	0.29
Item20	3.88	1.72	4.30	1.41	-0.42
Item21	3.84	1.57	4.00	1.35	-0.16
Item22	3.37	1.44	3.56	1.42	-0.19
Item23	4.07	1.72	4.17	1.49	-0.10
Item24	3.73	1.74	3.92	1.52	-0.19
Item25	3.15	1.68	3.56	1.38	-0.41
Item26	3.08	1.71	3.54	1.50	-0.46
Item27	4.22	1.51	4.22	1.38	0.00
Item28	3.15	1.54	3.38	1.34	-0.23
Item29	4.19	1.40	3.97	1.34	0.22
Item30	4.64	1.55	4.43	1.50	0.21
Item31	4.20	1.65	4.00	1.31	0.20
Item32	3.53	1.64	3.14	1.57	0.39
Item33	3.70	1.65	3.35	1.50	0.35
Item34	3.59	1.61	2.81	1.41	0.78
Item35	4.31	1.58	3.81	1.49	0.50

Reliability Analysis

Reliability estimates were calculated for the three subscales for both the student and faculty samples. Cronbach's alpha values indicate very high reliability, as seen in Table 5.

Table 5. Reliability Estimates by Subscale

Subscale	# Items	Perceived Preparation Student Cronbach's α	Expected Preparation Faculty Cronbach's α
Application of Procedures	13	0.974	0.984
Understanding of Statistical Concepts	17	0.969	0.979
Figure Interpretation	5	0.945	0.904

Group Differences

Demographic data was examined and a pattern did not emerge to indicate that there were differences between student scores based on: race, undergraduate major, the level of their graduate statistics course (introductory or intermediate), or the students' undergraduate statistics course grades. Therefore, statistical analyses were not conducted on these variables. Independent samples t tests showed that on average, male students reported being more prepared than female students across all three subscales (Table 6 provides t-test results for all three subscales as well as Cohen's *d* effect size estimates). The faculty sample results did not indicate any group differences, predicted or otherwise.

Table 6. t-test Results of Survey Subscale Differences Between Male and Female Students

Subscale	Male (N = 24)		Female (N = 68)					
	M	*SD*	*M*	*SD*	*df*	*t*	*p*	Cohen's *d*
Application of Procedures	3.95	1.49	3.19	1.45	90	2.20	0.03	0.52
Understanding of Statistical Concepts	4.25	1.38	3.65	1.18	90	2.06	0.04	0.47
Figure Interpretation	4.41	1.53	3.63	1.4	90	2.29	0.02	0.53

Results also show that there were significant differences between students based on their graduate major on the Application of Procedures subscale and the Figure Interpretation subscale. The student sample was comprised of nine categories of Psychology majors: Quantitative, Cognitive, Developmental, Social, Clinical, Experimental, School, Counseling, and Other. Table 7 provides the one-way ANOVA results for graduate major. Post hoc tests were also conducted to determine what groups in particular showed significant differences.

Table 7. One-Way Analysis of Variance of Graduate Major and Student Survey Subscales

Variable and Source	*SS*	*MS*	$F(8, 83)$	*p*	η^2
Application of Procedures					
Between	37.80	4.72	2.39	0.02	0.19
Within	164.27	1.98			
Understanding of Statistical Concepts					
Between	16.67	2.08	1.35	0.23	0.12
Within	128.07	1.54			
Figure Interpretation					
Between	38.92	4.87	2.56	0.02	0.20
Within	157.71	1.90			

Tukey post hoc tests revealed that students whose graduate major was Cognitive Psychology ($M = 4.9$, $SD = 1.0$) had significantly higher Figure Interpretation scores than Clinical Psychology majors ($M = 3.3$, $SD = 1.5$).

CHAPTER 5

Discussion

Determining if Psychology students feel prepared for graduate level statistics to the level that their graduate instructor expects, is an important issue to consider and was the main focus of this study. Students who wish to become psychologists must attend graduate school, as a bachelor's degree is rarely enough to be a researcher or practitioner.

A major component of graduate study is the thesis or dissertation, that is, the independent research project in which a student demonstrates his or her ability to design, conduct, and report original research findings. In addition, graduate students are often asked to contribute to research teams and publish articles. These types of projects require students to have more than just a superficial understanding of statistical concepts. Research projects require students to: be able to anticipate what kind of statistical technique they will need to analyze their results, be able to execute the statistical technique properly, and make appropriate inferences from the results. Therefore, it is essential that graduate students feel prepared for graduate study with respect to their understanding of statistical concepts and skills.

During the graduate application process, several indicators of success are considered, such as undergraduate transcripts, GRE scores, and statements of purpose are evaluated to determine if a particular student is prepared for a program. It is difficult to know if a grade of "A" in any undergraduate course is an effective predictor of preparation for graduate-level study. This is nearly impossible to control for, as undergraduate courses vary by institution, undergraduate major, and instructor. However, statistics knowledge is especially important to consider because it is integral to coursework, independent research, and success as both a student and as a practitioner beyond the attainment of a degree. Even if students do not go on to a career in research, statistical knowledge is imperative to understanding published work in the field. Therefore, statistics is arguably the most important subject for graduate students to understand. Because of this issue, it is even more important to determine if incoming students are prepared for the rigor of graduate-level statistics.

The foundation of students' preparation is the undergraduate statistics course. If this class does not adequately prepare students for graduate-level statistics, when they finally reach their masters or doctoral program they may become intimidated or overwhelmed. Students may become intimidated if graduate faculty make assumptions regarding what their students know upon entering their class that are based upon misconceptions about knowledge. In other words, instructors have to make some assumptions about what their students know in order to create a course plan, and in order to create it properly they should have an accurate estimator as to how prepared their students feel for the course.

Presently, there is no existing literature that compares students' perceived preparation for graduate-level statistics with faculty expectations of preparation. The purpose of this thesis project was to open the door to investigating this research question. This first step was to establish if discrepancies actually exist between perceived preparation of students and expected preparation of faculty. It was predicted that faculty expected preparation would be higher than student perceived preparation. Differences between student degree concentration and faculty academic background were also investigated. It was predicted that students enrolled in programs with more quantitative rigor would have higher levels of perceived preparation than students enrolled in programs with less quantitative rigor. It was also predicted faculty from a quantitatively focused background (e.g., Quantitative Psychology) would have higher levels of expected preparation than faculty from a less quantitatively focused background (e.g., Counseling Psychology). The following discusses the implications of results, limitations of the study, and ideas as to how to further the research in future studies.

Implications of Results

Results from the study did not support the hypotheses overall. There were no differences between the surveys' corresponding subscales of |.42| nor a difference between corresponding items that was greater than |.78| (see Tables 3 and 4). In other words, the faculty sample's expected level of preparation was about the same as the student sample's perceived level of preparation on all statistical concepts and skills, as most items had averages indicating that most respondents answered between "Somewhat Unprepared" and "Somewhat Prepared". Additionally, the results did not entirely support predicted group differences, such as students enrolled in programs with more quantitative rigor feeling more prepared than students enrolled in programs with less quantitative rigor, or faculty with a more quantitatively focused academic background expecting higher levels of preparation than faculty with a less quantitatively focused background. However, it is important to note that Cognitive majors on the student survey scored significantly higher than another Clinical majors on the Figure Interpretation subscale, but these results were not consistent throughout the survey subscales. There were also differences in student responses with respect to gender, that is, males endorsed feeling more prepared than females, but this difference was not predicted. Males have been shown to be more self-efficacious when it comes to statistics (Hi, Myint & Chieng, 2013) and mathematical problem-solving (Parjares & Miller, 1994), so it is reasonable that males would have higher levels of perceived preparation than females with respect to graduate statistics course material, assuming that the constructs are related.

Although there were no statistically significant differences between students levels perceived preparation and faculty levels of expected preparation, the results did show that students perceived to be more prepared with respect to items that were lower in the Bloom's Taxonomy hierarchy. On average, students felt "Somewhat Prepared" on questions that were designated at the "*Conceptual Knowledge*" level on the Knowledge Dimension and "*Understanding*" level of the Cognitive Dimension (i.e., the lowest possible levels for both dimensions). However, when answering questions designated at higher levels such as the "*Applying*" level of the Cognitive Dimension and "*Conceptual Knowledge*" of the Knowledge Dimension, students endorsed "Somewhat Unprepared" more than other response options. These results indicate that students perceive themselves to be more prepared for lower level concepts and skills, and less prepared for higher level concepts and skills. However, results for faculty expectations of preparation indicate that instructors expect their students to be "Somewhat Prepared" across all dimensions with the exception of the "*Analyzing*" and "*Conceptual Knowledge*" items, which were more endorsed as expecting students to be "Somewhat Unprepared". These findings may indicate that instructors may not have higher expectations for students with respect to their knowledge on specific statistical topics, but may have higher expectations with respect to cognitive thinking and how students approach answering questions. Further investigations into perceived preparation and expected preparation of statistical concepts and skills with an intention to focus on the Bloom's Taxonomy hierarchy with items designated in more dimensions would be worth investigating in follow up studies.

In addition to gender differences, results showed that on the Figure Interpretation subscale, Cognitive Psychology majors had significantly higher scores than Clinical Psychology majors, which may partially support the hypothesis that more quantitative focused concentrations endorse higher levels of preparation than non-quantitative focused concentrations. That is, Clinical Psychology tends to be more practitioner-based than research-based, especially when comparing those studying for Doctor of Psychology (PsyD) degrees with Doctor of Philosophy (PhD) degrees, and this may be reflected in the lower level of preparedness in Figure Interpretation scores. Unfortunately, the survey did not include a demographic question regarding the type of doctoral degree students were seeking. Future research should distinguish between types of doctoral degrees in order to get a more precise description of the sample. This demographic detail is one example of a limitation of the study. There were also additional limitations related to the sample as well as the methods used to collect the data.

Study Limitations

Limitations of the samples. There were small sample sizes for both the student and faculty groups, which may have obscured actual differences between groups. In fact, the faculty sample was too small to conduct an exploratory factor analysis on the faculty survey to determine if the factor structure was parallel to the student survey results. The issues of concern with small sizes in this study were that they did not provide the necessary diversity with respect to academic background of faculty, academic concentration of students, and that they did not encompass a representative sample of the target population (i.e., graduate psychology students/faculty in the United States) in that the goal of this study was to assess the perceived levels of preparation for graduate-level statistics and how they aligned with the expected levels of preparation of faculty across a variety of domains and academic levels of Psychology graduate programs. Unfortunately, the samples were not large or diverse enough to truly investigate the research questions in such a way that the findings would be deemed generalizable.

Limitations of the materials and method. Another possible contributor to the lack of observed differences is the survey itself. The items were constructed using information found in existing syllabi and introductory textbooks, and were also assessed by content experts. However, the reliability estimates were very high in both samples (see Table 5). This may indicate that the construct that was measured was too narrow and that there were redundant items (Briggs & Cheek, 1986; Clark & Watson, 1995). The survey was intentionally constructed to be brief in order to increase participation, and items were based on Bloom's Revised Taxonomy (Krathwohl, 2002) in order to get at perceived and expected levels of preparation that reflected a deeper level of understanding than surface level learning, such as rote memorization. However, this may have presented a content validity issue in that the items were merely assessing a small facet of introductory statistics knowledge and skills, for which there may be small differences between students and faculty. The survey items will need to be revised to include a wider breadth of content in order to both better represent the construct and to observe more discernible differences. This kind of revision may lengthen the survey, but it may make responses more meaningful.

The method and medium in which the survey was distributed may have also played a role in the outcome of the results. It was anticipated that there would be a somewhat low response rate due to the lack of incentive for respondents and the impersonal nature of the request to complete it via e-mail. The survey was delivered via a mass e-mail that included a Qualtrics link, and the recipients were not given advanced warning that the survey was forthcoming. There was most likely a lack of motivation to complete the survey because recipients may have failed to see the importance or relevance, even after a reminder e-mail was sent. The reasoning behind this choice of method was to obtain a nation-wide sample from a variety of disciplines, as the target population was all graduate psychology students. This goal may have been too lofty for the scope of this project, as more practical data collection protocols were sacrificed in the process, such as giving advanced notice to programs and even mailing written requests to department heads. More responses may have been collected if the sample pool had been scaled back to a select few universities or psychology disciplines in which more in depth communication procedures regarding the surveys could have been implemented.

Finally, there was a limitation with respect to the time in which the surveys were sent to respondents. Students were instructed to complete the survey if they had already completed their first graduate-level statistics course and to reflect on their perceptions of preparation after the fact. This was purposeful in the research design, as it would have been difficult for students to reflect on their graduate-level experience with statistics before it took place. However, administering the survey after the fact may have introduced bias, despite survey instructions to think about how they felt before the course. Had the student sample been administered the survey prior to their first statistics course, they would be reflecting on their undergraduate statistics experience and their expectations about their first graduate-level course. This type of design answers a slightly different research question, but is worth investigating as a follow up study.

Results may have supported the hypothesis if the aforementioned limitations were reduced. For example, scaling back the sample pool to fewer universities as well as improving communication with instructors and department heads would mitigate the problems associated with a time delay. Also, if the survey was provided to the programs at the beginning of the academic year, faculty could request the student survey link as soon as they were finished with their course and send it to the students. This would be the best way to control for the bias resulting from a long time lapse between course completion and the survey. Additionally, it would allow researchers to compare results from semester to semester. These recommendations are some examples as to how to improve upon this study. Careful consideration of methods is essential in order to provide future studies with the means to better answer the research question at hand and delve into the important issue of preparation with respect to transitioning from undergraduate-level to graduate-level study.

Future Research

Before the research questions of this project are investigated further, it is necessary to address the problems and limitations of the data collection methods and materials and revise the protocols. The following describes the next steps to improve the survey items and the survey medium in order to best conduct future studies, and discusses some next steps to further the research.

Improving the survey content. The first step that needs to be taken is revising the survey. There were issues of content coverage and brevity that may have adversely affected the results of this study. Since Cronbach's alpha estimates greater than .9 tend to indicate redundancy of items and tend to provide an inflated sense of reliability, it would be prudent to remove any redundant questions. This would also result in fewer questions, allowing for more items to be included that would add to the breadth of the construct. In order to include more items, it would be necessary to consult another panel of content experts to both contribute to the item-writing process, and evaluate the revised survey as was done with the original version. Additionally, condensing the existing items would allow for items from other existing instruments to be included in order to gather concurrent validity evidence. Some examples include the current statistics self-efficacy (CSSE) measure (Finney & Schraw, 2003), and the Survey of Attitudes Toward Statistics (SATS) measure (Schau, Stevens, Dauphinee, & Del Vecchio, 1995). These instruments assess self-efficacy to learn statistical concepts and attitudes about learning statistics respectively, and should be investigated to determine their relationship to student perceived preparation for graduate-level study in statistics. Additionally, including a direct measure such as actual statistics problems for respondents to solve would provide an assessment of knowledge and skills that could be compared to the students' levels of perceived preparation in order to examine the relationship between perceived preparation of knowledge and understanding and actual knowledge and understanding. It is clear that there are several benefits to revising the survey items, and this step is paramount before any future research is conducted.

Improving data collection procedures. As stated previously, the goal of obtaining a sizeable, nation-wide sample that was truly representative of the population of current graduate students in psychology via an online Qualtrics survey may have been an unrealistic expectation for this project. Once the survey is revised, it will be considerably longer, take a longer amount of time to complete, and will be more cognitively taxing on the respondent. It is logical to assume that if the response rate was too low for a brief, 5-minute survey that was designed to be easy to answer, using the same methods to collect data for a longer and more in depth version would be almost futile. Therefore, it is important to consider how to scale back the sample pool with the goals to both obtain a large enough sample, and ensure that the survey is representative in the sense that respondents take it seriously and provide accurate data.

In order to meet these goals and obtain a large sample with the revised survey, much more preparation prior to sending it out will be necessary. For this project, the survey links were sent to department heads with the hopes that it would be forwarded along to instructors and students. To improve upon this method, the researcher should contact programs weeks ahead of time explaining the purpose of the study and that the survey is forthcoming and ask if the departments are interested and answer any questions. Additionally, the existing sample pool will need to be reduced because it is unreasonable to try to communicate effectively with a sample pool as large as what was used for this study ($N = 352$). Therefore, future replications of this study must include an attempt to establish a rapport with the universities of interest. That way, it will not be a surprise when the survey is sent and participating departments may be more likely to pass it along to the desired faculty and students.

Finally, the survey scale may need to be adjusted or expanded upon. A six-point Likert scale was selected in order to prevent respondents from constantly selecting the middle option and encourage them to think about their answers. However, despite the effort most of the responses selected were still the middle two options (i.e., "Somewhat Unprepared" and "Somewhat Prepared"). Including anchors that give examples of why one might select a particular response option may help participants respond in a way that is truly reflective of their perceptions and expectations. An example of an anchor for the "Somewhat Prepared" option on the student survey may be "Select this option if you recall being exposed to this topic and/or could answer basic questions based on the concept, but may not be able to answer in depth questions regarding the topic". Additionally, including a "Not Applicable" option might provide a better response option for students who were not exposed to a particular concept or skill and instructors who do not teach a particular concept or skill. Once these improvements are made to the data collection process, then the next steps of the research process can be taken.

Next steps for future research. This study was designed to take the first step in asking questions about how consistent faculty expectations were with students' feelings of preparation for graduate-level statistics. The hope of this researcher is that future studies will begin to start conversations in the higher education community about the transition from undergraduate study in Psychology to graduate programs, especially related to statistics courses. If replications of this study support the hypothesis that faculty have greater expectations regarding statistical knowledge and understanding than students actually feel they have, steps can be taken in both undergraduate and graduate programs to close the gap.

Future studies should include investigating specific academic backgrounds of students, such as what was observed between Clinical Psychology and Cognitive Psychology students in the current study. More research on the quantitative rigor of specific programs would be helpful to determine what is considered a quantitatively based program and what is considered a non-quantitatively based program. The current study makes assumptions from generalizations regarding disciplines (e.g., Cognitive Psychology is more research-oriented and thus more quantitative than Clinical Psychology). However, the differences in quantitative rigor may lie in the program and university, not the field as a whole. It would be important to first determine which programs are considered to be quantitative in nature and which are not, and then investigate differences between the two groups.

Another direction for future studies is to examine the relationship between academic background of the instructor and expected preparation scores in depth. An attempt was made in the current study to investigate this question, but the faculty sample was not diverse enough to observe any patterns. However, it is still an important issue to consider because there is the potential for a Quantitative Psychology professor who teaches introductory statistics to have much higher expectations for their students than a Counseling Psychology professor has for their students. If evidence is gathered to support this hypothesis, it would be important to ascertain the fairness of those expectations. For example, it makes sense for a Quantitative Psychology professor to have certain expectations for a student in a Quantitative program, but if the same professor is teaching introductory statistics to a class with students from a variety of academic backgrounds, then a question of fairness is presented.

It is clear that there are many directions to take within this realm of study. It is important to investigate the questions and issues that have the potential to greatly impact student learning. Many lessons were learned throughout the course of this project that will be used to inform the next stage of research. This study opened the door to a new area that had yet to be investigated, and it is the goal of this author that the research continues in order to gather evidence in which informed decisions regarding programmatic change can be made.

References

Aliaga, M., Cobb, G., Cuff, C., Garfield, J., Gould, R., Lock, R., ... Witmer, J. (2005). Guidelines for assessment and instruction in statistics education (GAISE): College report. *The American Statistical Association,* 1-61.

American Psychological Association. (2013). The APA guidelines for the undergraduate psychology major version 2.0. *American Psychological Association,* 1-65.

Assessment of the understanding of statistical concepts. (1991). In *International Conference on Teaching Statistics* (pp. 461-466). ICOTS 8.

Ben-Zvi, D., & Garfield, J. B. (2004). *The challenge of developing statistical literacy, reasoning, and thinking*. The Netherlands: Kluwer Academic.

Bowman, A. (2010). Statistical cartoons: The role of graphics in understanding statistics. In *International Conference on Teaching Statistics*. Ljubljana, Slovenia: ICOTS 8.

Briggs, S. R., & Cheek, J. M. (1986). The role of factor analysis in the development and evaluation of personality scales. *Journal of Personality, 54*(1), 106-148.

Carmines, E. G., & Zeller, R. A. (1979). *Reliability and validity assessment*. Beverly Hills, CA: Sage Publications.

Chance, B. L. (2002). Componenets of statistical thinking and implications for instruction and assessment. *Journal of Statistics Education, 10*(3). Retrieved from http://www.amstat.org/publications/jse/v10n3/chance.html

Chervany, N. L., Collier, R. O., Fienberg, S. E., Johnson, P. E., & Neter, J. (1977). A framework for the development of measurement instruments for evaluating the introductory statistics course. *The American Statistician, 31*(1), 17-23.

Clark, L. A., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment, 7*(3), 309-319.

Cobb, G. (1992). Heeding the call for change: Suggestions for curricular action. *Teaching Statistics, 22*, 3-43.

Delmas, G., Garfield, J., Ooms, A., & Chance, B. (2007). Assessing students' conceptual understanding after a first course in statistics. *Statistics Education Research Journal, 6*(2), 28-58.

Delmas, R. C. (2002). Statistical literacy, reasoning, and learning: A commentary. *Journal of Statistics Education, 10*(3).

DeVellis, R. F. (1991). *Scale development: Theory and applications*. Newbury Park, CA: Sage.

Finney, S. J., & Schraw, G. (2003). Self-efficacy beliefs in college statistics courses. *Contemporary Educational Psychology, 28*(2), 161-186.

Forster, M., & MacGillivray, H. (2010). Student discovery projects in data analysis. In *International Conference on Teaching Statistics*. Ljubljana, Slovenia: ICOTS 8.

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

Garfield, J., & Ben-Zvi, D. (2007). How students learn statistics revisited: A current review of research on teaching and learning statistics. *International Statistical Review, 75*(3), 372-396.

Garfield, J. (1995). How students learn statistics. *International Statistics Review, 63*(1), 25-34.

Garfield, J. (2002). The challenge of developing statistical reasoning. *Journal of Statistics Education, 10*(3), 58-69.

Garfield, J. (2003). Assessing statistical reasoning. *Statistics Education Research Journal, 2*(1), 22-38.

Garfield, J., Delmas, R., & Zieffler, A. (2012). Developing statistical modelers and thinkers in an introductory, tertiary-level statistics course. *ZDM Mathematics Education, 44*(7), 883-898.

Garfield, J., Hogg, B., Schau, C., & Whittinghill, D. (2002). First courses in statistical science: The status of educational reform efforts. *Journal of Statistics Education, 10*(2), 456-467.

Gonzales, L. M., Allum, J. R., & Sowell, R. S. (2013). *Graduate Enrollment and Degrees: 2002 to 2012* (pp. 1-117, Publication). Washington, DC: Council of Graduate Schools.

Provides statistics on graduate enrollment and degrees

Grant, J. S., & Davis, L. L. (1997). Selection and use of content experts for instrument development. *Research in Nursing & Health, 20*(3), 269-274.

Halvorsen, K. T. (2010). Formulating statistical questions and implementing statistics projects in an introductory applied statistics course. In *International Conference on Teaching Statistics*. Ljublhana, Slovenia: ICOTS 8.

Hii, L., Myint, K. T., & Chieng, F. (2013). Students' self-efficacy in statistics and academic achievement. *International Higher Education Teaching and Learning Conference*.

Jannarone, R. J. (1986). Preparing incoming graduate students for statistics. *Teaching of Psychology, 13*(3), 156-157.

Krathwohl, D. R. (2002). A revision of Bloom's taxonomy: An overview. *Theory into Practice, 41*(4), 212-218.

Kuiper, S. R. (2010). Incorporating a research experience into an early undergraduate statistics course. In *International Conference on Teaching Statistics*. Ljubljana, Slovenia: ICTOS 8.

Lynn, M. R. (1986). Determination and quantification Of content validity. *Nursing Research, 35*(6), 382-386.

McGregor, D. (1937). Should there be academic prerequisites for graduate work in psychology? *The Psychological Bulletin, 34*(8), 501-509.

Moore, D. S. (1997). New pedagogy and new content: The case of statistics. *International Statistics Review, 65*(2), 123-137.

Norcross, J. C., Hanych, J. M., & Terranova, R. D. (1996). Graduate study in psychology: 1992-1993. *American Psychologist, 51*(6), 631-643.

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York, NY: McGraw-Hill.

Nunnally, J. C. (1978). *Psychometric theory*. New York, NY: McGraw-Hill.

Pajares, F., & Miller, M. D. (1994). Role of self-efficacy and self-concept beliefs in mathematical problem solving: A path analysis. *Journal of Educational Psychology, 86*, 193-203.

Parke, C. S. (2008). Reasoning and communicating in the language of statistics. *Journal of Statistics Education, 16*(1), 1-24.

Rumsey, D. J. (2002). Statistical literacy as a goal for introductory statistics courses. *Journal of Statistics Education, 10*(3). Retrieved from http://www.amstat.org/publications/jse/v10n3/rumsey2.html

Schau, C., Stevens, J., Dauphinee, T. L., & Vecchio, A. D. (1995). The development and validation of the survey of attitudes toward statistics. *Educational and Psychological Measurement, 55*(5), 868-875.

Society for the Teaching of Psychology (ToP). (2014). Project syllabus. Retrieved from http://teachpsych.org/otrp/syllabi/index.php

Tabachnik, B. G., & Fiddell, L. S. (2001). *Using multivariate statistics*.

Wade, B., & Goodfellow, M. (2009). Confronting statistical literacy in the undergraduate social science curriculum. *Sociological Viewpoints,* 75-89.

Wallman, K. K. (1993). Enhancing statistical literacy: Enriching our society. *Journal of the American Statistical Association, 88*(421), 1-8.

Wickham, H. (2010). Using visualisation to teaching data analysis and programming. In *International Conference on Teaching Statistics*. Ljubljana, Slovenia: ICOTS 8.

Appendix A

Student Survey

The following survey asks that you reflect upon your perceived level of preparation prior to enrolling in your first graduate-level statistics course. Please answer openly regarding your perceived level of preparation for the following statistical concepts and skills. Thank you in advance for your participation and effort on this survey.

Prior to enrolling in your first graduate-level statistics course, please indicate your perceived level of preparation for understanding statistical concepts. How prepared did you feel with respect to...

Explaining the following concepts:

	Very Unprepared	Unprepared	Somewhat Unprepared	Somewhat Prepared	Prepared	Very Prepared
frequency distributions	❍	❍	❍	❍	❍	❍
variance	❍	❍	❍	❍	❍	❍
standard deviation	❍	❍	❍	❍	❍	❍
statistical significance	❍	❍	❍	❍	❍	❍
effect sizes	❍	❍	❍	❍	❍	❍
confidence intervals	❍	❍	❍	❍	❍	❍

Differentiating between:

	Very Unprepared	Unprepared	Somewhat Unprepared	Somewhat Prepared	Prepared	Very Prepared
descriptive statistics and inferential statistics	❍	❍	❍	❍	❍	❍
variance and standard deviation	❍	❍	❍	❍	❍	❍
standard deviation and standard error	❍	❍	❍	❍	❍	❍

Explaining when to use the following analyses:

	Very Unprepared	Unprepared	Somewhat Unprepared	Somewhat Prepared	Prepared	Very Prepared
t-tests	❍	❍	❍	❍	❍	❍
ANOVA	❍	❍	❍	❍	❍	❍
regression	❍	❍	❍	❍	❍	❍
chi-square test	❍	❍	❍	❍	❍	❍
correlation	❍	❍	❍	❍	❍	❍

Interpreting a variance value
❍ Very Unprepared
❍ Unprepared
❍ Somewhat Unprepared
❍ Somewhat Prepared
❍ Prepared
❍ Very Prepared

Interpreting a standard deviation value
❍ Very Unprepared
❍ Unprepared
❍ Somewhat Unprepared
❍ Somewhat Prepared
❍ Prepared
❍ Very Prepared

Inferring statistical significance from a p-value
❍ Very Unprepared
❍ Unprepared
❍ Somewhat Unprepared
❍ Somewhat Prepared
❍ Prepared
❍ Very Prepared

Prior to enrolling in your first graduate-level statistics course, please indicate your perceived level of preparation in your ability to apply statistical software skills (e.g., SPSS, SAS, R).How prepared did you feel with respect to...

Generating the following:

	Very Unprepared	Unprepared	Somewhat Unprepared	Somewhat Prepared	Prepared	Very Prepared
frequency distribution	❍	❍	❍	❍	❍	❍
descriptive statistics	❍	❍	❍	❍	❍	❍
figures (plots, histograms, graphs)	❍	❍	❍	❍	❍	❍

Executing the following analyses:

	Very Unprepared	Unprepared	Somewhat Unprepared	Somewhat Prepared	Prepared	Very Prepared
t-tests	❍	❍	❍	❍	❍	❍
ANOVA	❍	❍	❍	❍	❍	❍
regression	❍	❍	❍	❍	❍	❍
chi-square test	❍	❍	❍	❍	❍	❍
correlation	❍	❍	❍	❍	❍	❍

Prior to enrolling in your first graduate-level statistics course, please indicate your perceived level of preparation of your ability to interpret analysis output generated by statistical software (e.g SPSS, SAS, R) How prepared did you feel with respect to...

Interpreting results from the following analyses:

	Very Unprepared	Unprepared	Somewhat Unprepared	Somewhat Prepared	Prepared	Very Prepared
t-tests	❍	❍	❍	❍	❍	❍
ANOVA	❍	❍	❍	❍	❍	❍
regression	❍	❍	❍	❍	❍	❍
chi-square test	❍	❍	❍	❍	❍	❍
correlation	❍	❍	❍	❍	❍	❍

Interpreting the following figures:

	Very Unprepared	Unprepared	Somewhat Unprepared	Somewhat Prepared	Prepared	Very Prepared
Histograms	❍	❍	❍	❍	❍	❍
Sampling Distribution	❍	❍	❍	❍	❍	❍
Sample Distribution	❍	❍	❍	❍	❍	❍
Boxplot	❍	❍	❍	❍	❍	❍
Scatterplot	❍	❍	❍	❍	❍	❍

Thank you for your responses. Please answer the following demographic questions:

How would you classify yourself?
- ❍ Arab
- ❍ Asian/Pacific Islander
- ❍ Black
- ❍ Caucasian/White
- ❍ Latino
- ❍ Multiracial
- ❍ Prefer not to say
- ❍ Other ____________________

What is your gender?
- ❍ Male
- ❍ Female

What is your age? _________

What is the highest level of education you have completed?
- ❍ Master's degree
- ❍ Doctoral degree
- ❍ Bachelor's degree

What was your undergraduate major?
- ❍ Psychology
- ❍ Sociology
- ❍ Business
- ❍ Other ____________________

What is your graduate major?
- ❍ Quantitative Psychology
- ❍ Cognitive Psychology
- ❍ Developmental Psychology
- ❍ Social Psychology
- ❍ Clinical Psychology
- ❍ Experimental Psychology
- ❍ Other ____________________
- ❍ School Psychology
- ❍ Counseling Psychology

What level was your first graduate-level statistics course?
- ❍ Introductory
- ❍ Intermediate
- ❍ Other ____________________

How long ago did you take your undergraduate-level statistics course?
- ❍ less than 1 year
- ❍ 1-2 years
- ❍ 3-5 years
- ❍ 6+ years

What grade did you receive in your undergraduate-level statistics course?
- ❍ A
- ❍ B
- ❍ C
- ❍ D
- ❍ F
- ❍ Prefer not to say

Your responses have been recorded. Thank you for your participation and effort on this survey!

Faculty Survey

The following survey asks that you reflect upon your expectations of students who enroll in your graduate-level statistics course. Please answer openly regarding your expected level of preparation for the following statistical concepts and skills. Thank you in advance for your participation and effort on this survey.

Prior to enrolling in your graduate-level statistics course, please indicate your expected level of preparation of your students for understanding statistical concepts. How prepared do you expect your students to be with respect to...

Explaining the following concepts:

	Very Unprepared	Unprepared	Somewhat Unprepared	Somewhat Prepared	Prepared	Very Prepared
frequency distributions	❍	❍	❍	❍	❍	❍
variance	❍	❍	❍	❍	❍	❍
standard deviation	❍	❍	❍	❍	❍	❍
statistical significance	❍	❍	❍	❍	❍	❍
effect sizes	❍	❍	❍	❍	❍	❍
confidence intervals	❍	❍	❍	❍	❍	❍

Differentiating between:

	Very Unprepared	Unprepared	Somewhat Unprepared	Somewhat Prepared	Prepared	Very Prepared
descriptive statistics and inferential statistics	❍	❍	❍	❍	❍	❍
variance and standard deviation	❍	❍	❍	❍	❍	❍
standard deviation and standard error	❍	❍	❍	❍	❍	❍

Explaining when to use the following analyses:

	Very Unprepared	Unprepared	Somewhat Unprepared	Somewhat Prepared	Prepared	Very Prepared
t-tests	❍	❍	❍	❍	❍	❍
ANOVA	❍	❍	❍	❍	❍	❍
regression	❍	❍	❍	❍	❍	❍
chi-square test	❍	❍	❍	❍	❍	❍
correlation	❍	❍	❍	❍	❍	❍

Interpreting a variance value
- ❍ Very Unprepared
- ❍ Unprepared
- ❍ Somewhat Unprepared
- ❍ Somewhat Prepared
- ❍ Prepared
- ❍ Very Prepared

Interpreting a standard deviation value
- ❍ Very Unprepared
- ❍ Unprepared
- ❍ Somewhat Unprepared
- ❍ Somewhat Prepared
- ❍ Prepared
- ❍ Very Prepared

Inferring statistical significance from a p-value
- ❍ Very Unprepared
- ❍ Unprepared
- ❍ Somewhat Unprepared
- ❍ Somewhat Prepared
- ❍ Prepared
- ❍ Very Prepared

Prior to enrolling in your graduate-level statistics course, please indicate your expected level of preparation in your students' ability to apply statistical software skills (e.g., SPSS, SAS, R). How prepared do you expect your students to be with respect to...

Generating the following:

	Very Unprepared	Unprepared	Somewhat Unprepared	Somewhat Prepared	Prepared	Very Prepared
frequency distribution	❍	❍	❍	❍	❍	❍
descriptive statistics	❍	❍	❍	❍	❍	❍
figures (plots, histograms, graphs)	❍	❍	❍	❍	❍	❍

Executing the following analyses:

	Very Unprepared	Unprepared	Somewhat Unprepared	Somewhat Prepared	Prepared	Very Prepared
t-tests	❍	❍	❍	❍	❍	❍
ANOVA	❍	❍	❍	❍	❍	❍
regression	❍	❍	❍	❍	❍	❍
chi-square test	❍	❍	❍	❍	❍	❍
correlation	❍	❍	❍	❍	❍	❍

Prior to enrolling in your graduate-level statistics course, please indicate your expected level of preparation of your students' ability to interpret analysis output generated by statistical software (e.g SPSS, SAS, R). How prepared do you expect your students to be with respect to...

Interpreting results from the following analyses:

	Very Unprepared	Unprepared	Somewhat Unprepared	Somewhat Prepared	Prepared	Very Prepared
t-tests	❍	❍	❍	❍	❍	❍
ANOVA	❍	❍	❍	❍	❍	❍
regression	❍	❍	❍	❍	❍	❍
chi-square test	❍	❍	❍	❍	❍	❍
correlation	❍	❍	❍	❍	❍	❍

Interpreting the following figures:

	Very Unprepared	Unprepared	Somewhat Unprepared	Somewhat Prepared	Prepared	Very Prepared
Histograms	❍	❍	❍	❍	❍	❍
Sampling Distribution	❍	❍	❍	❍	❍	❍
Sample Distribution	❍	❍	❍	❍	❍	❍
Boxplot	❍	❍	❍	❍	❍	❍
Scatterplot	❍	❍	❍	❍	❍	❍

Thank you for your responses. Please answer the following demographic questions:

How would you classify yourself?
❍ Arab
❍ Asian/Pacific Islander
❍ Black
❍ Caucasian/White
❍ Latino
❍ Multiracial
❍ Prefer not to say
❍ Other ____________________

What is your gender?
❍ Male
❍ Female

What is your age?_________

What is the highest level of education you have completed?
❍ Master's degree
❍ Doctoral degree

What was your undergraduate major?
❍ Psychology
❍ Sociology
❍ Business
❍ Other ____________________

What was your graduate major?
- ❍ Quantitative Psychology
- ❍ Cognitive Psychology
- ❍ Developmental Psychology
- ❍ Social Psychology
- ❍ Clinical Psychology
- ❍ Experimental Psychology
- ❍ Other ____________________
- ❍ School Psychology
- ❍ Counseling Psychology

In what field/domain do you primarily teach?
- ❍ Quantitative Psychology
- ❍ Cognitive Psychology
- ❍ Developmental Psychology
- ❍ Social Psychology
- ❍ Clinical Psychology
- ❍ Experimental Psychology
- ❍ School Psychology
- ❍ Counseling Psychology
- ❍ Other ____________________

In what field/domain do you primarily conduct research?
- ❍ Quantitative Psychology
- ❍ Cognitive Psychology
- ❍ Developmental Psychology
- ❍ Social Psychology
- ❍ Clinical Psychology
- ❍ Experimental Psychology
- ❍ School Psychology
- ❍ Counseling Psychology
- ❍ N/A
- ❍ Other ____________________

In what fields/domains are the students whom you primarily teach? Select all that apply
- ❑ Quantitative Psychology
- ❑ Cognitive Psychology
- ❑ Developmental Psychology
- ❑ Social Psychology
- ❑ Clinical Psychology
- ❑ Experimental Psychology
- ❑ School Psychology
- ❑ Counseling Psychology
- ❑ Other ____________________

What level statistics course do you teach first year graduate students?
- ❍ Introductory
- ❍ Intermediate
- ❍ Other ____________________

Are you willing to send a link to a student version of this survey to your previous students?

- ❍ Yes
- ❍ No

Please provide the e-mail address to which you would like the student survey link sent. The e-mail address you provide will not be linked to your survey data.

Your responses have been recorded. Thank you for your participation and effort on this survey!

Appendix B

Bloom's Taxonomy Table
Survey Items

	Cognitive Dimension				
Knowledge Dimension	*Remember*	*Understand*	*Apply*	*Analyze*	*Evaluate*
Factual Knowledge					
Conceptual Knowledge		1, 4, 5, 6		2	
Procedural Knowledge		3, 9, 10	7, 8		
Metacognitive Knowledge					

Krathwohl, D. R. (2002)