Primary Preservice Mathematics Teachers' Sense of Efficacy, Attitudes, and Anxiety About Statistics

A Dissertation submitted
to the Graduate School
Valdosta State University

in partial fulfillment of requirements
for the degree of

DOCTOR OF EDUCATION

in Curriculum and Instruction

in the Department of Leadership, Technology, and Workforce Development
of the Dewar College of Education and Human Services

October 2024

Andrew Kent Shealy, Jr.

MS Mathematics, College of Charleston, 2011
BS Mathematics, Georgia College and State University, 2010

This dissertation, "Primary Preservice Mathematics Teachers' Sense of Efficacy, Attitudes, and Anxiety About Statistics," by Andrew K. Shealy, Jr., is approved by:

Dissertation Committee Chair

Lantry L. Brockmeier
Lantry Brockmeier, Ph.D.
Professor of Leadership, Technology, and Workforce Development

Committee Member

Taralyn Hartsell
Taralynn Hartsell, Ph.D.
Professor of Leadership, Technology, and Workforce Development

Committee Member

Dae Kim
Daesang Kim, Ph.D.
Professor of Leadership, Technology, and Workforce Development

Associate Provost for Graduate Studies and Research

Becky K. da Cruz
Becky K. da Cruz, Ph.D.
Professor of Criminal Justice

Defense Date October 16, 2024

FAIR USE

This dissertation is protected by the Copyright Laws of the United States (Public Law 94553, revised in 1976). Consistent with fair use as defined in the Copyright Laws, brief quotations from this material are allowed with proper acknowledgement. Use of the material for financial gain without the author's expressed written permission is not allowed.

DUPLICATION

I authorize the Head of Interlibrary Loan or the Head of Archives at the Odum Library at Valdosta State University to arrange for duplication of this dissertation for educational or scholarly purposes when so requested by a library user. The duplication shall be at the user's expense.

Signature ____

I refuse permission for this dissertation to be duplicated in whole or in part.

Signature ____

Abstract

The increasing prevalence of data emphasizes the importance of statistical literacy. Educational systems are charged with developing students who are statistically literate before entering higher education or the workforce. Adequate teaching and learning of statistics in K-12 education faces challenges, due to limited statistical content knowledge and the effects of noncognitive factors on teaching.

The study employed structural equation modeling to analyze the interrelationships among statistics teaching efficacy, attitudes toward statistics (subscales: affect, cognitive competence, difficulty, and value), and anxiety about statistics. Results indicated primary preservice teachers with positive attitudes toward their statistical knowledge exhibited higher confidence in teaching statistics. Additionally, low anxiety about statistics significantly predicted positive attitudes across all subscales. Mediation analysis revealed teachers' low anxiety about statistics enhanced positive attitudes toward knowledge and intellectual skill in using statistical knowledge, subsequently boosting statistics teaching efficacy.

This research highlights the role of noncognitive factors in teaching statistics and provides insights for teacher preparation programs to enhance statistical literacy education. By addressing noncognitive aspects of teacher development, teacher preparation programs can better equip preservice mathematics teachers to foster statistical understanding in their future students. Focusing on the development of noncognitive factors can help cultivate competent, confident, and enthusiastic teachers about teaching statistics, fostering a culture of statistical literacy essential in today's data-driven world.

Table of Contents

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to my wife, Morgan, and my two boys, Andrew and Luke. Their unwavering support and understanding throughout this journey were invaluable. Their love and encouragement have kept me motivated and focused, and I am truly grateful for their presence in my life.

I would also like to extend my heartfelt thanks to Dr. Lantry Brockmeier for his guidance, expertise, and mentorship. His knowledge and insights have shaped my research and pushed me to achieve my best. I am indebted to Dr. Brockmeier for his unwavering commitment to my academic and personal growth.

Furthermore, I would like to acknowledge and express my appreciation to the faculty and staff of Valdosta State University for their dedication and support. Their commitment to excellence and their willingness to go above and beyond have contributed significantly to my academic success.

Lastly, I would like to thank all my family, friends, and colleagues who have supported me throughout the paper chase. Your encouragement, feedback, and camaraderie have made this journey more enjoyable and rewarding.

I am truly grateful to all those who have played a role in my academic journey and have helped me reach this milestone. Thank you all for your support and belief in me.

Chapter I

INTRODUCTION

Data continues to become more abundant and accessible in the world. It is found in many aspects of individuals' lives, such as education, health, and economics (Bargagliotti et al., 2020). Statistics and data analytics have become leading fields and integral parts of decision-making processes for organizations and individuals (Steen, 2001). According to the projections of the U.S. Bureau of Labor Statistics (2022), jobs related to statistics are expected to grow 35% between 2020 and 2030, much faster than the average expected growth rate for all occupations. In addition, jobs related to data analytics would increase by 31% (Bureau of Labor Statistics, 2022). Data analytics and statistics are not mutually exclusive and share many characteristics (Hassani et al., 2021). Data analysts focus much on understanding and describing data and patterns, while statisticians focus much on testing statistical hypotheses and drawing conclusions from observed data. Both fields are guided by statistical thinking. Statistical thinking refers to designing and conducting studies to collect meaningful data to draw conclusions by answering research questions.

Statistical thinking goes beyond the workplace (Franklin et al., 2015). Franklin et al. (2015) stated that statistical thinking was essential for all individuals in making informed decisions based on data. Educational systems are charged with developing students to become informed citizens with a working knowledge of statistics. Bargagliotti et al. (2020) summarized the importance of teaching and learning statistics in the K-12 curriculum:

> Each morning, the newspaper and other media confront us with statistical information on topics ranging from the economy to education, from movies to sports, from food to medicine, and from public opinion to social behavior. Such information guides decisions in our personal lives and enables us to meet our responsibilities as members of a community and society. Statistical literacy is a requirement for navigating today's world. (p. 5)

The learning of statistics guides students in developing critical and scientific thinking skills (Watson, 2001). Topics in statistics are introduced to students throughout the K-12 curriculum. Starting in kindergarten, students would develop a deeper understanding of the statistical investigative cycle as they progress through grade levels. Students would learn to decide what data is needed, design experiments, collect data, analyze and interpret data, and report conclusions (Georgia Department of Education, 2021). Learning statistics throughout the K-12 curriculum aims to develop statistically literate high-school graduates (Bargagliotti et al., 2020). These graduates would know how to analyze data, interpret charts, make decisions, and assess risk as they transition to postsecondary education or the workforce (National Council of Teachers of Mathematics, 2000). To develop statistically literate high school graduates, the teachers responsible for teaching statistical content must understand the content found in the curriculum. Many of these teachers are mathematics teachers. Statistical content is found across all grade levels, starting in kindergarten.

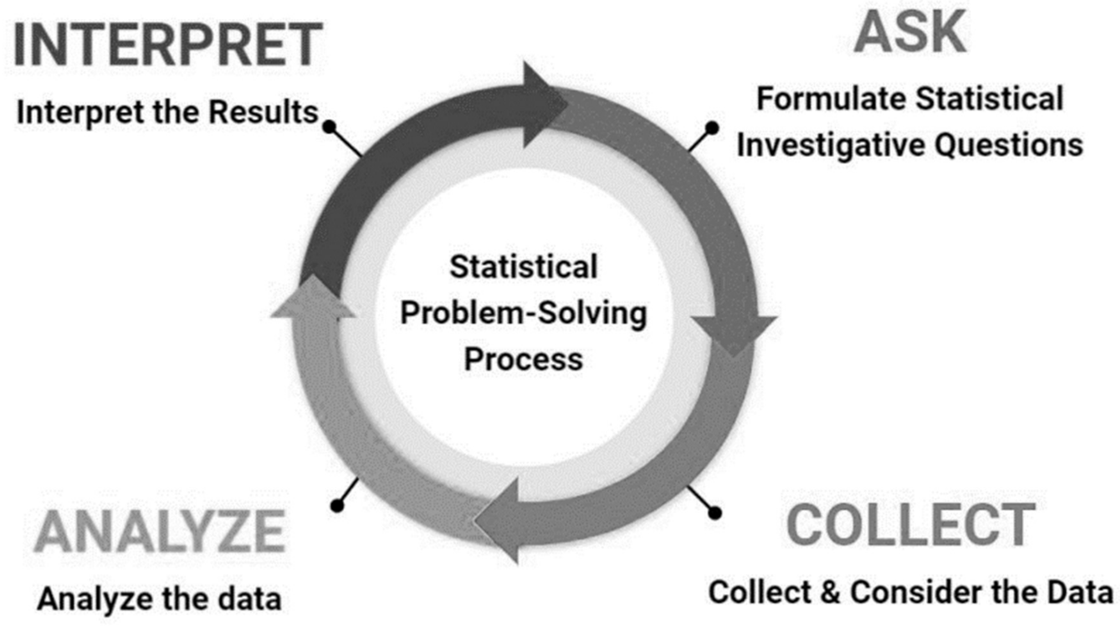
Educational systems have increased expectations in teaching and learning statistics as outlined in many K-12 mathematics curriculum standards (Schoen et al., 2019). Most states have adopted the Common Core State Standards for Mathematics (CCSSM) or adjusted their standards to align with CCSSM. The CCSSM emphasizes teaching and learning statistics across the K-12 curriculum (Capaldi, 2019). The state of Georgia has adopted a new K-12 mathematics curriculum in the 2023-2024 academic school year. Georgia's K-12 mathematics curriculum includes the Framework for Statistical Reasoning (see Figure 1). This framework would be used across all grade levels and mathematics courses to develop students' statistical literacy. Hence, teacher preparation programs should equip preservice mathematics teachers (PSMTs) with statistical content knowledge and pedagogical skills to reach various learners. Mathematics teachers are charged with linking statistics to other areas of academia by using real-life examples to engage and motivate students (Usiskin, 2015). Thus, mathematics teachers must be statistically literate (Ward-Penny, 2011).

Hannigan et al. (2013) reported that developing statistical literacy amongst students was the new norm in the mathematics curriculum. Correspondingly, Schoen et al. (2019) identified statistical content knowledge as the greater need for mathematics teachers. Many mathematics teachers have at least an undergraduate degree in mathematics, completed multiple college mathematics courses, or completed multiple mathematics education courses (Hannigan et al., 2013). However, statistics and mathematics are distinct disciplines (Zientek et al., 2011). In addition, Hannigan et al. (2013) found that mathematical content knowledge did not transfer to statistical content knowledge. Mathematics teachers across all grade levels must develop mathematical and statistical knowledge to teach mathematics curriculum effectively (Lovett & Lee, 2017b). They are required to teach statistical content that they have not experienced as students (Franklin et al., 2007).

Researchers showed that mathematics teachers lacked statistical content knowledge (Burrill & Biehler, 2011; Makar & Fielding-Wells, 2011). Makar and Confrey (2004) found that mathematics teachers taught statistics procedurally by focusing their instruction on computations of statistical measures. Similarly, Arnold (2008) found that mathematics teachers did not focus instruction on developing students' conceptual understanding of statistical ideas. Mathematics teachers struggled with sources of variation, cleaning data, re-categorizing data, sampling distributions, and confidence

Figure 1

Georgia Framework for Statistical Reasoning



Note. The Georgia Framework for Statistical Reasoning is based on the statistical investigative cycle (Georgia Department of Education, 2021). The four-step cyclic investigative cycle is used throughout each grade level and course to develop statistical literacy among students.

intervals (Arnold, 2008). In addition, Doerr and Jacob (2011) stated that mathematics teachers struggled with similar statistical content to their students. Hence, students enrolled in teacher preparation programs likely had little experience with statistics in their K-12 education (Lovett & Lee, 2017a). The National Council of Teacher Quality (2016) found that of 860 undergraduate primary (elementary) teacher education programs, only 13% required the coverage of the critical topics mathematics teachers needed. In addition, the National Council of Teacher Quality (2018) found that of 201 graduate primary teacher education programs, approximately one percent covered critical topics mathematics teachers needed. The critical topics were numbers and operations, geometry, data and probability (National Council of Teacher Quality, 2018). PSMTs lacked opportunities to develop conceptual understandings of statistical topics they were expected to teach compared to other areas of mathematics (Franklin et al., 2007).

Teachers' statistical content knowledge is essential when teaching statistics. However, researchers should consider noncognitive aspects and how they are related to a teachers’ statistical knowledge for teaching (Lancaster, 2008; Peters, 2014). These noncognitive factors represent affective and behavioral aspects of learning, such as teachers' beliefs, attitudes, and anxiety. Teachers' self-efficacy is a significant aspect of teachers' beliefs (Pajares, 1996). It refers to the teacher’s judgment in their ability to perform the duties of their profession to bring desired outcomes of student learning (Bandura, 1997; Finney & Schraw, 2003). Researchers showed that teachers with positive teaching efficacy and attitudes toward mathematics used innovative teaching strategies that positively influenced students' learning (Auzmendi, 1991; Gómez-Chacón, 2000). Teachers who lacked mathematics teaching efficacy or had negative attitudes toward mathematics used more procedure-based teaching strategies (Ross & Bruce, 2007). In addition, teachers' anxiety toward a subject had been shown to be a significant noncognitive aspect that affects the teaching and learning of the subject (Aydin, 2021). Teachers' self-efficacy, attitudes, and anxiety play a prominent role in teaching and learning. These noncognitive factors towards statistics are as important as teachers' statistical content knowledge.

There was limited research on the relationships between primary PSMTs' statistics teaching efficacy, attitudes toward statistics, and anxiety about statistics. Akin and Kurbanoglu (2011) determined the relationships between individuals' anxiety about mathematics, attitudes toward mathematics, and mathematics self-efficacy. Akin and Kurbanoglu (2011) administered a reliable 52-item survey to 372 university students to measure their sense of efficacy, attitudes, and anxiety about mathematics. The researchers found positive attitudes toward mathematics were predicted positively by self-efficacy and negative attitudes toward mathematics were predicted negatively by self-efficacy. In addition, self-efficacy significantly predicted anxiety about mathematics through the mediator attitudes toward mathematics. Since primary PSMTs are charged with teaching mathematical and statistical content, it might be expected that the results found by Akin and Kurbanoglu (2011) are transferrable to teaching statistical content.

Statement of Problem

The world continues to become more data-driven (Bargagliotti et al., 2020). Mathematics teachers are tasked with developing students' statistical content knowledge (Bargagliotti et al., 2020). Important is to understand if K-12 mathematics teachers can teach statistics effectively. Identifying if mathematics teachers can teach statistics requires understanding how teacher preparation programs develop teachers' cognitive and noncognitive factors that influence teaching and learning (Franklin et al., 2015; Hannigan et al., 2013; Hotaman, 2010). Researchers have studied teachers' statistical content knowledge extensively and found that teachers lacked the knowledge of statistics that they were required to teach (Hannigan et al., 2013; Harrell-Williams et al., 2015; Lovett & Lee, 2017b). The development of teachers' statistical content knowledge impacts the teaching and learning of statistics; however, it should not be considered the sole contributor (Groth, 2017; Hotaman, 2010; Kelly & Tomhave, 1985; Lee et al., 2020).

Studying noncognitive factors that influence teaching is important because teachers with positive beliefs, attitudes, and actions toward their instruction and content positively affect students' learning and attitudes toward the content (Gourneau, 2005; Williams, 2010; Zientek et al., 2011). Researchers found that teachers with high levels of self-efficacy, positive attitudes, and low anxiety toward mathematics positively impacted students' mathematics achievement scores (Harrell-Williams et al., 2015). Akin and Kurbanoglu (2011) stated that self-efficacy significantly predicted anxiety about mathematics and attitudes toward mathematics. Researchers found that high anxiety levels about mathematics and negative attitudes toward mathematics were negatively associated with self-efficacy, and low anxiety about mathematics and positive attitudes toward mathematics were positively associated with self-efficacy (Akin & Kurbanoglu, 2011).

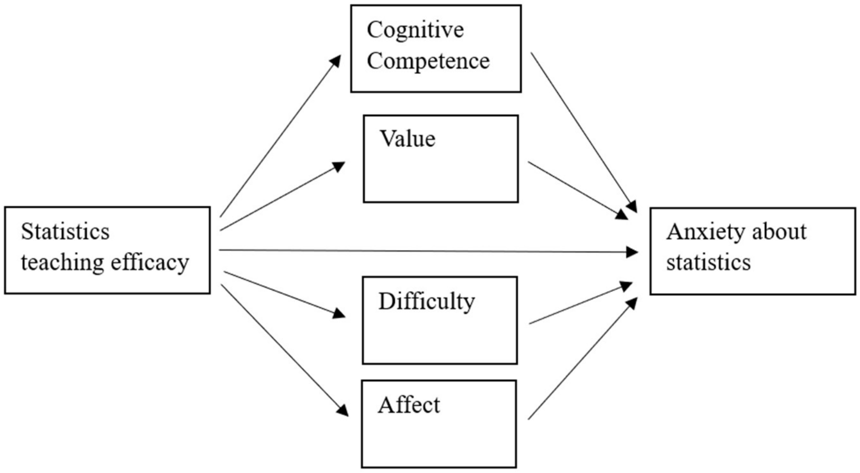
Purpose of Study

A need existed to better understand the relationship between primary PSMTs' statistics teaching efficacy, attitudes toward statistics, and anxiety about statistics to shed light on how teacher preparation programs were developing mathematics teachers' noncognitive factors related to statistics. The study's primary purpose was to examine how factors associated with statistics teaching efficacy and attitudes toward statistics impacted anxiety about statistics of primary PSMTs in Georgia. The individuals of the study were primary PSMTs, because students are introduced to statistical content starting in kindergarten. Thus, primary PSMTs are charged with developing the foundation of learning statistical content.

The study examined the overall fit of the hypothetical model (see Figure 2) and the relationships among statistics teaching efficacy and each subscale of attitudes toward statistics on anxiety about statistics. The four subscales of attitudes toward statistics measured are affect, cognitive competence, difficulty, and value. The hypothetical model was a path diagram that showed causal relationships through which the variables statistics teaching efficacy, affect, cognitive competence, difficulty, and value produced direct and indirect effects on anxiety about statistics. Since elementary PSMTs are charged with teaching mathematical and statistical content, the hypothesized model was created based on Akin and Kurbanoglu's (2011) findings. Patterns of correlations and variation of regions of the hypothetical model were examined.

Figure 2

Initial Hypothetical Path Model



Note. Affect, cognitive competence, difficulty, and value are subscales of attitudes toward statistics.

The secondary purpose of the study was to determine if the hypothetical model was a good fit across certain demographic characteristic variables. The demographic characteristic variables included gender, number of college mathematics courses completed (mathematics experience), number of college statistics courses completed (statistics experience), and number of completed years in a college teacher preparation program (college status).

Research Questions

The present study investigated primary PSMTs' statistics teaching efficacy, attitudes toward statistics, and anxiety about statistics. The study addressed the following research questions:

RQ1: Is the hypothetical path model, which describes the causal effects among PSMTs' attitudes toward statistics and statistics teaching efficacy on anxiety about statistics, consistent with the observed correlates among these variables?

RQ2: If the hypothetical path model is consistent, what are the estimated direct, indirect, and total effects among PSMTs' attitudes toward statistics and statistics teaching efficacy on anxiety about statistics?

RQ3: Is the specified path model equivalent across various demographic characteristic variables?

Research Methodology

This study employed a nonexperimental, correlational research design using structural equation modeling to test how the constructs were theoretically linked and the directionality of the significant relationships. Data collection and analysis began immediately after the Institutional Review Board (IRB) granted permission (see Appendix A). Permission must be granted from the IRB at Valdosta State University and from participating schools. Missing data were handled, and descriptive statistics were generated to provide information about the variables and highlight potential relationships between the variables. Correlations among demographic characteristic variables, statistics teaching efficacy, the four subscales that measure attitudes toward statistics, and anxiety about statistics were analyzed. Path analysis was used to examine the relationship between the exogenous variable, potential mediator variables, and the endogenous variable. The exogenous variable of interest was statistics teaching efficacy, and the four subscales measuring attitudes toward statistics were the potential mediator variables. These variables were investigated to determine the underlying structure of the endogenous variable anxiety about statistics. Finally, a multigroup analysis was conducted to determine if the model was a good fit across the demographic characteristic variables.

The target population for this research study was primary PSMTs enrolled in a primary teacher education program. The accessible population was primary PSMTs enrolled in a primary teacher education program at a 4-year accredited non-profit university or college in Georgia. Eligible primary PSMTs from such programs were recruited to participate voluntarily. Participation was confidential.

An email was sent to the primary teacher education program directors asking for the contact information of faculty members who teach primary education courses. The email explained the purpose and scope of the study, the requirements for participation, and included a written statement regarding the approval of the study by the IRB at Valdosta State University (see Appendix B). Faculty who taught primary education courses was sent an email explaining the study's purpose and scope, and the participation requirements. A follow-up phone call was made to each faculty member who agreed to participate to discuss when and how to disperse the survey packet.

Data analysis began immediately upon submission of all completed surveys. The open-source statistical programming language R was utilized for quantitative data analysis, including descriptive statistics and correlational analysis. An exploratory analysis began the process to summarize and visualize the data and identify missing values from the dataset. The data was cleaned and prepared for further analysis based on the exploratory data analysis. R software programs were employed to statistically analyze the fit of the observed data with the proposed structural equation model. The model in Figure 2 illustrates the proposed relationship between anxiety about statistics as determined by statistics teaching efficacy and attitudes toward statistics. Descriptive statistics were calculated from the self-reported factors to provide a sample profile. The self-reported factors were used in the analysis to identify factors that influenced statistics teaching efficacy, attitudes toward statistics, and anxiety about statistics.

Significance of Study

Teaching and learning statistics are in the mathematics curriculum standards (Hannigan et al., 2013; Lovett & Lee, 2017a; Schoen et al., 2019). Mathematics teachers are tasked with developing students' mathematical and statistical literacy. Students start learning about statistical reasoning in kindergarten by looking at measurements and data. They continue the development of statistical reasoning throughout their primary and secondary schooling. Researchers have studied PSMTs' statistical content knowledge, but research on elements of the affective domain that contribute to PSMTs' statistical knowledge for teaching is lacking. This study filled that knowledge gap on primary PSMT's attitudes toward statistics, statistics teaching efficacy, and anxiety about statistics.

The study aimed at benefiting primary PSMTs by shedding light on how teacher preparation programs are preparing primary PSMTs to teach statistics. Teacher preparation programs include colleges of education and Regional Education Service Agencies (RESA) centers. The knowledge obtained from this study would enable teacher preparation programs to understand better the elements of the affective domain that affect teachers' statistical knowledge for teaching. This knowledge may be used to influence training on the elements of the affective domain to develop statistical knowledge for teaching. Additionally, the study has the potential to benefit current mathematics teachers' statistical knowledge for teaching as the study may shed light on how to train in-service mathematics teachers through professional development opportunities.

Previous studies on PSMTs' statistical knowledge for teaching focused on elements: statistical content knowledge, statistics teaching efficacy, attitudes towards statistics, and anxiety about statistics (De Vetten et al., 2018; Harrell-Williams et al., 2015; Zientek et al., 2011). Some research determined the relationship between statistical content knowledge and statistics teaching efficacy (Harrell-Williams et al., 2013; Lee et al., 2020; Lovett and Lee, 2017a). Other research studies were descriptive in nature when studying PSMTs' anxiety about statistics and attitudes toward statistics (Estrada & Batanero, 2008; Estrada et al., 2011; Keeley et al., 2008; Yusuf et al., 2019). This study examined the relationship between PSMTs' statistics teaching efficacy, attitudes toward statistics, and anxiety about statistics using structural equation modeling. The findings contributed to the research on teachers' statistical knowledge for teaching.

Theoretical Framework

The leading theory used for this study is Groth's (2013) statistical knowledge for teaching (SKT) framework. Groth's (2013) SKT framework is a developmental structure for statistical knowledge for teaching. The foundation of the SKT framework is developing key developmental understandings and subject matter knowledge of statistics. Groth (2013) argued that teachers needed to lay the foundation before developing pedagogical skills to teach statistics. Lovett and Lee (2017b) stated that Groth's (2013) SKT framework explained the skills and knowledge teachers needed to effectively instruct statistics based on the investigative cycle (Groth, 2013).

The foundation of Groth's (2013) SKT framework includes two components: statistical content knowledge and pedagogical content knowledge. Statistical content knowledge refers to the knowledge teachers need to teach statistics at their grade level.

Researchers determined that mathematics teachers lacked the statistical content knowledge to teach the required statistical content at their grade levels (Hannigan et al., 2013; Harrell-Williams et al., 2015; Lovett & Lee, 2017b).

Pedagogical content knowledge refers to teachers applying their pedagogical knowledge to their knowledge of what they are trying to teach. Shulman (1987) identified pedagogical content knowledge as the most important aspect of teaching. Shulman (1987) and Garritz (2000) agreed that pedagogical content knowledge comprised noncognitive factors. These noncognitive factors are behavioral aspects and elements of the affective domain. The factors that will be measured in this study are statistics teaching efficacy, attitudes toward statistics, and anxiety about statistics. Hence, Groth's (2013) SKT framework ties into the research questions of this study.

Limitations

Limitations highlight potential weaknesses in the study and can affect the generalizability of a study (Creswell, 2009). A significant limitation of this study was voluntary participation. Primary PSMTs could choose not to participate, or they may have not felt encouraged to answer honestly. Participation from primary PSMTs depended on faculty members allowing the survey to take place, which was dependent upon the program director providing the faculty member's contact information.

The study was limited to primary education preservice teachers who teach mathematics in Georgia. The findings are generalizable to this population, and the ability to generalize results based on a nonrandom sample to the larger population of primary education preservice teachers is limited.

This study measured the relationship of statistics teaching efficacy and each subscale of attitudes toward statistics on anxiety about statistics. Correlational research can determine an association among variables but cannot determine a causal relationship, therefore interpretation is a potential limitation. Also, it was assumed that all participants in the study answered the survey questions truthfully. This is a limitation as participants self-reported.

Definition of Terms

For this study, the conceptual and/or operational definitions are included. These terms are used throughout the dissertation and knowing what they mean can help in the comprehension of any discussion:

Preservice teacher – student in a teacher education program at an institution pursuing a teaching license.

In-service teacher – a teacher currently employed in grades K-12 that holds a teaching license.

Primary school – an accredited educational system for grades K-5.

Secondary school - an accredited educational system for grades 6-12.

Postsecondary school – an accredited educational system for students who completed high school that leads to a degree or certification.

Preservice mathematics teacher (PSMT) – student in a mathematics teacher education program at an institution pursuing a teaching license. PSMTs may be called primary PSMT or secondary PSMT depending on the grade level they teach (primary school or secondary school).

In-service mathematics teacher (ISMT) – a mathematics teacher who is currently employed. ISMTs may be referred to as primary ISMT or secondary ISMT depending on the grade level they teach (primary school or secondary school).

Statistical content knowledge – knowledge of statistics across three levels of development and four phases of the statistical investigative cycle.

Pedagogical statistical content knowledge – teacher's knowledge of potential student difficulties with statistics, developing strategies to support student's learning, teaching strategies to engage students in the statistical investigative cycle, and vertical and horizontal knowledge of the statistics and mathematics curricula (Ball et al., 2008; Groth, 2013).

Statistical knowledge for teaching – statistical knowledge and pedagogical statistical knowledge teachers need to teach statistics based on the statistical knowledge for teaching framework (Groth, 2007).

Statistics teaching efficacy – a teacher's belief in his/her ability to teach statistics to bring about student learning measured by *reading the data* and *reading between the data*.

Reading the data – a teacher's self-efficacy to teach statistics found in Level A of the GAISE report.

Reading between the data – a teacher's self-efficacy to teach statistics found in Level B of the GAISE report.

Anxiety about statistics – a teacher's anxiety about the field statistics.

Attitudes toward statistics – a teacher's attitudes towards the field of statistics measured by *Affect*, *Cognitive Competence*, *Difficulty*, and *Value*.

Affect – a teacher's expression towards statistics.

Cognitive competence – a teacher's attitude toward the knowledge and intellectual skill in using statistics knowledge.

Difficulty – a teacher's attitude toward difficulties in understanding statistical content.

Value – a teacher's attitude towards a statistics course.

Endogenous variable – a variable that is caused by one or more variables in the model.

Exogenous variable – a variable not caused by another variable in the model.

Structural model – the set of structural equations.

Path diagram – a diagram that pictorially represents a structural equation model.

Structural Coefficient – a measure of the amount of change in the effect variable expected given a one unit change in the causal variable and no change in any other variable.

Organization of the Study

Chapter 1 described the introduction and importance of understanding the factors of primary PSMTs' SKT. The factors included statistics teaching efficacy, attitudes toward statistics, and anxiety about statistics. This chapter contained the research questions, methodology, significance of the study, and theoretical framework. In Chapter 2, a literature review is presented on PSMTs' SKT. The chapter explains the SKT frameworks and the components that make up the frameworks important for teachers to possess to teach statistics effectively. The chapter discusses research on factors that influence PSMTs' SKT, which includes statistical content knowledge, statistics teaching efficacy, attitudes toward statistics, and anxiety about statistics. Chapter 3 presents the study's methodology, including an identification and description of the survey participants, survey instruments, and the process and procedures for collecting and analyzing data that helped answer the research questions. Chapter 4 includes an analysis of the data collected in terms of the research questions. Chapter 5 presents a discussion of the results, conclusions, recommendations, and implications for further research.

Chapter II

LITERATURE REVIEW

The review of literature provides an examination of factors that influence preservice mathematics teachers (PSMTs') and in-service mathematics teachers' (ISMTs') statistical knowledge for teaching. The chapter begins with the discussion of knowledge for teaching frameworks, which includes the mathematical knowledge for teaching (MKT) and statistical knowledge for teaching (SKT) frameworks. Finally, the review of literature concludes with describing studies that focus on teachers' statistical content knowledge, statistics teaching efficacy, attitudes toward statistics, and anxiety about statistics.

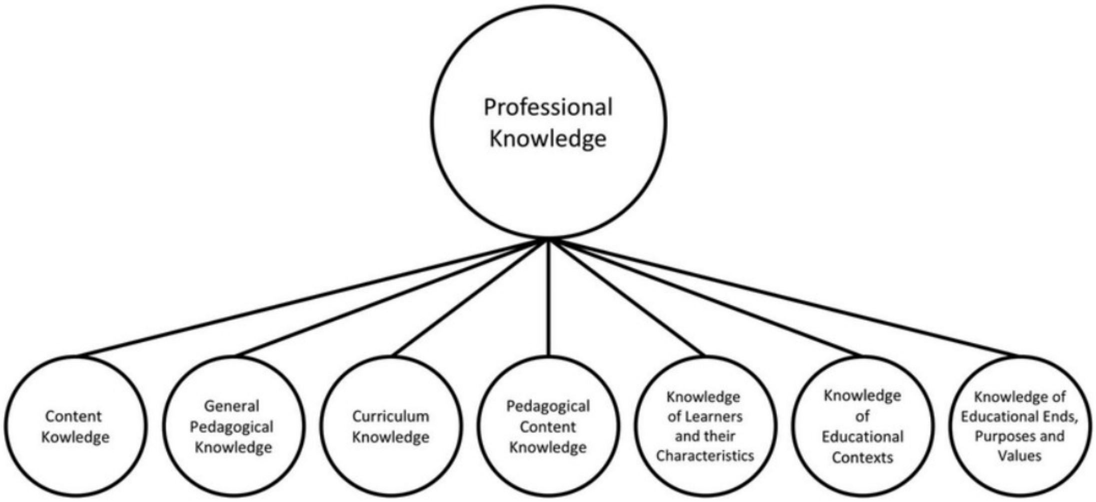
Knowledge for Teaching

A teacher cannot teach what they cannot understand. Teachers need knowledge of the content they are required to teach. Shulman (1986) identified teacher's content knowledge important but included a new way to think about knowledge for teaching. Shulman (1987) classified teacher knowledge into seven components: (a) content knowledge; (b) general pedagogical knowledge; (c) curriculum knowledge; (d) pedagogical content knowledge; (e) knowledge of learners and their characteristics; (f) knowledge of educational contexts; and (g) knowledge of educational ends, purposes, and values (see Figure 3). Content knowledge included knowledge of the content and its' systemizing structures (Shulman, 1987). It was considered the foundation of teacher knowledge; however, the most significant component was pedagogical content knowledge (Burgess, 2006). Shulman called pedagogical content knowledge the missing standard. Burgess (2008) stated Shulman's framework described pedagogical content knowledge "as one of three categories of content knowledge and the one that differentiates the teacher from the mathematician" (p. 1).

Pedagogical content knowledge is how teachers relate their teaching knowledge to their content knowledge (Shulman, 1986). Groth (2017) defined pedagogical content knowledge as "the knowledge teachers need in order to make subject matter comprehensible to students" (p. 376). This knowledge is unique to teachers (Burgess, 2006; Groth, 2017). Pedagogical content knowledge combined with subject matter knowledge and curricular knowledge comprises Shulman's (1986) knowledge for teaching framework. *Content knowledge* refers to teachers' knowledge of the content and

Figure 3

Shulman (1987) Teaching Knowledge Framework



Note. From "Probing the amalgam: the relationship between science teachers' content, pedagogical and pedagogical content knowledge" by K. Neumann, V. Kind and U. Harms, 2017, *International Journal of Science Education* p. 3. Copyright 2019 by Informa UK Limited.

how well they organize that knowledge (Groth, 2017). *Curricular knowledge* refers to the scope and sequence of content (Burgess, 2006). Researchers must look past content knowledge to pedagogical content knowledge (Burgess, 2006; Groth, 2007). Baumert et al. (2010) stated that teachers need content knowledge and pedagogical content knowledge to provide quality instruction to improve students' achievement. In addition, Shulman (2007) identified the affective domain of teaching as a component of pedagogical content knowledge by stating that an effective teacher understands the content matter and holds genuine affection for the subject. The affective domain consists of factors that play a crucial role in teaching and learning success; these factors include attitudes and anxiety towards the subject (Fishbein & Ajzen, 1975; Tasnimi, 2009). Epstein and Hundert (2002) stated that someone's professional aptitude included emotions and values, which were part of the affective domain. A teacher who does not have the affective domain of teaching will never teach the subject effectively (Shulman, 2007). Shulman's knowledge for teaching framework was foundational for teachers' knowledge of mathematics frameworks (Burgess, 2006; Groth, 2007).

Mathematical Knowledge for Teaching

Teachers' knowledge of mathematics has been a hot topic over the past 50 years (Lovett & Lee, 2017a). Over the years, researchers have attempted to understand teachers' mathematical knowledge for teaching based on content knowledge and the number of courses completed at the university level. For example, Hill et al. (2005) found that students' mathematics achievement scores were positively associated with teachers' mathematical content knowledge. Through probability sampling, the researchers recruited graduate students (*N* = 2963) and in-service teachers (*N* = 699) from 115 primary schools. Students' mathematical achievement was measured using the CTB/McGraw-Hill's Terra Nova Complete Battery and Basic Battery instrument. Teachers' mathematical content knowledge was measured using a questionnaire of 30 items covering three topics: number concepts (13 items), operations (13 items), and prealgebra (4 items). The researchers found that students' mathematical achievement scores were positively associated with teachers' mathematical content knowledge ($r = .39$, $p < .001$). In addition, results showed that the average number of mathematical content courses teachers completed in their preservice training was positively associated with student achievement ($r = .26$, $p < .001$). The researchers focused on students and teachers' content knowledge, leaving out teachers' pedagogical content knowledge. Researchers began to include teachers' pedagogical content knowledge in mathematical knowledge for teaching frameworks.

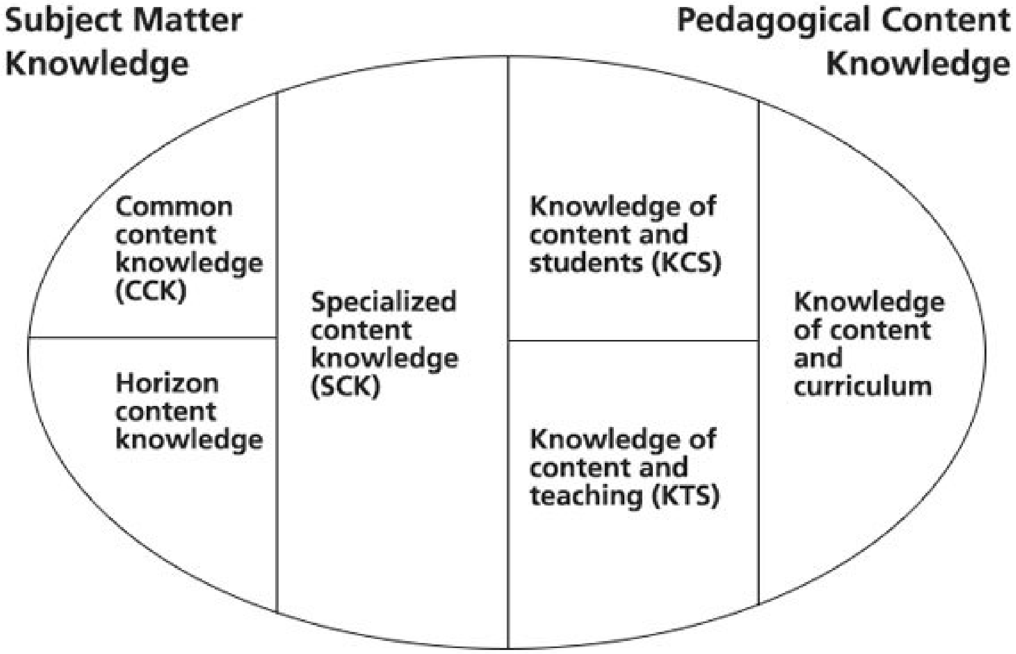
Ball et al. (2008) developed a widely used mathematical knowledge for teaching framework (see Figure 4). Ball et al. identified content knowledge as mathematical knowledge, contrasting with Shulman's framework (Lovett & Lee, 2017a). The two multi-faceted parts of the MKT framework are subject matter knowledge and pedagogical content knowledge (Ball et al., 2008).

Subject matter knowledge consists of (a) common content knowledge, (b) horizon content knowledge, and (c) specialized content knowledge. *Common content knowledge* refers to mathematical knowledge used in various settings not unique to teaching (Ball et al., 2008). *Horizon content knowledge* is teachers' knowledge of mathematics curriculum scope and sequence (Ball et al., 2008). *Specialized content knowledge* is mathematics knowledge beyond that expected of any well-educated adult but not requiring students' knowledge (Ball et al., 2008). Ball et al. (2008) identified specialized content knowledge as mathematical knowledge. In contrast, Shulman would identify specialized content knowledge as pedagogical content knowledge (Hill et al., 2005).

Pedagogical content knowledge has three components: knowledge of (a) content and students, (b) content and curriculum, and (c) content and teaching. *Knowledge of content and students* provides teachers with insight into how students think about the discipline, which allows teachers to anticipate students' thinking (Ball et al., 2008). *Knowledge of content and curriculum* enables teachers to sequence lessons and tasks to support student learning (Groth & Meletiou-Mavrotheris, 2018). *Knowledge of content*

Figure 4

Ball et al. (2008) MKT Framework



Note. From “Content knowledge for teaching: What makes it special?”, by D. L. Ball, M. H. Thames, and G. Phelps, 2008, *Journal of Teacher Education*, 59(5), p. 403. Copyright 2008 by SAGE Publications.

and teaching provides instructional strategies for making specific concepts understandable (Groth & Meletiou-Mavrotheris, 2018). Table 1 provides a quick reference for each component of Ball et al.'s (2008) MKT framework.

Ball et al. (2008) studied the work mathematics teachers did in the classroom and centralized their research on one question: "What do teachers need to know and be able to do to teach effectively?" (p. 394). The researchers conducted a qualitative study using a National Science Foundation database that documented third-grade teachers teaching mathematics from 1989 to 1990. Included in the database were videotapes and audiotapes of the mathematics instruction, transcripts, copies of student work, notes, and reflections. The researchers found that teachers with mathematical and pedagogical content knowledge could teach mathematics more effectively and understand students' mathematical difficulties and thought processes.

Similarly, Groth and Meletiou-Mavrotheris (2018) stated that teachers who understood the content knowledge and teaching practices provided students with an overall better quality of instruction. Because mathematics and statistics are distinct disciplines, these results cannot transfer to the teaching and learning of statistics. With an increase in statistics content in Pre-K–12 curriculum, mathematics teachers need to understand statistics and possess statistical knowledge for teaching to be effective in teaching statistics.

Table 1

Ball et al. (2008) MKT Framework Components Defined

Component	***Definition***
Common content knowledge	Mathematical knowledge and skills used in settings other than teaching.
Specialized content knowledge	Mathematical knowledge is unique to teaching.
Horizon knowledge	Knowledge of how mathematical topics are related over the span of mathematics and statistics.
Knowledge of content and students	Knowledge to anticipate students' thinking and misconceptions in mathematics.
Knowledge of content and teaching	Knowledge needed to sequence instruction and appropriate teaching practices for mathematics.
Knowledge of content and curriculum	Vertical and horizontal knowledge of mathematics content and curriculum.

Note. From “Content knowledge for teaching: What makes it special?”, by D. L. Ball, M. H. Thames, and G. Phelps, 2008, *Journal of Teacher Education*, 59(5), p. 405. Copyright 2008 by SAGE Publications.

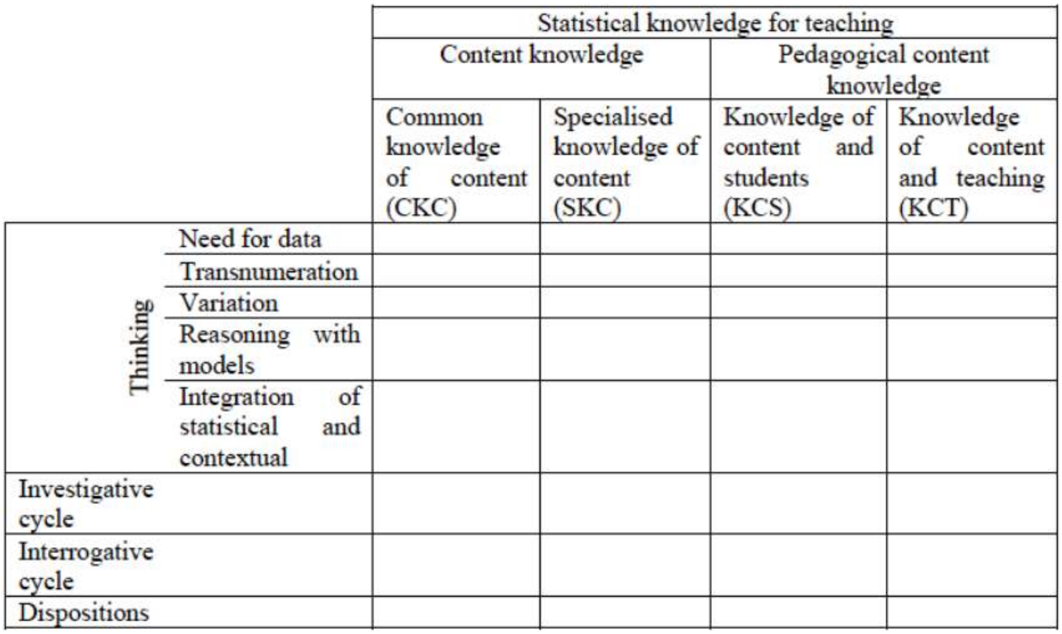
Statistical Knowledge for Teaching

Mathematics and statistics are distinct disciplines (Schoen et al., 2019; Zientek et al., 2011). Thus, MKT differs from statistical knowledge for teaching (SKT). Burgess (2006) and Groth (2007) agreed that mathematics teachers needed mathematical and statistical knowledge to advance students' learning in secondary mathematics. Therefore, SKT is purposeful for mathematics teachers and researchers.

Burgess (2006) examined which SKT mathematics teachers needed to develop. Burgess (2006) developed an SKT framework based on the work of Hill et al. (2004), Ball et al. (2005), and Wild and Pfannkuch (1999) (see Figure 5). Similar to Ball et al.'s (2008) MKT framework, Burgess' (2006) SKT framework consisted of two components: (a) content knowledge and (b) pedagogical content knowledge. Common knowledge and specialized knowledge of content are within the content knowledge component. The pedagogical content knowledge component comprises knowledge of (a) content and students and (b) content and teaching. Groth (2007) constructed a framework acknowledging the differences between Ball et al.'s (2008) MKT framework and Burgess' (2006) SKT framework using two components: (a) common content knowledge

Figure 5

Burgess (2006) SKT Framework



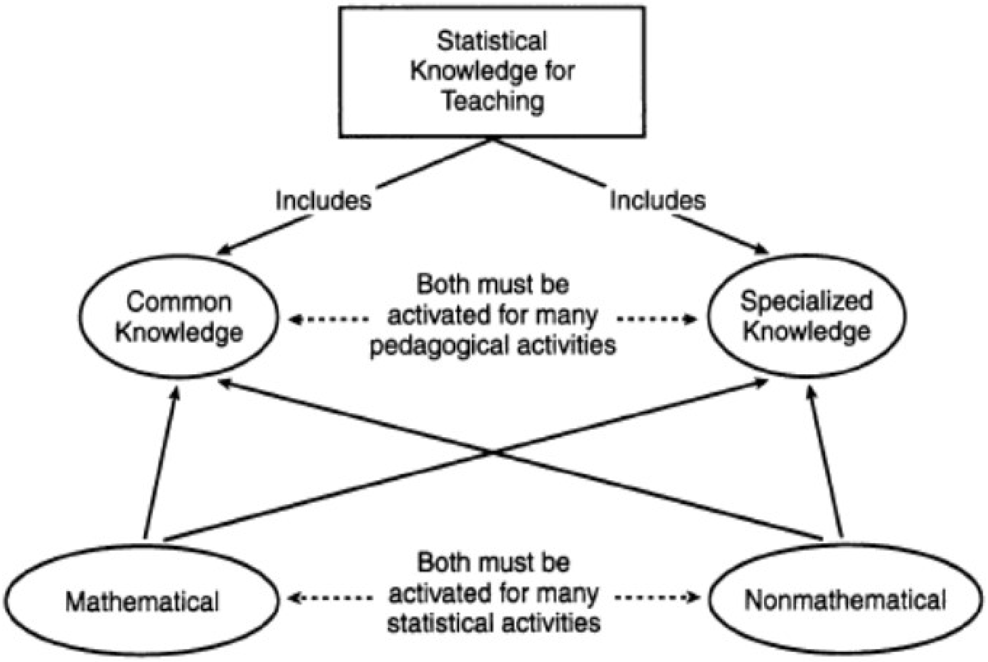
		Statistical knowledge for teaching			
		Content knowledge		Pedagogical content knowledge	
		Common knowledge of content (CKC)	Specialised knowledge of content (SKC)	Knowledge of content and students (KCS)	Knowledge of content and teaching (KCT)
Thinking	Need for data				
	Transnumeration				
	Variation				
	Reasoning with models				
	Integration of statistical and contextual				
Investigative cycle					
Interrogative cycle					
Dispositions					

Note. From "A framework for examining teacher knowledge as used in action while teaching statistics" by T. Brugess, 2006, *Proceedings of the Seventh International Conference on Teaching Statistics*, p. 4. Copyright 2006 by Mathematics Education Research Group of Australasia Incorporated.

and (b) specialized content knowledge (see Figure 6). Groth used the statistical investigation cycle to identify components of common content knowledge and specialized content knowledge. The statistical investigative cycle has four phases: (a) formulating questions, (b) collecting data, (c) analyzing data, and (d) interpreting results. Groth described each phase using common mathematics knowledge, common statistics knowledge, specialized mathematics knowledge, and specialized statistics knowledge. Unlike Burgess, Groth did not include pedagogical content knowledge. Groth expanded his 2007 framework in 2013.

Figure 6

Groth (2007) SKT Framework



Note. From “Toward a conceptualization of statistical knowledge for teaching” by R. E. Groth, 2007, *Journal for Research in Mathematics Education*, *38*(5), p. 429. Copyright 2007 by The National Council of Teachers of Mathematics Incorporated.

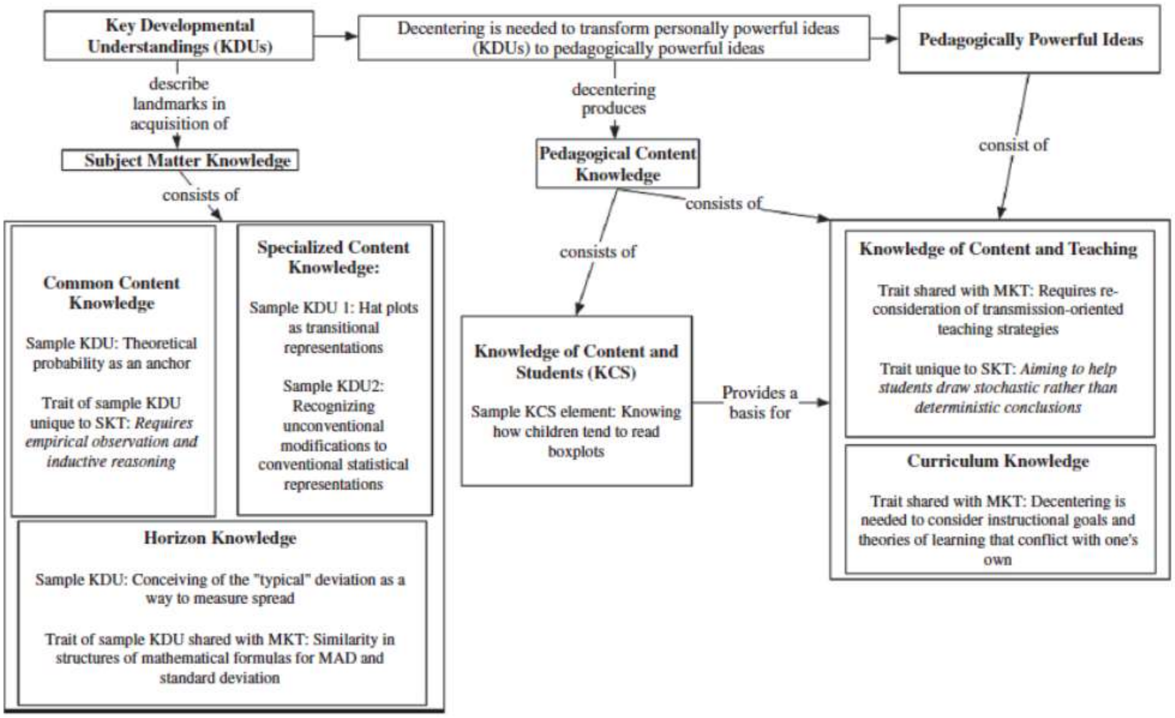
Groth (2013) developed a new SKT framework including aspects of pedagogical content knowledge (see Figure 7). The foundation of Groth's (2013) framework was key developmental understandings of statistical knowledge. Key developmental understandings were landmarks in subject matter knowledge development (Groth, 2013). Teachers need to reach these landmarks to develop subject matter knowledge before developing powerful pedagogical statistical knowledge (Groth, 2013). Groth (2013) explained that once a teacher gained personal experiences with key developmental understandings of statistics, they developed powerful pedagogical ideas. These powerful pedagogical ideas are landmarks in pedagogical content knowledge. After a teacher develops key developmental understandings and transforms personal, powerful ideas into pedagogically powerful ideas, Groth's (2013) framework resembles Ball et al.'s (2008) MKT framework. Some aspects of Groth's (2007) framework were included to distinguish between MKT and SKT.

Groth's (2007, 2013) SKT frameworks focused on teachers' understanding of statistics and the statistical investigation cycle. These frameworks included aspects of the affective domain, such as attitudes and anxiety. Much of the statistics curriculum found in Pre-K–12 was developed from the *Guidelines for Assessment and Instruction in Statistics Education* (GAISE) Pre-K–12 reports (Franklin et al., 2007). Groth's (2007, 2013) SKT frameworks centered on the GAISE report.

The first GAISE Pre-K–12 report published in 2007 to provide a framework for the statistics education curriculum (Groth, 2007, 2013). The GAISE report enhanced statistics standards in the National Council of Teachers of Mathematics (2000) *Principles*

Figure 7

Groth (2013) SKT Framework



Note. From "Characterizing key developmental understandings and pedagogically powerful ideas within a statistical knowledge for teaching framework" by R. E. Groth, 2013, *Mathematical Thinking and Learning*, *15*(2), p. 143. Copyright 2013 by Taylor & Francis Group, LLC.

and Standards for School Mathematics (Franklin et al., 2007; Froelich, 2011). The GAISE report provided direction for teachers on how to include statistical content in their mathematics curriculum (Franklin et al., 2007; Groth, 2007). The statistical content included the study of qualitative and quantitative variables, experimental designs on small data sets, and descriptive statistics (Franklin et al., 2007). Since the 2007 release date, statistical content within the mathematics curriculum has expanded.

In 2020, a second GAISE Pre-K–12 report (GAISE II) was published, including new statistical content skills needed for developing students' statistical literacy (Bargagliotti et al., 2020). Like the original GAISE report, the GAISE II report had three levels: A, B, and C. The levels increased with statistical sophistication from A to C. Content in Level A was mainly found in elementary school; Level B content was mainly found in middle school; and Level C content was mainly found in high school (Lovett & Lee, 2017a). Franklin et al. (2007) identified that students should engage in the investigative cycle at each level (see Appendix C), which agreed with Groth's (2013) SKT framework. Knight (2021) recommended that students should have an introduction to the investigative cycle by third grade, and statistical thinking should be included in all mathematics courses for high school students. These recommendations align with the GAISE II report. Hence, mathematics teachers need statistical knowledge of the four phases of the statistical investigative cycle across the three GAISE levels to teach statistics (Hannigan et al., 2013; Harrell-Williams et al., 2015; Lovett & Lee, 2017a). In addition, the Georgia Framework for Statistical Reasoning includes the four phases of the statistical investigative cycle (Georgia Department of Education, 2021). The cycle is used throughout each grade level to help learners develop statistical reasoning and literacy. Hence, the Georgia Framework for Statistical Reasoning aligns with the GAISE II report. Much research on PSMTs' SKT used Groth's (2013) SKT framework as it incorporated the statistical investigative cycle found in the GAISE II report.

Research on Statistical Knowledge for Teaching

Teacher training and professional development programs need to develop primary and secondary mathematics teachers who possess MKT and SKT (Hannigan et al., 2013; Harrell-Williams et al., 2015; Lovett & Lee, 2017a). To meet the challenges of preparing K-12 PSMTs to teach statistics, the National Council of Teachers of Mathematics (2022) stated, "faculty who teach statistics need to work together with education faculty to provide coursework that emphasizes stronger conceptual knowledge of statistics and the essential ideas of statistical thinking and problem solving" (p. 2). Franklin et al. (2007) stated one challenge was that primary and secondary mathematics teachers were required to teach statistical content that they did not experience as students. In addition, some primary teachers have never completed a stand-alone statistics course. Teaching and learning statistics are influenced by cognitive and noncognitive factors (Yusuf et al., 2019). Tchoshanov (2011) stated teachers' cognitive factors that may influence their instruction include mathematical and statistical ability and pedagogical knowledge. The noncognitive factors include teaching efficacy, attitudes towards statistics, and anxiety about statistics. Hence, research is needed on the factors that may influence teachers' statistical knowledge for teaching to better prepare K-12 PSMTs.

Researchers have analyzed many factors of the affective domain on PSMTs' statistical knowledge for teaching, with much focus given to teaching efficacy, attitudes, and anxiety about statistics (Estrada et al., 2011). PSMTs' attitudes toward teaching and learning statistics have been measured on various instruments, but results are inconsistent (Estrada et al., 2011). Yusuf et al. (2019) acknowledged that research on PSMTs' anxiety about teaching and learning statistics was scarce. In addition, researchers have determined factors that correlated to PSMTs' statistical knowledge for teaching (Lovett & Lee, 2017a).

Statistical Content Knowledge and Statistics Teaching Efficacy

Over the last 25 years, there has been much research regarding teachers' statistical knowledge for teaching. Researchers examined PSMTs' statistics teaching efficacy and their content knowledge of statistics. Lovett and Lee (2017a) examined 236 PSMTs' statistics teaching efficacy by administering the validated Self-Efficacy to Teach Statistics (SETS) instrument. The SETS instrument was developed under the GAISE report framework and measures teacher efficacy in teaching statistical concepts to K-12 students by collecting quantitative and qualitative data (Franklin et al., 2007). There are 44 six-point Likert scale items: 11 items align to GAISE level A, 15 items to GAISE level B, and 18 items to GASIE level C. The instrument provides an overall score and subscale score for each GAISE level. In addition, the instrument produces an overall confidence score to teach statistics and a confidence score to teach statistics for each level. PSMTs were asked to identify a statistical concept that they felt least confident and most confident in teaching for each GAISE level, which resulted in six open-ended response questions.

The PSMTs were undergraduate juniors and seniors or graduate students in a school of education from a university in the U.S. Many PSMTs have completed one or two statistics courses, and the majority were female (70%). The PSMTs were administered the SETS instrument at the end of their final mathematics education course. Results showed that PSMTs, on average, were between somewhat confident and confident in teaching statistics, where the highest confidence levels were found on level A items. Level A items cover content taught in primary school; thus, secondary PSMTs should be highly confident in teaching these items. Lovett and Lee (2017a) used repeated measures ANOVA and found that the mean confidence scores across the three levels differed significantly ($F(2, 470) = 66.54$, $p < .001$). A post hoc test using a Bonferroni correction showed PSMTs' confidence was significantly lower, moving from level A to level C ($p < .001$). Hence, PSMTs' confidence in teaching statistics was lower for statistical content at the secondary level compared to the primary level. For level A items, PSMTs were most confident in teaching the creation of graphical representations and recognizing statistical results may differ from group to group. PSMTs were least confident in teaching the selection of appropriate graphic displays and generalizing results from a small to a large group. For level B, PSMTs were most confident in teaching the creation of histograms and computing the interquartile range and least confident in teaching interpreting measures of association. For level C, PSMTs were most confident in teaching characteristics of a normal distribution and identifying slope and y-intercept in a regression equation. PSMTs were least confident in teaching finding conditional and marginal frequencies using two-way tables and formal ideas of inference using simulations, such as margin of error and testing for statistical significance.

Lovett and Lee (2017a) provided knowledge on PSMTs' statistics teaching efficacy relating to the investigative cycle across the three GAISE report levels. However, there were limitations to the study. The researchers used a purposeful sample of PSMTS with a lack of randomization. Also, PSMTs self-reported on the open-ended questions, which may result in them showing favorable images of themselves. A similar study was performed on middle school PSMTs (Harrell-Williams et al., 2015).

Harrell-Williams et al. (2015) administered the SETS instrument to 309 middle school PSMTs enrolled in an introductory statistics course or designated mathematics education course. The researchers collected a sample of PSMTs from four large size public universities. Many PSMTs (67%) completed one college-level statistics course. The SETS instrument was administered at the end of the introductory statistics course or designated mathematics education course and was voluntary. Because the sample was middle school PSMTs, Harrell-Williams et al. (2015) administered the SETS items belonging to level A (11 items) and level B (15 items). The researchers scaled the data to the Multidimensional Random Coefficients Multinomial Logit Model to obtain person and item parameters to understand which statistical concepts were easier and difficult for PSMTs to state they were completely confident to teach using the item difficulty estimates. Harrell-Williams et al. (2015) identified six items with a lower level of teaching efficacy. The six items, starting with the highest item difficulty estimates, were:

1. Interpret measures of association (0.74).
2. Distinguish between a question based on data that vary and a question based on a deterministic model (0.61).
3. Distinguish between association and causation (0.47).
4. Describe the strength of the association between two variables using linear models (0.43).
5. Use interquartile range, five number summaries, and boxplots for comparing distribution (0.36).
6. Explain the differences between two or more groups concerning center, variability, and shape (0.31) (Harrell-Williams et al., 2015).

All six items were level B items. Harrell-Williams et al. (2015) identified five items with a high level of teaching efficacy, four level A items and one level B item. The five items, including the corresponding item difficulty measure, were:

1. Use dotplot, stem and leaf plot, and tables for describing distributions (-0.34).
2. Recognize that a sample may or may not be a representative of a larger population (-0.41).

3. Recognize that there will be natural variability between observations for individuals (-0.48).
4. Create dotplot, stem and leaf plot, and tables for summarizing distributions (-0.50).
5. Recognize that statistical results may differ in another class or group (-0.90) (Harrell-Williams et al., 2015).

Lovett and Lee (2017a) and Harrell-Williams et al. (2015) have had similar results showing PSMTs lacked confidence in teaching statistical concepts at their corresponding grade level. The results from both studies may be because PSMTs' needed more content knowledge of statistics and needed to develop key understandings of statistics, as outlined in Groth's (2013) framework. The researchers of the two studies recommended further studies on teachers' statistical and pedagogical content knowledge.

Lovett and Lee (2017b) administered the Levels of Conceptual Understanding of Statistics (LOCUS) instrument to 221 secondary PSMTs who were undergraduate or graduate students enrolled in their last mathematics education course before student teaching. Approximately 71% of the PSMTs were female, and almost all (93.4%) completed a college statistics course or high school Advanced Placement Statistics. The researchers collected their sample by identifying 57 universities that had some faculty attend the NSF-funded program, *Preparing to Teach Mathematics with Technology* (PTMT, n.d.) or the ASA-funded program *Math/Stat Teacher Education: Assessment, Methods, and Strategies* (TEAMS, n.d.) conference between 2002-2014. Lovett and Lee (2017b) contacted the universities to determine if the mathematics education program was interested in participating. Many showed interest, but only eighteen universities agreed to participate. The number of PSMTs participating from each institution ranged from 2 to 31 ($M = 12$). The PMSTs were recruited by their mathematics teaching methods instructor to complete the LOCUS instrument online at the end of the semester. The instructors agreed to include the LOCUS instrument as an assignment as part of their course. It was unknown if the LOCUS instrument was proctored.

The LOCUS instrument assesses understanding of statistics within each phase of the statistical investigative cycle across the three GAISE levels. The PSMTs took the 30 multiple choice Intermediate/Advanced Statistical Literacy versions of the assessment designed for secondary students. There are two level A items, 11 level B items, and 17 level C items. The LOCUS instrument is not intended as a high stakes knowledge assessment but represents the statistics content PSMTs are expected to teach their students (Lovett & Lee, 2017b). The LOCUS instrument provides an overall score and subscores for each GAISE level B and C and each part of the investigative cycle (*formulating questions, collecting data, analyzing data, and interpreting results*).

Lovett and Lee (2017b) found that the PSMTs needed to conceptualize the statistical content to teach statistics to students at the secondary level. The LOCUS mean overall score was 68.61 ($SD = 14.06$), where 25% of PSMTs scored above 77 and 25% scored below 57. The mean score for level B items was 70.85 ($SD = 17.69$), and for level C items was 64.87 ($SD = 14.16$). PSMTs scored significantly higher, on average, on level B items compared to level C items ($t(220) = 5.772$, $p < .001$), implying that PSMTs have less understanding of statistical concepts as items increase in sophistication. In addition, results relating to the investigative cycle showed PSMTs scored higher, on average, on *formulating questions* ($M = 80.37$, $SD = 21.51$) and lower as the cycle progressed. The mean for *collect data* was 70.40 (*SD* = 63.34), and for *analyze data* was 63.34 (*SD* = 22.22). The PSMTs scored lowest on *interpreting results* (*M* = 60.48, *SD* = 16.25). A repeated measures ANOVA showed the mean scores of the investigative cycle differed significantly ($F(3, 648) = 64.73$, $p < .001$). The researchers found no significant difference between mean scores for *analyze data* and *interpret results* ($p = .32$). From the results of post hoc tests using a Bonferroni correction, PSMTs scored significantly lower as the investigative cycle progressed ($p < .001$). Results showed that most PSMTs needed to develop key understandings of common statistical knowledge to provide a foundation for teaching students key concepts across the investigative cycle.

Lovett and Lee (2017b) did an item analysis of the LOCUS assessment as classified by the investigative cycle. Results showed no common misunderstandings for *formulating questions*. For *collecting data*, the researchers found that PSMTs' strengths were in identifying ways to improve a study design given a study and measurements, identifying which study design would be best based on a question of interest, and determining a data collection plan based on a study description. The researchers found a PSMTs' weakness to be identifying how to choose a sample to minimize bias. Approximately 65% were able to select a correct sampling method. In addition, PSMTs' struggled with answering questions based on what conclusions could be drawn from a specific study design. Regarding one question on drawing conclusions, approximately 58% of PSMTs responded that a researcher could make generalizations about a population based on a sample of volunteers (Lovett & Lee, 2017b).

For *analyzing data*, the researchers found that PSMTs demonstrated an understanding in selecting appropriate measures of central tendency, how measures of center and variation change when data values were changed, and justification of an association from a two-way table. The PSMTs showed weaknesses in *analyze data* items that involved understanding variation in data. Only 43% of PSMTs could identify a histogram containing data that varied the least from its mean. Instead, 30% of PSMTs chose a uniform distribution, and about 20% thought variability from the mean was the same for all three distributions.

For *interpret results*, PSMTs could compare distributions in a context using the center and spread, demonstrate an understanding of the effect of sample size on a sample mean, and interpret survey results with a given margin of error. PSMTs struggled most with statistical significance, identifying and interpreting a *p*-value, and explaining confidence intervals. About half (48%) of PSMTs could correctly interpret results given a significant *p*-value and failed to reject the null hypothesis. However, 40% of PSMTs chose a conclusion that a significant *p*-value meant they could leave the null hypothesis. On another item regarding *p*-value, PSMTs were asked to reason if a *p*-value would be large or small for comparing means of two distributions given data on a dot-plot. Only 35% of PSMTs were able to correctly identify that the *p*-value would be small due to the large gap between distributions. Almost 47% incorrectly answered that the *p*-value would be large due to a large distribution gap. These findings demonstrated that PSMTs, on average, needed an understanding of what it meant to be statistically significant and what a *p*-value represents. In addition, PSMTs needed to develop an understanding of statistics content knowledge found in high school curricula.

In short, the researchers found that PSMTs' strengths were in identifying appropriate statistical questions, selecting the proper study design, identifying data collection plans, identifying the right measure of center, and providing justification of an association from a two-way table. In addition, PSMTs' strengths were in comparing distributions in a context using the center and spread, demonstrating an understanding of the effect of sample size on a sample mean, and interpreting survey results with a given margin of error. The researchers found PSMTs' weaknesses in understanding variation in data, identifying histogram properties, understanding the expected variation in sample means, and formal inferences. In addition, the PSMTs struggled most with statistical significance, identifying and interpreting a *p*-value, and explaining confidence intervals. Lovett and Lee (2017b) recommended that educational training programs focus on developing PSMTs' knowledge of variability, sampling distributions, and formal inference as they relate to the investigative cycle.

De Vetten et al. (2018) used a pretest-posttest design on a class of primary PSMTs from a Netherlands teacher college. There were 21 second year PSMTs in the class with an average age of 20.95 (*SD* = 2.19), where 81% were female. All PSMTs experienced learning basic descriptive statistics. The instrument used for the pretest and posttest had two tasks in the form of a questionnaire. The two tasks were based on 12 learning objectives. There were six categories for the learning objectives: data as evidence (one learning objective), generalization beyond the data (three learning objectives), sampling variability (one learning objective), sampling method (four learning objectives), sample size (one learning objective), and uncertainty (two learning objectives). The first task had PSMTs investigate the selection of a representative sample, and the second task asked PSMTs to compare two sample distributions and to make generalizations from the sample. The intervention was short (180 minutes) and PSMTs worked through learning activities that included topics on inferential reasoning, random sampling, statistical investigative cycle, and the equiprobability bias. The pretest was administered digitally at the beginning of the course, and the posttest was administered at the end. It was unknown if the instruments were proctored. Twenty-one students took the pretest. Of the 21 students, five students did not take the posttest and were excluded from the analysis.

De Vetten et al. (2018) attempted to understand PSMTs' statistical content knowledge by comparing the items of the pretest and posttest after the intervention. The researchers found that most PSMTs (94%) agreed that commonly held beliefs were not valid evidence for a conclusion and 94% of PSMTs used descriptive statistics to compare two sample distributions. The number of PSMTs who acknowledged that making generalizations from a sample was possible increased from 12 in the pretest to 14 in the posttest. Eight PSMTs understood sampling variability in the pretest, which increased to 12 in the posttest. The number of PSMTs who agreed that random sampling was appropriate rose from nine in the pretest to 11 in the posttest. A large majority (94% in the pretest and 81% in the posttest) incorrectly agreed that distributed sampling was an appropriate sampling method. In the pretest, five participants thought a convenience sample could represent the population well, which decreased to none in the posttest. Two PSMTs in the pretest and six in the posttest agreed that the sample was a good representative of the population when an appropriate sampling method was used. The researchers found minimum development relating to concepts on sample size. The researchers found that no PSMTs met all learning objectives after the intervention. The small size of the sample implies results were limited. However, the study gave more evidence of the lack of statistical content knowledge of PSMTs.

Moody and Gardner (2014) found similar results in their small-scale study of 19 secondary PSMTs. The secondary PSMTs were enrolled in a teacher preparation program at a university in the southeastern U.S. The PSMTs participated in a capstone research project as part of the requirements for teacher certification to engage undergraduate mathematics majors in thinking independently, synthesizing mathematical content, and writing and presenting formal mathematics or a mathematical idea, concept, or topic. At the end of the capstone project, students were required to submit a written report and present their project during a mathematics department colloquium. The PSMTs in the study designed a statistical investigation based on the popular video game, *Call of Duty: Modern Warfare 3*. The 19 secondary PSMTs played 40 *Call of Duty: Modern Warfare 3* video games and recorded the end-of-game statistics. The video game provides the end-of-game statistics for each player on four variables: number of kills, number of deaths, number of assisted kills, and kills-to-deaths ratio. The 19 PSMTs based their capstone project on the following questions: (1) to what extent do particular variables, namely kills, assisted kills, deaths, and the kills to deaths ratio, affect a player's overall game score, and (2) is there a significant difference between specific variables given a particular game mode? The 19 PSMTs discovered some linear relationship between particular variables and a player's overall game score, but they needed help to use the collected data to investigate this relationship. Also, the PSMTs did not attempt to use the sample data to determine if correlations existed among the four variables. Instead, they tried to develop a system of equations to use as a predictor model. Through investigation, Moody and Gardner (2014) found that the PSMTs' statistical knowledge and understanding of statistical methods existed but were compartmentalized such that they did not relate one to the other. The PSMTs began the project by identifying a response variable and several explanatory variables but needed help understanding the connections. In addition, the PSMTs made appropriate assumptions that would predict an overall game score, but statistical methods to analyze the data collected were not used. PSMTs' observations and game experience directed much of their predictions. Moody and Gardner (2014) acknowledged that PSMTs would have benefitted from more explicit guidance to help make the connections among statistical concepts, measures, and representations of data. Results are limited because of sampling techniques.

Leavy and O’Loughlin (2006) studied 263 undergraduate students enrolled in their first year of primary teacher education program in southwest Ireland. All students studied five years of mathematics at the high school level and 76% were female. The study aimed to gain insight into primary PSMTs’ understanding of the mean. The researchers used multiple data sources and data collection methods to provide insights into the PSMTs’ statistical understandings. The PSMTs completed a questionnaire consisting of six tasks where they were asked to provide a justification when solving the task. The results showed that 57% of PSMTs correctly utilized the mean to compare two data sets, 21% provided an accurate answer to a weighted means problem, and 88% constructed a data set to reflect a pre-given mean value. Approximately 25% of the PSMTs indicated some form of conceptual understanding of the mean. In addition, 25% of the PSMTs confused the mean with the mode in narrative description of the mean or through identification of the mean on a graphical representation. Overall, the primary

PSMTs were found to not possess a conceptual understanding of the measures of central tendencies.

Zientek et al. (2011) conducted a study to determine if the number of mathematics courses affects PSMTs' statistical achievements. The sample consisted of 95 PSMTs enrolled in two sections of a statistics course explicitly designed for PSMTs. Most (89%) of the PSMTs completed at least two mathematics courses. The researchers examined statistical differences between two classes. The two classes were PSMTs' statistical achievements in the statistics courses and the number of mathematics courses PSMTs completed. Statistical achievements were the student's course grade at certain intervals during the semester. A Kruskal-Wallis test indicated no statistically significant difference existed between the two classes on statistical achievement ($H(95) = 1.98$, $p = .16$). The researchers found no statistical significant linear relationship between achievement and number of mathematics courses ($r(85) = .115$, $p = .287$). The researchers concluded that completing numerous mathematics courses did not affect PSMTs' statistical achievements. The studies on PSMTs' statistical content knowledge point to a lack of statistical knowledge and statistics teaching efficacy.

Arnold (2008) used action research on ISMTs from 11 different secondary schools ranging in size from 500 to 2500 students in New Zealand. Arnold (2008) aimed to use three action research cycles as outlined by Wilson and Berne (1999) and Goldman (2001) to determine the statistical content knowledge primary ISMTs need, the benefits of professional development relating to statistical content knowledge, and the learning experiences required to improve statistical content knowledge. The first cycle was the initial workshop, where ISMTs identified areas they wanted to work on in subsequent workshops. In addition, the ISMTs self-reported their statistical knowledge of statistics achievement objectives formulated by Wild and Pfannkuch's (1999) Problem, Plan, Data, Analysis, and Conclusion cycle. The second cycle involved three 2-hour workshops that discussed sources of variation and had the ISMTs work in small groups to discuss the meaning of considering sources of variation. The final cycle was a self-reflection on learning experiences that helped ISMTs understand new statistical content knowledge. There were between 12 to 19 ISMTs at each cycle. The ISMTs had a wide range of experience, from two years of experience to thirty years. Some teachers completed a statistics major, and some had never met a statistics course. Data from the teachers' responses to open-ended questionnaires, emails, and journal questions were collected. Spreadsheets were used to categorize and analyze the data qualitatively. Results showed ISMTs needed to gain knowledge in sources of variation, cleaning data, and recategorizing data. Also, ISMTs were found to slightly improve their new statistical content knowledge after the professional development workshop. Arnold (2008) acknowledged that the depth and breadth of the ISMTs' knowledge were varied. The limitations of this study included study design, researcher bias, and sample size. There was no evidence that the workshops improved ISMTs' statistical content knowledge, but their perceptions were that their understanding of statistics improved after the workshop.

Doerr and Jacob (2011) studied 11 secondary mathematics teachers who enrolled in a statistics education professional development course, in which they found teachers struggled with similar difficulties with statistics as students. Eight teachers were PSMTs, and three were ISMTs. Ten teachers have completed a college-level introductory to statistics course. The researchers used a pretest-posttest design. The researchers administered an instrument that measured statistical knowledge at the beginning and end of professional development. The instrument included 20 items across six categories: *graphical representation*, *sampling variation*, *inference*, *data collection and design*, *bivariate data*, and *probability*. The items were drawn from the validated Comprehensive Assessment of Outcomes in a First Statistics course (CAOS) instrument. The results showed an overall improvement in the teachers' understanding of statistical knowledge across the six levels.

The mean pretest score was 11.09, while the mean posttest score was 12.54. Results from the paired samples *t*-test showed a statistically significant gain ($p = .027$). The only category of the test with a significant gain from the pretest to the posttest was graphical representations ($p = .033$). Graphical representations had six items; the mean on the pretest was 4.09, and the posttest was 5.00. The first item was describing the distribution of a histogram. Approximately 82% of teachers answered correctly. There was no change in this item from the pretest to the posttest. The second item was on interpreting normal distributions. Approximately 64% of teachers answered correctly on the pretest and 73% of teachers answered correctly on the posttest. Items three and four were on interpreting measures of central tendency and quartiles in a box plot. Approximately 91% of teachers answered item 3 correctly on the posttest compared to 73% on the pretest. Item 4 correct response rate went from 55% to 82% correct. Item five was on the effects central tendencies have on normal distributions. The correct response increased from 64% to 82%. Item six addressed the ability to estimate standard deviations for different histograms. Approximately 73% of teachers answered correctly on the pretest and 91% on the posttest. The researchers were able to identify common misunderstandings the teachers exhibited regarding variability and shape. The intervention was shown to improve teachers’ overall scores and scores on the graphical representations category. However, the sample size was small implying results were limited. The findings suggested that further quantitative studies were needed on larger samples to determine if teachers understand the statistical content they must teach.

Engledowl and Tarr (2020) used a task-based clinical interview design to understand secondary ISMTs' statistical knowledge structures. The researchers used a stratified purposeful sample of 9 ISMTs. Recruitment extended across four strata:

- Teachers who had taught an Advanced Placement (AP) statistics course ($N = 2$),
- A non-AP statistics course ($N = 2$),
- At least one unit of statistics in the context of a secondary mathematics course ($N = 2$),
- At least one unit of statistics in the context of a middle-level mathematics course ($N = 3$).

Data came from two sources: statistical content knowledge instruments and open-ended constructed response questions. The Goals and Outcomes Associated with Learning Statistics (GOALS-2) was the statistical content knowledge instrument. The GOALS-2 was developed from items from the psychometrically valid and reliable Comprehensive Assessment of Outcomes in Statistics (CAOS) assessment (delMas et al., 2007). The instrument included 20 forced-choice items on content that covered covariation, samples, sampling techniques, hypothesis testing, and study designs (Sabbag & Zieffler, 2015). The constructed response questions were from the four released questions from the

LOCUS assessment (Jacobbe, 2016). Engledowl and Tarr (2020) developed statistical knowledge structure maps for each ISMT. The researchers analyzed the maps using cross-case analysis. They found secondary teachers' knowledge structures could be described as being in three categories: compatible-connected (*N* = 3), incompatible-connected (*N* = 4), and incompatible-disconnected (*N* = 2). The results of this study were limited due to the sample size.

Umugiraneza et al.'s (2018) studied 75 primary and secondary ISMTs who attended a statistical education workshop in 2015 at the University of KwaZulu-Natal. The workshop aimed to improve the statistical content knowledge of teachers. The 75 ISMTs volunteered for the study, which makes the results limited. The researchers administered a questionnaire to the ISMTs after the workshop. The questionnaire consisted of open-ended and closed questions relating to teaching and learning statistics and was adapted from the instrument developed by Beswick et al. (2012). Umugiraneza et al.'s (2018) used quantitative methods after they coded teachers' demographic factors and counted the number of approaches reported by teachers to improve statistics learning. The researchers found that approximately 57% of female and 32% of male teachers suggested multiple strategies for improving statistics. About 68% who studied at the postgraduate level or more suggested multiple strategies for improving statistics compared to 29% who only studied at the undergraduate level.

The researchers analyzed the ISMTs' responses to the open-ended questions for emerging themes using qualitative methods. First, the researchers independently coded all ISMTs' responses regarding improving teaching and learning statistics into broad categories. Secondly, the researchers reviewed the responses together to determine where differences occurred. The categories included: *increasing the motivation and interest of learners, teachers' explanations, linking teaching to real life, using practical activities and concrete examples, attending professional workshops,* and *using investigations and projects*.

Increasing the motivation and interest of learners had the highest number of suggestions. Approximately 29% of teachers suggested strategies to increase motivation and interest of learners. For *teachers' explanations*, 20% of teachers suggested improving teacher explanations by preparing lessons in advance, going back to basic computational skills, and providing more practice opportunities. Approximately 15% of teachers made suggestions on linking teaching to real life, 13% of teachers suggested strategies relating to practical or concrete examples, and 13% of teachers suggested investigations and projects for teaching and learning statistics. Only 10% of teachers made suggestions on professional development in teaching and learning statistics. Umugiraneza et al. (2018) used binary logistic regression to model teachers' suggestions and demographic factors.

The independent variables were gender, age, teaching experience, use of the curriculum, and teachers' level of education. The response variables used in the models were teaching strategies. The researchers coded the response variables with 0 indicating a single strategy and 1 indicating multiple strategies, making it binary. The results showed that 56.8% of female teachers suggested multiple strategies for improving mathematics compared to 31.6% of male teachers, 52.3% of young teachers suggested multiple strategies for improving mathematics compared to 32.3% of older teachers, 55.6% of those who integrate curriculum suggested multiple strategies for improving mathematics compared to 30.0% of those who did not use it, 67.5% of those who studied at postgraduate level or more suggested multiple strategies for improving mathematics and statistics compared to 28.6% of those who only studied at undergraduate level, and 64.4% of those who attended previous mathematics workshops suggested multiple strategies for improving statistics compared to 26.7% who did not attend them.

In addition, the researchers found that female teachers were four times more likely than men to suggest multiple strategies in relation to improving statistics, and teachers 40 years old and younger are five times more likely to express multiple strategies for improving statistics than the case for teachers over 40 years old. The results showed a significant difference in teachers that did not use the curriculum and teachers who did use the curriculum on suggesting multiple strategies in improving statistics teaching and learning than those (p = .043). The researchers stated that these findings supported the importance of teachers engaging with the statistics curriculum because it provided a roadmap for the teachers as they prepared learning experiences. Also, the researchers stated that teachers who used the statistics curriculum seem to be more aware of strategies for improving the teaching of mathematics and statistics. Unknown was if the strategies mentioned were research-based. The study used a purposeful sample which made results limited.

Lee et al. (2020) observed that 489 secondary ISMTs' perspectives and beliefs about statistics changed after a 15-week online professional development course. The course consisted of five units covering various teaching and learning statistics topics. The ISMTs were from 46 different states and 29 different countries. Most ISMTs were from the U.S. (N = 380) and New Zealand (N = 48). The majority of the ISMTs were female (68%). Approximately 73% of the ISMTs had a master's degree or higher. The researchers used multiple quantitative and qualitative data collected from six data sources: registration form, click logs, discussion forum posts, end-of-unit surveys, an end-of-course survey, and a follow-up survey six months after the course to participants who engaged in any aspect of the course. Many of the ISMTs (76%) had an active click log by engaging in various aspects of the course by accessing a page, viewing a video, downloading a document, and posting in a forum. The participation decreased sharply from Unit 1 (60%) to Unit 5 (19%). Approximately 58% of ISMTs posted to a discussion forum. Almost one-third of the ISMTs finished the course. The qualitative analysis from the discussion forums, end-of-unit surveys, and follow-up surveys indicated that the ISMTs viewed statistics as more than computations and procedures and felt more confident in teaching statistics. Understanding in-service secondary mathematics teachers' knowledge of statistics would suggest how to approach future professional development workshops. Lee et al. (2020) recommended including interviews or conducting classroom observations in future studies to understand if teachers' experience changes due to the professional development course. Understanding ISMTs' knowledge of statistics would shed light on how to approach future professional development workshops. Other studies exist on ISMTs' statistical content knowledge,

Schoen et al. (2019) used a pretest-posttest design with 180 Florida secondary ISMTs. The intervention was a professional development course on teaching and learning statistics. The secondary ISMTs were teachers from 35 of Florida's 67 regular public-school districts and one university. Almost 85% of the ISMTs were predominantly regular classroom teachers, approximately 10% being teachers of advanced classes, about 5% being special education or intervention teachers, and less than 3% with the primary role of math coach. All ISMTs worked at the secondary level. The researchers measured the ISMTs' statistical content knowledge using probability and statistics from the Diagnostic Teacher Assessment of Mathematics and Science (DTAMS) instrument. The instrument measured teachers' content knowledge and pedagogical content knowledge at the middle school level. The DTAMS' validity was established by reviewing national standards for content and the use of expert question writing teams and reviewers (Ronau et al., 2010). Saderholm et al. (2010) reported Cronbach's alpha coefficient of .90 after administering the DTAMS probability and statistics scale to 543 mathematics teachers.

The ISMTs completed parallel forms for the pretest, posttest, and delayed posttest. The researchers calculated Cronbach's alpha of .82, .81, and .81 for each of the three forms used in the study. The pretest was completed at the beginning of the professional development course, the posttest at the end of the professional development course (Fall 2014), and the delayed posttest months later (Spring 2015). The ISMTs completed all assessments on their own time during a testing window. Once completed, the ISMTs returned the assessments to the evaluation team by mail. The test scorers were not aware of the ISMTs' intervention.

Schoen et al. (2019) used statistical models to fit the data to investigate:

- the effect of the professional development program,
- the moderation of treatment effect by baseline knowledge and years of teaching experience, and
- the persistence of effects on the delayed posttest.

The researchers identified multiple covariates to be included in the models: grade levels taught, having a college degree in mathematics or science, having a teaching certificate in secondary mathematics, gender, ethnicity, and school district urbanicity. A stepwise backward elimination approach to covariate selection was used for the statistical models and showed the first-order predictors to be treatment, pretest, experience, and rural, and the higher-order predictor to be treatment-by-experience. The researchers tested the effect of treatment without any extraneous variables, with extraneous variables pretest and rural, and with extraneous variables pretest, rural, and years of experience. The researchers determined how the treatment-by-experience moderator was operating by conducting a *t*-test to compare the means of two groups: teachers with 10 or more years of experience and teachers with less than 10 years of experience. Ten years was approximately the sample mean of teaching years. The mean of the pretest scores for teachers with less than 10 years of experience was 23.46 (*SD* = 6.53), and the mean of the pretest scores for teachers with 10 or more years of experience was 22.05 (*SD* = 6.88). The researchers found that the pretest mean scores between teachers with less than 10 years of experience and teachers with 10 or more years of experience were not significant ($t(173) = 1.41$, $p = .178$, Hedges' $g = 0.25$). Although there was a small effect size, the researchers stated that teachers' baseline statistical knowledge appeared to be a factor in explaining the moderation of treatment by years of experience.

Results showed the treatment had an effect size of $g = 0.31$ and significance level of $p = .07$ without including for covariates. The treatment was statistically significant ($p = .009$) and had an effect size of $g = 0.25$ when including covariates pretest and rural. Adding years of experience, the treatment was statistically significant ($p = .005$) and had an effect size of $g = 0.26$. The results showed the treatment-by-experience interaction was statistically significant ($p= .001$) with a standardized effect of $\beta = 0.24$. The standardized effect indicated an increase of 0.24 standard deviations in the effect of treatment for each standard deviation increase in teaching experience. The results found by the researchers may indicate that professional development activities increase teachers' statistical content knowledge. However, noncognitive factors may influence teachers' motivation to attend professional development activities regarding teaching and learning statistics. Understanding the noncognitive factors that influence teaching and learning statistics may shed light on how to approach teachers with professional learning opportunities in teaching statistics.

Attitudes Toward Statistics

The affective domain includes noncognitive factors influencing teaching and learning, such as emotions, attitudes, and beliefs (McLeod, 1992; Olson & Zanna, 1993). Emotions can change rapidly and lead to the development of attitudes toward a subject (Gal et al., 1997). Mcleod (1992) defined attitudes as "affective responses that involve negative or positive feelings of moderate intensity" (p. 581). Students' attitudes towards a subject can influence their performance by being an obstacle (negative attitudes) or an advantage (positive attitudes) for learning.

Philipp (2007) identified that people's attitudes were shown through their acting, feeling, or thinking on their disposition or opinion. The researcher described attitudes that students may experience toward a subject as "liking, disliking, being curious about, or being bored by the subject" (p. 261). As another example, Philipp (2007) described that a student who struggled with a mathematical concept might feel discomfort and develop negative attitudes toward the concept, transitioning to mathematics in general.

Estrada et al. (2011) stated that attitudes were "…relatively stable, resistant to change, and comprise a larger cognitive component and less emotional intensity than emotions" (p. 2). Teachers should possess positive attitudes toward mathematics and mathematical knowledge to promote student mathematical literacy and positive attitudes towards mathematics (Wilson & Cooney, 2002). Students and teachers with negative attitudes toward mathematics transfer these attitudes toward statistics (Gal & Ginsburg, 1994). Hence, PSMTs with negative attitudes toward mathematics would be more likely to transfer these attitudes toward statistics.

Attitudes toward statistics are comprised of someone's feelings towards and perceptions of statistics in terms of relevance and value (Evans, 2007). Teachers' positive attitudes toward statistics influence the teaching and learning of statistics (Gal & Ginsberg, 1994; Schau, 2003). Positive attitudes toward statistics may influence students' attitudes toward statistics, which are typically negative (Nassar, 2004; Roberts & Bilderback, 1980). Important is to understand teachers' attitudes towards statistics and how they affect other areas of teachers' statistical knowledge for teaching, providing direction for teacher preparation and professional development programs. Currently, little research on teachers' attitudes towards statistics exists (Estrada et al., 2011). Much of the research used Likert scale survey instruments and interviews.

Estrada et al. (2011) stated that the scarce research on teachers' attitudes toward statistics focused on three themes: global attitudes toward statistics, attitudes toward professional development in statistics, and attitudes concerning teaching statistics. Begg and Edwards (1999) focused on the former theme in a small-scale study on 22 ISMTs and 12 PSMTs. Begg and Edwards (1999) used unstructured, semi-structured, and clinical interviews to understand teachers' global attitudes toward statistics. The interviews' results indicated that most of the sample considered statistics important. Four themes emerged from the interviews: statistics (1) helps people make sense of our world, (2) helps people compare and organize things, shows trends and enables people to predict, (3) assists in planning for the future, and (4) helps in summarizing a lot of information. Most teachers related their attitudes toward statistics to their prior experiences, which indicated a negative attitude toward statistics. Overall, teachers held negative attitudes toward statistics. Results showed that teachers feared statistics, were not interested in statistics and considered themselves not good with statistics. In addition, Begg and Edwards (1999) found that teachers believed someone did not need a good understanding of mathematics to learn basic statistical concepts. Details on how the interviews were coded and analyzed were left out. Thus, results are limited.

Chick and Pierce (2008) studied how 27 primary preservice teachers used real-world statistical data to understand their attitudes toward statistics education. The teachers were enrolled in a mathematics education course for preservice primary school teachers. The teachers had completed three semesters of mathematics education courses that included topics of statistics at the primary level. The preservice teachers were asked to create a lesson plan for a sixth-grade mathematics class that included real-world data. Chick and Pierce (2008) used qualitative methods to analyze teachers' lesson plans using real-world data in a mathematics classroom. The results showed that the most common feature was class discussions and students sharing their findings in small groups. They mentioned that "While they are certainly suitable pedagogical strategies (class and small group discussions) . . . the efficacy will only be maximized if the teacher has a clear idea of the important points that the discussion needs to address" (Chick & Pierce, 2008, p. 5). The researchers found three lesson plans emphasized data and interpreting results, two lesson plans focused on creating graphical representations of data, and the remaining lesson plans lacked statistical facts and skills. In addition to the lesson plans, Chick and Pierce (2008) administered a 17-item five-point Likert scale survey. The researchers added ten items from the Statistics Course Attitudes Scale (SCAS) and seven items specifically related to teaching statistics at primary school and the student's background in mathematics. The SCAS instrument explores students' understanding, attitudes, and beliefs about statistics, including its relevance to functioning in society (Garfield, 1996). Approximately 30% of the PSMTs gave positive responses across all items, and 52% of PSMTs expressed neutrality or agreement when asked if they trusted statistics. In addition, approximately 44% of the PSMTs expressed neutrality or disagreement when asked if they felt they had sufficient knowledge of statistics.

Estrada et al. (2005) and Hannigan et al. (2013) used the Survey of Attitudes Toward Statistics (SATS) instrument in their studies. The SATS instrument utilizes a 7-point Likert scale to measure 36 items on statistical content found in primary and secondary school across six subscales: affect, cognitive competence, difficulty, value, effort, and interest. The subscale affect measures the individuals' feelings concerning statistics, cognitive competence measures the individuals' attitudes about their intellectual knowledge and skills when applied to statistics, difficulty measures individuals' attitudes about the difficulty of statistics as a subject, value measures individuals attitudes about the usefulness, relevance, and worth of statistics in personal and professional life, effort measures amount of work the individual expends to learn statistics and interest measures individuals' level of individual interest in statistics. The validation of the subscales affect, cognitive competence, difficulty, and value was established by Schau et al. (1995) and Cashin and Elmore (2005). Schau et al. found the subscales exhibited reasonably high Cronbach's coefficient alpha values: affect (.80 - .89), cognitive competence (.77 - .88), value (.74 - .90), and difficulty (.64 - .81). Hilton et al. (2004) obtained a sample of 5,360 undergraduate students enrolled in an introductory statistics course at Brigham Young University over a four-semester period. The mean age of students was 21.4 (*SD* = 2.6) where 51% male and 49% female. The SATS instrument was administered at the beginning and end of each semester. Researchers found that Cronbach's coefficient alpha values were high and remained consistent across subscales from the beginning of semester to end of semester: affect (.83 to .89), cognitive competence (.84 to .84), value (.87 to .90), and difficulty (.72 to .76). For the other two subscales, Interest and Effort, Tempelaar et al. (2007) used confirmatory factor analysis to determine that the six-factor model fits the data well and that the two new components, interest and effort, are valuable additions to the original instrument.

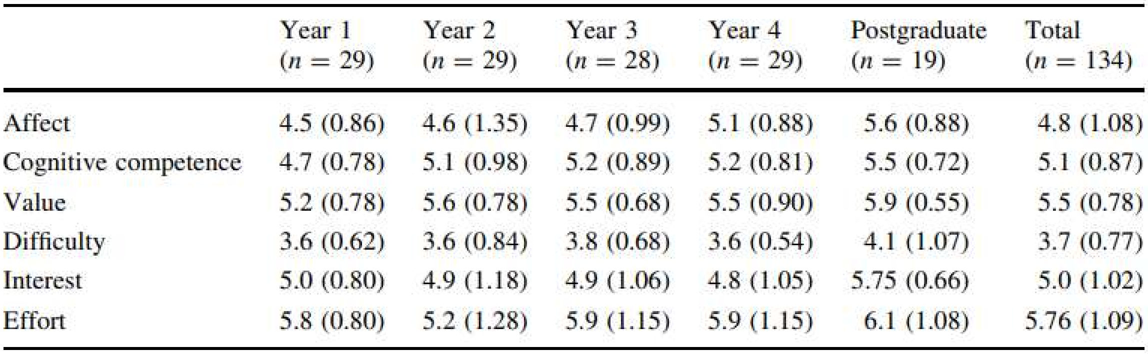
Estrada et al. (2005) administered the instrument to 367 primary preservice teachers to understand their attitudes toward statistics. The preservice teachers were studying different specialties at a teacher college in Spain. All preservice teachers were enrolled in a mathematics methods course and were administered the SATS instrument before they were taught the statistics unit of the course. The researchers used exploratory factor analysis on the data set. The first factor was affect and cognitive components that explained 26.2% of the total variance and included most items in the affective and cognitive components. This may suggest teachers' affection towards statistics was conditioned by their understanding of the topic. The second factor was value, which groups all items that present statistics as an important tool in the different domains. This factor represents teachers' attitudes on the relevance of statistics and explained 7.9% of the total variance. The third factor was affect and cognitive components, which was repeated from the first factor with slight variations. The fourth and fifth factor was difficulty and grouped most statements related to sources of difficulty in statistics. Results from this factor show that many students associate the difficulty of statistics with mathematical technical features, and professional relevance of statistics was perceived in inverse relation to degree of difficulty of the subject matter. Overall, results from the study showed positive moderate correlations between affect and cognitive (r = .78), and affect and value (r = .64). In addition, Estrada et al. (2005) found that teachers who were not confident in learning statistics had negative attitudes toward statistics.

Estrada and Batanero (2008) expanded on Estrada et al.'s (2005) study to understand why teachers yielded high or low scores on the SATS instrument. The researchers compared the SATS results of 66 ISMTs and 74 PSMTs. The independent variables used were gender, group (ISMT and PSMT), number of mathematics courses completed where statistics was included, topic in which the prospective teachers were specializing or topic the teachers taught, and number of years of teaching experience in mathematics. Analysis of Variance informed teachers' attitudes seemed to get worse with the actual practice of teaching. There was found to be a positive correlation between the number of previous mathematics courses with a statistics component and positive attitudes. The researchers found a negative correlation between the number of years of teaching (ISMTs only) and positive attitudes (r = -0.37); that was, teachers' attitudes were worse as teaching experience increased. *Z*-scores were computed for overall and subscale scores by subtracting the theoretical mean (position of indifference) from the mean and dividing it by the standard deviation. The *z*-scores showed that the teachers had positive attitudes overall ($z = 1.4$) and for each subscale: Affect ($z = 0.88$), Cognitive ($z = 1.53$), Value ($z = 0.91$), and Difficulty ($z = 0.85$). In short, Estrada and Batanero (2008) found that both groups (PSMTs and ISMTs) had moderate or positive attitudes toward statistics.

Hannigan et al. (2013) administered the SATs instrument to 134 PSMTs and found most teachers had positive attitudes toward statistics. Approximately 86% of the PSMTs were undergraduates at a teacher college and the remaining ($N = 19$) were postgraduate students on a 1-year graduate diploma program in mathematics education. There were 29 students who were in their first year, 29 students in their second year, 28 students in their third year, and 29 students in their fourth year. The researchers selected the SATS instrument because it has been used more frequently in studies of PSMTs and ISMTs. The researchers also asked demographics and prior mathematical achievement questions. The researchers used one-way analysis of variance to compare the means across the five groups (Year 1, Year 2, Year 3, Year 4, and Postgraduate students). Post hoc analysis was carried out using Tukey's HSD multiple comparison test. Independent samples tests were used to compare means for group Year 1 and Year 2 and also for males and females. Of the 134 students, 116 completed the SATS survey. The mean scores for each of the six components reflected positive attitudes (see Table 2).

Table 2

Mean (SD) for Each Component by Year Group



	Year 1 ($n = 29$)	Year 2 ($n = 29$)	Year 3 ($n = 28$)	Year 4 ($n = 29$)	Postgraduate ($n = 19$)	Total ($n = 134$)
Affect	4.5 (0.86)	4.6 (1.35)	4.7 (0.99)	5.1 (0.88)	5.6 (0.88)	4.8 (1.08)
Cognitive competence	4.7 (0.78)	5.1 (0.98)	5.2 (0.89)	5.2 (0.81)	5.5 (0.72)	5.1 (0.87)
Value	5.2 (0.78)	5.6 (0.78)	5.5 (0.68)	5.5 (0.90)	5.9 (0.55)	5.5 (0.78)
Difficulty	3.6 (0.62)	3.6 (0.84)	3.8 (0.68)	3.6 (0.54)	4.1 (1.07)	3.7 (0.77)
Interest	5.0 (0.80)	4.9 (1.18)	4.9 (1.06)	4.8 (1.05)	5.75 (0.66)	5.0 (1.02)
Effort	5.8 (0.80)	5.2 (1.28)	5.9 (1.15)	5.9 (1.15)	6.1 (1.08)	5.76 (1.09)

Note. From “An investigation of prospective secondary mathematics teachers’ conceptual knowledge of and attitudes towards statistics” by A. Hannigan, O. Gill, and A. M. Leavy, 2013, *Journal of Mathematics Teacher Education*, *16*(6), p. 439. Copyright 2013 by Springer Science+Business Media Dordrecht.

The researchers found a statistically significant difference between the means for the five groups for Affect ($F(114) = 4.04$, $p = 0.004$) and Value ($F(113) = 3.01$, $p = 0.02$). A post hoc analysis revealed that the main differences were between the postgraduate students and the first-, second- and third-year students for Affect, and between postgraduate students and first-year students for Value (Hannigan et al., 2013). In general, postgraduate students had the most positive attitudes of the five groups, and mean scores tended to increase from group Year 1 to Year 2 (Hannigan et al., 2013). There were no statistically significant differences between the means for any of the six components for students in Year 1 and Year 2. Researchers concluded that the PSMTs placed value on statistics, showed interest, and had confidence in statistical knowledge. Also, the researchers found that students tended to disagree with statements such as statistics is a subject quickly learned by everyone and agreed with statements such as learning statistics requires a great deal of discipline. Overall, Hannigan et al. (2013)

found positive attitudes toward mathematics correlated with positive attitudes toward statistics.

Onwuegbuzie (1998) administered the Attitudes Toward Statistics (ATS) instrument to 222 in-service teachers who were enrolled in an introductory graduate course in statistics and research methods. The majority (63%) of teachers had not taken an undergraduate statistics course. The ATS instrument utilizes a 5-point Likert scale on 29 items. There were two subscales: Field (20 items), and Course (19 items). The Cronbach's alpha coefficients for Field and Course were .92 and .90, respectively (Wise, 1985). In addition, Wise (1985) used factor analysis showing two factors explained 49% of total variance indicating construct validity. Onwuegbuzie (1998) administered the ATS on the first day of class to compare the findings to that of Shultz and Koshino's (1998) study. Schultz and Koshino (1998) administered the ATS instrument to 38 psychology graduate students. Results showed teachers scored lower on the Field subscale compared to graduate students ($t(258) = 4.9$, $p < .001$, $d = 0.86$). Teachers had a mean of 72.4 ($SD = 10.2$) and the graduate students had a mean of 81.1 ($SD = 9.2$).

In addition, the teachers scored lower on the Course subscale than the graduate students ($t(258) = 3.1$, $p < .001$, $d = 0.55$). The teachers had a mean of 25.3 ($SD = 8.1$) and graduate students had a mean of 29.8 ($SD = 8.9$). In addition, a statistically significant correlation was found between the number of courses in statistics teachers completed and teachers' global attitudes toward statistics ($r = .28$, $p < .001$). Onwuegbuzie (1998) also found that teachers who completed at least one undergraduate statistics course had statistically significantly less positive global attitudes toward statistics than did graduate students ($t(139) = 2.1$, $p < .001$, $d = .58$). The researcher acknowledged that replicated studies are needed to assess external validity of the findings.

Begg and Edwards (1999) and Watson (2001) found that teachers had poor to neutral attitudes toward teaching statistics. As mentioned earlier, Begg and Edwards (1999) interviewed and surveyed 22 in-service and 12 preservice primary teachers. Approximately 75% of in-service teachers felt confident in teaching statistics. Teachers were confident in their statistical ability and showed no interest in professional development in teaching and learning statistics.

Watson (2001) administered a survey to 15 primary teachers and 28 secondary teachers to determine if professional development was needed to keep up with the changes in the mathematics curriculum. Watson (2001) designed the survey, which included Likert scale items and open-ended questions over 10 sections: (1) significant factors for teaching chance and data, (2) preparing to teach a unit in chance and data, (3) preparing to teach a lesson in chance and data, (4) teaching practices, (5) the topic of sampling, (6) confidence, (7) beliefs about statistics in everyday life, (8) student survey items, (9) teacher background, and (10) professional development. The researcher found that the teachers identified previous experience, current experience, and pedagogy as factors for teaching chance and data. Approximately 53% of teachers selected the same topics regarding enjoyable topics related to chance and data. Most teachers (75%) acknowledged using concrete materials, materials for exploring chance outcomes, various sources of data, and computers or calculators.

For the confidence section, the teachers rated their level of confidence in their ability to teach nine topics on chance and data using a 5-point rating scale ranging from 1

= *low confidence* to 5 = *high confidence*. The researcher found teachers were confident in teaching basic statistical concepts, such as graphical representations, and not confident in teachng topics on probability and odds. The primary teachers' mean scores ranged from 3.00 for *median* to 3.92 for *data collection*. The secondary teachers' range was from 3.68 for *odds* to 4.59 for *median*. All means for secondary teachers, except for *odds*, were higher than the highest mean (3.92) for primary teachers.

For the professional development section, teachers were asked to respond to three questions. The questions asked the teachers if they participated in professional development on chance and data, what types of professional development opportunities they preferred, and what qualifications the leaders of professional development should possess. Additionally, the teachers had the opportunity to make comments about professional development in chance and data. The results showed that at least 47% of teachers had not attended a professional development on chance and data. Regarding who should lead professional development on chance and data, approximately 21% suggested another teacher at the school, 21% suggested a regional curriculum officer, and 51% suggested an outside expert. Many teachers indicated more than one choice. Many teachers expressed negative attitudes when discussing professional development and identified the need for professional development in teaching and learning statistics. In addition, teachers had the opportunity to do a volunteer interview. The format was structured, and the teachers were asked about various statistical topics over 45 minutes. The interviews informed that the teachers had positive attitudes toward attending professional development opportunities but negative attitudes toward leading professional development opportunities.

More research needs to be conducted on attitudes toward statistics to motivate teachers to participate in professional development opportunities (Gould & Peck, 2004). Another noncognitive factor of the affective domain that might influence teachers' statistical knowledge and willingness for professional development in teaching and learning statistics is their anxiety about statistics. However, little research exists on mathematics teachers' anxiety about statistics. Understanding mathematics teachers' anxiety about statistics will shed light on how teacher preparation programs are introducing statistical content and pedagogical skills needed to teach statistics effectively.

Anxiety about Statistics

Much research on anxiety about statistics is conducted on undergraduate or graduate students (Keeley et al., 2008; Yusuf et al., 2019). Anxiety about statistics is the feeling of anxiety experienced by students when taking a statistics course or performing statistical analysis (Cruise et al., 1985; Yusuf et al., 2019). Onwuegbuzie (2003) identified anxiety towards statistics as an apprehension occurring when the student encounters statistics, which could lead to low performance in a statistics course. Anxiety about statistics occurs when someone "experiences fears, dislikes, mental disorders, tensions, and mental emotions when dealing with concepts, problems, learning, and statistical evaluation" (Yusuf et al., 2019, p. 697). In summary, Yusuf et al. (2019) stated anxiety about statistics occurs when someone who was learning and applying statistics becomes worrisome and fearful.

Research has shown that students' anxiety about a subject was negatively associated with their learning and performance. Bostani et al. (2014) administered a mental health questionnaire to 200 undergraduate students. The questionnaire consisted of 28 multiple choice items classified into four subscales: Depression, Anxiety, Physical Symptoms, and Disorder Social Performance. Each subscale consisted of seven questions. The Cronbach's alpha coefficient was .96 overall, and .90 for the subscale Anxiety. In addition, the questionnaire asked students to provide age, gender, year of university entrance, major, academic degree, and total grade-point average. Total grade-point average was used as the measure for academic performance. An independent-samples *t*-test was conducted to compare academic performance between students who experienced anxiety about statistics and students who did not experience anxiety about statistics. The results showed a significant difference in academic performance for the two groups ($t(199) = 2.70$, $p = .04$).

Vitasari et al. (2010) administered a questionnaire to 770 undergraduate students to identify possible study anxiety sources. Study anxiety refers to anxiety that affects students' academic performance while they pursue a college major. The questionnaire utilized a 5-point Likert scale to measure 40 items on students' study anxiety across five subscales: Exam, Presentation, Mathematics, Language, and Social. The questionnaire asked students about their experiences, feelings and thoughts related with anxieties during study process. The Cronbach's alpha coefficient was .93. The researchers used descriptive statistics to explain the five potential sources of study anxiety. The results showed Exam anxiety was the leading source ($M = 1870.29$, $SD = 99.52$). Presentation anxiety was the second leading source ($M = 1715.20$, $SD = 108.99$). Mathematics anxiety was the third leading source ($M = 1694$, $SD = 152.49$). The last two sources were, in order, Language ($M = 1672.20$, $SD = 90.67$) and Social ($M = 1463.43$, $SD = 97.34$) anxieties.

Macher et al. (2013) administered the Statistical Anxiety Rating Scale (STARS) instrument on 284 undergraduate psychology students. Majority (79%) were female. The STARS instrument "distinguishes between six factors: (1) *test and class anxiety*: anxiety experienced when taking a statistics examination or when attending a statistics class, (2) *interpretation anxiety*: anxiety when being face with making decisions from or interpreting statistical data; (3) *fear of asking for help*: anxiety when intending to ask for help on a statistical problem, (4) *attitudes toward statistics teachers*, (5) *computational self-concept*: self-rated ability to master statistical tasks, and (6) *worth of statistics*: perceived usefulness of statistics. These six factors form the six subscales of the STARS instrument. The instrument utilizes a 5-point Likert scale to measure 51 items across six subscales: Worth of Statistics, Interpretation Anxiety, Test and Class Anxiety, Computation Self-Concept, Fear of Asking for Help, and Fear of the Statistics Instructor. In addition to the STARS instrument, Macher et al. (2011) asked students their interests in statistics through a questionnaire and found that there was a statistically significant relationship between statistics anxiety and academic performance ($r = -.416$, $p < .001$), statistics anxiety and interest in statistics ($r = -.364$, $p < .001$), and statistics anxiety and computation self-concept ($r = -.155$, $p < .001$).

Onwuegbuzie's (2003) investigated 130 graduate students' anxiety towards statistics and how it related to their statistics achievement. The students were enrolled in an introductory quantitative research design course and were from different educational disciplines: early childhood education, elementary education, special education, and psychology. Results were limited as participation in the study was voluntary. The mean age of students was 26.0 ($SD = 6.8$) and majority (97%) were female. The researcher administered the STARS instrument. The subscale Computation Self-Concept was the only subscale administered to the graduate students. This subscale measures the anxiety experienced when attempting to solve and understand statistics problems. A high score on this subscale implied a high level of anxiety associated with poor Computation Self-Concept. Cronbach's alpha coefficient for Computation Self-Concept was .86. Students' statistical achievement was measured through an exam given in the course. The results showed that the mean for Computation Self-Concept was 18.41 ($SD = 6.27$) and the mean for the statistics achievement exam was 89.08 ($SD = 11.19$). The researcher found that there was a significant relationship between Computation Self-Concept and students' statistics achievement ($r = -.24$, $p < .001$).

Zare et al. (2011) conducted a similar study on undergraduate students. Zare et al. (2011) measured 323 undergraduate students' anxiety about statistics and statistics teaching efficacy. The researchers administered the STARS instrument to measure students' anxiety about statistics and a questionnaire to measure students' self-efficacy to learn statistics. Both instruments were given at the start of statistics courses. The final grade for the statistics course was used as the educational achievement index. The results showed a statistically significant relationship between students' self-efficacy to learn statistics and academic performance ($r = .47$, $p =.01$), and students' anxiety about statistics and academic performance ($r = -.41$, $p = .01$).

Williams (2010) used a pretest-posttest control group design on 76 graduate students enrolled in a statistics course to determine the relationship between instructor immediacy and graduate students' anxiety towards statistics. There were four sections of the statistics courses; two served as the treatment group ($N = 38$), and two served as the control group ($N = 38$). Christophel and Gorham (1995) defined instructor immediacy as "verbal and nonverbal behaviors that reduce psychological, physical, or both distance between teachers and students" (p. 292). Instructor immediacy served as the independent variable, and students' anxiety toward statistics was the dependent variable. The instructor of the treatment group practiced instructor immediacy throughout the course, with the instructor of the control group not knowing. The STARS instrument was administered to measure students' anxiety towards statistics at the beginning of the course (pretest) and the end of the course (posttest). In addition to the STARS posttest, the Instructor Immediacy scale was administered as part of the posttest. The Instructor Immediacy scale measures verbal and nonverbal indicators of instructor immediacy utilizing a 5-point Likert scale to measure 23 questions across two subscales: Verbal (17 items) and Nonverbal (6 items). High scores indicate higher levels of instructor immediacy. Cronbach's alpha coefficients were .89 for Verbal items and .73 for Nonverbal items. Results showed that the treatment group had significantly higher levels of instructor immediacy ($M = 4.26$) compared to the control group ($M = 3.75$) ($t(74) = 4.48$, $p < .001$).

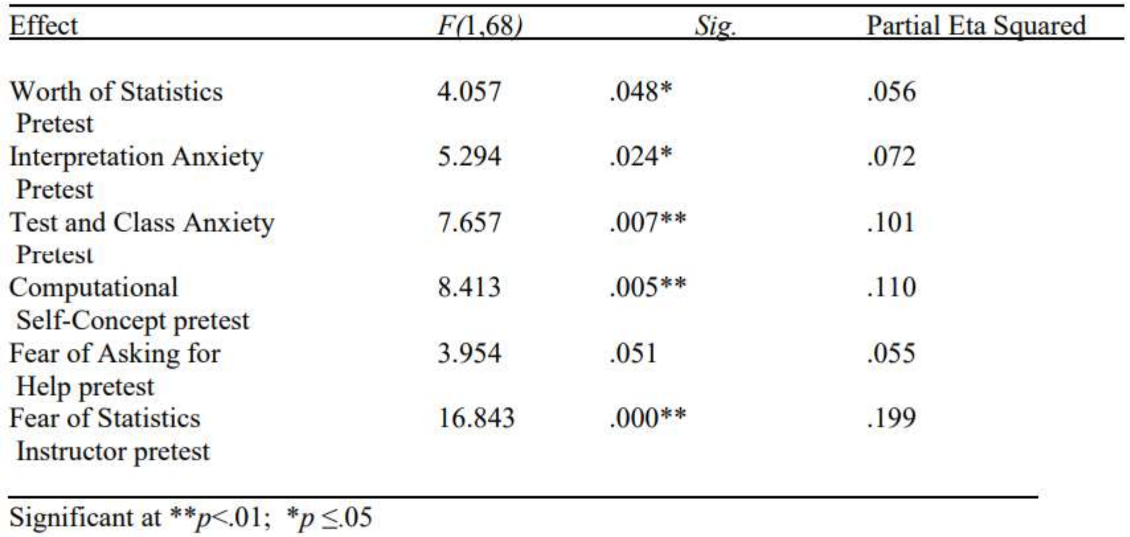
Williams (2010) conducted a one-way multiple analysis of covariance (MANCOVA) test to test the prediction that students who reported higher levels of instructor immediacy would also report lower levels of anxiety towards statistics. There were two groups (control and treatment). The STARS subscales (Worth of Statistics, Interpretation Anxiety, Test and Class Anxiety, Fear of Asking for Help, Computation Self-Concept, Fear of Statistics Teacher) posttest scores were used as the continuous dependent variables. The STARS subscales pretest scores were used as the six continuous covariates. The results showed a statistically significant difference between the treatment and control groups on the combined dependent variables (p = .011). The covariates of subscales Worth of Statistics (p < .001), Interpretation Anxiety (p = .002), Test and Class Anxiety (p = .002), and Computational Self-Concept (p = .002) significantly influenced the combined dependent variable. The covariates of subscales Fear of Asking for Help (p = .116), and Fear of Statistics Teachers (p = .724) did not influence the dependent variables. The researchers conducted an analysis of covariance (ANCOVA) test on each dependent variable as a follow-up test to MANCOVA. Results showed that immediacy differences were significant for all six STARS subscales (see Table 3).

In summary, the researcher found that instructor immediacy was significantly related to students' anxiety toward statistics as outlined by each subscale from the STARS instrument (p = .01). This implied that the instructor had an influence on students' anxiety toward statistics which influenced academic performance.

Primary teachers have high levels of anxiety towards mathematics (Hembree, 1990; Kelly & Tomhave, 1985). In Onwuebuzie et al.’s (1997) phenomenological study on factors that may relate to teachers’ anxiety about statistics, the researchers hypothesized that anxiety about mathematics may be related to anxiety about statistics (Onwuebuzie et al., 1997). Thus, primary education students may have high levels of anxiety about statistics. The anxiety about statistics and mathematics felt by primary education students can affect their ability to teach the subject with confidence (Brady & Bowd, 2005). There is one study regarding teachers’ anxiety on teaching statistics. More research studies are needed to guide professional development and teacher education programs.

Table 3

ANCOVA for each Dependent Variable



Effect	*F*(1,68)	*Sig.*	Partial Eta Squared
Worth of Statistics Pretest	4.057	.048*	.056
Interpretation Anxiety Pretest	5.294	.024*	.072
Test and Class Anxiety Pretest	7.657	.007**	.101
Computational Self-Concept pretest	8.413	.005**	.110
Fear of Asking for Help pretest	3.954	.051	.055
Fear of Statistics Instructor pretest	16.843	.000**	.199

Significant at $**p<.01$; $*p \leq .05$

Note. From “Statistics anxiety and instructor immediacy” by A. S. Williams, 2010, *Journal of Statistics Education, 18*, p. 5. Copyright 2010 by Informa UK Limited.

Yusuf et al. (2019) conducted a mixed-method research study on 33 PSMTs teachers, where 70% were female. The PSMTs were enrolled in a basic statistics course. The researchers administered an anxiety about statistics questionnaire and a statistical reasoning test to the PSMTs. The anxiety about statistics questionnaire had three subscales: Taking the Course, Learn and Practice, and Exam. The researchers followed their quantitative analysis of the assessments with interviews and observations. Results found that the overall anxiety about statistics level was moderate in both men ($M = 20.5$) and women ($M = 21.30$). The subscale Taking the Course was moderate in both men ($M = 2.7$) and women ($M = 3.5$), Learn and Practice was moderate in both men ($M = 14.7$) and women ($M = 10.7$), and Exam was high in both men ($M = 3.1$) and women ($M = 2.9$). The results were confirmed by interviews and observations during learning. The PSMTs felt comfortable during statistical learning and felt confident in understanding and completing assignments with the guidance of the course. The PSMTs stated their anxiety levels were high when thinking about the exam because they were afraid of not doing well on the exam. Yusuf et al. (2019) stated that "students feel tense when they are about to face exams and have to prepare better for the exam, this means statistical anxiety becomes the motivation for students to learn" (p. 698).

Yusuf et al. (2019) analyzed the relationship between PSMTs' anxiety about statistics and statistical reasoning. A one-way ANOVA demonstrated that PSMTs' anxiety about statistics did not directly affect the ability of statistical reasoning ($F(1, 33) = .021$, $p = .886$). Follow up interviews showed that PSMTs were confused about some items on the statistical reasoning exam that focused on analyzing and interpreting data. The PSMTs did not feel worrisome or fearful but were confused on some items. Observations throughout the course informed the researchers of indicators that influenced statistical anxiety and reasoning. These indicators included: making conclusions based on mathematical solutions, explaining statistical findings, and interpreting results. In summary, the researchers found that PSMTs experienced anxiety when thinking about their performance in the course, such as exams. Limitations of the study include small sample size and convenience sampling. In addition, there was lack of details regarding the items of the questionnaire and reasoning instrument. Although this is the only study regarding PSMTs' anxiety about statistics, results are limited. More research needs to be conducted on teacher's anxiety about statistics to provide teacher preparation programs with knowledge on if different approaches are needed to introduce statistical content and pedagogical content knowledge to ease the anxiety felt among teachers regarding statistics.

Summary

Teacher training and professional development programs are challenged to develop mathematics teachers with mathematical and statistical knowledge for teaching. Educational systems increased students' expectations of learning statistics in primary and secondary school. The teaching and learning of statistics revolve around the statistical investigative cycle. The investigative cycle is the primary focus on Groth’s (2013) SKT framework, which is the theoretical framework of this study. Although there exist other SKT frameworks, Groth’s (2013) framework is aligned to the GAISE II report. The GASIE II report was used to develop statistical content found in K-12 mathematics curriculum across many educational systems (Bargagliotti et al., 2020). Developing key understandings of statistical content was the foundation of Groth’s (2013) framework. The foundation leads to two components: statistical content knowledge and pedagogical content knowledge.

Researchers have shown how cognitive and noncognitive factors influence teachers' statistical knowledge of teaching. Teachers lack the statistical content knowledge that they are required to teach (Harrel-Williams et al., 2015; Lovett & Lee, 2017b). Researchers suggested that PSMTs across all levels receive more training in statistics during their teacher preparation programs. This training would include more courses on statistics and statistical education, and workshops on teaching and learning statistics. Although development of teachers’ statistical content knowledge would impact the teaching and learning of statistics, it should not be considered the sole contributor (Groth, 2017; Hotaman, 2010; Kelly & Tomhave, 1985; Lee et al., 2020).

Teachers with positive beliefs, attitudes, and actions toward their instruction and content positively affect students' learning and attitudes toward the content (Gourneau, 2005; Williams, 2010; Zientek et al., 2011). Research has shown that mathematics teachers are overly confident in teaching statistics (Harrel-Williams et al., 2015; Lovett & Lee, 2017b). Estrada et al. (2011) stated that PSMTs' attitudes toward teaching and learning statistics have been measured on various instruments, but results were inconsistent. Researchers have analyzed many factors of the affective domain on PSMTs' statistical knowledge for teaching, with much focus given to teaching efficacy, attitudes, and anxiety statistics (Estrada et al., 2011). Akin and Kurbanoglu (2011) found that high self-efficacy in mathematics predicted positive attitudes toward mathematics and low self-efficacy predicted negative attitudes toward mathematics. In addition, the researchers found that positive attitudes toward mathematics predicted mathematics anxiety. A similar study on statistics would shed light on how to develop noncognitive factors for PSMTs across all levels in teacher preparation programs, and guide how to provide professional development opportunities for in-service mathematics teachers.

Chapter III

METHODOLOY

This chapter contains a description of the quantitative methods that were used in this study. This chapter is divided into six sections, starting with the study's research design, including the exogenous, potential mediator, and endogenous variables. The second section discusses the participants of this study, including the target population and accessible population. The third section focuses on the instrumentation of the research and the accuracy of the data by discussing the instruments' reliability and validity. The fourth section deals with the collection of data. The fifth section describes the data analysis. The chapter ends with a summary.

The research questions framing this study are:

RQ1: Is the hypothetical path model, which describes the causal effects among the PSMTs' attitudes toward statistics and statistics teaching efficacy on anxiety about statistics, consistent with the observed correlates among these variables?

RQ2: If the hypothetical path model is consistent, what are the estimated direct, indirect, and total effects among the PSMTs' attitudes toward statistics and statistics teaching efficacy on anxiety about statistics?

RQ3: Is the specified path model equivalent across various demographic characteristic variables?

Research Design

This study employed a nonexperimental, correlational research design using structural equation modeling to test how the constructs are theoretically linked and the directionality of the relationships. Path analysis examined the relationship between the observed variables (see Figure 2). Path analysis is a type of structural equation modeling that employs a single indicator for each observed construct, evaluating the effects through various causal pathways of the observed constructs impacting a particular dependent variable (Mueller & Hancock, 2019). Schumaker and Lomax (2016) explained that path analysis extends multiple regression by incorporating several regression equations and identifying direct, indirect, and total effects among the observed variables in the model. Unlike multiple regression, path analysis evaluates the overall fit of the model, indicating how well the model accounts for the data (Schumaker & Lomax, 2016). Path analysis examined the relationship between the exogenous variable, potential mediator variables, and the endogenous variable. The exogenous variable of interest was statistics teaching efficacy and the four subscales that measure attitudes toward statistics (affect, cognitive competence, value, and difficulty) were the potential mediator variables. These variables were investigated to determine the underlying structure of the endogenous variable anxiety about statistics.

Validated psychometric instruments measured the exogenous and potential mediator variables. Exogenous variables are defined as variables not influenced by another variable and are often referred to as independent variables. The exogenous variable in the hypothetical model was statistics teaching efficacy. The potential mediator variables in the hypothetical model were the four subscales of attitudes toward statistics (affect, cognitive competence, value, and difficulty). All variables were measured on the interval level scale.

The demographic characteristic variables included gender, mathematics experience, statistics experience, and college status. Gender is on the nominal level of measurement. Mathematics experience refers to the number of college-level mathematics courses that the student has completed. This includes Advanced Placement (AP), International Baccalaureate (IB), and dual-enrollment mathematics courses completed during high school. Statistics experience refers to the number of college-level statistics courses that the student has completed, including AP, IB, and dual-enrollment statistics courses completed during high school. College status refers to the number of years the student has completed in a college teacher preparation program. The variables mathematics experience, statistics experience, and college status were measured on the ordinal level scale.

A validated psychometric instrument measured the endogenous variable. Endogenous variables are variables that are influenced by other variables and are often referred to as dependent variables. The endogenous variable in this study was anxiety about statistics. The measurement of this variable was on the interval level scale.

Participants

The target population for this research study were primary preservice mathematics teachers (PSMTs) enrolled in a primary teacher education program in Georgia. There are over 45 colleges in Georgia, public and private, that offer elementary education program studies. This does not include colleges belonging to the Technical College System of Georgia. Some elementary education programs do not lead to a bachelor's degree, which is a requirement to teach elementary education in Georgia. Instead, these programs require students to transfer to another school to complete the remaining studies of the program. Some private schools offering elementary education studies are for-profit colleges with no accreditation. For these reasons, the accessible population were primary PSMTs enrolled in a primary teacher education program at a 4-year accredited non-profit university or college in Georgia. All primary education programs in Georgia will be contacted to participate in the study. The size of the programs varies, ranging from 20 primary PSMTs to over 100 primary PSMTs across all college grade-levels. Faculty of mathematics education courses were the primary target for distributing the survey packet.

Eligible primary PSMTs from such programs were recruited to participate voluntarily. An email was sent to primary teacher education program directors requesting contact information for faculty teaching primary education courses. It outlined the study's purpose, scope, participation requirements, and IRB approval from Valdosta State University (see Appendix A). Faculty received a similar email and a follow-up call to arrange the distribution of the survey packet. Faculty who agreed to participate as survey proctors shared the Qualtrics® survey packet link with their classes via email or their learning management system.

There are many guidelines regarding how many participants are needed for structural equation modeling. Schreiber et al. (2006) suggested 20 participants per variable for structural equation modeling (Kline, 2015; Suhr, 2006). Other researchers recommended at least 300 participants. (Comrey & Lee, 2013; Tabachnick & Fidell, 2013). This study aimed for 300 primary PSMTs enrolled in a primary teacher education program at a 4-year accredited non-profit university or college in Georgia.

Instrumentation

The study consisted of three instruments and four demographic characteristic variables. The demographic characteristic variables in the study were gender, mathematics experience, statistics experience, and college status. The three instruments measured PSMTs' statistics teaching efficacy, attitudes toward statistics across four subscales, and anxiety about statistics. The demographic characteristic variables and instruments were made into one survey packet (see Appendix D). The demographic characteristic variables were listed at the end of the survey packet following the three instruments.

Statistical Anxiety Scale

Primary PSMTs' anxiety about statistics was measured using the Statistical Anxiety Scale (SAS) instrument. Developed by Pretorius and Norman (1992), the SAS is a 10-item instrument revised version of the Mathematics Anxiety Scale used by Betz (1978). The instrument is measured on a 5-point Likert scale ranging from 1 = *strongly disagree* to 5 = *strongly agree*, where low scores indicate high levels of anxiety about statistics. Pretorius and Norman (1992) developed the SAS instrument by changing all words relating to mathematics in the validated Mathematics Anxiety Scale to words relating to statistics. For example, the statement "I usually don't worry about my ability to solve mathematical problems" was changed to "I usually don't worry about my ability to solve statistics problems." As another example, "I get a sinking feeling when I think of trying hard math problems" was changed to "I get a sinking feeling when I think of trying hard statistics problems."

Validity.

Pretorius and Norman (1992) administered the SAS instrument to 337 undergraduate students enrolled in an elementary statistics course at the University of Western Cape, South Africa. The mean age of the sample was 22.08, where 42% were male and 58% were female. The researchers administered the instrument during class. In addition, the researchers administered the validated State-Trait Anxiety Inventory (STAI), which measures generalized anxiety. The STAI instrument is a 20-item 4-point rating scale survey ranging from 1 = *almost never* to 4 = *almost always*, where high scores indicate high levels of trait anxiety. Both instruments were administered back-to-back in the same setting. The student's achievement in the statistics courses (pass or fail) was recorded at the semester's end.

Pretorius and Norman (1992) used a principal factor analysis with iterations and varimax rotation. The principal factor analysis showed the existence of a single factor. The single factor had an eigenvalue greater than one and accounted for approximately 91% of the variance of the scale. In addition, the researchers established the validity of the instrument by confirming that the SAS score was related to achievement in statistics by using a *t*-test that indicated a significant difference in terms of statistics anxiety ($t_{(266)} = 2.24$, $p < .05$) exists between students who passed and failed statistics. The researchers showed that the SAS instrument was significantly related to general anxiety as measured by the STAI ($r = 0.26$, $p < 0.5$). Individuals who tend to be anxious in various situations are more likely to report math and statistics anxiety (Betz, 1978).

Reliability.

Reliability was assessed in two ways. First, the Cronbach's alpha coefficient for the scale was found to be 0.90. Secondly, the researchers readministered the SAS instrument to 196 individuals. Pretorius and Norman (1992) found the test-retest reliability was 0.75 over a 3-month interval. In addition, Pretorius and Norman (1992) found the average intercorrelation between the items of the SAS was generally high (0.15 to 0.86). Approximately 75% of the item-total correlations were greater than 0.50.

Survey of Attitudes toward Statistics

Primary PSMTs' attitudes toward statistics were measured using the Survey of Attitudes toward Statistics (SATS) instrument. Developed by Schau et al. (1995), the SATS is a 28-item instrument to assess four subscales of attitudes toward statistics (affect, cognitive competence, value, and difficulty). The instrument is measured on a 7-point Likert scale ranging from 1 = *strongly disagree* to 7 = *strongly agree*. Affect measures students' expression towards the statistics course across six items. Items included interest statements, not feeling threatened, not disappointed, fun, and not stressed in solving statistical problems. Cognitive competence measures attitude toward knowledge and intellectual skill in using statistical knowledge across six items. Items included statements regarding individuals not having difficulties in understanding statistical concepts based on their way of thinking, learning statistics by making the least errors in calculation, and understanding statistical formulas and concepts. Value measures attitude towards a statistics course across nine items. Items included statements showing that statistics are useful, necessary, and relevant in individuals' daily lives and careers. Difficulty measures attitude towards the difficulties in understanding the statistics across seven items. Items included statements regarding how easy it is to understand the statistical formula and statistical computations.

The researchers developed the SATS instrument using the nominal group technique with a panel of two graduate and two undergraduate students enrolled in a statistics course and two introductory statistics teachers. The panel identified 92 words and phrases regarding attitudes toward statistics, 21 items from previous surveys on attitudes toward statistics, and comments on attitudes toward statistics from other students in an introductory statistics course. The panel sorted the data into four dimensions (subscales) and reached a consensus on 60 items across the four dimensions. The 60-item survey was administered to 132 introductory college statistics students. The researchers used correct item-total correlation coefficients and squared multiple correlation coefficients to find the best 32 items. The new 32-item survey was administered to 1402 introductory college statistics students. An item analysis was used to identify problem items. The researchers found two problematic items and removed them from the survey. One item from the subscale value and one from the subscale difficulty.

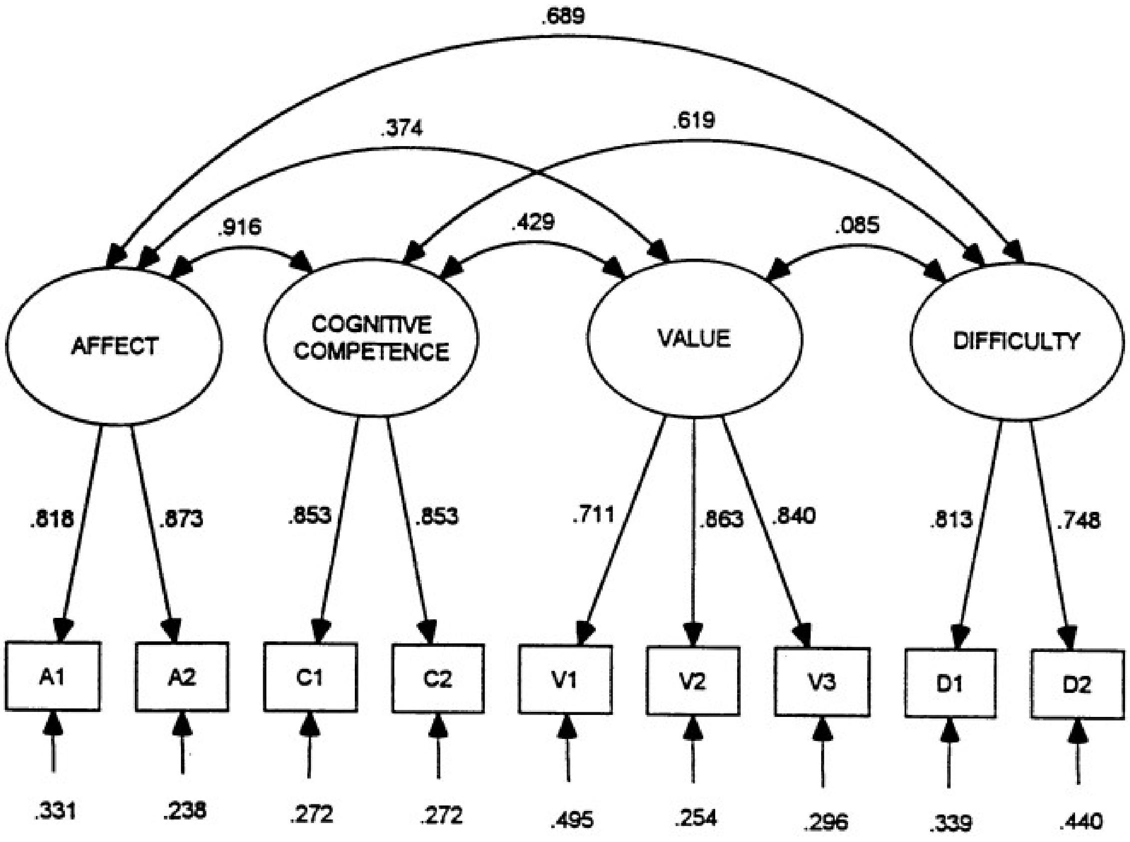
Validity.

An item-based confirmatory-factor analysis (CFA) identified two more problematic items, one from affect and one from cognitive competence. Both items had low factor loading on their hypothesized factor and were excluded from the survey. There were no other problematic items. The researchers concluded with a 28-item survey.

The researchers grouped items within each subscale from the 28-item survey into parcels: two parcels for affect items (A1 and A2), two parcels for cognitive competence items (C1 and C2), three parcels for value items (V1, V2, and V3), and two parcels for difficulty items (D1 and D2). Each parcel consisted of three items, except parcel D1 had four items. The parcels were balanced with respect to the number of negatively worded items and the size of each parcel's mean, standard deviation, and skewness. The researchers used data screening on the univariate and multivariate distributions of the parcel scores identifying eight outliers. The outliers were excluded

Figure 8

Four-Factor Model of Attitudes toward Statistics.



Note. From "The development and validation of the survey of attitudes towards statistics" by C. Schau, J. Stevens, T. Dauphine, and A. del Vecchio, 1995, *Educational and Psychological Measurement, 55*(5), p. 874. Copyright 1995 by SAGE Publications.

from the data analysis. A Chi-square goodness of fit test was performed to determine whether the proportion was equal between ($X^2(21,798) = 48.68$). The hypothesized structure (see Figure 8) provided a good model fit with adjusted goodness of fit of .97 and a root mean square residual of .03. These results showed that the structure of the SATS demonstrated a four-factor structure that provided a good fit to the observed interrelationships among parcels of items.

Reliability.

Cronbach's alpha coefficients were determined to indicate the reliability of the subscales. The initial 32-item SATS instrument resulted with the following Cronbach's alpha coefficients: .81 for the affect items, .77 for the cognitive competence items, .80 for the value items, and .64 for the difficulty items. After the removal of items to obtain the 28-item SATS instrument, the Cronbach's alpha coefficients increased across each subscale: .81 to .85 for the affect items, .77 to .83 for the cognitive competence items, .80 to .85 for the value items, and .64 to .70 for the difficulty items.

Statistics Teaching Efficacy

Harrell-Williams et al. (2013) developed the Self-Efficacy to Teach Statistics (SETS) instrument that was utilized for this study. The SETS measures levels of statistics teaching efficacy corresponding to GAISE level A (11 items) and level B (15 items). The SETS instrument is a 26-item instrument assessing two subscales: reading data and reading between data. The instrument is measured on a 6-point rating scale ranging from 1 = *not at all confident* to 6 = *completely confident.* The subscale reading data represents the GAISE level A items, and the subscale reading between data represents the GAISE level B items. GAISE level A covers topics found in primary school, and GAISE level B covers topics found in middle school. The subscale reading data represents items about graphical representations, and reading between data represents advanced knowledge on interpretation and integration of the data in a graph. Since this study is on primary PSMTs, the subscale reading between data was excluded from the instrument as it pertains to middle-school PSMTs.

A pilot study was carried out to confirm the reliability of measuring primary PSMTs' statistics teaching efficacy using the 11 items from subscale reading data. The 11 items from subscale reading data were uploaded to Qualtrics®. A faculty member who taught primary PSMTs at Valdosta State University was contacted via email to participate in the pilot study. The email provided details on the purpose and scope of the pilot study. The faculty member agreed to participate in the pilot study. The faculty member was emailed the link to the electronic survey and asked to disperse it to their class. After administering the survey, the data was formatted into an Excel file and imported into the R DataFrame. The 11-item survey was completed by 29 primary PSMTs. The subscale reading data was found to be highly reliable (11 items; $\alpha = .92$).

Validity.

Harrell-Williams et al. (2013) developed the SETS instrument by comprising a list of behaviors representing state mathematics standards and students' statistics learning objectives related to GASIE levels A and B. The researchers developed a preliminary survey of items from the list of behaviors. The researchers recruited 12 in-service middle and primary school teachers to comment on the survey's wording. The teachers were enrolled in a specialist K-8 graduate program in teacher education. The teachers' teaching experience ranged from 2 to 19 years, with a median of 6 years. The teachers provided revisions and submitted the revised survey to the researchers. The revised survey was administered to 23 middle school PSMTs and 21 noneducation majors at a public university who were enrolled in an introductory statistics course.

The researchers used confirmatory factor analysis to analyze the data by scaling the data to the multidimensional random coefficients multinomial logit model using two correlated latent traits. In addition, the data were scaled to a unidimensional rating scale model to model statistics teaching efficacy as a single latent trait. The researchers estimated item and teacher parameters for the factor structures, and Linacre's (2004) criteria for rating scale analysis was performed. The researchers found that evidence for content, substantive evidence, and structural aspects of validity was present in the analysis (Harrell-Williams et al., 2013).

Structural evidence was confirmed using the Rasch framework and confirmatory factor analysis. The researchers found that the Akaike Information Criteria (AIC) and Bayesian Information Criteria (BIC) were smaller for the two-dimensional model (AIC = 21541, BIC = 21558) compared to the unidimensional model (AIC = 21765, BIC = 21780). Hence, the results support the two-dimensional model. Item-measure correlations and mean square fit indices were used to evaluate the quality of items. This provided evidence of content validity (Wolfe & Smith, 2007).

Reliability.

The researchers found that the direction and magnitude of the Fisher transformation of the point-polyserial correlations of each item were greater than .67, which indicates a strong positive relationship between the polytomous item score and subscale score based on remaining items (Harrell-Williams et al., 2013). Mean square statistics for each item were in the acceptable range of 0.5 to 1.5 (Linacre, 2011). Thus, the researchers found no problematic items. In addition, the researchers found that the reliability of separation for the subscales reading data and reading between data were .87 and .91, respectively, and the overall scale was .94. Harrell-Williams et al. (2013) stated that the reliability of separation coefficients showed little measurement error as they are analogous to Cronbach's alpha. Batur et al. (2020) used Cronbach's alpha coefficients to indicate the reliability of the subscales in his study on the reliability of the SETS instrument. The Cronbach's alpha coefficients were .94 for the whole instrument, .87 for reading data, and .91 for reading between data.

Data Collection

Data was collected once the IRB granted permission (see Appendix A). The survey packet was uploaded to Qualtrics® by copying and pasting each item from the three instruments and the four demographic characteristics to ensure no typographical errors. A cover letter at the beginning of the survey explained the purpose and scope of the study. Participants' responses were recorded in Qualtrics® and transferred to an Excel file. All students' responses and information were kept confidential.

An email was sent to program directors of primary teacher education programs at 4-year accredited non-profit universities and colleges in Georgia asking for contact information for faculty members who teach primary education courses. The email explained the purpose and scope of the study, requirements for participation, and included the written statement regarding the approval of the study by the IRB at Valdosta State University. A follow-up phone call was made seven days after emailing nonrespondents. A list of faculty members and their contact information was compiled and stored for reference. An email was sent to the contact list explaining the study's purpose and scope, and the participation requirements. In addition, a phone call was made to each faculty member to discuss when and how to disperse the survey packet. Follow-up phone calls were made as needed.

The faculty who agreed to participate as proctors of the survey packet emailed their classes or posted in their learning management system the link to the survey packet on Qualtrics®. Participation from PSMTs was voluntary. There was no penalty for not participating. Once the survey closed, the data was stored on Qualtrics® and exported to an Excel file. The Excel file was saved and stored on the researcher's computer hard drive.

Data Analysis

Data analysis was conducted using R statistical packages. Tidyverse packages were used for the exploratory data analysis. The collected data were formatted into an Excel file and imported into R DataFrame. There are 19 negatively worded items in the SATS instrument and five negatively worded items in the SAS instrument that were recoded by reversing the responses (1 becomes 7, 2 becomes 6, et cetera). The demographic characteristic question on gender was transformed to data points (male will be assigned 0 and female will be assigned 1). An exploratory data analysis began the process by identifying missing values from the dataset and summarizing and visualizing the data. Participants were eliminated from the study if they did not complete the three instruments, spent less than 5 minutes on the survey packet, or responded to each question with the same response.

Missing Data.

The structural equation modeling statistical consideration for missing data was examined on each variable. The number of missing values for each variable was counted. There was expected to be missing data for each variable, but the missing data should constitute for less than 1% implying not systemic (Kline, 2015). Little' Missing Completely at Random (MCAR) test was employed to test the null hypothesis that the missing data was missing completely at random. Median imputation was used for imputing missing data that would not largely impact the overall results of the data analysis.

Descriptive Statistics.

After handling missing data, exploratory data analysis was performed to obtain summary statistics and graphical representations of the data to understand patterns within the data, identify obvious errors, and detect outliers. The dataset was visualized using histograms, boxplots, and bar graphs. Descriptive statistical techniques were used to calculate frequency statistics for nominal and ordinal data, and to calculate measures of central tendency and dispersion for interval data. Exploratory data analysis provides the foundation for subsequent analysis and could result in poorly fitted models or inadmissible models if neglected (Field, 2013). Each variable from the instruments was summarized using descriptive techniques in R. The total score for each PSMTs' subscale was calculated by adding the values of each item together. In addition, Cronbach's alpha coefficients were calculated on each variable to show reliability.

Inferential Statistics.

Correlations among the demographic characteristic variables, statistics teaching efficacy, the four subscales that measure attitudes toward statistics, and anxiety about statistics were analyzed using statistical methods. The R statistical packages mixedcor and polycor were employed to calculate Pearson's, polychoric, and tetrachoric correlation coefficients and their corresponding significance. The dataset was visualized using correlation matrices. Structural equation modeling (SEM) was employed to test the overall fit of the observed data with the proposed hypothetical model. The model to be tested includes one exogenous variable (statistics teaching efficacy) and four potential mediator variables (affect, cognitive competence, difficulty, and value). These variables were tested as predictors of the endogenous variable anxiety about statistics. A path coefficient was calculated for each path, and overall model fit indices were calculated to determine the overall fit of the theorized model with the data collected for the study. Finally, a multigroup analysis was conducted to determine if the model fits across the demographic characteristics.

Statistical Assumptions.

Statistical assumptions must be met before data analysis. If assumptions are not met, the results are not of value (Field, 2013). Structural equation modeling requires certain assumptions be satisfied to ensure accurate inferences (Cohen et al., 2003). The assumptions for structural equation modeling include sufficiently large sample size, no systematic missing data, univariate normality, multivariate normality, and correct model specification. These assumptions pertain to sample size, distributional issues, and complete data.

There are various guidelines for sample size in structural equation modeling. Kline (2015) recommended 20 participants for each variable used in the study. This would imply a sample size of 200 for this study. Comrey and Lee (2013) and Tabachnick and Fidell (2013) recommended a sample size of 300 for not so complex models. This study aimed for a sample size of 300.

The data was examined for univariate normality by examining the variables' skewness and kurtosis values, and visually inspecting histograms and QQ-plots. Additionally, the Jarque-Bera test was employed to assess the normality of each variable. If there was a departure from normality, then the variable would be examined for univariate outliers. Univariate outliers would be handled in a way that preserved the size of the data set. After assessing univariate normality, multivariate normality would be assessed. Multivariate outliers were assessed by using Mahalanobis distance. Mahalanobis distance detects scores that deviate from the mean of a group of variables. All data cases that resulted in a Mahalanobis distance greater than the critical value was omitted from the data set. After handling multivariate outliers, multivariate normality was assessed using Mardia's test. Mardia's test would check whether the multivariate skewness and kurtosis are consistent with a multivariate normal distribution (Korkmaz et al., 2014). Assumptions were checked before data analysis began. If assumptions were not met, other analysis methods might be employed.

Summary

This chapter outlined the research design and methodology for studying primary PSMTs' SKT as outlined by Groth's (2013) SKT framework. A nonexperimental, correlational research design using structural equation modeling is appropriate to investigate factors related to PSMTs' SKT, namely, pedagogical content knowledge. These factors included statistics teaching efficacy, the four subscales of attitudes toward statistics, and anxiety about statistics. The exogenous variable and potential mediators, statistics teaching efficacy and the four subscales of attitudes toward statistics, were chosen based on a thorough literature review of PSMTs' statistical knowledge for teaching. They served as predictors for the endogenous variable, PSMTs' anxiety about statistics. This study included primary PSMTs enrolled in a primary education teaching program at a 4-year public or private non-profit, accredited college or university in Georgia. Details on the instruments in this study, including validity and reliability, were provided in this chapter. These instruments were part of a survey packet including four demographic characteristics.

Participating schools received details on administering the survey packet to their primary PSMTs electronically. Participation is voluntary. Data was collected, cleaned, and prepared for data analysis. Data analysis and statistical considerations, and assumptions for structural equation modeling were discussed in this chapter. Structural equation modeling tested the overall fit of the observed data with the proposed hypothetical model outlined in Chapter 1. Multigroup analysis was conducted to determine if the model was a goof fit across the four demographic characteristic variables.

Chapter IV

RESULTS

This chapter presents the results of the nonexperimental, correlational study that employed structural equation modeling to explore the relationships among primary preservice mathematics teachers' (PSMTs) anxiety about statistics, statistics teaching efficacy, and attitudes toward statistics. This chapter also examines the directionality of the relationships between the variables, providing insights into how these constructs interact and influence each other. Understanding these dynamics aims to inform and improve primary teacher training programs, enabling the cultivation of educators who can teach statistics with confidence and effectiveness. After establishing a model that fits the observed data well, the secondary purpose of the study was to determine if the model fits across certain demographic characteristic variables. Data were collected over four weeks using a self-report questionnaire, which participants completed using an online version posted on Qualtrics®. The open-source statistical programming language R was employed in the analysis. The current research was designed to answer the following research questions:

> RQ1: Is the hypothetical path model, which describes the causal effects among the PSMTs' attitudes toward statistics and statistics teaching efficacy on anxiety about statistics, consistent with the observed correlates among these variables?

RQ2: If the hypothetical path model is consistent, what are the estimated direct, indirect, and total effects among the PSMTs' attitudes toward statistics and statistics teaching efficacy on anxiety about statistics?

RQ3: Is the specified path model equivalent across various demographic characteristic variables?

This chapter presents the data analysis and results of the models. The first section of this chapter describes the demographic characteristics of the sample. The second section reports the item-level descriptive statistics, and the third section reports scale descriptive statistics along with Pearson's, polychoric, and tetrachoric correlation coefficients with the original data. Statistical considerations and assumptions of structural equation modeling are discussed in the fourth section. After identifying and handling outliers, the following section provides scale descriptive statistics of the data. The next section describes the model identification process and the direct, indirect, and total effects. Lastly, the final section presents measurement invariance across selected demographic variables.

Demographic Characteristics

This section describes and presents the demographic characteristics of the sample of primary PSMTs used in this study. Forty-five universities or colleges in Georgia were approached to participate in the study, and 11 schools agreed. Of the 11 schools, approximately 1,200 primary PSMTs were approached to participate, and 358 primary PSMTs (29.83%) responded to the survey. Two hundred ninety-seven data cases met the requirements for data analysis.

The participants were asked to answer four demographic questions related to their position as students in a primary education program: gender, number of college-level mathematics courses completed (mathematics experience), number of college-level statistics courses completed (statistics experience), and number of years enrolled in a primary education training program (college experience). Table 4 presents the individual characteristics of the sample.

Table 4

Number and Percentage of Georgia Primary PSMTs by Demographic Characteristic

Demographic Characteristics	Value	*n*	%
Gender	Male	26	8.75
	Female	271	91.25
How many college-level mathematics courses have you completed?	0	0	0.00
	1	33	11.11
	2	16	5.38
	3	44	14.81
	4	80	26.93
	5	124	41.75
How many college-level statistics courses have you completed?	0	0	0.00
	1	98	33.00
	2	136	45.79
	3	33	11.11
	4	12	4.04
	5	18	6.06
How many years have you completed in your primary education program?	0	0	0.00
	1	41	13.80
	2	72	24.24
	3	77	25.93
	4	61	20.53
	5	46	15.49

Note. n = 297.

Consistent with the literature review, 91.25% of primary PSMTs were female (n = 271) (National Center for Education Statistics, 2023). All primary PSMTs responding to the questionnaire completed a college-level mathematics and statistics course. The largest number of respondents (n = 124, 41.75%) completed more than four mathematics courses and the smallest (n = 16, 5.39%) completed two mathematics courses. The median number of mathematics courses completed was four mathematics courses (Interquartile Range (IQR) = 2). The largest number of respondents (n = 136, 45.79%) completed two statistics courses and the smallest (n = 12, 4.04%) completed four statistics courses. The median number of statistics courses completed was two statistics courses (IQR = 1). The college-level mathematics (or statistics) courses included Advanced Placement (AP), International Baccalaureate (IB), and dual-enrollment mathematics (or statistics) courses completed during high school.

Primary PSMTs responded to one question regarding the years completed in a primary education training program. The median number of years completed in a primary education training program that primary PSMTs reported was 3 years (IQR= 2). The largest number of respondents (n = 77, 25.92%) completed 3 years and the smallest (n = 41, 13.80%) completed 1 year.

Item-Level Descriptive Statistics

The data set was explored for missing data by examining the frequency distribution of each item. Thirty-nine cases had over 50% of missing data and were removed from the data set. Twenty-two cases had greater than 50% of missing data from two or more subscales of the survey instrument and were removed from the data set. No cases had greater than 50% of missing data from one subscale. In the remaining 297 cases, there was less than 1% of missing data. Of the missing data, 18 cases were missing one data point, ten were missing two data points, three were missing three data points, and four were missing four data points.

Little’s Missing Completely at Random (MCAR) test was employed to test the mean differences on all items sharing the same missing data pattern. The MCAR test compares the observed variable means for each pattern of missing data with the expected population means estimated using the expectation-maximization algorithm (Little’s Missing Completely at Random (MCAR) test, 2019). Under the null hypothesis, the missing data was missing completely at random. Little’s MCAR test results ($X^2(297) = 820$, $df = 854$, $p = .79$) implied that the missing data was independent of the observed and unobserved data and was missing completely at random. There were no systematic differences between individuals with missing and complete data. As the data were on the ordinal level scale, the 35 missing data points were replaced by the median values of the corresponding item. Median imputation provides robustness against outliers (Alam et al., 2023).

The variables in this study were measured using the Teacher’s Efficacy, Attitudes, and Anxiety about Statistics Survey (see Appendix D). The survey contained 49 items from three previously validated surveys. The variables of interest included statistics teaching efficacy, the four subscales that measured attitudes toward statistics (affect, cognitive competence, difficulty, and value), and anxiety about statistics. The measurement of each variable’s total score was on the interval level scale, whereas the item-level statistics were on the ordinal level scale.

The statistics teaching efficacy variable was measured using 11 items from Harrell-Williams’ et al. (2013) Self-Efficacy to Teach Statistics (SETS) instrument. The instrument is measured on a 6-point rating scale ranging from 1 = *not at all confident* to 6 = *completely confident*. Table 5 shows the number of responses for each item and the relative frequency, mean, median, and standard deviation for statistics teaching efficacy items. The median of the items ranged from 3 to 4. There were 109 respondents (36.70%) who reported very confident or completely confident in recognizing that statistical results may differ in another class or group (item 10). There were 90 respondents (30.30%) who reported very confident or completely confident in recognizing that there would be natural variability between observations for individuals (item 2). These results indicate the sample of primary PSMTs lacked a foundational understanding of statistical variation among individual and group observations. Ninety-eight respondents (33.00%) reported they were not at all confident or only a little confident in creating boxplots for summarizing distributions (item 4), suggesting a need for enhanced educational focus on graphical data representation. There were 85 respondents (28.62%) who reported not at all confident or only a little confident in using dotplots, stem and leaf plots, and tables (using counts) for describing distributions (item 5), pointing to a gap in practical skills essential for effective data analysis and interpretation. These insights underscore the importance of targeted training for primary PSMTs to boost confidence and competence in teaching statistical content found across primary school.

Table 5

Number of Responses and Descriptive Statistics by Item for Statistics Teaching Efficacy

Item	1	2	3	4	5	6	*Mdn*	*M*	*SD*
1. Collect data to answer a posed statistical question in contexts of interest to high school students.	21 (7%)	39 (13%)	12 (38%)	78 (26%)	19 (6%)	28 (9%)	3	3.40	1.28
2. Recognize that there will be natural variability between observations.	11 (4%)	19 (6%)	62 (21%)	115 (39%)	46 (15%)	44 (15%)	4	4.00	1.25
3. Select appropriate graphical displays and numerical summaries to compare individuals to each other and an individual to a group.	12 (4%)	36 (12%)	101 (34%)	86 (29%)	41 (14%)	21 (7%)	3	3.58	1.20
4. Create dotplot, stem and leaf plot, and tables (using counts) for summarizing distributions.	35 (12%)	37 (12%)	90 (30%)	77 (26%)	36 (12%)	22 (7%)	3	3.36	1.37
5. Use dotplot, stem and leaf plot, and tables for describing distributions.	38 (13%)	47 (16%)	91 (31%)	74 (25%)	27 (9%)	20 (7%)	3	3.22	1.36
6. Create boxplots for summarizing distributions.	44 (15%)	54 (18%)	90 (30%)	69 (23%)	22 (7%)	18 (6%)	3	3.08	1.36
7. Use boxplots, median, and range for describing distributions.	29 (10%)	37 (12%)	89 (30%)	83 (28%)	35 (12%)	24 (8%)	3	3.44	1.34

Note. 1 (*Not at All Confident*), 2 (*Only a Little Confident*), 3 (*Somewhat Confident*), 4 (*Confident*), 5 (*Very Confident*), 6 (*Completely Confident*).

Table 5 (continued)

Number of Responses and Descriptive Statistics by Item for Statistics Teaching Efficacy

Item	1	2	3	4	5	6	*Mdn*	*M*	*SD*
8. Identify the association between two variables from scatterplots.	21 (7%)	48 (16%)	93 (31%)	84 (28%)	30 (10%)	21 (7%)	3	3.39	1.27
9. Generalize a statistical result from a small group to a larger group such as the whole class.	22 (7%)	35 (12%)	101 (34%)	83 (28%)	34 (11%)	22 (7%)	3	3.36	1.27
10. Recognize that statistical results may be different in another class or group.	11 (4%)	21 (7%)	65 (22%)	91 (31%)	57 (19%)	52 (18%)	4	4.07	1.31
11. Recognize the limitation of making inferences from a classroom dataset to any population beyond the classroom.	13 (4%)	45 (15%)	85 (29%)	84 (28%)	35 (12%)	35 (12%)	4	3.63	1.32

Note. 1 (*Not at All Confident*), 2 (*Only a Little Confident*), 3 (*Somewhat Confident*), 4 (*Confident*), 5 (*Very Confident*), 6 (*Completely Confident*).

The four subscales that measured primary PSMTs' attitudes toward statistics were affect, cognitive competence, difficulty, and value. These variables were measured using the Survey of Attitudes toward Statistics (SATS) instrument developed by Schau et al. (1995). The instrument is measured on a 7-point Likert scale ranging from 1 = *strongly disagree* to 7 = *strongly agree*. The variable affect measured individuals' positive and negative feelings concerning statistics using six items. Table 6 shows the number of responses for each item and the relative frequency, mean, median, and standard deviation for affect items. The median of the items ranged from 3 to 5. There were 93 respondents (31.31%) who reported agreeing or strongly agreeing to being under stress during a statistics class (item 25), highlighting a common challenge. Seventy-nine respondents (26.60%) reported feeling insecure when tackling statistical problems (item 13). There were 130 respondents (43.77%) who reported disagreeing or strongly disagreeing with enjoying a statistics course (item 26), indicating a general lack of enthusiasm. In addition, 133 respondents (44.78%) expressed frustration when reviewing statistics tests in class (item 22). On a positive note, although 81 respondents (27.27%) did not favor statistics (item 12), 105 (35.35%) expressed a liking for statistics, reflecting a divided but significant interest in the subject. The responses regarding fear of statistics were mixed: 78 respondents (26.26%) reported a fear of statistics (item 32), matching the number who disagreed or strongly disagreed (n = 60, 20.20%), and an additional 60 respondents (20.20%) remained neutral.

Table 6

Number of Responses and Descriptive Statistics by Item for Affect

Item	1	2	3	4	5	6	7	*Mdn*	*M*	*SD*
12. I will like statistics.	36 (12%)	45 (15%)	47 (16%)	64 (22%)	64 (22%)	26 (9%)	15 (5%)	4	3.72	1.67
13. I will feel insecure when I have to do statistical problems.	15 (5%)	23 (8%)	38 (13%)	60 (20%)	82 (28%)	56 (19%)	23 (8%)	5	4.45	1.56

Note. 1 (*Strongly Disagree*), 2 (*Disagree*), 3 (*Somewhat Disagree*), 4 (*Neither Disagree nor Agree*), 5 (*Somewhat Agree*), 6 (*Agree*), 7 (*Strongly Agree*).

Table 6 (continued)

Number of Responses and Descriptive Statistics by Item for Affect

Item	1	2	3	4	5	6	7	*Mdn*	*M*	*SD*
22. I will get frustrated going over statistics tests in class.	11 (4%)	33 (11%)	37 (12%)	83 (28%)	77 (26%)	33 (11%)	23 (8%)	4	4.26	1.50
25. I will be under stress during statistics class.	8 (3%)	26 (9%)	17 (6%)	63 (21%)	90 (30%)	61 (21%)	32 (11%)	5	4.72	1.49
26. I will enjoy taking statistics courses.	45 (15%)	85 (28%)	97 (33%)	45 (15%)	16 (5%)	9 (3%)	0 (0%)	3	2.76	1.22
32. I am scared of statistics.	26 (9%)	34 (11%)	32 (11%)	67 (23%)	60 (20%)	41 (14%)	37 (12%)	4	4.25	1.78

Note. 1 (*Strongly Disagree*), 2 (*Disagree*), 3 (*Somewhat Disagree*), 4 (*Neither Disagree nor Agree*), 5 (*Somewhat Agree*), 6 (*Agree*), 7 (*Strongly Agree*).

The variable cognitive competence measured individuals' attitudes about intellectual knowledge and skills applied to statistics using six items. Table 7 shows the number of responses for each item and the relative frequency, mean, median, and standard deviation for cognitive competence items. The median of the items ranged from 4 to 6. There were 158 respondents (53.20%) who reported agreeing or strongly agreeing with item 34, "I can learn statistics". This demonstrates the respondents' confidence in their ability to grasp statistical concepts. A small fraction, 15 respondents (5.05%), disagreed with the statement, indicating a lack of self-efficacy in learning statistics. For concerning the understanding of statistical equations (item 35), 85 respondents (28.62%) agreed or strongly agreed, whereas only 16 respondents (5.39%) felt they struggled with this aspect of statistics. Additionally, 80 respondents (26.94%) remained neutral regarding their difficulties in understanding statistics due to their cognitive approaches (item 14). Almost half of the respondents (n = 143, 48.15%) anticipated difficulties in comprehending statistical concepts (item 38). This highlights a perceived challenge among nearly half of the participants. In addition, about half of the respondents, 147 (49.49%), admitted to frequently making mathematical errors in statistics, indicating a common area of difficulty. Conversely, 78 respondents (26.26%) disagreed with experiencing frequent mathematical errors, showing a contrast in proficiency and comfort with statistics.

Table 7

Number of Responses and Descriptive Statistics by Item for Cognitive Competence

Item	1	2	3	4	5	6	7	*Mdn*	*M*	*SD*
14. I will have trouble understanding statistics because of how I think.	24 (8%)	42 (14%)	43 (14%)	80 (27%)	54 (18%)	37 (12%)	17 (6%)	4	3.93	1.62
20. I will have no idea of what is going on in statistics.	22 (7%)	60 (20%)	51 (17%)	69 (23%)	55 (19%)	21 (7%)	19 (6%)	4	3.72	1.62
31. I will make a lot of math errors in statistics.	8 (3%)	28 (9%)	42 (14%)	72 (24%)	75 (25%)	41 (14%)	31 (10%)	4	4.43	1.53
34. I can learn statistics.	11 (4%)	4 (1%)	10 (3%)	45 (15%)	69 (23%)	100 (34%)	58 (20%)	6	5.32	1.42

Note. 1 (*Strongly Disagree*), 2 (*Disagree*), 3 (*Somewhat Disagree*), 4 (*Neither Disagree nor Agree*), 5 (*Somewhat Agree*), 6 (*Agree*), 7 (*Strongly Agree*).

Table 7 (continued)

Number of Responses and Descriptive Statistics by Item for Cognitive Competence

Item	1	2	3	4	5	6	7	*Mdn*	*M*	*SD*
35. I will understand statistics equations.	11 (4%)	5 (2%)	37 (12%)	66 (22%)	93 (31%)	58 (20%)	27 (9%)	5	4.71	1.39
38. I will find it difficult to understand statistics concepts.	10 (3%)	28 (9%)	40 (13%)	76 (26%)	83 (28%)	39 (13%)	21 (7%)	4	4.33	1.47

Note. 1 (*Strongly Disagree*), 2 (*Disagree*), 3 (*Somewhat Disagree*), 4 (*Neither Disagree nor Agree*), 5 (*Somewhat Agree*), 6 (*Agree*), 7 (*Strongly Agree*).

The variable value measured individuals' attitudes about the usefulness, relevance, and worth of statistics in personal and professional life using nine items. Table 8 shows the number of responses for each item and the relative frequency, mean, median, and standard deviation for value items. The median of the items ranged from 2 to 4. There were 179 respondents (60.27%) who disagreed or strongly disagreed with the belief that statistics are worthless (item 16). This indicated a significant appreciation for the value of statistics. Only 11 respondents (3.70%) viewed statistics as worthless. Additionally, 112 respondents (56.85%) disagreed or strongly disagreed with the idea that statistics have no application in their profession (item 30), suggesting most respondents see the relevance in their professional fields. However, responses were mixed regarding the impact of statistical skills on employability; 96 respondents (32.32%) were neutral about whether statistical skills enhance employability (item 19), and 91 respondents (30.64%) remained neutral about the utility of statistics in typical professions (item 21). Meanwhile, 104 respondents (35.02%) disagreed with the statement that statistics are not useful in typical professions. Ninety-eight respondents (33.00%) disagreed with the idea that statistics are irrelevant in their lives (item 27), further affirming the perceived personal and professional importance of statistics. Despite this recognition, only 39 respondents (13.13%) agreed they use statistics in their everyday life (item 24), with 57 (19.19%) strongly disagreeing, highlighting a gap between recognizing the importance of statistics and applying statistical knowledge in daily activities. This discrepancy highlights a potential area for educational improvement to bridge the gap between theoretical understanding and practical application of statistics.

Table 8

Number of Responses and Descriptive Statistics by Item for Value

Item	1	2	3	4	5	6	7	*Mdn*	*M*	*SD*
16. Statistics is worthless.	73 (25%)	106 (36%)	27 (9%)	67 (23%)	13 (4%)	6 (2%)	5 (2%)	2	2.59	1.43
18. Statistics should be required as part of my professional training.	22 (7%)	32 (11%)	32 (11%)	90 (30%)	66 (22%)	43 (14%)	12 (4%)	4	4.09	1.54
19. Statistical skills make me more employable.	10 (3%)	26 (9%)	17 (6%)	96 (32%)	69 (23%)	58 (20%)	21 (7%)	4	4.50	1.45
21. Statistics is not useful to the typical professional.	27 (9%)	77 (26%)	54 (18%)	91 (31%)	35 (12%)	6 (2%)	1 (0%)	3	3.26	1.38

Note. 1 (*Strongly Disagree*), 2 (*Disagree*), 3 (*Somewhat Disagree*), 4 (*Neither Disagree nor Agree*), 5 (*Somewhat Agree*), 6 (*Agree*), 7 (*Strongly Agree*).

Table 8 (continued)

Number of Responses and Descriptive Statistics by Item for Value

Item	1	2	3	4	5	6	7	*Mdn*	*M*	*SD*
23. Statistical thinking is not applicable in my life outside my job.	26 (9%)	59 (20%)	54 (18%)	87 (29%)	36 (12%)	22 (7%)	13 (4%)	4	3.56	1.55
24. I use statistics in my everyday life.	20 (7%)	37 (12%)	46 (15%)	79 (27%)	76 (26%)	23 (8%)	16 (5%)	4	3.97	1.52
27. Statistics conclusions are rarely presented in everyday life.	16 (5%)	52 (18%)	63 (21%)	103 (35%)	39 (13%)	14 (5%)	10 (3%)	4	3.60	1.37

Note. 1 (*Strongly Disagree*), 2 (*Disagree*), 3 (*Somewhat Disagree*), 4 (*Neither Disagree nor Agree*), 5 (*Somewhat Agree*), 6 (*Agree*), 7 (*Strongly Agree*).

The variable difficulty measured individuals' attitudes about the difficulty of statistics as a subject using seven items. Table 9 shows the number of responses for each item and the relative frequency, mean, median, and standard deviation for difficulty items. The median of the items ranged from 3 to 5. There were 123 respondents (41.41%) that indicated a difficulty in understanding statistical formulas by disagreeing or strongly disagreeing that they are easy to understand (item 15). Similarly, 120 respondents (40.40%) disagreed or strongly disagreed with the statement that statistics is a subject quickly learned by most people (item 28). This highlights the perceived complexity of the subject. Aligning with these perspectives, 104 respondents (35.02%) agreed or strongly agreed that statistics is a complicated subject (item 17). In addition, 98 respondents (33.00%) agreed that learning statistics requires a great deal of discipline (item 29), which suggests that mastering statistics is seen as a challenging endeavor. A substantial number, 138 respondents (46.46%), agreed that statistics involve massive computations (item 33). There were 156 respondents (52.53%) who recognized statistics as a highly technical field (item 37). Collectively, these responses highlight the consensus among respondents regarding the complex, technical nature of statistics and the significant effort required to become proficient in this discipline.

Table 9

Number of Responses and Descriptive Statistics by Item for Difficulty

Item	1	2	3	4	5	6	7	*Mdn*	*M*	*SD*
15. Statistics formulas are easy to understand.	27 (9%)	96 (32%)	75 (25%)	64 (22%)	26 (19%)	9 (3%)	0 (0%)	3	2.98	1.24
17. Statistics is a complicated subject.	6 (2%)	8 (3%)	26 (9%)	51 (17%)	102 (34%)	67 (23%)	37 (12%)	5	4.97	1.35
28. Statistics is a subject quickly learned by most people.	38 (13%)	82 (28%)	66 (22%)	80 (27%)	23 (8%)	6 (2%)	2 (1%)	3	2.98	1.29
29. Learning statistics requires a great deal of discipline.	4 (1%)	9 (3%)	20 (7%)	72 (24%)	94 (32%)	72 (24%)	26 (9%)	5	4.90	1.26
33. Statistics involves massive computations.	7 (2%)	13 (4%)	37 (12%)	100 (34%)	81 (27%)	36 (12%)	21 (7%)	4	4.43	1.32

Note. 1 (*Strongly Disagree*), 2 (*Disagree*), 3 (*Somewhat Disagree*), 4 (*Neither Disagree nor Agree*), 5 (*Somewhat Agree*), 6 (*Agree*), 7 (*Strongly Agree*).

Table 9 (continued)

Number of Responses and Descriptive Statistics by Item for Difficulty

Item	1	2	3	4	5	6	7	*Mdn*	*M*	*SD*
37. Statistics is highly technical.	3 (1%)	8 (3%)	28 (9%)	102 (34%)	88 (30%)	40 (13%)	28 (9%)	5	4.67	1.24
39. Most people have to learn a new way of thinking to do statistics.	2 (1%)	10 (3%)	24 (8%)	104 (35%)	83 (28%)	47 (16%)	27 (9%)	5	4.70	1.23

Note. 1 (*Strongly Disagree*), 2 (*Disagree*), 3 (*Somewhat Disagree*), 4 (*Neither Disagree nor Agree*), 5 (*Somewhat Agree*), 6 (*Agree*), 7 (*Strongly Agree*).

Anxiety about statistics was measured using the 10-item Statistical Anxiety Scale (SAS) instrument developed by Pretorius and Norman (1992). The items were measured on a 5-point Likert scale ranging from 1 = *strongly disagree* to 5 = *strongly agree*. Table 10 presents the number of responses for each item and the relative frequency, mean, median, and standard deviation for anxiety about statistics items. The median of the items ranged from 2 to 4. One hundred sixty-four respondents (55.22%) expressed openness to taking more statistics courses by disagreeing or strongly disagreeing with the statement of being bothered by additional statistics coursework (item 40). In contrast, 152 respondents (51.18%) disagreed or strongly disagreed with feeling at ease during a statistics test (item 41), and 134 respondents (45.12%) expressed discomfort during statistics courses (item 42). Likewise, 136 respondents (45.79%) agreed or strongly agreed that they feel uptight while taking statistics tests (item 45). Thes findings highlight the anxiety associated with exams. A significant number of participants, 146 (49.16%), reported a sinking feeling when faced with challenging statistical problems (item 46). One hundred thirty-eight respondents (46.46%) agreed that statistics made them feel uncomfortable and nervous (item 48), with a similar number (138, 46.46%) also feeling uneasy and confused about the subject (item 49). However, 67 respondents (22.56%) and 63 respondents (21.21%) disagreed with these sentiments, respectively. This suggests a split in the perceived difficulty and emotional response to statistics among the participants.

Table 10

Number of Responses and Descriptive Statistics by Item for Anxiety About Statistics

Item	1	2	3	4	5	*Mdn*	*M*	*SD*
40. It wouldn't bother me at all to take more statistics courses.	58 (20%)	106 (36%)	71 (24%)	40 (13%)	22 (7%)	2	2.54	1.17
41. I have usually been at ease during statistics tests.	50 (17%)	102 (34%)	92 (31%)	44 (15%)	9 (3%)	2	2.53	1.03
42. I have usually been at ease in statistics courses.	41 (14%)	93 (31%)	102 (34%)	52 (18%)	9 (3%)	3	2.65	1.02
43. I usually don't worry about my ability to solve statistics problems.	39 (13%)	89 (30%)	93 (31%)	64 (22%)	12 (4%)	3	2.73	1.07
44. I almost never get uptight while taking statistics tests.	38 (13%)	83 (28%)	104 (35%)	40 (13%)	7 (2%)	3	2.58	0.98
45. I get really uptight during statistics tests.	17 (6%)	43 (14%)	101 (34%)	100 (34%)	36 (12%)	3	3.32	1.05
46. I get a sinking feeling when I think of trying hard statistics problems.	15 (5%)	44 (15%)	86 (29%)	111 (37%)	41 14%)	4	3.40	1.06
47. My mind goes blank, and I am unable to think clearly when working statistics.	16 (5%)	54 (18%)	100 (34%)	91 (31%)	36 (12%)	3	3.26	1.06

Note. 1 (*Strongly Disagree*), 2 (*Disagree*), 3 (*Neither Disagree nor Agree*), 4 (*Agree*), 5 (*Strongly Agree).*

Table 10 (continued)

Number of Responses and Descriptive Statistics by Item for Anxiety About Statistics

Item	1	2	3	4	5	*Mdn*	*M*	*SD*
48. Statistics makes me feel uncomfortable and nervous.	13 (4%)	54 (18%)	92 (31%)	98 (33%)	40 (13%)	3	3.33	1.07
49. Statistics makes me feel uneasy and confused.	17 (6%)	46 (15%)	96 (32%)	100 (34%)	38 (13%)	3	3.32	1.05

Note. 1 (*Strongly Disagree*), 2 (*Disagree*), 3 (*Neither Disagree nor Agree*), 4 (*Agree*), 5 (*Strongly Agree).*

Scale Descriptive Statistics

Negatively worded items were reverse coded to ensure a positive continuum. In the statistics anxiety scale, items 45, 46, 47, 48, and 49 were reverse coded; in the affect scale, items 13, 22, 25, and 32 were reverse coded; in the cognitive competence scale, items 14, 20, 31, and 38 were reverse coded; in the difficulty scale, items 17, 29, 33, 37, and 39 were reverse coded; and in the value scale, items 16, 21, 23, 27, 30 and 36 were reverse coded. After reverse coding, the total scores were calculated by adding the items of each scale together to derive a total score. Cronbach's alpha coefficient was used to assess the internal consistency of the scales. Cronbach's alpha ranged from .71 to .93 indicating acceptable to excellent reliability of each scale: statistics teaching efficacy (α = .93), affect (α = .85), cognitive competence (α = .83), difficulty (α = .71), value (α = .82), and anxiety about statistics (α = .92). Important to note that variable difficulty has been known to have moderate, acceptable reliability (Schau et al., 1995).

The variables' mean total scores ranged from 20.79 to 40.91 (see Table 11). Descriptive statistics indicated the total score of statistics teaching efficacy ranged from 11 to 66, with a mean total score of 38.65 (*SD* = 10.96). The total score of affect ranged from 6 to 41 with a mean total score of 20.79 (*SD* = 7.02), cognitive competence ranged from 6 to 42 with a mean total score of 25.61 (*SD* = 6.70), difficulty ranged from 7 to 43 with a mean total score of 22.29 (*SD* = 5.41), and value ranged from 9 to 61 with a mean total score of 40.91 (*SD* = 8.97). The total score of anxiety about statistics ranged from 10 to 50 with a mean total score of 26.39 (*SD* = 8.01).

Table 11

Descriptive Statistics of Each Variable's Total Scale Score

Scale	Min	Max	*Mdn*	*M*	*SD*	Skewness	Kurtosis
Cognitive Competence	6	42	25	25.61	6.70	-0.22	0.48
Value	9	61	40	40.91	8.97	-0.18	0.38
Difficulty	7	43	23	22.29	5.41	-0.19	1.10
Affect	6	41	21	20.79	7.02	0.25	0.18
Statistics Teaching Efficacy	11	66	38	38.65	10.96	0.04	-0.01
Anxiety About Statistics	10	50	27	26.39	9.01	0.15	-0.22

Note. n = 297.

A correlation matrix was generated to examine the correlations among the variables: statistics teaching efficacy, affect, cognitive competence, difficulty, value, anxiety about statistics, gender, mathematics experience, statistics experience, and college status. The relationships among the study's interval level variables (statistic teaching efficacy, affect, cognitive competence, difficulty, value, and anxiety about statistics) were assessed using the Pearson's product-moment correlation coefficient. The relationships among the ordinal level variables (mathematics experience, statistics experience, and college status) were assessed using polychoric correlation coefficients, and the relationship among the nominal, dichotomous level variable (gender) was assessed using tetrachoric correlation coefficients. Correlation coefficients are presented in Table 12.

Table 12

Correlations among Variables Total Scale Scores

Variable	1	2	3	4	5	6	7	8	9	10
1. Cognitive	1.00									
2. Value	.61**	1.00								
3. Difficulty	.57**	.32**	1.00							
4. Affect	.85**	.57**	.65**	1.00						
5. Efficacy	.39**	.20**	.16*	.35**	1.00					
6. Anxiety	.70**	.40**	.62**	.74**	.26**	1.00				
7. Gender	-.14	.04	-.03	-.17	.00	-.06	1.00			
8. Math	.12	.11	-.01	.15	.16	.11	.03	1.00		
9. Stat	.10	.10	.08	.15	.21	.06	.26	.29	1.00	
10. College	.12	.13	.14	.13	.14	.10	.10	.21	.11	1.00

Note. $^{*}p< .01$, $^{**}p< .001$.

Correlations among the interval level variables ranged from .16 to .85. The analysis revealed that significant, positive correlations exist amongst all interval-level variables. The strongest, positive correlation was found between affect and cognitive competence, $r(297) = .85$, $p < .001$, indicating as the affect frequency total score increased, the cognitive competence frequency total score increased. The weakest, positive correlation was found between statistics teaching efficacy and difficulty, $r(297) = .16$, $p < .01$. In addition, a weak, positive correlation was found between statistics teaching efficacy and value ($r(297) = .20$, $p < .001$), and between difficulty and value ($r(297) = .32$, $p < .001$).

The variable, anxiety about statistics, had the strongest, positive correlation with affect, $r(297) = .74, p < .001$. Anxiety about statistics also had a strong, positive correlation with cognitive competence ($r(297) = .70, p < .001$). Anxiety about statistics had a moderate, positive correlation with difficulty ($r(297) = .62, p < .001$) and value ($r(297) = .40, p < .001$). Anxiety about statistics, had the weakest, positive correlation with statistics teaching efficacy, $r(297) = .26, p < .001$. The polychoric and tetrachoric correlations for the nominal and ordinal level variables were nonsignificant.

SEM Model Assumptions

Prior to model identification, the basic assumption requirements for structural equation modeling (SEM) include missing data, adequate sample size, outliers, and univariate and multivariate normality. There were missing data points for each variable, but the missing data constituted less than 1%, implying missing data was not systemic (Kline, 2011). Little's MCAR test confirmed the missing data was missing completely at random, implying no systematic differences between individuals with missing data and those with complete data.

There are many guidelines regarding how many participants are needed for SEM. Schreiber et al. (2006) suggested 20 participants per variable for SEM, with an optimal sample size of more than 200 participants (Kline, 2015; Suhr, 2006). In agreement, Loehlin (1992) recommended collecting at least 100 cases, with 200 cases being more optimal for an adequate sample size in SEM. Other researchers recommended at least 300 participants (Comrey & Lee, 2013; Tabachnick & Fidell, 2013). This study had a final data set of 297 cases, which is a minimally adequate sample size for SEM.

Each variable was evaluated for univariate outliers by converting total scale scores to z-scores to mathematically assess for outliers. Summary statistics of the variables' z-scores are found in Table 13. The z-scores were analyzed for outliers using a cutoff value of 3.29. Difficulty had one data point (43, $z = 3.83$) deemed an outlier, and value had one data point (9, $z = -3.56$) deemed an outlier. The outliers occurred from distinct individuals. Before handling outliers, univariate normality was assessed on each variable.

Table 13

Variables' Z-scores Summary Statistics

Scale	Min	Max	*Mdn*	*M*	*SD*	Skewness	Kurtosis
Cognitive Competence	-2.93	2.44	-0.09	0	1	-0.22	0.43
Value	-3.56	2.24	-0.10	0	1	-0.18	0.33
Difficulty	-2.83	3.83	0.13	0	1	0.23	-0.18
Affect	-2.11	2.99	0.03	0	1	0.25	0.14
Statistics Teaching Efficacy	-2.52	2.50	-0.06	0	1	0.04	-0.04
Anxiety About Statistics	-2.05	2.95	0.08	0	1	0.23	-0.18

Note. n = 297.

Univariate normality posits that a single variable conforms to a normal distribution. All variables demonstrated normal skewness and kurtosis values except for difficulty. Difficulty's kurtosis value was greater than one, suggesting a departure from normality (Mishra et al., 2019). The distribution of each variable was visually assessed using histograms and Q-Q plots (see Appendix E). The histogram and QQ-plot for difficulty confirmed the departure from normality. To further assess univariate normality,

the Jarque-Bera (JB) test was employed. The Jarque-Bera (JB) test is a robust goodness of fit test used to assess the distributional structure of data (Aslam et al., 2021).
The Jarque-Bera test indicated normal distributions for anxiety about statistics ($X^2(297) = 2.93$, $p = .23$), statistics teaching efficacy ($X^2(297) = 0.11$, $p = .94$), affect ($X^2(297) = 5.06$, $p = .08$), cognitive competence ($X^2(297) = 5.00$, $p = .08$), and value ($X^2(297) = 2.50$, $p = .29$). The Jarque-Bera test indicated difficulty violated the assumption of normality ($X^2(297) = 18.17$, $p < .001$). The departure from normality of variable difficulty improved after handling outliers.

Instead of removing univariate outliers, they were replaced with the data point that had the nearest z-score. For difficulty, 43 ($z = 3.83$) was replaced by 37 ($z = 2.72$). For value, 9 ($z = -3.56$) was replaced by 15 ($z = -2.89$). The decision to replace the data points was to keep the individual data cases intact within the data set while also improving normality. The individual who scored 43 for difficulty had positive attitudes about the difficulty of statistics as a subject. With the replacement, the individual now has a value of 37, which still represents similar attitudes toward the difficulty of statistics. The individual who scored 9 for value had negative attitudes about the usefulness, relevance, and worth of statistics in personal and professional life. With the replacement, the individual now has a value of 15, which still represents similar attitudes toward the value of statistics. After replacement, normality was reassessed for each variable.

The distribution shape of each variable was analyzed visually by histograms and Q-Q plots. The Jarque-Bera test was used to assess univariate normality, and the results stayed consistent. The Jarque-Bera test indicated normal distributions for anxiety about statistics ($X^2(297) = 1.73$, $p = .42$), statistics teaching efficacy ($X^2(297) = 0.09$, $p = .96$), affect ($X^2(297) = 5.59$, $p = .06$), cognitive competence ($X^2(297) = 5.31$, $p = .07$), and value ($X^2(297) = 0.51$, $p = .77$). The Jarque-Bera test indicated difficulty violated the normality assumption ($X^2(279) = 14.79$, $p < .001$). However, after handling outleirs, the skewness (-0.32) and kurtosis (0.70) values of difficulty improved. The QQ-plot and histogram for difficulty also improved (Figure 9). Thus, univariate normality assumption was met after visually inspecting and assessing the variables.

Multivariate normality describes the joint distribution of all variables in the sample. Multivariate normality was examined after univariate outliers were replaced and univariate normality was determined. Multivariate outliers were assessed by examining Mahalanobis distance, which measures the observations farthest from the centroid. One case had a Mahalanobis distance greater than the critical value of 22.46 ($p < .001$) and was removed from the data set. The case represented a female student who had completed four mathematics courses, two statistics courses, 4 years in an education program, and had the following scale total scores: statistics teaching efficacy = 34, affect = 6, cognitive competence = 19, difficulty = 15, value = 19, and anxiety about statistics = 34. The removal of this case leaves a data set containing 296 cases.

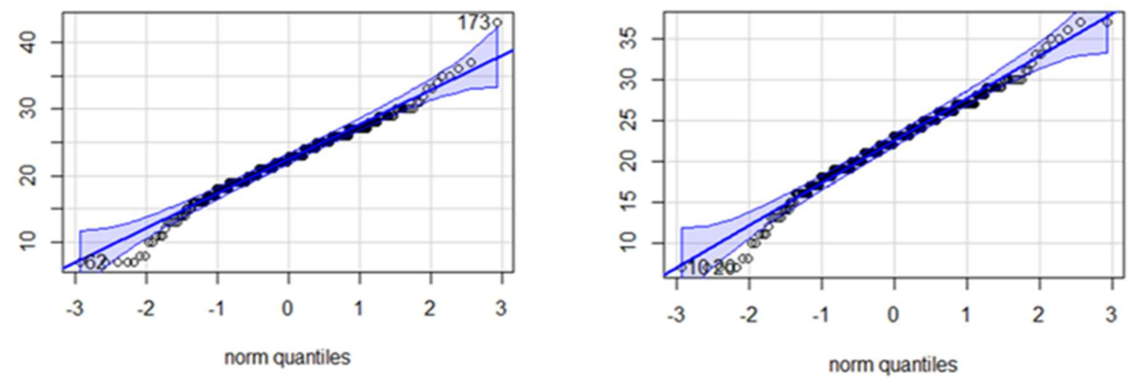
Multivariate normality was examined by Mardia's coefficients. Both Mardia's skewness estimate of 148.71 ($p < .001$) and Mardia's kurtosis estimate of 5.32 ($p < .001$) indicated multivariate normality was not met. Therefore, according to Mardia's multivariate normality test, the data set does not follow a multivariate normal distribution. The multivariate normality assumption is made when using maximum likelihood estimation in path analysis (Geiser, 2023). Because the assumption of multivariate normality was not met, parameter standard errors and fit measurements may be incorrect and misleading (Geiser, 2023). Thus, for path analysis, the maximum likelihood estimation with robust standard errors and a Satorra-Bentler scaled test statistic were employed to provide better parameter standard errors and fit measurements.

Figure 9

Difficulty Histogram and QQ-Plot Before and After Handling Outliers

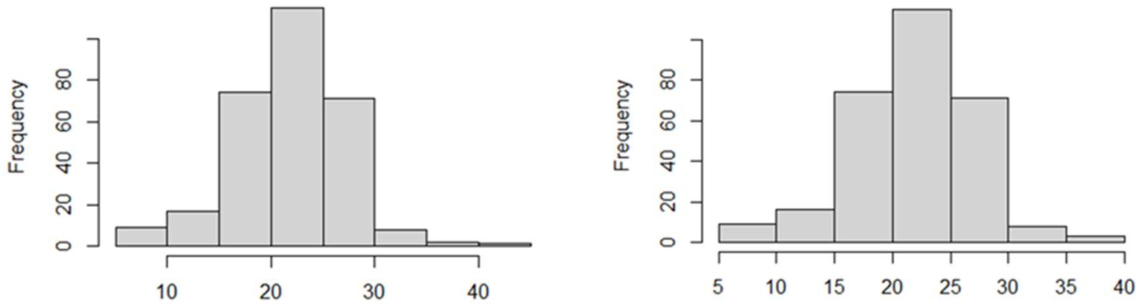
QQ-Plot Before Handling Outlier

QQ-Plot After Handling Outlier



Histogram Before Handling Outlier

Histogram After Handling Outlier



Descriptive Statistics and Correlations after Handling Outliers

An analysis was performed after the replacement of univariate outliers and the removal of the multivariate outlier. Once again, Cronbach's alpha coefficient was used to assess the internal consistency of the scales and results stayed consistent: statistics teaching efficacy (α = .93), affect (α = .85), cognitive competence (α = .83), difficulty (α = .71), value (α = .82), and anxiety about statistics (α = .92). The variables mean total scores ranged from 20.84 to 41.01 (see Table 14). Descriptive statistics indicated affect ranged from 6 to 41 with a mean total score of 20.84 (*SD* = 6.98), cognitive competence ranged from 6 to 42 with a mean total score of 25.63 (*SD* = 6.70), difficulty ranged from 7 to 37 with a mean total score of 22.30 (*SD* = 5.33), and value ranged from 15 to 61 with a mean total score of 41.01 (*SD* = 8.83). The total score of statistics teaching efficacy ranged from 11 to 66 with a mean total score of 38.66 (*SD* = 10.98), and anxiety about statistics ranged from 10 to 46 with a mean total score of 27.01 (*SD* = 6.76). The skewness and kurtosis values for all variables were less than one.

Table 14

Descriptive Statistics of Each Variable's Total Scale Score after Handling Outliers

Scale	Min	Max	*Mdn*	*M*	*SD*	Skewness	Kurtosis
Cognitive Competence	6	42	25	25.63	6.70	-0.23	0.44
Value	15	61	40	41.01	8.83	-0.10	0.11
Difficulty	7	37	23	22.30	5.33	-0.32	0.31
Affect	6	41	21	20.84	6.98	0.26	0.15
Statistics Teaching Efficacy	11	66	38	38.66	10.98	0.04	-0.05
Anxiety About Statistics	10	46	27	27.01	6.76	0.16	-0.25

Note. n = 296.

A correlation matrix was generated to examine the correlations among the variables: statistics teaching efficacy, affect, cognitive competence, difficulty, value, anxiety about statistics, gender, mathematics experience, statistics experience, and college status. The relationships among the study's interval level variables (statistics teaching efficacy, affect, cognitive competence, difficulty, value, and anxiety about statistics) were assessed using the Pearson's product-moment correlation coefficient, relationships among the ordinal level variables (mathematics experience, statistics experience, and college status) were assessed using polychoric correlation coefficients, and the relationship among the nominal, dichotomous level variable (gender) was assessed using tetrachoric correlation coefficients (see Table 15).

Table 15

Correlations among Variables Total Scale Scores after Handling Outliers

Variable	1	2	3	4	5	6	7	8	9	10
1. Cognitive	1.00									
2. Value	.61**	1.00								
3. Difficulty	.57**	.30**	1.00							
4. Affect	.85**	.56**	.63**	1.00						
5. Efficacy	.39**	.20**	.15*	.35**	1.00					
6. Anxiety	.70**	.42**	.62**	.75**	.26**	1.00				
7. Gender	-.20	-.21	-.06	-.22	.07	-.19	1.00			
8. Math	-.11	-.14	-.03	-.05	-.06	-.04	.08	1.00		
9. Stat	-.11	-.14	-.08	-.13	-.08	-.14	.29	.29	1.00	
10. College	-.05	-.22	.05	-.04	-.04	.05	.10	.20	.13	1.00

Note. $^*p < .01$, $^{**}p < .001$.

Correlations among the interval level variables were positive and ranged from .15 to .85. As before handling outliers, the analysis revealed significant, positive correlations amongst all interval level variables. The strongest, positive correlation was found between cognitive competence and affect, $r(296) = .85$, $p < .001$. The weakest, positive correlation was found between statistics teaching efficacy and difficulty, $r(296) = .15$, $p < .01$. Both results agreed with the correlation findings before handling outliers.

When comparing the correlations of the original data set to the data set after handling outliers, the strength of the relationship of statistics teaching efficacy to all variables increased and all other correlations decreased in strength. The variable statistics teaching efficacy had weak, positive correlations with all subscales of attitudes toward statistics (affect, $r(296) = .35$, $p < .001$; cognitive competence, $r(296) = .39$, $p < .001$; difficulty, $r(296) = .15$, $p < .01$; value, $r(296) = .20$, $p < .001$). The variable anxiety about statistics had the strongest, positive correlation with affect, $r(296) = .75$, $p < .001$. Anxiety about statistics also had a strong, positive correlation with cognitive competence ($r(296) = .70$, $p < .001$). Anxiety about statistics had a moderate, positive correlation with difficulty ($r(296) = .62$, $p < .001$). Anxiety about statistics had a weak, positive correlation with value ($r(296) = .42$, $p < .001$), and the weakest, positive correlation was with statistics teaching efficacy, $r(296) = .26$, $p < .001$. The strongest, positive correlation among the subscales of attitudes toward statistics was between affect and cognitive competence ($r(296) = .85$, $p < .001$). The weakest, positive correlation among the subscales of attitudes toward statistics was between difficulty and value ($r(296) = .30$, $p < .001$). The polychoric and tetrachoric correlations for the nominal and ordinal level variables were nonsignificant. These results are consistent with the correlations before handling outliers.

SEM Model Analysis

A model based on previous research was developed to describe the relationships between the variables under examination. The variables in this study were evaluated by using the total scale scores after handling outliers. The hypothesized model illustrates the proposed path diagram that depicts the relationship between statistics teaching efficacy, affect, cognitive competence, difficulty, value, and anxiety about statistics. This section is guided by the study's first research question:

> RQ1: Is the hypothetical path model, which describes the causal effects among the PSMTs' attitudes toward statistics and statistics teaching efficacy on anxiety about statistics, consistent with the observed correlates among these variables?

Various indices were assessed to evaluate the fit of the hypothesized model in the data analysis. The model fit criteria considered in this study included the chi-square (X^2), the root-mean-square error of approximation (RMSEA), and the standardized root mean square residual (SRMR). Model comparison was based on the comparative fit index (CFI) and Tucker-Lewis index (TLI). A non-significant chi-square (X^2) value indicates the model effectively reproduces the variance-covariance relationships in the data matrix (Schumaker & Lomax, 2016). Because the chi-square test statistic is influenced by the sample size, alternative indices are used for model evaluation.

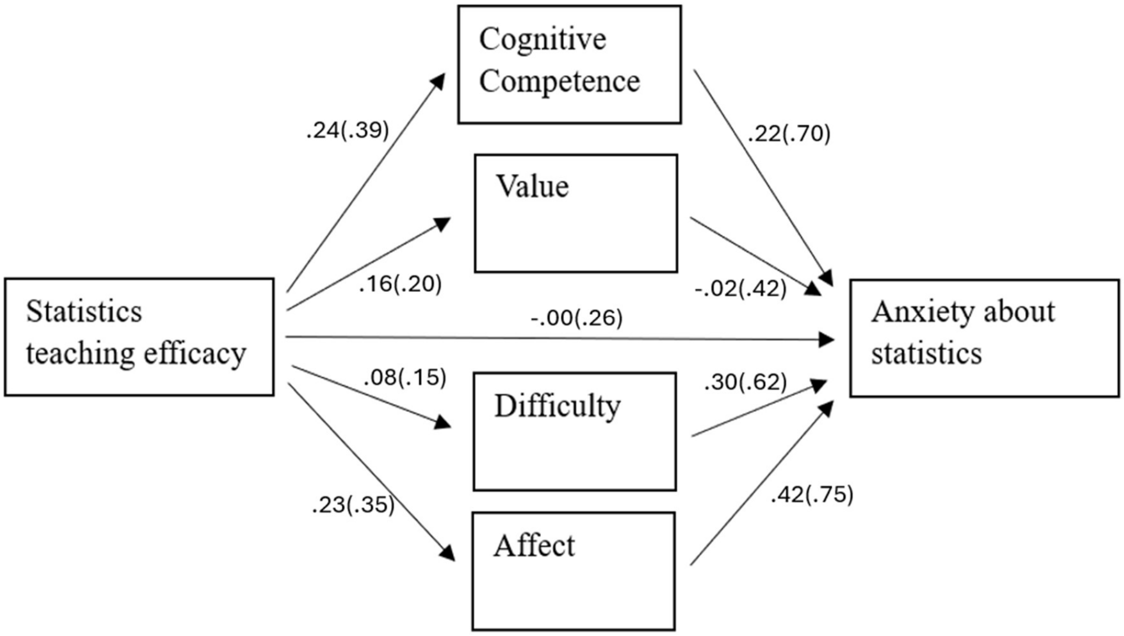
The RMSEA value represents the average discrepancies expected in the model when applied to population data. An RMSEA value between .10 and .08 suggests a model of moderate fit, between .08 and .05 indicates a good fit, and less than .05 signifies an excellent fit (Schumaker & Lomax, 2016). The SRMR is an absolute fit index. An SRMR value less than or equal to .08 indicates good fit for the data (Schumaker & Lomax, 2016). The CFI and TLI served as model comparison indices, measuring the improvement in fit between the proposed model and a null model that establishes a baseline for comparing alternative models (Schumacker & Lomax, 2016). Generally, CFI and TLI values above .95 indicate an acceptable fit of the model to the data.

Likelihood ratio tests were conducted to assess the goodness of fit of the model, with statistical significance set at the .05 level of significance. The model parameters, including path coefficients and variance estimates, were estimated using the maximum likelihood (MLM) estimation. The Satorra-Bentler scaled test statistic and scaled, robust fit indices are reported for each model. In addition, the original X^2, RMSEA, SRMR, CFI, and TLI are reported for comparison. Figure 10 displays the results of the path analysis on Model 1, including standardized coefficients and correlation coefficients for each pair of variables.

The Satorra-Bentler chi-square results, $X^2(df = 6) = 621.92$, $p < .001$, indicated Model 1 did not fit the observed data. The analysis produced scaled CFI and scaled TLI values of 0.47 and -0.31, respectively. The robust RMSEA value was 0.43 and the scaled SRMR value was 0.33, which suggested a poor fitting of the hypothetical model (Model 1) to the observed data. In addition, the chi-square results ($X^2(df = 6) = 621.92$, $p < .001$), CFI (0.38), TLI (-0.55), RMSEA (0.59), and SRMR (0.33) agreed with Model 1 not fitting the observed data.

Figure 10

Model 1 with Standardized Coefficients



Note. Correlation coefficients are displayed in parentheses.

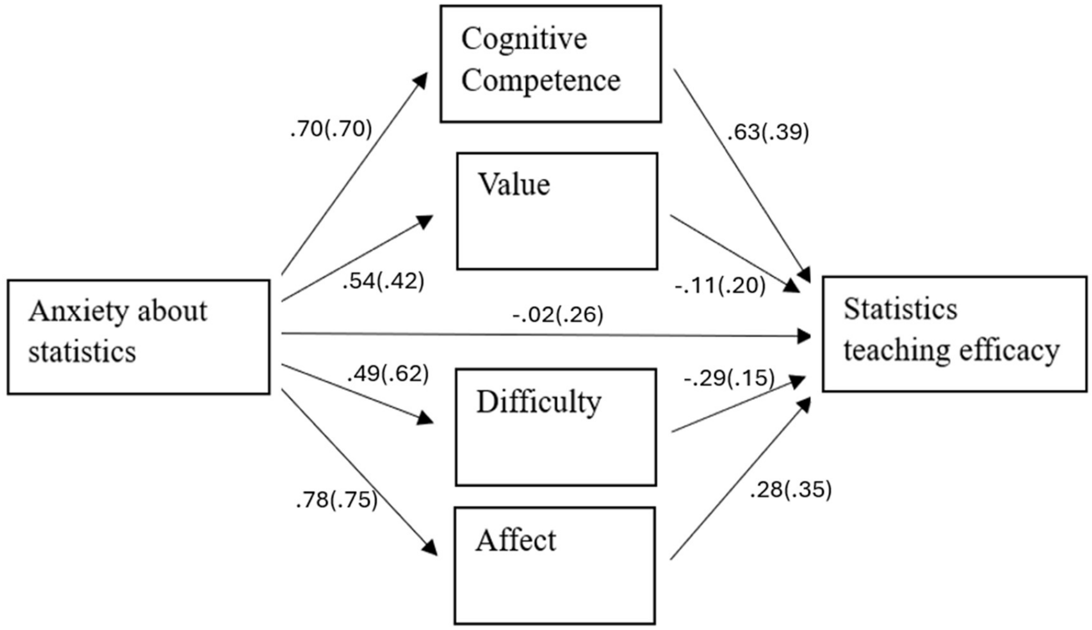
Conducting modifications to Model 1 became necessary because of poor fit of the data to the model. Path analysis is an extension of multiple regression. Model 1 was hypothesized based on the research of Akin and Kurbanoglu's (2011), who found attitudes toward mathematics were predicted by self-efficacy and self-efficacy significantly predicted anxiety about mathematics through the mediator attitudes toward mathematics. Rationally, anxiety about statistics should be predictive of the subscales of attitudes toward statistics and statistics teaching efficacy, and the subscales of attitudes toward efficacy should be predictive of statistics teaching efficacy. Modification indices were examined for Model 1 and agreed with the rationale. Guided by modification indices and rational thought, the model fit would improve if the model reversed direction; that was, anxiety about statistics would have a direct path to statistics teaching efficacy and have an indirect path to statistics teaching efficacy through affect, cognitive competence, difficulty, and value. This resulted in a second model, Model 2 (see Figure 11), that provided a better fit compared to Model 1. Figure 11 displays the results of the path analysis on Model 2, including standardized coefficients and correlation coefficients for each pair of variables.

The Satorra-Bentler chi-square results, $X^2(df = 6) = 621.92$, $p < .001$, and the chi-square results, $X^2(df = 6) = 305.49$, $p < .001$, indicated Model 2 did not fit the observed data. The scaled CFI and TLI estimates improved from 0.47 in Model 1 to 0.65 in Model 2 and -0.31 in Model 1 to 0.13 in Model 2, respectively. In addition, CFI and TLI estimates improved from 0.38 in Model 1 to 0.70 in Model 2 and -0.55 in Model 1 to 0.25 in Model 2, respectively. The robust RMSEA estimate and RMSEA estimate improved to 0.35 and 0.41, respectively. The scaled SRMR and SRMR both improved to 0.13, as well. Although the fit measures showed improvement, Model 2 did not fit the data well.

Conducting modifications to Model 2 became necessary to improve fit. The addition of the covariances of the potential mediator variables (affect, cognitive competence, difficulty, and value) was added to the model to determine if it resulted in a better fit of the data. The model with covariances added between the potential mediator variables was a saturated just-identified model, which had zero degrees of freedom and resulted in a probability level that could not be calculated. Additional analysis was performed by examining the covariances to determine if a modified model could be estimated that would result in a better fit of the data. The chi-square difference test suggested the covariance between difficulty and value could be dropped from the

Figure 11

Model 2 with Standardized Coefficients



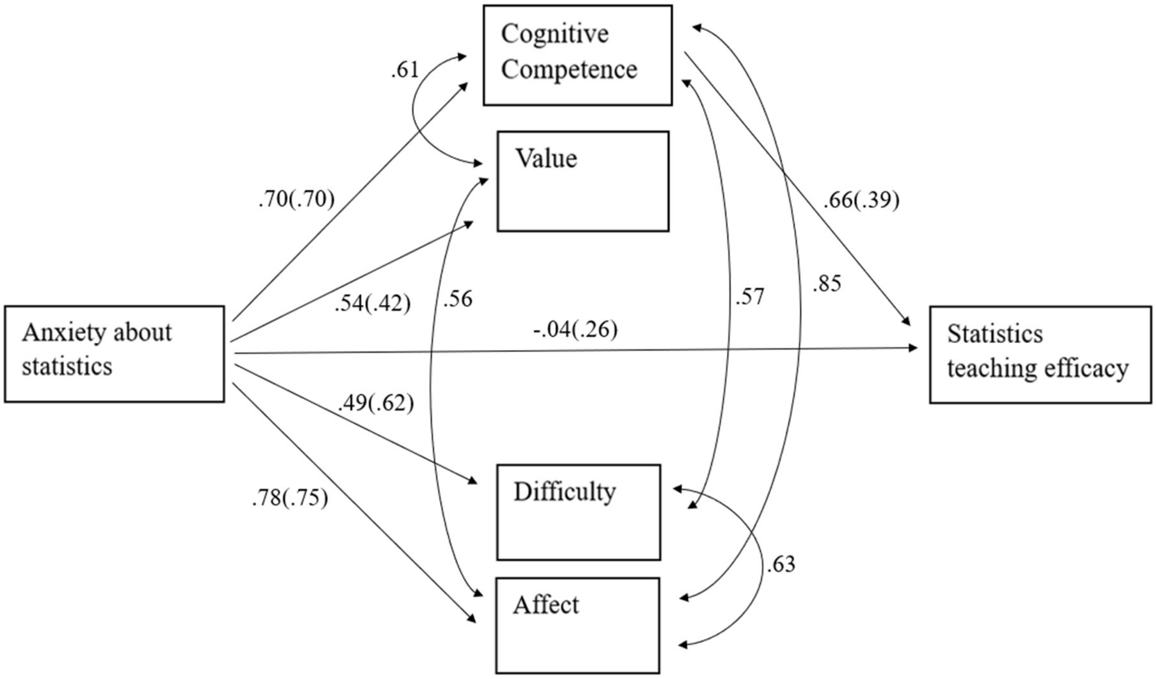
Note. Correlation coefficients are displayed in parentheses.

the analysis to improve fit. In addition, path coefficients of affect, difficulty, and value were insignificant when regressed on statistics teaching efficacy. These paths were dropped from the analysis to improve fit. This resulted in a third model, Model 3 (see Figure 12), that provided a good fit of the observed data. Figure 12 displays the results of the path analysis on Model 3, including standardized coefficients and correlation coefficients for each pair of variables.

In the final model, the Satorra-Bentler chi-square results, $X^2(df = 4) = 6.24$, $p = .18$, and chi-square results, $X^2(df = 4) = 6.88$, $p = .14$, indicated the model fits the observed data. The scaled CFI and TLI estimates improved from 0.65 in Model 2 to 1.00 in Model 3 and 0.13 in Model 2 to 0.99 in Model 3, respectively. The CFI and TLI estimates improved from 0.70 in Model 2 to 1.00 in Model 3 and 0.25 in Model 2 to 0.99

Figure 12

Model 3 with Standardized Coefficients



Note. Correlation coefficients are displayed in parentheses. The correlation coefficients are only displayed for covarying mediator variables.

in Model 3, respectively. In addition, the robust RMSEA estimate and RMSEA estimate improved to 0.04 and 0.05, respectively. The scaled SRMR and SRMR both improved to .02, as well. The robust RMSEA and scaled SRMR indicated an excellent fit of Model 3 to the observed data. All fit measures agreed that Model 3 fits the observed data well. Table 16 presents a summary of the model fit indices across the three models.

Bootstrapping was conducted to verify if Model 3 was a good fit of the data. Since the data did not meet multivariate normality, the fit indices may appear higher than their normal values (Kim & Millsap, 2010). The Bollen-Stine bootstrapping method was employed to impose Model 3 on the sample data so that bootstrapping is done under Model 3. It rescales the observed data, so they are consistent with the null hypothesis of perfect fit. Five thousand bootstraps were performed, and results did not reject the null hypothesis of perfect fit ($X^2(df = 4) = 6.88, p = .18$), suggesting Model 3 fits the data well.

Table 16

Summary of the Model Fit Indices for Each Path Model

Model	X^2	df	*p* value	CFI	TLI	RMSEA	SRMR
Model 1	621.92 (328.30)	6	.00 (.00)	0.38 (0.47)	-0.55 (-0.31)	0.59 (0.43)	0.33 (0.33)
Model 2	305.49 (219.68)	6	.00 (.00)	0.70 (0.65)	0.25 (0.13)	0.41 (0.35)	0.13 (0.13)
Model 3	6.88 (6.24)	4	.14 (.18)	1.00 (1.00)	0.99 (0.99)	0.05 (0.04)	0.05 (0.04)

Note. X^2 = chi-square test; CFI = comparative fit index; TLI = Tucker-Lewis index; RMSEA = root-mean-square-error of approximation, SRMR = standardized root mean square residual. Robust, scaled indices are listed in parentheses.

Direct, Indirect, and Total Effects

The hypothesized model did not fit the data. Modifications did not suggest using the path directions of the hypothesized model to make modifications to improve fit. Instead, suggested modifications led to Model 3. Model 3 was a good fit of the observed data. This section is guided by the study's second research question but with the use of Model 3 instead of the hypothetical path model.

> RQ2: If the hypothetical path model is consistent, what are the estimated direct, indirect, and total effects among the PSMTs' attitudes toward statistics and statistics teaching efficacy on anxiety about statistics?

The final model (Model 3) supported the data among primary PSMTs in Georgia, and the respecified model was presented in Figure 12. Table 17 presents the indirect,

direct, and total effects among the variables. After determining the model fits the data and is theoretically consistent, the parameter estimates and individual tests of significance of each parameter estimate were interpreted. The process allowed for estimation of causal relations among variables and mediating effects of direct and indirect effects of exogenous through potential mediator variables in the prediction of the endogenous variable (Kline, 2015).

Table 17

Indirect, Direct, and Total Effects

Variable	Indirect Effect	Direct Effect	Total Effect
Statistics Teaching Efficacy			
Cognitive Competence	–	.66*	.15*
Anxiety about Statistics	.46*	-.04	.42*
Affect			
Anxiety about Statistics	–	.78*	.78*
Cognitive Competence			
Anxiety about Statistics	–	.70*	.70*
Difficulty			
Anxiety about Statistics	–	.49*	.49*
Value			
Anxiety about Statistics	–	.54*	.54*

Note. **p* < .001, n = 296

The standardized parameter estimates and standard error for each path coefficient are displayed in Table 18. The results show evidence of a statistically nonsignificant path between anxiety about statistics and statistics teaching efficacy ($\beta = -.04$, $Z = -0.23$, $p = .82$). Cognitive competence had a significant positive path coefficient in its prediction on statistics teaching efficacy ($\beta = .66$, $Z = 4.47$, $p < .001$). This implies PSMTs' positive attitudes toward knowledge and intellectual skill in using statistical knowledge (cognitive competence) significantly predicted high confidence in teaching statistics (see Appendix F). Anxiety about statistics had significant positive path coefficients in its' prediction of affect (β = .78, Z = 17.14, p < .001), cognitive competence (β = .70, Z = 14.90, p < .001), difficulty (β = .49, Z = 11.48, p < .001), and value (β = .54, Z = 7.40, p < .001). This implies PSMTs' who possessed low anxiety about statistics significantly predicted positive attitudes toward statistics through affect, cognitive competence, difficulty, and value (see Appendix G).

Table 18

Direct Effects

Predictor Variable	Response Variable	β	Std. Error	*Z*
Anxiety	Affect	.78	.05	17.14*
Anxiety	Value	.54	.07	7.40*
Anxiety	Difficulty	.49	.04	11.48*
Anxiety	Cognitive competence	.70	.05	14.90*
Anxiety	Statistics teaching efficacy	-.04	.15	-0.23
Cognitive competence	Statistics teaching efficacy	.66	.15	4.47*

Note. *p < .001, n = 296.

Mediation analysis was conducted to test the effect of cognitive competence as a mediator in the prediction between anxiety about statistics and statistics teaching efficacy. Cognitive competence was significantly predictive of statistics teaching efficacy (β = .66, Z = 4.47, p < .001). Anxiety about statistics was not significantly predictive of statistics teaching efficacy in the absence of mediator cognitive competence (β = -.04, Z = -0.23, p = .82). However, anxiety about statistics was significantly predictive of statistics teaching efficacy through mediator cognitive competence (β = .46, Z = 4.15, p < .001).

This implies low anxiety about statistics significantly predicted positive attitudes toward knowledge and intellectual skill in using statistical knowledge (cognitive competence), which, in turn, significantly predicted high confidence in teaching statistics (see Appendix H). The total effect of anxiety about statistics on statistics teaching efficacy through mediator cognitive competence was significant ($\beta = .43$, $Z = 6.09$, $p < .001$). Total effect was calculated by taking the sum of the direct and indirect effects of anxiety about statistics on statistics teaching efficacy.

Multigroup Invariance

RQ3: Is the specified path model equivalent across various demographic characteristic variables?

The previous analyses were performed for a single group. To answer this research question, it was necessary to determine if the model was equivalent across two or more groups. The model was tested for the property of invariance by examining the model differences across groups which determines if the model was a good fit for the data of one group as it was for another group (Schumacker & Lomax, 2016). The participants were asked to answer four demographic questions: gender, number of college-level mathematics courses completed (mathematics experience), number of college-level statistics courses completed (statistics experience), and number of years enrolled in a primary education training program (college experience).

A series of model comparisons were performed to test for measurement invariance (configural, weak, strong, and strict invariance) among the groups by defining more and more stringent equality constraints (Schumacker & Lomax, 2016). Model A was the configural (baseline) model in which the same factor structure was imposed on each group. Model B was the weak (metric) invariance model which was a constrained version of the configural model where the factor loadings are assumed to be equal across groups (Bulut, 2020). Model C included all components of Model A and Model B with the addition of the structural weight to be equal across all groups and measured for strong (scalar) measurement invariance by comparing the model against the weak measurement invariance model. Model D included all components of the first three models with the addition of the residual variances fixed across all groups (Bulut, 2020) and measured for strict measurement invariance by comparing the model against the strong measurement invariance model.

The chi-square statistic was utilized in the chi-square difference test of nested models in the analysis of factorial invariance. The multigroup analyses produced a single chi-square for each model tested independent of the number of groups compared in the analyses. The changes in two models were assessed and compared among a default model and different constrained models and evaluated for statistically significant differences. The comparative fit indices difference test (ΔCFI) was used for comparison as it is a robust statistic for testing the between-group invariance (Chen, 2007). Weak invariance was tested by comparing the Model A against Model B; strong invariance was tested by comparing Model B against Model C; and strict invariance was tested by comparing Model C against Model D (Bulut, 2020). Starting with testing for weak invariance, if the test is not significant at the .05 level of significance, then weak invariance is established, and we can test for strong invariance. If strong invariance is not significant at the .05 level of significance, then strong invariance is established, and we can test for strict invariance. If strict invariance is not significant at the .05 level of significance, then strict invariance is established (Bulut, 2020). At least strong invariance must be established to interpret latent means and regressions across groups (Van de Schoot et al., 2012). If strong invariance is not established, but weak invariance is, then partial measurement invariance can be established by adjusting factor loadings and intercepts (Steenkamp & Baumgartner, 1998).

Participants responded to the gender question with two responses: male or female. The results showed an imbalanced data set in terms of gender, 270 females and 16 males. The technique of upsampling was employed to handle the imbalanced data set and improve model performance by providing more data for training. Upsampling created a new data set (n = 540) with 270 females and 270 males. The comparison between the configural invariance model and weak invariance model was not statistically significant ($\Delta X^2 = 6.88$, $df = 4$, $p = 1.00$); suggesting the weak invariance model fits the data equally well. The comparison between the weak invariance model and strong invariance model was statistically significant ($\Delta X^2 = 77.08$, $df = 6$, $p < .001$); suggesting a lack of strong invariance. Table 19 presents the results of measurement invariance across groups. Partial measurement invariance was attempted.

Partial measurement invariance was established by deciding which fixed parameters in the model should be released to improve fit for the strong invariance model. The release of the intercept parameters on statistics teaching efficacy, anxiety about statistics, cognitive competence, and value estimated for male and female was suggested to improve model fit ($p < .05$). The comparison between the weak invariance model and adjusted strong invariance model was not statistically significant ($\Delta X^2 = 5.17$, $df = 10$, $p = .08$); suggesting the adjusted strong invariance model fits the data equally well and partial strong invariance was established. The strict invariance model was built from the adjusted strong invariance model by constraining the residuals to be equal for male and female participants (Bulut, 2020). The comparison between the adjusted strong invariance model and strict invariance model was not statistically significant ($\Delta X^2 = 10.11$, $df = 16$, $p = .12$), thus; suggesting the adjusted strict invariance model fitted the data equally well. Although partial strict invariance was established, it was not reasonable to indicate the two groups, male and female, did not differ on the construct of the model.

Table 19

Model Fit Indices for Measurement Invariance

	X^2	df	*p* value	RMSEA	CFI	ΔCFI
Male or Female (n = 540)						
Model A: Baseline	29.27	8	–	.10	0.99	–
Model B: Weak Invariance	29.27	8	1.00	.10	0.99	.00
Model C: Strong Invariance	104.93	14	.00*	.16	0.95	.04
Model D: Strict Invariance	116.43	20	.04*	.13	0.95	.00
Mathematics Experience (n = 296)						
Model A: Baseline	10.02	8	–	.04	1.00	–
Model B: Weak Invariance	10.02	8	1.00	.04	1.00	.00
Model C: Strong Invariance	14.21	14	.63	.01	1.00	.01
Model D: Strict Invariance	20.92	20	.34	.02	1.00	.00
Statistics Experience (n = 296)						
Model A: Baseline	9.87	8	–	.04	1.00	–
Model B: Weak Invariance	9.87	8	1.00	.04	1.00	.00
Model C: Strong Invariance	18.58	14	.20	.05	1.00	.00
Model D: Strict Invariance	20.29	20	.94	.01	1.00	.00
College Experience (n = 296)						
Model A: Baseline	19.54	16	–	.06	1.00	–
Model B: Weak Invariance	19.54	16	1.00	.06	1.00	.00
Model C: Strong Invariance	45.37	34	.11	.07	0.99	.01
Model D: Strict Invariance	73.76	52	.08	.08	0.98	.01

Note. $^*p < .05$

Participants responded to mathematics experience question with five responses: one course, two courses, three courses, four courses, and more than four courses. Most teacher education programs that participated in the study require students to complete three mathematics courses for their major requirement. Thus, two groups were created: Group 1 (n = 92) consisted of participants who completed three or less mathematics courses, and Group 2 (n = 204) consisted of participants who completed more than three mathematics courses. The comparison between the configural invariance model and weak invariance model was not statistically significant ($\Delta X^2 = 6.88$, $df = 4$, $p = 1.00$); suggesting the weak invariance model fits the data equally well. The comparison between the weak invariance model and strong invariance model was not statistically significant ($\Delta X^2 = 4.39$, $df = 6$, $p = .63$); suggesting the strong invariance model fits the data equally well. Lastly, comparison between the strong invariance model and strict invariance model was not statistically significant ($\Delta X^2 = 6.51$, $df = 6$, $p = .36$), thus; suggesting the strict invariance model fits the data equally well. Strict invariance indicated the two groups did not differ on the construct of the model, and the latent means and correlations could be compared across the groups.

Participants responded to statistics experience question with five responses: one course, two courses, three courses, four courses, and more than four courses. Most teacher education programs that participated in the study require students to complete one statistics course for their general core requirement. Statistical content is also found in most mathematics courses teachers need to complete for graduation. Two groups were created: Group 1 (n = 234) consisted of participants who completed one or two statistics courses, and Group 2 (n = 62) consisted of participants who completed more than two statistics courses. The comparison between the configural invariance model and weak invariance model was not statistically significant ($\Delta X^2 = 6.88$, $df = 4$, $p = 1.00$), thus;

suggesting the weak invariance model fitted the data equally well. The comparison between the weak invariance model and strong invariance model was not statistically significant ($\Delta X^2 = 8.71$, $df = 6$, $p = .19$), and thus; suggested the strong invariance model fitted the data equally well. The comparison between the strong invariance model and strict invariance model was not statistically significant ($\Delta X^2 = 1.71$, $df = 6$, $p = .94$), thus; suggesting the strict invariance model fitted the data equally well. Strict invariance indicated the two groups did not differ on the construct of the model, and the latent means and correlations could be compared across the groups.

Participants responded to the college experience question with five responses: one year, two years, three years, four years, and more than four years. Teacher education programs that participated in the study require students to complete four years of coursework for graduation. Four groups were created based on years of completion: Group 1 (n = 41) one year, Group 2 (n = 72) two years, Group 3 (n = 77) three years, and Group 4 (n = 106) at least four years. The comparison between the configural invariance model and weak invariance model was not statistically significant ($\Delta X^2 = 6.88$, $df = 4$, $p = 1.00$); suggesting the weak invariance model fits the data equally well. The comparison between the weak invariance model and strong invariance model was not statistically significant ($\Delta X^2 = 25.83$, $df = 18$, $p = .10$); suggesting the strong invariance model fits the data equally well. The comparison between the strong invariance model and strict invariance model was not statistically significant ($\Delta X^2 = 28.38$, $df = 18$, $p = .06$), thus; suggesting the strict invariance model fitted the data equally well. Strict invariance indicated the four groups did not differ on the construct of the model, and the latent means and correlations could be compared across the groups.

Summary

This chapter described the quantitative data analysis using structural equation modeling and its results. A total of 358 primary PSMTs enrolled in an accredited 4-year teacher education program at a college or university in the state of Georgia responded to the survey. After handling missing data, 297 primary PSMTs remained. The primary PSMTs who participated in the study were primarily female (91.25%). The majority of primary PSMTs completed at least four mathematics courses (68.68%), two statistics courses (45.79%), and at least four years in a teacher education program (36.02%). The median number of mathematics courses completed was four mathematics courses, median number of statistics courses completed was two statistics courses, and median number years completed in a primary education program was 3 years.

The results of this study revealed that most primary PSMTs reported moderate to elevated levels of statistics teaching efficacy. Primary PSMTs were most confident in recognizing that statistical results may differ from group to group and that there will be natural variability between observations for individuals. The primary PSMTs reported moderate levels of anxiety about statistics. Primary PSMTs reported not feeling at ease during a statistics assessment and when enrolled in a statistics course. In general, PSMTs reported feeling uncomfortable, nervous, and confused when working with statistical content.

Regarding attitudes toward statistics, most PSMTs reported positive attitudes toward the value of statistics, and moderate-level attitudes toward the affect, cognitive competence, and difficulty of statistics. Primary PSMTs reported statistics are not worthless and are helpful to the typical professional. They reported neutral attitudes to liking statistics and taking future statistics courses, and positive attitudes to learning new statistical content.

The basic assumption requirements for structural equation modeling were assessed. It was found that variable difficulty violated the univariate normality assumption. There were two outliers in the data set, one from value and one from difficulty. After handling outliers, normality was reassessed. The departure from normality of difficulty improved. Univariate normality for all variables was met after visually inspecting histograms and QQ-Plots and examining skewness and kurtosis values.

Multivariate outliers were assessed using Mahalanobis distance. One case was deemed an outlier and removed from the data set, leaving a data set of 296 cases. Multivariate normality was examined by Mardia's coefficients and results showed the data set did not follow a multivariate normal distribution. Since the multivariate normality assumption failed, the maximum likelihood estimation with robust standard errors and a Satorra-Bentler scaled test statistic was employed for path analysis to provide better parameter standard errors and fit measurements.

The initial hypothesized model, Model 1, of variables did not fit the observed data. After examining modification indices and rationale thought, modifications of Model 1 produced a new model, Model 2. Model 2 did not fit the observed data. Covariances of the potential mediator variables (affect, cognitive competence, difficulty, and value) were added to Model 2, resulting in a saturated just-identified model with zero degrees of freedom. To improve fit, insignificant covariances and paths were dropped. This resulted in a final model. The final model, Model 3, of variables included covariances between various mediator variables and fit the observed data appropriately. Model 3 revealed the variables that were statistically significant for predicting statistics teaching efficacy of primary PSMTs in Georgia.

The results of the final model supported the prediction of affect, cognitive competence, difficulty, and value through the predictive variable anxiety about statistics. Anxiety about statistics was not significantly predictive of statistics teaching efficacy. Cognitive competence was significantly predictive of statistics teaching efficacy. Mediation analysis was conducted to test the effect of cognitive competence as a mediator in predicting anxiety about statistics and statistics teaching efficacy. Interestingly, anxiety about statistics was significantly predictive of statistics teaching efficacy through mediator cognitive competence. In addition, anxiety about statistics had a statistically significant positive correlation with all scales of attitudes toward statistics and statistics teaching efficacy. The variable anxiety about statistics had the strongest correlation with affect and the weakest correlation with statistics teaching efficacy. The strongest correlation among the subscales of attitudes toward statistics was between affect and cognitive competence. Statistics teaching efficacy had the strongest correlation with cognitive competence and the weakest correlation with difficulty.

Furthermore, multigroup analysis was utilized to analyze parameters specific for PSMTs according to gender, mathematics experience, statistics experience, and college experience. Strict measurement invariance was found for mathematics experience, statistics experience, and college experience indicating that the constructs of the model were the same for the groups. Upsampling was employed to handle the unbalanced data set for gender. Partial strict invariance was found for gender; however, comparisons should not be made as most survey variables were released in terms of their intercept parameters. Although upsampling was employed on gender, the small group of 26 males might not represent the broader population of male primary PSMTs.

Chapter V

SUMMARY AND DISCUSSION

Starting in kindergarten, statistics topics are introduced to students throughout the K-12 curriculum. A sizable portion of this statistical content is integrated within the mathematics curriculum, making mathematics teachers responsible for fostering both mathematical and statistical literacy among students. Primary teacher education programs aim to develop preservice mathematics teachers (PSMTs) who can effectively teach both disciplines. These programs focus on equipping primary PSMTs with both content knowledge and pedagogical content knowledge in mathematics and statistics. While the statistical content knowledge of PSMTs has been extensively studied, their pedagogical content knowledge has not received the same level of attention. Noncognitive factors such as teaching efficacy, attitudes, and anxiety are considered key components of pedagogical content knowledge. This study's primary purpose was to evaluate the interrelationships among these noncognitive factors that influence primary PSMTs' statistical knowledge for teaching. Structural equation modeling was employed to explore the roles of statistics teaching efficacy, the four subscales of attitudes towards statistics (affect, cognitive competence, difficulty, and value), and anxiety about statistics. In addition, this study aimed to assess whether the final model was consistent across various demographic characteristics. The final model benefits primary teacher education programs in understanding the factors affecting primary PSMTs' statistical knowledge for teaching, namely, pedagogical content knowledge. The model would also benefit in-service teachers by shedding light on how to use professional development opportunities in developing teachers' statistical knowledge for teaching by focusing on elements of the affective domain.

The research comprised 11 primary education programs across universities and colleges in Georgia, surveying 359 PSMTs. Participants completed a survey comprised of three previously validated instruments and a section on demographic characteristics. Data collection occurred in the spring of 2024. An analysis was shortly followed and was structured around three research questions. The study employed structural equation modeling to explore the directional influences of noncognitive factors on PSMTs' statistical knowledge for teaching.

Related Literature

Researchers have extensively explored the impact of noncognitive factors on PSMTs' statistical knowledge for teaching, mainly focusing on exploratory studies on teachers' teaching efficacy, attitudes, and anxiety about statistics (Estrada et al., 2011). Despite the abundance of research in these areas, there remains a notable gap in studies specifically targeting elementary PSMTs in Georgia and examining how these factors influence their statistical knowledge for teaching. This study not only supports but also builds on the empirical findings from previous research, offering new insights for this distinct group. This chapter will discuss the implications of these findings for enhancing statistical knowledge in teaching among primary PSMTs and suggest directions for future research to explore the influences of these noncognitive factors further. Additionally, the chapter compares the study’s findings of primary PSMTs’ teaching efficacy, attitudes, and anxiety toward statistics to previous research findings.

Statistics Teaching Efficacy

Teachers' self-efficacy is pivotal to their belief systems (Pajares, 1996). Studies have demonstrated that higher levels of teaching efficacy are positively correlated with improved student learning outcomes (Auzmendi, 1991; Gómez-Chacón, 2000). Lovett and Lee (2017a) explored secondary PSMTs' statistics teaching efficacy using the validated Self-Efficacy to Teach Statistics (SETS) instrument after their final mathematics education course. Lovett and Lee (2017a) found PSMTs felt somewhere between somewhat confident and confident in their ability to teach statistics. The highest levels of confidence observed across all PSMTs were found in teaching elementary grades, which was no surprise as the PSMTs were secondary teachers. Surprisingly, the PSMTs’ confidence in teaching statistical content at the secondary level was low. The teachers felt most confident in teaching the mechanics of creating graphical representations and recognizing statistical results may vary between different groups. The secondary PSMTs were least confident in selecting appropriate graphical displays and generalizing results from a small to a large group.

Harrell-Williams et al. (2015) also administered the SETS instrument to middle school PSMTs either in an introductory statistics course or a designated mathematics education course. Both Lovett and Lee (2017a) and Harrell-Williams et al. (2015) found similar trends, with PSMTs lacking confidence in teaching statistical concepts appropriate to their teaching levels. Based on these findings, the researchers recommended further investigation into teachers' statistical and pedagogical content knowledge. A deeper understanding of PSMTs' statistics teaching efficacy could illuminate pathways to enhance their preparation through targeted professional learning opportunities and coursework before they commence their teaching careers.

Attitudes Toward Statistics

Teachers' attitudes toward statistics significantly impacted teaching and learning statistics (Gal & Ginsberg, 1994; Schau, 2003). Estrada et al. (2005) and Hannigan et al. (2013) used the Survey of Attitudes Toward Statistics (SATS) instrument and found contrasting findings. Estrada et al. (2005) found most teachers with negative attitudes toward statistics lacked confidence in learning statistics. Hannigan et al. (2013) reported most teachers held positive attitudes toward statistics. The researchers found that teachers valued statistics and exhibited an interest in it. Furthermore, Hannigan et al. (2013) found teachers held elevated confidence levels in their statistical knowledge. Estrada and Batanero (2008) extended the research on attitudes toward statistics by researching factors that influenced SATS scores. The researchers compared responses from in-service mathematics teachers (ISMTs) and PSMTs and found attitudes toward statistics declined with the experience of teaching statistics. They also noted a negative correlation between years of teaching among ISMTs and positive attitudes towards statistics.

Begg and Edwards (1999) found that teachers held negative attitudes towards statistics through a qualitative study. The teachers expressed fear, disinterest, and incompetence toward statistics. Watson (2001) found that teachers held elevated confidence levels in their statistical abilities but lacked interest in pursuing professional development in teaching and learning statistics. Both studies highlight the need for more research into teachers' attitudes toward statistics. Gould and Peck (2004) found a lack of interest in teaching and learning statistics. The researchers encouraged both ISMTs and PSMTs to engage in professional development in teaching and learning statistics. Further research studies are crucial for developing and understanding strategies to enhance PSMTs' statistical knowledge for teaching.

Anxiety About Statistics

Anxiety about a subject can negatively impact students' learning and performance. Additionally, teachers' anxiety about a subject can negatively impact their teaching and learning. Bostani et al. (2014) assessed undergraduate students' anxiety using a mental health questionnaire. The researchers compared the academic performance between two groups: students anxious about statistics (Group 1) and students not anxious about statistics (Group 2). The results indicated Group 1's performance was significantly lower than Group 2's performance. Similarly, Macher et al. (2011) found negative correlations between anxiety about statistics and academic performance. Vitasari et al. (2010) further confirmed these findings, observing that students experienced anxiety during statistical assessments and presentations.

Primary teachers also exhibit elevated levels of mathematics anxiety, which can extend to statistics (Hembree, 1990; Kelly & Tomhave, 1985). Onwuebuzie et al. (1997) in their phenomenological study, suggested a possible link between math anxiety and statistics anxiety among teachers. Additionally, Yusuf et al. (2019) in a mixed-method study of PSMTs in a basic statistics course found elevated anxiety levels associated with assessments, although these findings are somewhat limited. Gaining a deeper understanding of PSMTs’ anxiety about statistics could inform strategies to enhance their readiness to teach, integrating professional development and tailored coursework effectively before they commence their teaching careers.

Methodology

This study employed a nonexperimental, correlational research design to analyze primary PSMTs' sense of efficacy, attitudes, and anxiety about statistics. The study utilized survey data to investigate insights from primary PSMTs across 11 colleges and universities in Georgia. Structural equation modeling was employed to determine how the constructs were theoretically linked and the directionality of the relationships. The research examined the structural relationships within a hypothetical model that hypothesized the impact of statistics teaching efficacy and attitudes toward statistics on anxiety about statistics among these primary PSMTs. Initial descriptive analyses provided a preliminary understanding of the sample. These results were refined through structural path analysis. Modifications to the hypothetical model were needed to improve the model fit for the data. This led to the development of a final model that effectively fit the observed data. The final model explained the theoretical causal effects of anxiety about statistics and the four subscales of attitudes (affect, cognitive competence, difficulty, and value) on statistics on teaching efficacy. This approach provided a deeper understanding of the factors influencing primary PSMTs' statistical knowledge for teaching.

Participants

The target population for this study comprised primary PSMTs enrolled in primary teacher education programs. The accessible population specifically included primary PSMTs at 43 four-year accredited non-profit universities and colleges in

Georgia. Out of these, 11 schools agreed to participate. An estimated 1,200 primary PSMTs were eligible to complete the survey, and 359 primary PSMTs (29.92%) participated in the survey. After data cleaning, 296 primary PSMTs met the inclusion criteria for the study. Approximately 91% of the participants were female, consistent with literature findings (National Center for Education Statistics, 2023). This demographic detail is significant as it supports the study's validity in representing this population.

Instrumentation

The study consists of three instruments and four demographic characteristic variables. The demographic characteristic variables in the study were gender, mathematics experience, statistics experience, and college experience. The three instruments measured primary PSMTs' anxiety about statistics, attitudes toward statistics across four subscales, and statistics teaching efficacy. The four subscales of attitudes toward statistics included affect, cognitive competence, difficulty, and value. The demographic characteristic variables and instruments were made into one survey packet (see Appendix D). The demographic characteristic variables were listed at the end of the survey packet following the three instruments. The final questionnaire consisted of 53 items.

Data Collection and Analysis

The participants completed the final questionnaire online using Qualtrics®. The survey took approximately 15 minutes to complete, and results were formatted into an Excel file before being imported into R for analysis (see Appendix I). Missing data was addressed by first removing cases with more than 50% missing data across all subscales and those with more than 50% missing data from two or more subscales. Little's Missing

Completely at Random (MCAR) test indicated that the remaining missing data points were indeed missing completely at random. Median imputation was employed for these remaining missing data points. Item-level descriptive statistics were computed for all subscales to provide details on the sample of primary PSMTs.

Descriptive statistics and correlation coefficients were computed for all measures utilizing R software packages. Pearson's product-moment correlation coefficients assessed the interval level variables, polychoric correlation coefficients assessed the ordinal level variables, and tetrachoric correlation coefficients assessed the nominal, dichotomous level variable. Questionnaire data from participants was converted to total scale scores for variables except demographic characteristic variables. Statistical considerations and assumptions were checked, and univariate and multivariate outliers were replaced and removed, respectively, to meet the assumptions. The final sample included 296 cases. Structural equation modeling was used to explore the hypothesized paths among the constructs via R software packages.

Summary of Findings

A structural equation model (SEM) was developed and tested to evaluate the relationships among variables within a sample of 296 primary PSMTs from teacher education programs across Georgia. This model aimed to determine direct, indirect, and total effects of various factors influencing statistics teaching efficacy. The sample predominantly comprised females who completed at least three college-level mathematics courses, two or fewer statistics courses, and enrolled in their respective programs for at least three years.

Initial findings indicated significant Pearson’s correlations among interval level variables, specifically between anxiety about statistics, attitudes toward statistics, and statistics teaching efficacy. In contrast, tetrachoric and polychoric correlations among nominal and ordinal variables showed no significant relationships. Notably, anxiety about statistics demonstrated moderate to high positive correlations with all subscales of attitudes toward statistics, while statistics teaching efficacy had only weak positive correlations with these subscales.

The analysis of the final model revealed that anxiety about statistics was not a significant predictor of statistics teaching efficacy. Cognitive competence, which reflects PSMTs' positive attitudes towards their knowledge and intellectual skills in statistics, significantly predicted statistics teaching efficacy. Anxiety about statistics significantly predicted positive attitudes towards statistics across several of the four subscales: affect, cognitive competence, difficulty, and value. Notably, the study found that lower anxiety about statistics enhances cognitive competence, which, in turn, significantly boosts confidence in teaching statistics. The total effect of anxiety on teaching efficacy, mediated by cognitive competence, was also significant.

Furthermore, multigroup analysis was employed to explore differences between groups based on gender, mathematics experience, statistics experience, and college experience. The findings indicated strict measurement invariance across mathematics experience, statistics experience, and college experience, confirming that the model's constructs were consistent across these groups. Thus, latent means and correlations can be compared across groups. Partial strict invariance was found for gender. However, four of the six variables were released, so comparisons across the gender groups should be made with caution.

Discussion of Findings

This study sought to understand the relationship between primary PSMTs' teaching efficacy, attitudes, and anxiety about statistics to shed light on how teacher preparation programs were developing mathematics teachers' statistical knowledge for teaching. Structural equation modeling was employed to test a hypothetical model, with initial findings indicating that the hypothesized model did not fit the observed data. Rationale thought and modification indices guided model modification to refine the hypothetical model to achieve a model of good fit. A revised final model was subsequently developed, successfully fitting the observed data. This research aimed to understand primary PSMTs' statistical knowledge for teaching by studying noncognitive factors that fall under pedagogical content knowledge, which was one of the two components.

This section integrates the results of the research variables within the broader framework of existing literature and theoretical foundations. This section highlights how the study's findings compare to previous studies and explores the underlying reasons for these outcomes. Following this analysis, the implications of the findings are addressed in terms of practical applications for educational practice, contributions to theoretical understanding, and recommendations for future research. This comprehensive approach ensures that the study contributes to academic knowledge and has practical relevance for improving teacher preparation programs in statistics education.

Statistics Teaching Efficacy

This study measured primary PSMTs' statistics teaching efficacy using the Self-Efficacy to Teach Statistics (SETS) survey instrument. Lovett and Lee (2017a) and Harrell-Williams et al. (2015) used the SETS instrument in their studies. The instrument consists of three subscales: reading the data (elementary grades), reading between the data (middle grades), and reading beyond the data (high-school grades). Lovett and Lee (2017a) used all three subscales as their sample consisted of secondary PSMTs. Harrell-Williams et al. (2015) used reading the data and reading between the data subscales as their sample consisted of middle-grades PSMTs. This study used reading the data subscale as the sample consisted of primary PSMTs. In agreement with Lovett and Lee (2017a), this study found PSMTs, on average, were between somewhat confident and confident in teaching statistics in the primary grades. Lovett and Lee (2017a) found PSMTs were most confident in teaching the creation of graphical representations and recognizing statistical results may differ from group to group, which was like the findings of this study. The results of this study showed PSMTs were most confident in recognizing statistical results may differ from group to group and recognizing there will be natural variability between observations for individuals. Lovett and Lee (2017a) found PSMTs were least confident in teaching the selection of appropriate graphical displays and generalizing results from a small to a large group, which are found outside the reading the data subscale.

Harrell-Williams et al. (2015) found PSMTs were least confident on items found in the middle-grades. The researchers found PSMTs were most confident in using dotplot, stem and leaf plot, and tables for describing distributions, recognizing there will be natural variability between observations for individuals, creating dotplot, stem and leaf plot, and tables for summarizing distributions, and recognizing that statistical results may differ in another class or group. All items are found in the primary grades. These results agree with the results of this study. The PSMTs of this study were least confident in creating boxplots for summarizing distributions and using dotplot, stem and leaf plot, and tables for describing distributions.

Regarding correlations, this study found that statistics teaching efficacy had weak, positive correlations with all subscales of attitudes toward statistics (affect, $r(296) = .35$, $p < .001$; cognitive competence, $r(296) = .39$, $p < .001$; difficulty, $r(296) = .15$, $p < .01$; value, $r(296) = .20$, $p < .001$), and a weak, positive correlation with anxiety about statistics ($r(296) = .26$, $p < .001$). Regarding path analysis, this study found that anxiety about statistics was nonsignificant in predicting statistics teaching efficacy ($\beta = -.04$, $Z = -0.23$, $p = .82$). Cognitive competence was the only subscale from attitudes toward statistics that was significantly predictive of statistics teaching efficacy ($\beta = .66$, $Z = 4.47$, $p < .001$); implying PSMTs' positive attitudes toward knowledge and intellectual skill in using statistical knowledge (cognitive competence) significantly predicted high confidence in teaching statistics.

Attitudes Toward Statistics

Estrada et al. (2005) and Hannigan et al. (2013) used the Survey of Attitudes Toward Statistics (SATS) instrument in their exploratory studies of PSMTs' attitudes toward statistics. Estrada et al. (2005) found positive moderate correlations between affect and cognitive competence ($r(367) = .78$, $p < .001$) and affect and value ($r(367) = .64$), $p < .001$. This study found a positive strong correlation between affect and cognitive competence ($r(296) = .85$, $p < .001$), and a positive moderate correlation between affect and value ($r(296) = .56$, $p < .001$). The strongest, positive correlation among the subscales of attitudes toward statistics was between affect and cognitive competence, and the weakest, positive correlation was between difficulty and value ($r(296) = .30$, $p < .001$).

Hannigan et al. (2013) found most teachers had positive attitudes toward statistics across the four subscales of attitudes toward statistics. The researchers concluded that the PSMTs placed value on statistics and had confidence in statistical knowledge. Likewise, this study showed PSMTs placed value in statistics and had confidence in statistical knowledge. For instance, in this study the majority of the PSMTs disagreed with the statements “Statistics is worthless”, and “Statistics is not useful to the typical profession”. The PSMTs in this study reported they understand what is going on in a statistics course and they can learn statistics.

Regarding path analysis, anxiety about statistics was significantly predictive of affect ($\beta = .78$, $Z = 17.14$, $p < .001$), cognitive competence ($\beta = .70$, $Z = 14.90$, $p < .001$), difficulty ($\beta = .49$, $Z = 11.48$, $p < .001$), and value ($\beta = .54$, $Z = 7.40$, $p < .001$); implying PSMTs’ who possessed low anxiety about statistics significantly predicted positive attitudes toward statistics through affect, cognitive competence, difficulty, and value. The final model of the path analysis reported only a path from cognitive competence to statistics teaching efficacy. The path from affect, difficulty, and value to statistics teaching efficacy was not significant and removed to obtain a final model that fits the observed data. Cognitive competence was significantly predictive of statistics teaching efficacy ($\beta = .66$, $Z = 4.47$, $p < .001$), and served as a significant mediator from anxiety about statistics to statistics teaching efficacy ($\beta = .46$, $Z = 4.15$, $p < .001$); implying low anxiety about statistics significantly predicted positive attitudes toward knowledge and intellectual skill in using statistical knowledge (cognitive competence), which, in turn, significantly predicted high confidence in teaching statistics.

Anxiety About Statistics

Yusuf et al. (2019) analyzed the relationship between PSMTs' anxiety about statistics and statistical reasoning and found that PSMTs experienced anxiety when thinking about their performance in the course, such as exams. This study found PSMTs do not feel at ease during a statistics assessment when enrolled in a statistics course. PSMTs reported feeling uptight when taking statistics course assessments and getting a sinking feeling when attempting statistical problems. In general, PSMTs reported feeling uncomfortable, nervous, and confused when working with statistical content. This feeling may relate to their lack of statistical content knowledge. Bostani et al. (2014) found that students (non-PSMTs) in a statistics course who possessed high levels of anxiety about statistics performed lower on assessments compared to students who possessed low to moderate levels of anxiety about statistics. In addition, Macher et al. (2011) found that students in a statistics course who had high levels of anxiety about statistics did not perform well on assessments and lacked interest in statistics. Onwuegbuzie's (2003) investigated graduate students' anxiety towards statistics and found similar results to Bostani et al. (2014) and Macher et al. (2011). This may suggest that the primary PSMTs in the study may lack statistical content knowledge as they reported moderate to high levels of anxiety about statistics.

Zare et al. (2011) measured undergraduate students' anxiety about statistics, statistics teaching efficacy, and academic performance. The results showed a statistically significant relationship between students' self-efficacy to learn statistics and academic performance ($r(323) = .47$, $p =.01$), and students' anxiety about statistics and academic performance ($r(323) = -.41$, $p = .01$). The correlation between self-efficacy to learn statistics and anxiety about statistics was not reported. In this study on primary PSMTs, statistics teaching efficacy and anxiety about statistics had a significant positive, weak correlation (($r(296) = .26$, $p < .001$). In addition, anxiety about statistics had significant, positive, moderate to strong correlations with the four subscales of attitudes toward statistics (affect ($r(296) = .75$, $p < .001$), cognitive competence ($r(296) = .70$, $p < .001$), difficulty ($r(296) = .62$, $p < .001$), and value ($r(296) = .42$, $p < .001$)).

Regarding path analysis, anxiety about statistics had significant positive path coefficients in its' prediction of affect ($\beta = .78$, $Z = 17.14$, $p < .001$), cognitive competence ($\beta = .70$, $p < .001$), difficulty ($\beta = .49$, $Z = 11.48$, $p < .001$), and value ($\beta = .54$, $Z = 7.40$, $p < .001$); implying primary PSMTs' who possessed low anxiety about statistics significantly predicted positive attitudes toward statistics through affect, cognitive competence, difficulty, and value. Anxiety about statistics by primary PSMTs can affect their ability to teach the subject with confidence (Brady & Bowd, 2005). This study found primary PSMTs' anxiety about statistics was not statistically predictive of statistics teaching efficacy. However, anxiety about statistics was significantly predictive of statistics teaching efficacy through mediator cognitive competence ($\beta = .46$, $Z = 4.15$, $p < .001$). This implies low anxiety about statistics significantly predicted positive attitudes toward knowledge and intellectual skill in using statistical knowledge (cognitive competence), which, in turn, significantly predicted high confidence in teaching statistics.

Limitations of the Study

This study aimed to evaluate the interrelationships among anxiety about statistics, attitudes toward statistics, and statistics teaching efficacy that influence primary PSMTs' statistical knowledge for teaching, however, limitations existed that restrict the generalization of the findings. The recruitment of primary PSMTs was contingent upon the cooperation of faculty members, which in turn depended on program directors or deans to provide contact details. Busy schedules and high volumes of emails might have deterred responses from program leaders, thereby reducing the overall response rate. Moreover, the voluntary nature of participation might have introduced biases. Primary PSMTs could opt out or provide false responses as they might not have felt compelled to answer honestly. Primary PSMTs are students; thus, their busy schedules and demanding coursework could have influenced decisions to participate genuinely or at all. These concerns about the authenticity of the data collected, as all responses were self-reported, were considered.

The study's scope was limited to primary PSMTs within teacher education programs in Georgia, which restricts the generalizability of the findings to this specific demographic. Using a nonrandom sample means extending these results to the broader population of primary PSMTs, which should be done cautiously. Additionally, this research explored correlations between the constructs; thus, causation cannot be established.

Demographic variables considered in this study included gender, the number of college mathematics completed, the number of college statistics courses completed, and the number of years completed in a teacher education program. Responses to these items may have needed to be more accurate; for instance, the distinction between mathematics and mathematics education courses was not emphasized, potentially confusing participants. Similarly, courses perceived as related to statistics might have been misreported as statistics courses. A more effective approach might have been asking whether participants completed an introductory or AP statistics course. Additionally, the omission of an ethnicity question in the survey represented a missed opportunity for richer demographic analysis, which should be considered in future research to enhance the understanding of diverse educational impacts.

Structural equation modeling (SEM) introduces several potential constraints. SEM assumes model relationships are linear, and variables are measured without error. This does not reflect the complexity of real-world data. Misspecification of the model, whether through omitted variables or incorrect paths, could lead to biased results. SEM also requires sufficiently large sample sizes to achieve stable and reliable estimates. Although the sample size in this study met SEM assumptions, it was still a small sample and could have led to overfitting and less precise estimates. Additionally, SEM's reliance on covariance structures assumes that data are normally distributed, which, if violated and uncorrected, can significantly impact the findings. These inherent limitations of SEM necessitate cautious interpretation of results. Therefore, supplementary analyses or alternative statistical methods should be employed in future studies to confirm the study’s findings.

Implications for Future Research

Educational systems face many challenges. A significant teacher shortage in mathematics and elementary grades exists in Georgia (Georgia Department of Education, 2024). This highlights a need for effective strategies to enhance current teachers' retention and attract new candidates. This study's findings on primary PSMTs offer suggestions for future research, educational practices, and professional development. First, the study highlights the need to address primary PSMTs’ anxiety about statistics to improve their statistics teaching efficacy. Future research could explore the effectiveness of targeted interventions such as workshops and seminars designed to reduce anxiety about statistics. The reduction in teacher’s anxiety could lead to teachers who are more confident in teaching statistics, which, in turn, could significantly improve student learning. In addition, if successful, these interventions could be integrated into teacher training programs.

The study found a positive correlation between primary PSMTs' cognitive competence (subscale of attitudes toward statistics) and statistics teaching efficacy. The results from the study can guide professional development opportunities for enhancing teachers' statistical content knowledge. Enhancing teachers’ content knowledge would enhance their positive attitudes to their statistical knowledge. Professional development opportunities should focus on the knowledge that teachers need to teach statistics effectively. Therefore, statistical content should be restricted to the K-12 mathematics curriculum content. Content outside this curriculum should be introduced to develop teachers' statistical content knowledge further, but not required.

Additionally, preservice teachers could enhance their statistical content knowledge by taking stand-alone courses in statistics education during their teacher training program. Many universities do not offer such a course. This would allow primary PSMTs to learn more statistical content knowledge and develop pedagogical skills to teach statistical content knowledge effectively. Future research could investigate content knowledge and pedagogical strategies that most effectively boost primary PSMTs' statistical teaching efficacy. For instance, studies could examine whether a stand-alone statistics education course improved statistical content knowledge and the efficacy of statistics teaching for primary PSMTs. Similar studies could be performed on ISMTs regarding professional development opportunities.

Furthermore, the study's findings suggest a need for continuous professional development that helps teachers stay abreast of new teaching methodologies and curricular advancements in statistics education to enhance their pedagogical content knowledge. Research could explore the long-term benefits of regular, structured professional development opportunities for in-service teachers and how these sessions influence their teaching practices and student outcomes. For instance, a professional development series on the four subscales of attitude toward statistics could be implemented where, for each subscale, there is a stand-alone workshop. Attitudes toward statistics would be measured before and after the professional development series. Additionally, qualitative methods could be employed to understand the change in teachers' pedagogical content knowledge, adding to the body of knowledge on teachers' statistical knowledge for teaching. Also, such studies could assess why many teachers chose not to participate in these professional development opportunities and determine ways to incentivize more sustained engagement.

Although this study focused on primary PSMTs, this research could be extended to other populations. A similar study can be conducted on secondary PSMTs, primary ISMTs, and secondary ISMTs. Longitudinal studies could track the transition of primary (or secondary) PSMTs to primary (or secondary) ISMTs. The study could examine how early interventions and teacher training could impact teachers' teaching practices, pedagogical content knowledge, statistical content knowledge, anxiety about statistics, statistics teaching efficacy, and attitudes toward statistics. Research could also investigate the needs of in-service teachers and how various stages of their career might require different support systems to maintain or improve statistics teaching efficacy.

In summary, the implications of this study highlight the importance of comprehensive and customized professional development for both preservice and in-service teachers across all grade levels.

Conclusion

The findings and conclusions from this research study highlight the complexity and significance of noncognitive factors that influence primary PSMTs’ statistical knowledge for teaching. The study determined the directionality of the significant relationships between primary PSMTs’ teaching efficacy, attitudes toward statistics, and anxiety about statistics. These noncognitive factors interplay and shape primary PSMTs’ pedagogical content knowledge, which is one of the two components of the statistical knowledge for teaching framework. By analyzing these noncognitive factors through structural equation modeling, this research verified previous research findings and offered fresh insights particularly relevant to Georgia's educational landscape.

This study addressed a gap in research by studying the interrelationships of noncognitive factors that contribute to primary PMSTs' statistical knowledge for teaching. Noncognitive factors that influence teaching are elements in the pedagogical content knowledge component. The findings from the study provided insight into how teacher training programs are developing mathematics teachers, enabling one to understand better the affective factors that impact teachers' statistical knowledge for teaching. The results provide guidance on how to train PSMTs through stand-alone statistics education courses and train ISMTs through professional development opportunities. Future educational opportunities should highlight the importance of teachers' statistics teaching efficacy, attitudes toward statistics, and anxiety about statistics. These factors influence the teaching and learning of statistics and shape teachers' teaching practices and instructional design. These components directly impact students' learning and motivation toward a subject.

It is known that elevated anxiety levels, negative attitudes, and low teaching efficacy can lead to teacher burnout and job satisfaction. Additionally, these factors can lead to mathematics teachers skipping statistical content found in their mathematics curriculum. Mathematics and statistics are two different subjects, and teachers with negative feelings and beliefs about statistics are more prone to skip the content. Hence, these factors directly influence students' learning of statistics. Therefore, continuing research on these factors are important to the educational community.

Future research should continue to explore noncognitive factors and other mediators that influence teachers' statistical knowledge for teaching. Expanding the scope to include in-service and secondary teachers could provide insights into other factors that affect teaching and learning statistics. Additionally, longitudinal studies could provide valuable data on the long-term impacts of the proposed interventions. This could further inform teacher education programs and policy decisions, ideally leading to a more robust, effective, and responsive educational system that can meet the needs of teachers and students in Georgia and beyond. In addition, better survey instruments are needed to study teachers' pedagogical content knowledge. These instruments must meet psychometric properties and be convenient for time. This will allow more quantitative studies in the field of K-12 statistics education.

In summary, this research contributes to a deeper understanding of the factors influencing how statistics is taught at the primary level, offering evidence-based strategies to enhance teaching outcomes. The recommendations hoped to foster better statistical teaching practices and inspire continued scholarly inquiry into this critical aspect of statistics education.

References

Akin, A., & Kurbanoglu, I. N. (2011). The relationships between math anxiety, math attitudes, and self-efficacy: A structural equation model. *Studia Psychologica*, *53*(3), 263–273.

Alam, S., Ayub, M. S., Arora, S., & Khan, M. A. (2023) An investigation of the imputation techniques for missing values in ordinal data enhancing clustering and classification analysis validity. *Decision Analytics Journal*, *9*. https://doi.org/10.1016/j.dajour.2023.100341

Arnold, P. (2008). *Developing new statistical content knowledge with secondary school mathematics teachers* [Paper presentation]. ICMI Study 18 Conference and IASE 2008 Round Table Conference, Monterrey, Mexico. https://iase-web.org/documents/papers/rt2008/T4P5_Arnold.pdf?1402524990

Aslam, M., Sherwani, R., & Saleem, M. (2021). Vague data analysis using neutrosophic Jarque-Bera test. *PLoS One, 16*(12) https://doi.org/10.1371/journal.pone.0260689

Auzmendi, E. (1991). *Factors related to statistics: A study with a Spanish sample*. [Paper presentation]. Annual meeting of the American Educational Research Association. Chicago, Il.

Aydin, S. (2021). A systematic review of research on teaching anxiety. *International Online Journal of Education and Teaching (IOJET), 8*(2). 730-761.

Ball, D. L., Thames, M. H., & Phelps, G. (2005). *Articulating domains of mathematical knowledge for teaching*. www.personal.umich.edu/~dball/

Ball, D. L., Thames, M. H., & Phelps, G. (2008). Content knowledge for teaching: What makes it special? *Journal of Teacher Education*, *59*(5), 389–407. https://doi.org/10.1177%2F0022487108324554

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

Bargagliotti, A., Franklin, C., Arnold, P., Gould, R., Johnson, S., Perez, L., & Spangler, D. A. (2020). *Pre-K-12 Guidelines for assessment and instruction in statistics education II (GASIE II)*. American Statistical Association.

Batur, A., Kontas, H., & Ozcan, B. (2020). Adapting sources of middle school mathematics self-efficacy scale to Turkish culture. *International Journal of Evaluation and Research in Education (IJERE)*, *6*(4), 288. https://doi.org/10.11591/ijere.v6i4.10771

Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A., Tsai, Y. M. (2010). Teachers' mathematical knowledge, cognitive activation in the classroom, and student progress. *American Educational Research Journal, 47*(1), 133–180.

Begg, A., & Edwards, R. (1999). Teachers' ideas about teaching statistics. *Proceedings of the 1999 combined conference of the Australian Association for Research in Education and the New Zealand Association for Research in Education*. Melbourne: AARE & NZARE. www.aare.edu.au/99pap/beg99082.htm.

Beswick, K., Callingham, R., & Watson, J. (2012). The nature and development of middle school mathematics teachers' knowledge. *Journal of Mathematics Teacher Education, 15*(2), 131–157.

Betz, N. E. (1978). Prevalence, distribution and correlates of math anxiety in college students. *Journal of Counseling Psychology, 25*, 441-448.

Bostani, M., Nadri, A., & Nasab, A. R. (2014). A study of the relation between mental health and academic performance of students of the Islamic Azad University Ahvaz branch. *Procedia Social and Behavioral Sciences, 116*, 163–165. https://doi.org/10.1016/j.sbspro.2014.01.186

Brady, P., & Bowd, A. (2005). Mathematics anxiety, prior experience, and confidence to teach mathematics among pre-service education students. *Teachers and Teaching: Theory and Practice, Philadelphia, 11*(1), 37-46.

Bulut, O. (2020). *Testing for measurement invariance in R*. Okan Bulut. https://okan.cloud/posts/2020-12-21-testing-for-measurement-invariance-in-r/

Burgess, T. (2006). A framework for examining teacher knowledge as used in action while teaching statistics. In A. Rossman & B. Chance (Eds.), *Proceedings of the Seventh International Conference on Teaching Statistics*. International Statistics Institute.

Burgess, T. (2008). Teacher knowledge for teaching statistics through investigations. In C. Batanero, G. Burrill, C. Reading & A. Rossman (Eds.), *Joint ICMI/IASE Study: Teaching Statistics in School Mathematics. Challenges for Teaching and Teacher Education. Proceedings of the ICMI Study 18 and 2008 IASE Round Table Conference*. International Statistics Institute.

Burrill, G., & Biehler, R. (2011). Fundamental statistical ideas in the school curriculum and in training teachers. *Teaching Statistics in School Mathematics-Challenges for Teaching and Teacher Education*, 57-69. https://doi.org/10.1007/978-94-007-1131-0_10

Capaldi, M. B. (2019). Mathematics versus statistics. *Journal of Humanistic Mathematics*, *9*(2), 149-156. https://doi.org10.5642/jhummath.201902.10

Cashin, F., & Elmore, P. (2005). The Survey of Attitudes Toward Statistics scale: A construct validity study. *Educational and Psychological Measurement*, *65,* 509-524.

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, 14, Article 464504.

Chick, H. L., & Pierce, R. U. (2008). Teaching statistics at the primary school level: Beliefs, affordances, and pedagogical content knowledge. *Joint ICMI/IASE Study: Teaching Statistics in School Mathematics. Challenges for Teaching and Teacher Education. Proceedings of the ICMI Study 18 and 2008 IASE Round Table Conference*. International Statistics Institute.

Christophel, D. M., & Gorham, J. (1995). A test-retest analysis of student motivation, teacher immediacy, and perceived sources of motivation and demotivation in college classes. *Communication Education*, *44*(4), 292–306.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Mahwah, NJ: Erlbaum.

Comrey, A. L., & Lee, H. B. (2013). *A first course in factor analysis*. Psychology Press.

Creswell, J. W. (2009). *Research design: Qualitative, quantitative, and mixed methods approaches* (3rd ed.). Sage Publications, Inc.

Cruise, R. J., Cash, R. W., & Bolton, D. L. (1985). Development and validation of an instrument to measure statistical anxiety. In *Annual Meeting of the American Statistical Association: Statistics Education Section*. Proceeding of the American Statistical Association Dublin: OCLC, 417 – 425.

delMas, R., Garfield, J. B., Ooms, A., & Chance, B. (2007). Assessing students' conceptual understanding after a first-year course in statistics. *Statistics Education Research Journal*, *6*(2), 28-58. https://doi.org/10.52041/serj.v6i2.483

De Vetten, A., Schoonenboom, J., Keijzer, R., & van Oers, B. (2018). The development of informal statistical inference content knowledge of pre-service primary school teachers during a teacher college intervention. *Educational Studies in Mathematics*, *99*(2), 217–234. https://doi.org/10.1007/s10649-018-9823-6

Doerr, H., & Jacob, B. (2011). Investigating secondary teachers' statistical understandings. In M. Pytlak, T. Rowland, & E. Swoboda (Eds.), *Proceedings of the Seventh Congress of the European Society for Research in Mathematics Education*, 776–786.

Engledowl, C. & Tarr, J. E. (2020). Secondary teachers' knowledge structures for measures of center, spread & shape of distribution supporting their statistical reasoning. *International Journal of Education in Mathematics, Science and Technology (IJEMST), 8*(2), 146-167.

Epstein, R. M., & Hundert, E. M. (2002). Defining and assessing professional competence. *JAMA, 287*(2), 226–235. https://doi.org/10.1001/jama.287.2.226

Estrada, A., & Batanero, C. (2008). Explaining teachers' attitudes towards statistics. *Joint ICMI/IASE Study: Teaching Statistics in School Mathematics. Challenges for*

Teaching and Teacher Education. Proceedings of the ICMI Study 18 and 2008 IASE Round Table Conference. International Statistics Institute.

Estrada, A., Batanero, C., Fortuny, J. M., & Diaz, C. (2005). A structural study of future teachers' attitudes towards statistics. In M. Bosch (Ed.), *Proceedings of the Fourth Congress of the European Society for Research in Mathematics Education* (508- 517). Barcelona: IQS Fundemi. www.ermeweb.free.fr/CERME4/

Estrada, A., Batanero, C., & Lancaster, S. (2011). Teachers' Attitudes Towards Statistics. *Teaching Statistics in School-Mathematics Challenges for Teaching and Teacher Education: A Joint ICMI/IASE Study*. International Statistics Institute.

Evans, B. (2007). Student attitudes, conceptions, and achievement in introductory undergraduate college statistics. *The Mathematics Educator, 17*(2), 24–30.

Field, A. P. (2013). *Discovering statistics using SPSS* (4th ed.). Sage publications.

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

Fishbein, M., & Ajzen, I. (1975). *Belief, attitude, intention and behaviour: An introduction to theory and research.* Reading, MA: Addison-Wesley.

Franklin, C. A., Bargagliotti, A. E., Case, C. A., Kader, G. D., Scheaffer, R. L., & Spangler, D. A. (2015). *The statistical education of teachers*. American Statistical Association website.

Franklin, C., Kader, G., Mewborn, D., Moreno, J., Peck, R., Perry, M., & Scheaffer, R. (2007). *Guidelines for assessment and instruction in statistics education (GAISE) report: A Pre-K–12 curriculum framework*. American Statistical Association.

Froelich, A. G. (2011). *Developing a Statistics Curriculum for Future Secondary Mathematics Teachers*. New ICMI Studies Series. https://doi.org/10.1007/978-94-007-1131-0_5

Gal, I., & Ginsburg, L. (1994). The role of beliefs and attitudes in learning statistics: Towards an assessment framework. *Journal of Statistics Education, 2*(2). www.amstat.org/publications/jse/

Gal, I., Ginsburg, L., & Schau, C. (1997). Monitoring attitudes and beliefs in statistics education. In I. Gal & J. B. Garfield (Eds.)*, The assessment challenge in statistics education*. Voorburg, Netherlands: IOS Press.

Garfield, J. B. (1996). Assessing student learning in the context of evaluating a chance course. *Communications in Statistics: Theory and Methods*, *25*(11), 2863-2873.

Garritz, A. (2000). Pedagogical content knowledge and the affective domain of scholarship of teaching and learning. *International Journal for the Scholarship of Teaching and Learning*, *4*(2), 1-6. https://doi.org/10.20429/ijsotl.2010.040226

Geiser, C. (2023). Structural equation modeling with the Mplus and lavaan programs. In *R. H. Hoyle: Handbook of structural equation modeling* (241-258). Guilford Press.

Georgia Department of Education (2021). *Framework for statistical reasoning*. https://lor2.gadoe.org/gadoe/file/5e835b39-307f-4d61-aa50-6e3f58edbf22/1/K-12-Statistical-Reasoning-Framework.pdf

Georgia Department of Education (2024). *High needs subject area.* https://www.gadoe.org/External-Affairs-and-Policy/Educator-Support-Services/Documents/HNSA%202024-2025%20FINAL.pdf

Goldman, S. (2001). Professional development in a digital age: Issues and challenges for standards-based reform. *Interactive Educational Multimedia, 2*, 19-46.

Gómez-Chacón, I. (2000). Affective influences in the knowledge of mathematics. *Educational Studies in Mathematics, 43*(2), 149-168.

Gould, R., & Peck, R. (2004). *Preparing secondary teachers to teach statistics* [Paper presentation]. The 10th International Congress on Mathematics Education. Copenhagen, Denmark.

Gourneau, B. (2005). Five attitudes of effective teachers: Implications for teacher training. *Essays in Education, 13*(5).

Groth, R. E. (2007). Toward a conceptualization of statistical knowledge for teaching. *Journal for Research in Mathematics Education*, *38*(5), 427–437. https://doi.org/10.2307/30034960

Groth, R. E. (2013). Characterizing key developmental understandings and pedagogically powerful ideas within a statistical knowledge for teaching framework. *Mathematical Thinking and Learning*, *15*(2), 121–145. http://dx.doi.org/10.1080/10986065.2013.770718

Groth, R. E. (2017). Developing statistical knowledge for teaching during design-based research. *Statistics Education Research Journal*, *16*(2), 376–396. http://dx.doi.org/10.52041/serj.v16i2.197

Groth, R., & Meletiou-Mavrotheris, M. (2018). Research on statistics teachers' cognitive and affective characteristics. In D. Ben-Zvi, K. Makar, & J. Garfield (Eds.), *International handbook of research in statistics education*. Springer. https://doi.org/10.1007/978-3-319-66195-7_10

Hannigan, A., Gill, O., & Leavy, A. M. (2013). An investigation of prospective secondary mathematics teachers’ conceptual knowledge of and attitudes towards statistics. *Journal of Mathematics Teacher Education*, *16*(6), 427–449. http://dx.doi.org/10.1007/s10857-013-9246-3

Harrell-Williams, L. M., Sorto, M. A., Pierce, R. L., Lesser, L. M., & Murphy, T. J. (2015). Identifying statistical concepts associated with high and low levels of self-efficacy to teach statistics in middle grades. *Journal of Statistics Education*, *23*(1), 1–10. http://dx.doi.org/10.1080/10691898.2015.11889724

Harrell-Williams, L. M., Sorto, M. A., Pierce, R. L., Lesser, L. M., & Murphy, T. J. (2013). Validation of scores from a new measure of preservice teachers’ self-efficacy to teach statistics in the middle grades. *Journal of Psychoeducational Assessment*, *32*(1), 40–50. http://dx.doi.org/10.1080/10691898.2015.11889724

Hassani, H., Beneki, C., Silva, E., Vandeput, N., & Madesn, D. O. (2021). The science of statistics versus data science: What is the future? *Technological Forecasting and Social Change*, *173*(2). https://doi.org/10.1016/j.techfore.2021.12111

Hembree, R. (1990). The nature, effects, and relief of mathematics anxiety. *Journal for Research in Mathematics Education, 21*(1), 33–46.

Hill, H. C., Rowan, B., & Ball, D. L. (2005). Effects of teachers' mathematical knowledge for teaching on student achievement. *American Educational Research Journal, 42*(2), 371-406.

Hill, H. C., Schilling, S., & Ball, D. L. (2004). Developing measures of teachers’ mathematics knowledge for teaching. *Elementary School Journal*, *105*(1), 11-30.

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

Hotaman, D. (2010). The teaching profession: Knowledge of subject matter, teaching skills and personality traits. *Procedia Social and Behavioral Sciences*, *2*(1), 1416-1420. https://doi.org/10.1016/j.sbspro.2010.03.211

Jacobbe, T. (2016). The Levels of Conceptual Understanding in Statistics (LOCUS) project: Results of the pilot study. *Numeracy*, *8*(2). http://dx.doi.org/10.5038/1936-4660.8.2.3

Keeley, J., Zayac, R., & Correia, C. (2008). Curvilinear relationships between statistics anxiety and performance among undergraduate students: Evidence for optimal anxiety. *Statistics Education Research Journal, 7*(1), 4–15.

Kelly, W. P., & Tomhave, W. K. (1985). A study of math anxiety and math avoidance in preservice elementary teachers. *The Arithmetic Teacher, 32*, 51–53

Kim, H., & Millsap, R. (2010). Using the Bollen-Stine bootstrapping method for evaluating approximate fit indices. *Multivariate Behavior Research*, *49*(6), 581-596. http://dx.doi.org10.1080/00273171.2014.947352

Kline, R. B. (2015). *Principles and practice of structural equation modeling*. Guilford publications.

Knight, R. (2021). *SC mathematics academic standards cyclical review*. Academic Standards and Assessment Subcommittee.

Korkmaz, S., Goksuluk, D., & Zararsiz, G. (2014). MVN: An R package for assessing multivariate normality. *The R Journal*, *6*(2).

Lancaster, S. (2008). A study of pre-service teachers' attitudes toward their role as students of statistics and implications for future professional development in statistics. *Joint ICMI/IASE Study: Teaching Statistics in School Mathematics. Challenges for Teaching and Teacher Education. Proceedings of the ICMI Study 18 and 2008 IASE Round Table Conference*. International Statistics Institute.

Lee, H. S., Mojica, G. M., & Lovett, J. N. (2020). Examining how online professional development impacts teachers’ beliefs about teaching statistics. *Online Learning Journal*, *24*(1), 5–27. http://dx.doi.org/10.24059/olj.v24i1.1992

Leavy, A., & O‘Loughlin, N. (2006). Preservice teachers understanding of the mean: Moving beyond the arithmetic average. *Journal of Mathematics Teacher Education, 9*(1), 53–90.

Linacre J. M. (2004). Optimal rating scale category effectiveness. In Smith E. V., Smith R. M. (Eds.), *Introduction to Rasch measurement*. Maple Grove, MN: JAM Press.

Loehlin, J. C. (1992). *Latent variable models: An introduction to factor, path, and structural analysis* (2nd ed.). Lawrence Erlbaum Associates, Inc.

Lovett, J. N., & Lee, H. S. (2017a). New standards require teaching more statistics: Are pre-service secondary mathematics teachers ready? *Journal of Teacher Education*, *68*(3), 299–311. https://doi.org/10.1177/0022487117697918

Lovett, J. N., & Lee, H. S. (2017b). Pre-service secondary mathematics teachers' statistical knowledge: A snapshot of strengths and weaknesses. In E. Galindo & J. Newton (Eds.), *Proceedings of the 39th annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (1048–1055).

Macher, D., Paechter, M., Papousek, I., Ruggeri, K., Freudenthaler, H. H., & Arendasy, M. (2013). Statistics anxiety, state anxiety during an examination, and academic achievement. *The British Journal of Educational Psychology, 83*, 535–549. https://doi.org/10.1111/j.2044-8279.2012.02081

Makar, K., & Confrey, J. (2004). Secondary Teachers' Statistical Reasoning in Comparing Two Groups. In Ben-Zvi, D., Garfield, J. (Eds.), *The Challenge of Developing Statistical Literacy, Reasoning and Thinking*. Springer, Dordrecht. https://doi.org/10.1007/1-4020-2278-6_15

Makar, K., & Fielding-Wells, J. (2011). Teaching Teachers to Teach Statistical Investigations. In Batanero, C., Burrill, G., Reading, C. (Eds.), *Teaching Statistics in School Mathematics-Challenges for Teaching and Teacher Education: New ICMI Study Series*. Springer, Dordrecht. https://doi.org/10.1007/978-94-007-1131-0_33

Math/Stat Teacher Education: Assessment, Methods, and Strategies (TEAMS). (n.d.) American Statistical Association. www.amstat.org/sections/educ/newsletter/v9n1/TEAMS.html

McLeod, D. B. (1992). Research on affect in mathematics education: A reconceptualization. In D. A. Grows (Eds.), *Handbook of Research on Mathematics Teaching and Learning* (575-596). New York: Macmillan and National Council of Teachers of Mathematics.

Mishra, P., Pandey, C. M., Singh, U., Gupta, A., Sahu, C., & Keshri, A. (2019). Descriptive statistics and normality tests for statistical data. *Ann Card Anaesth.*, *22*(1), 67-72. https://doi.org/10.4103/aca.ACA_157_18

Moody, V. R., & Gardner, L. (2014). An investigation of a pre-service secondary mathematics teacher's conceptual understanding of statistical content and analysis. *American Journal of Educational Research*, *2*(8), 617–623. https://doi.org/10.12691/education-2-8-10

Mueller, R. O., & Hancock, G. R. (2019). Structural equation modeling. In G. R. Hancock, L. M. Stapleton, & R. O. Mueller (Eds.), *The reviewer's guide to quantitative methods in the social sciences* (2nd ed., 445–456). Routledge/Taylor & Francis Group. https://doi.org/10.4324/9781315755649-33

Nasser, F. M. (2004). Structural model of the effects of cognitive and affective factors on the achievement of Arabic-speaking prospective teachers in introductory statistics. *Journal of Statistics Education, 12*(1).

National Center for Education Statistics. (2023). *Characteristics of public-school teachers: Condition of education*. U.S. Department of Education, Institute of Education Sciences. https://nces.ed.gov/programs/coe/indicator/clr.

National Council of Teachers of Mathematics (2022). *Joint NCTM-ASA Position Statement on Preparing PK–12 Teachers of Statistics and Data Science.* https://www.nctm.org/Standards-and-Positions/Position-Statements/Preparing-Pre-K-12-Teachers-of-Statistics/

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. https://www.nctm.org/standards/

National Council on Teacher Quality. (2016). *A closer look at elementary mathematics*. https://www.nctq.org/dmsView/UE_2016_Math_Findings

National Council on Teacher Quality. (2018). *A closer look at elementary mathematics*. https://www.nctq.org/dmsView/2018_Math_Findings

Neumann, K., Kind, V., & Harms, U. (2017). Probing the amalgram: The relationship between science teachers' content, pedagogical and pedagogical content knowledge. *International Journal of Science Education*, *41*(7). https://doi.org/10.1080/09500693.2018.1497217

Olson, J. M., & Zanna, M. P. (1993). Attitude and attitude change. *Annual Review of Psychology, 44,* 117-154.

Onwuegbuzie, A. J. (1998). Teachers' attitudes toward statistics. *Psychological Reports, 83*, 1008-1010.

Onwuegbuzie, A. J. (2003). Modelling statistics achievement among graduate students. *Educational and Psychological Measurement, 63*, 1020-1038.

Onwuegbuzie, A. J., Daros, D., & Ryan, J. (1997). The components of statistics anxiety: A phenomenological study. *Focus on Learning Problems in Mathematics, 19*(4), 11-35.

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

Peters, S. A. (2014). Developing understanding of statistical variation: Secondary statistics teachers' perceptions and recollections of learning factors. *Journal of Mathematics Teacher Education*, *17*(6), 539–582. http://dx.doi.org/10.1007/s10857-013-9242-7

Philipp, R. A. (2007). Mathematics teachers' beliefs and affects. In F. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (257-315).

Charlotte, NC: Information Age Publishing and National Council of Teachers of Mathematics.

Pretorius, T. B., & Norman, A. M. (1992). Psychometric data on the Statistics Anxiety Scale for a sample of South African students. *Educational and Psychological Measurement, 52*(4), 933–937. https://doi.org/10.1177/0013164492052004015

Roberts, D. M., & Bilderback, E. W. (1980). Reliability and validity of a statistics attitude survey. *Educational and Psychological Measurement, 40*, 235-238.

Ronau, R., Brown, E. & Collins, G. (2010). Validation of the Diagnostic Teacher Assessment of Mathematics and Science (DTAMS) Instrument. *School Science and Mathematics, 110*, 180 - 192. https://doi.org/10.1111/j.1949-8594.2010.00021.x

Ross, J. A., & Bruce, C. (2007). Self-assessment and professional growth: The case of a grade 8 mathematics teacher. *Teaching and Teacher Education, 23*(2), 146-159.

Sabbag, A. G., & Zieffler, A. (2015). Assessing learning outcome: An analysis of the GOALS-2 instrument. *Statistics Education Research Journal*, *14*(2), 93-116.

Saderholm, J., Ronau, R., Brown, E. T., & Collins, G. (2010). Validation of the diagnostic teacher assessment of mathematics and science (DTAMS) instrument. *School Science and Mathematics, 110*(4), 180–192. https://doi.org/10.1111/ssm.2010.110.issue-4

Schau, C. (2003). Students' attitudes: The other important outcome in statistics education. *Joint Statistical Meeting of the American Statistical Association*. San Francisco, CA.

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

Schoen, R. C., LaVenia, M., Chick, E., Razzouk, R., & Kisa, Z. (2019). Increasing secondary-level teachers' knowledge in statistics and probability: Results from a randomized controlled trial of a professional development program. *Cogent Education*, *6*(1), Article 1613799. https://doi.org/10.1080/2331186X.2019.1613799

Schreiber, J., Nora, A., Stage, F., Barlow, E., & King, J. (2006). Reporting structural equation modeling and confirmatory factor analysis results: A review. *Journal of Educational Research*, *99*, 323-338. https://doi.org/10.3200/JOER.99.6.323-338

Schumaker, R. E., & Lomax, R. G. (2016). *A Beginners Guide to Structural Equation Modeling*. New York, NY: Routledge.

Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, *15*(2), 4–14. http://dx.doi.org/10.3102/0013189X015002004

Shulman, L. S. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review*, 57(1).

Shulman, L. S. (2007). Interview conducted at the annual meeting of the American Educational Research Association, Chicago, April. Part of the text is reproduced in Berry, Loughran & van Driel (2008) reference.

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

Steen, L. A. (2001). *Mathematics and democracy: The case for quantitative literacy*. Washington, DC: National Council on Education and the Disciplines.

Steenkamp, J. B. E., & Baumgartner, H. (1998). Assessing measurement invariance in cross-national consumer research. *Journal of Consumer Research, 25*(1), 78–90.

Suhr, D. (2006). *The basics of structural equation modeling*. http://www.scribd.com/doc/205955716/tut-suhr-pdf

Tabachnick, B. G., & Fidell, L. S. (2013). *Using Multivariate Statistics* (6th ed.). Great Britain, England: Clays Ltd.

Tasnimi, M. (2009). Affective factors: Anxiety. *Journal of Pan-Pacific Association of Applied Linguistics*, *13*(2), 117-124.

Tchoshanov, M. A. (2011). Relationship between teacher knowledge of concepts and connections, teaching practice, and student achievement in middle grades mathematics. *Educational Studies in Mathematics, 76*(2), 141–164.

Tempelaar, D. T., Gijselaers, W. H., van der Loeff, S. S., & Nijhuis, J. F. H., (2007). A structural equation model analyzing the relationship of student achievement motivations and personality factors in a range of academic subject-matter areas. *Contemporary Educational Psychology, 32*(1), 105–131. https://doi.org/10.1016/j.cedpsych.2006.10.004

Umugiraneza, O., Bansilal, S., & North, D. (2018). Examining teachers' perceptions about improving the teaching and learning of mathematics and statistics. *Statistics Education Research Journal*, *17*(2), 239-254.

U.S. Bureau of Labor Statistics. (2022). *U.S. Department of labor, occupational outlook handbook* (2021 - 2022 ed.). https://www.bls.gov/ooh/math/mathematicians-and-statisticians.htm

Usiskin, Z. (2015). Mathematical modeling and pure mathematics. *Mathematics Teaching in the Middle School*, *20*(8), 476-482. https://doi.org/10.5951/mathteacmiddscho.20.8.0476

Van de Schoot, R., Lugtig, P., & Hox, J. (2012). A checklist for testing measurement invariance. *European Journal of Developmental Psychology, 9*(4), 486–492.

Vitasari, P., Abdul Wahab, M. N., Othman, A., & Awang, M. G. (2010). A research for identifying study anxiety sources among university students. *International Education Studies, 3*(2), 189-196.

Ward-Penny, R. (2011). Cross-curricular teaching and learning in the secondary school: Mathematics (1st ed.). Routledge. https://doi.org/10.4324/9780203835630

Watson, J. M. (2001). Profiling teachers' competence and confidence to teach particular mathematics topics: The case of chance and data. *Journal of Mathematics Teacher Education, 4*(4), 305-337.

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

Williams, A. S. (2010). Statistics anxiety and instructor immediacy. *Journal of Statistics Education, 18*, 1–18.

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

Wilson, M. S., & Cooney, T. J. (2002). Mathematics teacher change and development: The role of beliefs. In G. C. Leder, E. Pehkonen, & G. Torner (Eds.), *Beliefs: A hidden variable in mathematics education?*. London: Kluwer Academic Publishers.

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

Wolfe, E. W. & Smith, E. V., Jr. (2007). Instrument development tools and activities for measure validation using Rasch models: Part II Validation activities. In Smith E. V., Smith R. M. (Eds.), *Rasch measurement: Advanced and specialized applications* (243-290). Maple Grove, MN: JAM Press.

Yusuf, Y., Suyitno, H., Sukestiyarno, Y. L., & Isnarto, A. (2019). The influence of statistical anxiety on statistic reasoning of preservice mathematics teachers. *BOLEMA, 33*(64), 694-706. https://doi.org/10.1590/1980-4415v33n64a12

Zare, H., Rastegar, A., & Hosseini, S. M. D. (2011). The relation among achievement goals and academic achievement in statistics: The mediating role of statistics anxiety and statistics self-efficacy. *Procedia - Social and Behavioral Sciences, 30*, 1166–1172. https://doi.org/10.1016/j.sbspro.2011.10.227

Zientek, L. R., Carter, T. A., Taylor, J. M., & Capraro, R. M. (2011). Preparing prospective teachers: An examination of attitudes toward statistics. *Journal of Mathematical Sciences & Mathematics Education*, *5*(1), 25–38.

Zhu, Y., & Kaiser, G. (2022). Impacts of classroom teaching practices on students' mathematics learning interest, mathematics self-efficacy and mathematics test achievements: A secondary analysis of Shanghai data from the international video study Global Teaching InSights. *ZDM Mathematics Education, 54*, 581–593. https://doi.org/10.1007/s11858-022-01343-9

Appendix A:

Institutional Review Board (IRB) Approval Form

Institutional Review Board (IRB)
For the Protection of Human Research Participants

Protocol Exemption Report

Protocol Number: 04470-2023 **Responsible Researcher(s):** Andrew Shealy

Supervising Faculty: Dr. Lantry Brockmeier

Project Title: *Primary Preservice Mathematics Teachers' Sense of Efficacy, Attitudes, and Anxiety About Statistics.*

INSTITUTIONAL REVIEW BOARD DETERMINATION:

This research protocol is **exempt** from Institutional Review Board (IRB) oversight under 45 CFR 46.101(b) of the federal regulations, **category 2**. If the nature of the research changes such that exemption criteria no longer apply, please consult with the IRB Administrator (irb@valdosta.edu) before continuing your research study.

ADDITIONAL COMMENTS:

☒ ***If this box is checked, please submit any documents you revise to the IRB Administrator at tmwright@valdosta.edu to ensure an updated record of your exemption.***

Elizabeth Ann Olphie 11.15.2023

Elizabeth Ann Olphie, IRB Administrator

Thank you for submitting an IRB application.

Please direct questions to irb@valdosta.edu or 229-253-2947.

Revised: 06.02.16

Appendix B:

Letter to Director or Dean

Hello,

I need your assistance. My dissertation topic is primary (elementary) preservice teachers' sense of efficacy, attitudes, and anxiety about statistics. As a mathematics educator, I am aware of the issues teachers are faced with daily in teaching mathematics. I plan to conduct a confidential, online questionnaire in hopes of discovering a set of predictor variables for primary preservice teachers' anxiety about statistics. The online questionnaire will take approximately 10-15 minutes to complete.

I would like to request a letter of cooperation to conduct research by surveying primary preservice teachers enrolled in your elementary program. I am required to submit letters of cooperation from participating schools before I am granted final IRB approval. If you are unable to write a letter of cooperation before final IRB approval and are willing to participate in the study, then please let me know.

I will be contacting you in the next two weeks to request your assistance with contacting primary preservice teachers within your program and/or would like to speak with you directly about your thoughts or opinions related to this survey. The anticipated time of data collection will be the Spring semester of 2024. Participation in this study will help to improve our understanding of how teacher preparation programs are preparing teachers to teach statistical content.

Thank you for your time and consideration. If you have further questions regarding this study, you may contact me at akshealy@valdosta.edu. I look forward to speaking to you soon.

Sincerely,

Andrew Shealy
Valdosta State University
Curriculum and Instruction
Doctoral Student

Appendix C:

GAISE II Framework

Process Component	Level A	Level B	Level C
I. Formulate Statistical Investigative Questions	Understand when a statistical investigation is appropriate. Pose statistical investigative questions of interest to students where the context is such that students can collect or have access to all required data. Pose summary (or descriptive) statistical investigative questions about one variable regarding small, well-defined groups (e.g., subset of a classroom, classroom, school, town) and extend these to include comparison and association statistical investigative questions between variables. Experience different types of questions in statistics: those used to frame an investigation, those	Recognize that statistical investigative questions can be used to articulate research topics and that multiple statistical investigative questions can be asked about any research topic. Understand that statistical investigative questions take into account context as well as variability present in data. Pose summary, comparative, and association statistical investigative questions about a broader population using samples taken from the population. Pose statistical investigative questions that require looking at a variable over time. Understand that there are different types of	Formulate multivariable statistical investigative questions and determine how data can be collected and analyzed to provide an answer. Pose summary, comparative, and association statistical investigative questions for surveys, observational studies, and experiments using primary or secondary data. Pose inferential statistical investigative questions regarding causality and prediction.

Appendix C (continued)

	used to collect data, and those used to guide analysis and interpretation.	questions in statistics: those used to frame an investigation, those used to collect data, and those used to guide analysis and interpretation. Pose statistical investigative questions for data collected from online sources and websites, smartphones, fitness devices, sensors, and other modern devices.	
Process Component	**Level A**	**Level B**	**Level C**
II. Collect Data/ Consider Data	Understand that data are information; recognize that to answer a statistical investigative question, a person may collect data themselves specifically for that purpose, or a person may use data that have been collected by other people for another purpose. Understand how to collect and record information from the group of interest	Understand that data are information collected and recorded with a purpose and can be organized and stored in a variety of structures (e.g., spreadsheets). Understand that a sample can be used to answer statistical investigative questions about a population. Recognize the limitations and scope	Word as: Apply an appropriate data collection plan when collecting primary data or selecting secondary data for the statistical investigative question of interest. Distinguish between surveys, observational studies, and experiments. Understand what constitutes good practice in designing a sample survey, an

Appendix C (continued)

	using surveys and measurements collected from observations and simple experiments. Understand that a variable measures the same characteristic on several individuals or objects and results in data values that may fluctuate. Understand that within a data set there can be different types of variables (e.g., categorical or quantitative). Interrogate the data set to understand the context of the variables as they may relate to statistical investigative questions. Understand that data are not always pristine but may contain errors, have missing values, etc., and that decisions have to be made about how to account for these issues.	of the data collected by describing the group or population from which the data are collected. Understand that data can be used to make comparisons between different groups at one point in time and the same group over time. Recognize that data can be collected using surveys and measurements, and develop a critical attitude in analyzing data collection methods. Understand that quantitative variables may be either discrete or continuous. Understand how to interrogate the data to determine how the data were collected, from whom they were collected, what types of variables are in the data, how the variables were measured (including units used), and the	experiment, and an observational study. Understand the role of random selection in sample surveys and the effect of sample size on the variability of estimates. Understand the role of random assignment in experiments and its implications for cause-and-effect interpretations. Understand the issues of bias and confounding variables in observational studies and their implications for interpretation. Understand practices for handling data that enhance reproducibility and ensure ethical use, including descriptions of alterations, and an understanding of when data may contain sensitive information.

Appendix C (continued)

		possible outcomes for the variables. Understand that data can be collected (primary data) or existing data can be obtained from other sources (secondary data). Understand how random assignment in comparative experiments is used to control for characteristics that might affect responses.	Understand how concerns about privacy and human subjects may affect the collection and distribution of data. Understand that in some circumstances, the data collected or considered may not generalize to the desired population, or this data may be the entire population.
Process Component	**Level A**	**Level B**	**Level C**
III. Analyze the Data	Understand that the distribution of a categorical variable or quantitative variable describes the number of times a particular outcome occurs. Represent the variability of categorical variables or quantitative variables using appropriate displays (e.g., tables, picture graphs, dotplots, bar graphs).	Represent the variability of quantitative variables using appropriate displays (e.g., dotplots, boxplots). Learn to use the key features of distributions for quantitative variables, such as: - center: mean as a balance point, and median as the middle-ordered value.	Use technology to subset and filter data sets and transform variables, including smoothing for time series data. Identify appropriate ways to summarize quantitative or categorical data using tables, graphical displays, and numerical summary statistics, which includes using standard deviation as a measure of

Appendix C (continued)

	Describe key features of distributions for quantitative variables, such as: - center: mean as the equal share, and median as the middle-ordered value of the data. - variability: range as the difference between the greatest and least value, and dispersion as how many units from the equal share value. - shape: number of clusters, symmetric or not, and gaps Recognize distributions can be used to compare two groups. Observe whether there appears to be an association between two variables	- variability: interquartile range and mean absolute deviation - shape: symmetric or asymmetric and number of modes. Use reasoning about distributions to compare two groups based on quantitative variables. Explore patterns of association between two quantitative variables or two categorical variables: - measures of correlation: quadrant count ratio (QCR). - comparison of conditional proportions across categorical variables.	variability and a modified boxplot for identifying outliers. Summarize and describe relationships among multiple variables. Understand how sampling distributions (developed through simulation) are used to describe the sample-to-sample variability of sample statistics. Develop simulations to determine approximate sampling distributions and compute p-values from those distributions. Describe associations between two categorical variables using measures such as difference in proportions and relative risk. Describe the relationship between two quantitative variables by

Appendix C (continued)

		-	interpreting Pearson’s correlation coefficient and a least-squares regression line. Use simulations to investigate associations between two categorical variables and to compare groups.

Process Component	**Level A**	**Level B**	**Level C**
IV. Interpret Results	Use statistical evidence from analyses to answer the statistical investigative questions and communicate results through structured answers with teacher guidance. Make statements about the group or population from which the data were collected, recognizing that conclusions are limited to these groups and cannot be generalized to other groups. Describe the difference between	Use statistical evidence from analyses to answer the statistical investigative questions and communicate results with comprehensive answers and some teacher guidance. Acknowledge that looking beyond the data is feasible. Generalize beyond the sample providing statistical evidence for the generalization and including a statement of uncertainty and plausibility when needed.	Use statistical evidence from analyses to answer the statistical investigative questions and communicate results through more formal reports and presentations. Evaluate and interpret the impact of outliers on the results. Understand what it means for an outcome or an estimate of a population characteristic to be plausible or not plausible compared to chance variation.

Appendix C (continued)

	two groups with different conditions	Recognize the uncertainty caused by sample to sample variability. State the limitations of sample information (e.g., a sample may or may not be representative of the larger population, measurement variability). Compare results for different conditions in an experiment	Interpret the margin of error associated with an estimate of a population characteristic. Acknowledge the presence of missing values and understand how missing values may add bias to an analysis. Use multivariate thinking to understand how variables impact one another. Communicate statistical reasoning and results to others in a variety of formats (verbal, written, visual). Understand how to interpret simulated *p*-values appropriately.

Appendix D:

Teacher’s Efficacy, Attitudes, and Anxiety about Statistics Survey

Purpose: This research is to examine how factors associated with statistics teaching efficacy and attitudes toward statistics impact anxiety about statistics of primary preservice mathematics teachers in Georgia.

Consent: Submission of this survey indicates your consent for participation. All responses will be kept strictly confidential, and only group-level results will be reported.

Directions: Select the number in each column that best reflects your opinion as accurately as possible.

A. Self-Efficacy to Teach Statistics Survey

Directions: Using a scale of (1, 2, 3, 4, 5, 6) where 1 = *not at all confident*, 2 = *only a little confident*, 3 = *somewhat confident*, 4 = *confident*, 5 = *very confident*, 6 = *completely confident*, please rate your confidence in teaching students the skills necessary to complete the following tasks successfully:

		Not at All Confident	Only a Little Confident	Somewhat Confident	Confident	Very Confident	Completely Confident
1	Collect data to answer a posed statistical question in contexts of interest to high school students.	1	2	3	4	5	6
2	Recognize that there will be natural variability between observations for individuals.	1	2	3	4	5	6

Appendix D (continued)

		Not at All Confident	Only a Little Confident	Somewhat Confident	Confident	Very Confident	Completely Confident
3	Select appropriate graphical displays and numerical summaries to compare individuals to each other and an individual to a group.	1	2	3	4	5	6
4	Create dotplot, stem and leaf plot, and tables (using counts) for summarizing distributions.	1	2	3	4	5	6
5	Use dotplot, stem and leaf plot, and tables (using counts) for describing distributions.	1	2	3	4	5	6
6	Create boxplots for summarizing distributions.	1	2	3	4	5	6
7	Use boxplots, median, and range for describing distributions.	1	2	3	4	5	6

Appendix D (continued)

		Not at All Confident	Only a Little Confident	Somewhat Confident	Confident	Very Confident	Completely Confident
8	Identify the association between two variables from scatterplots.	1	2	3	4	5	6
9	Generalize a statistical result from a small group to a larger group such as the whole class.	1	2	3	4	5	6
10	Recognize that statistical results may be different in another class or group.	1	2	3	4	5	6
11	Recognize the limitation of making inference (i.e. generalization) from a classroom dataset to any population beyond the classroom.	1	2	3	4	5	6

Appendix D (continued)

B. Attitudes Toward Statistics Survey								
Directions: Using a scale of (1, 2, 3, 4, 5, 6, 7) where 1 = *strongly disagree*, 2 = *disagree*, 3 = *somewhat disagree*, 4 = *neither disagree nor agree*, 5 = *somewhat agree*, 6 = *agree*, 7 = *strongly agree*, please rate your attitudes toward statistics as it relates to the statement:								
		Strongly Disagree	Disagree	Somewhat Disagree	Neither Disagree nor Agree	Somewhat Agree	Agree	Strongly Agree
12	I will like statistics.	1	2	3	4	5	6	7
13	I will feel insecure when I have to do statistical problems.	1	2	3	4	5	6	7
14	I will have trouble understanding statistics because of how I think.	1	2	3	4	5	6	7
15	Statistics formulas are easy to understand.	1	2	3	4	5	6	7
16	Statistics is worthless.	1	2	3	4	5	6	7
17	Statistics is a complicated subject.	1	2	3	4	5	6	7
18	Statistics should be required as part of my professional training.	1	2	3	4	5	6	7
19	Statistical skills make me more employable.	1	2	3	4	5	6	7

Appendix D (continued)

		Strongly Disagree	Disagree	Somewhat Disagree	Neither Disagree nor Agree	Somewhat Agree	Agree	Strongly Agree
20	I will have no idea of what's going on in statistics.	1	2	3	4	5	6	7
21	Statistics is not useful to the typical professional.	1	2	3	4	5	6	7
22	I will get frustrated going over statistics tests in class.	1	2	3	4	5	6	7
23	Statistical thinking is not applicable in my life outside my job.	1	2	3	4	5	6	7
24	I use statistics in my everyday life.	1	2	3	4	5	6	7
25	I will be under stress during statistics class.	1	2	3	4	5	6	7
26	I will enjoy taking statistics courses.	1	2	3	4	5	6	7

Appendix D (continued)

		Strongly Disagree	Disagree	Somewhat Disagree	Neither Disagree nor Agree	Somewhat Agree	Agree	Strongly Agree
27	Statistics conclusions are rarely presented in everyday life.	1	2	3	4	5	6	7
28	Statistics is a subject quickly learned by most people.	1	2	3	4	5	6	7
29	Learning statistics requires a great deal of discipline.	1	2	3	4	5	6	7
30	I will have no application for statistics in my profession.	1	2	3	4	5	6	7
31	I will make a lot of math errors in statistics.	1	2	3	4	5	6	7
32	I am scared of statistics.	1	2	3	4	5	6	7
33	Statistics involves massive computations.	1	2	3	4	5	6	7
34	I can learn statistics.	1	2	3	4	5	6	7
35	I will understand statistics equations.	1	2	3	4	5	6	7

Appendix D (continued)

		Strongly Disagree	Disagree	Somewhat Disagree	Neither Disagree nor Agree	Somewhat Agree	Agree	Strongly Agree
36	Statistics is irrelevant in my life.	1	2	3	4	5	6	7
37	Statistics is highly technical.	1	2	3	4	5	6	7
38	I will find it difficult to understand statistics concepts.	1	2	3	4	5	6	7
39	Most people have to learn a new way of thinking to do statistics.	1	2	3	4	5	6	7

Note. From "CS Consultants, LLC," by C. Schau (https://www.evaluationandstatistics.com/). Copyright 1995 by CS Consultants, LLC. Reprinted with permission.

C. Anxiety About Statistics Survey

Directions: Using a scale of (1, 2, 3, 4, 5, 6, 7) where 1 = *strongly disagree*, 2 = *disagree*, 3 = *neither disagree nor agree*, 4 = *agree*, 5 = *strongly agree*, please rate your anxiety about statistics as it relates to the statement:

		Strongly Disagree	Disagree	Neither Disagree nor Agree	Agree	Strongly Agree
40	It wouldn't bother me at all to take more statistics courses.	1	2	3	4	5
41	I have usually been at ease during statistics tests.	1	2	3	4	5
42	I have usually been at ease in statistics courses.	1	2	3	4	5

Appendix D (continued)

		Strongly Disagree	Disagree	Neither Disagree not Agree	Agree	Strongly Agree
43	I usually don’t worry about my ability to solve statistics problems.	1	2	3	4	5
44	I almost never get uptight while taking statistics tests.	1	2	3	4	5
45	I get really uptight during statistics tests.	1	2	3	4	5
46	I get a sinking feeling when I think of trying hard statistics problems.	1	2	3	4	5
47	My mind goes blank, and I am unable to think clearly when working statistics.	1	2	3	4	5
48	Statistics makes me feel uncomfortable and nervous.	1	2	3	4	5
49	Statistics makes me feel uneasy and confused.	1	2	3	4	5

D. Demographic Information

Directions: Check the box that reflects your answer.

50 What is your gender?

( ) Male
( ) Female

51 How many college-level mathematics courses have you completed? This includes Advanced Placement (AP), International Baccalaureate (IB), and dual-enrollment mathematics courses completed during high school?

( ) 0
( ) 1
( ) 2
( ) 3
( ) 4
( ) Other: ______

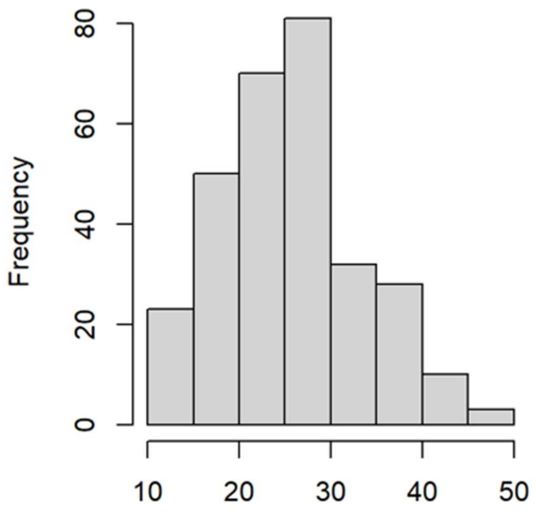
Appendix D (continued)

52	How many college-level statistics courses have you completed? This includes Advanced Placement (AP), International Baccalaureate (IB), and dual-enrollment mathematics courses completed during high school?
	( ) 0 ( ) 1 ( ) 2 ( ) 3 ( ) 4 ( ) Other: ______
53	How many years have you completed in your primary education program?
	( ) 0 ( ) 1 ( ) 2 ( ) 3 ( ) 4 ( ) Other: ______

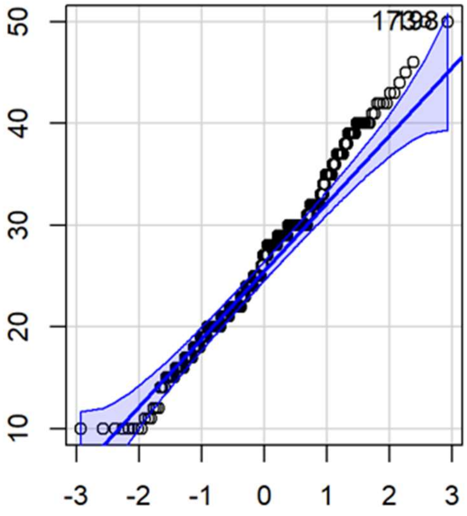
Appendix E:

Histograms and QQ-Plots of Variables Before and After Handling Outliers

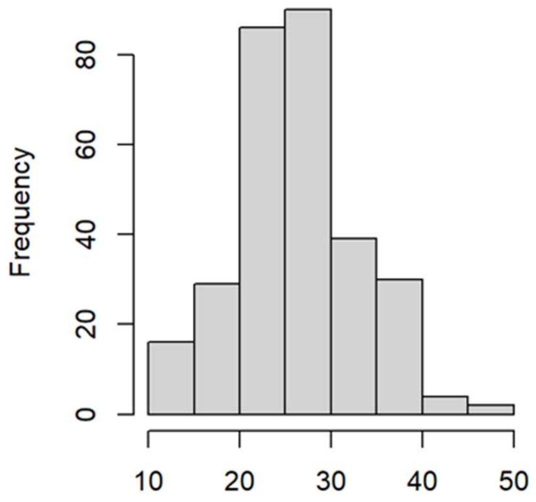
Histogram of Anxiety before Handling Outliers



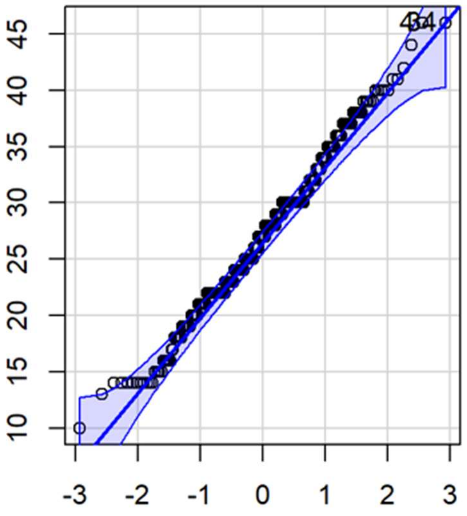
QQ-plot of Anxiety before Data Transformation



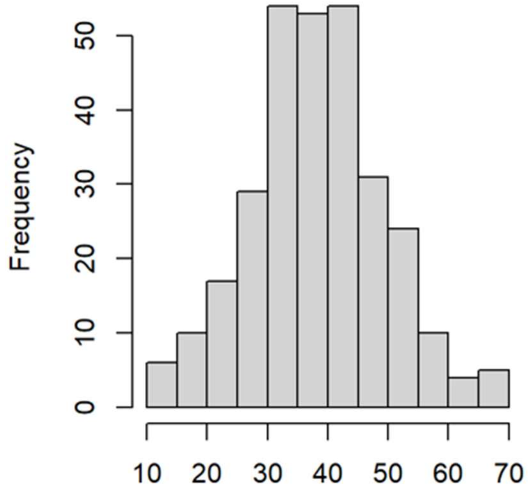
Histogram of Anxiety after Handling Outliers

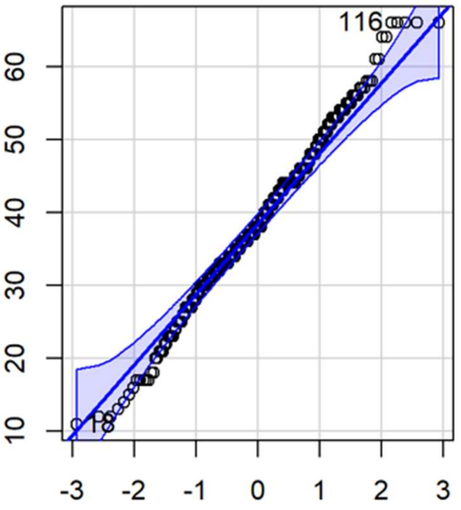


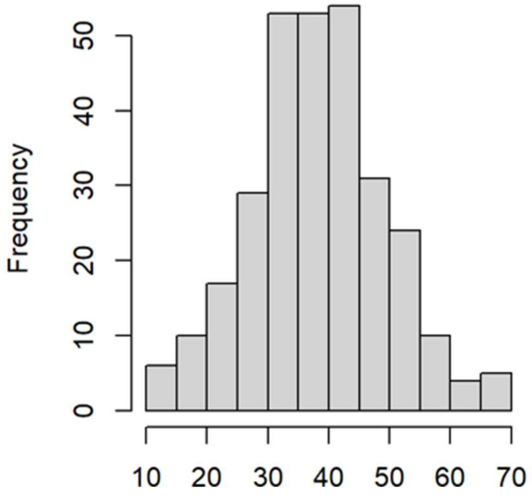
QQ-plot of Anxiety after Handling Outliers

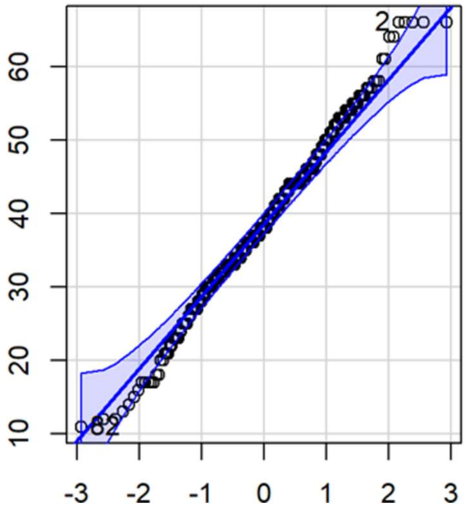


Appendix E (continued)

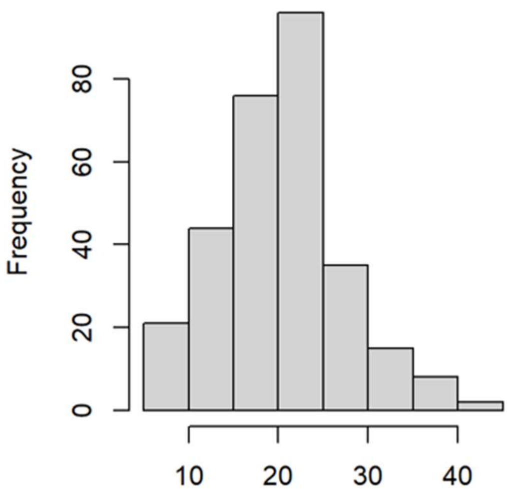

Histogram of Efficacy before Handling Outliers


QQ-plot of Efficacy before Handling Outliers

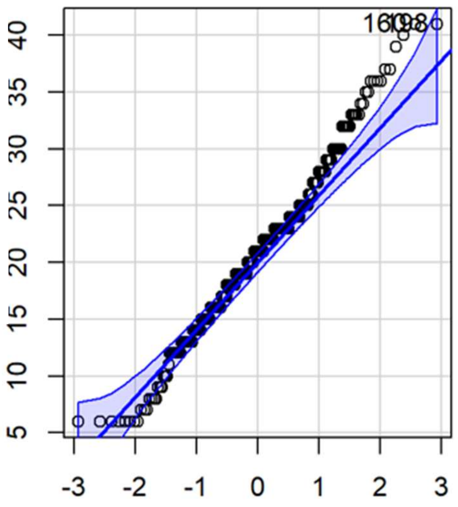

Histogram of Efficacy after Handling Outliers


QQ-plot of Efficacy after Handling Outliers

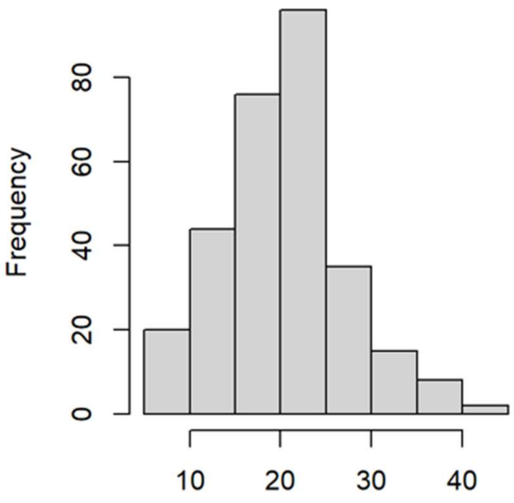
Appendix E (continued)



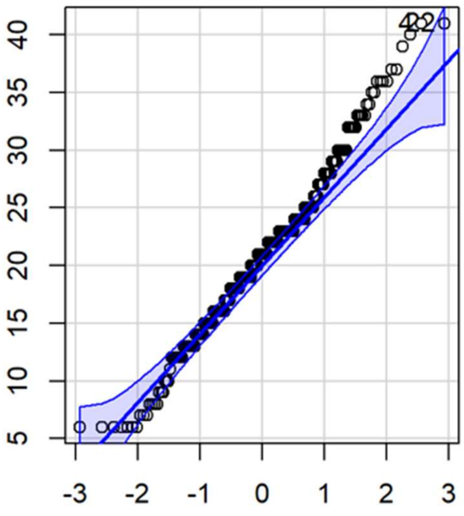
Histogram of Affect before Handling Outliers



QQ-plot of Affect before Handling Outliers



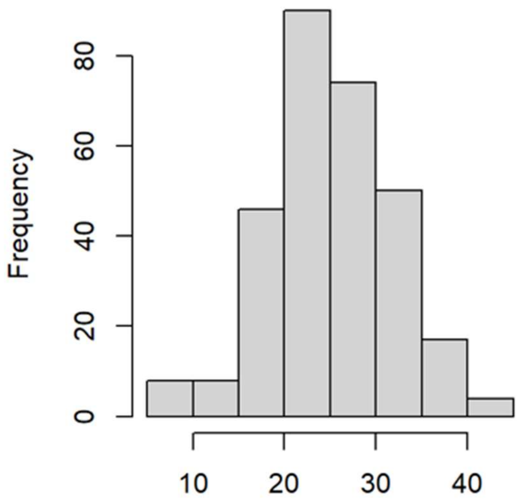
Histogram of Affect after Handling Outliers



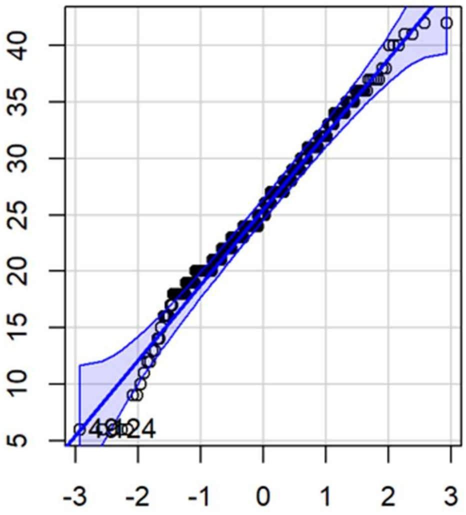
QQ-plot of Affect after Handling Outliers

Appendix E (continued)

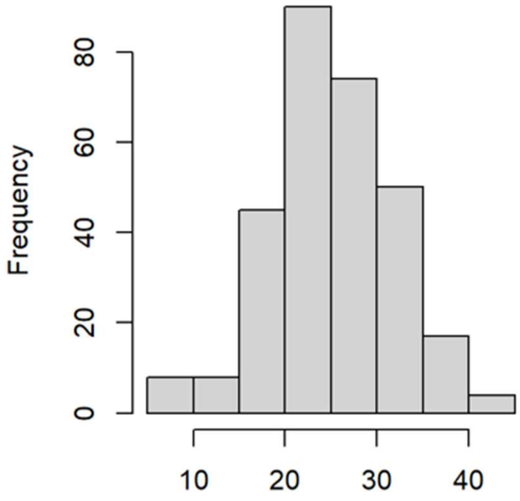
Histogram of Cognitive before Handling Outliers



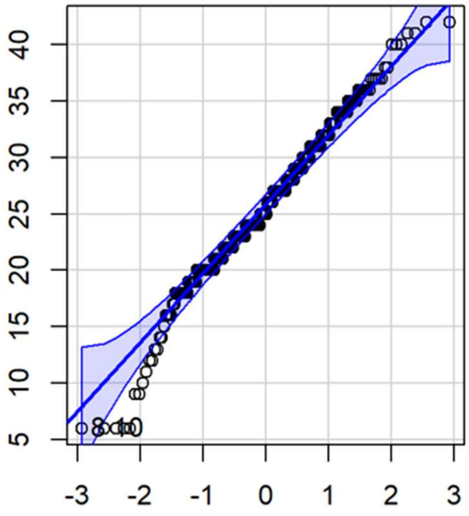
QQ-plot of Cognitive before Handling Outliers



Histogram of Cognitive after Handling Outliers

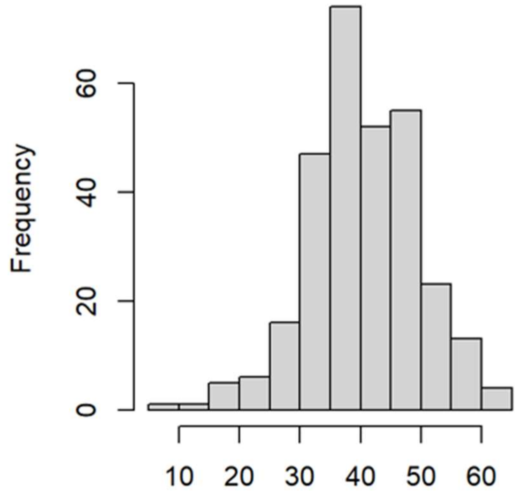


QQ-plot of Cognitive after Handling Outliers

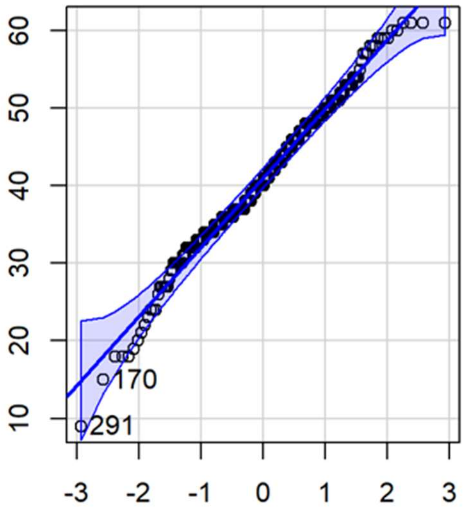


Appendix E (continued)

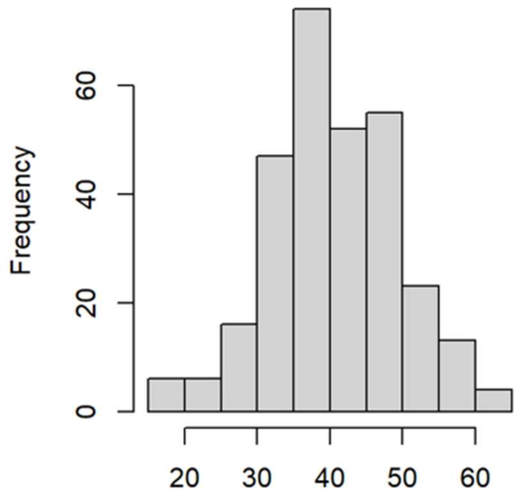
Histogram of Value before Handling Outliers



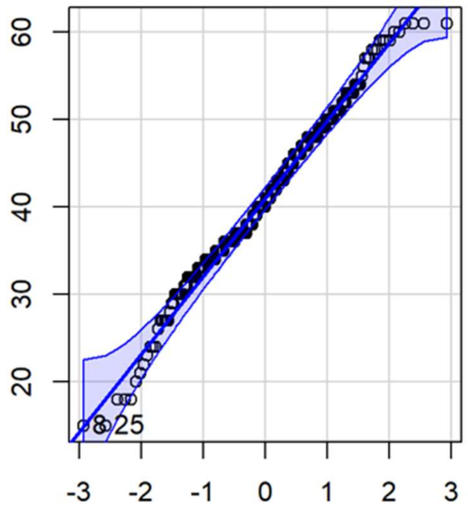
QQ-plot of Value before Handling Outliers



Histogram of Value after Handling Outliers

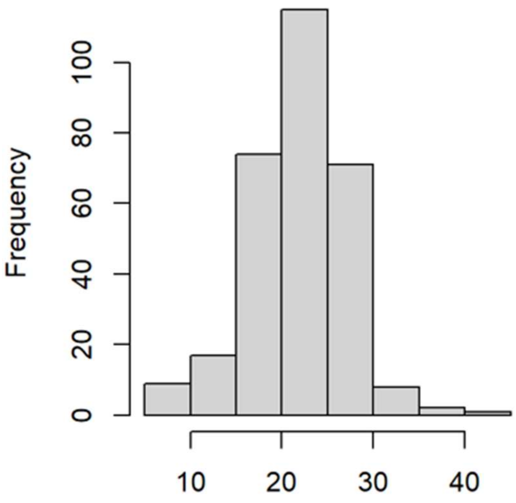


QQ-plot of Value after Handling Outliers

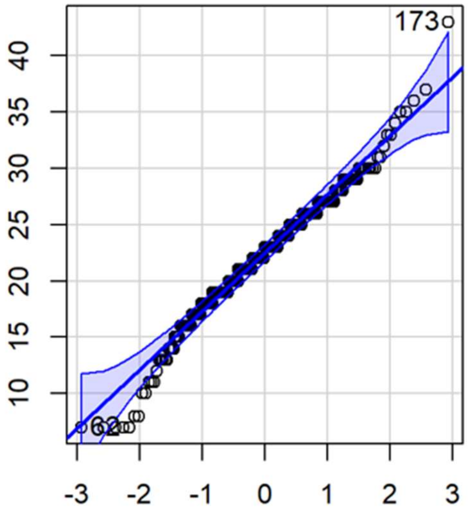


Appendix E (continued)

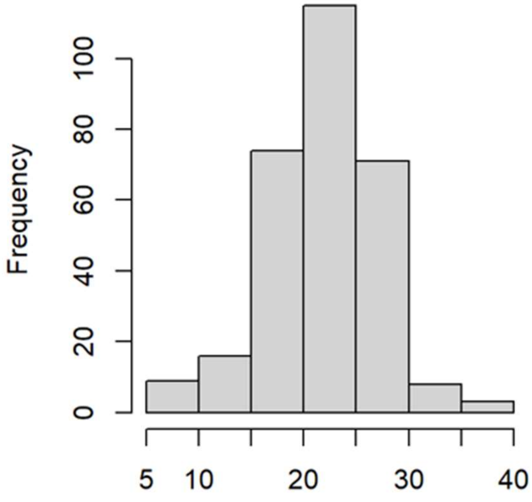
Histogram of Difficulty before Handling Outliers



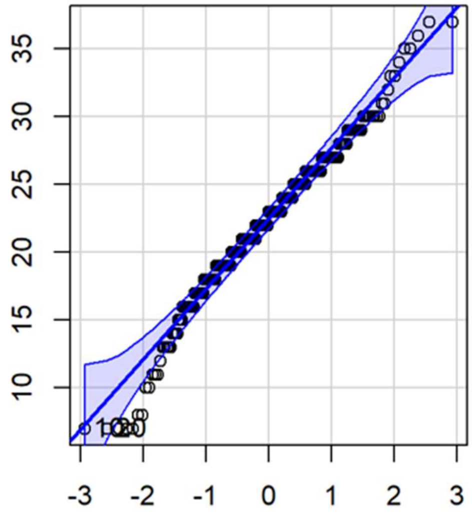
QQ-plot of Difficulty before Handling Outliers



Histogram of Difficulty after Handling Outliers



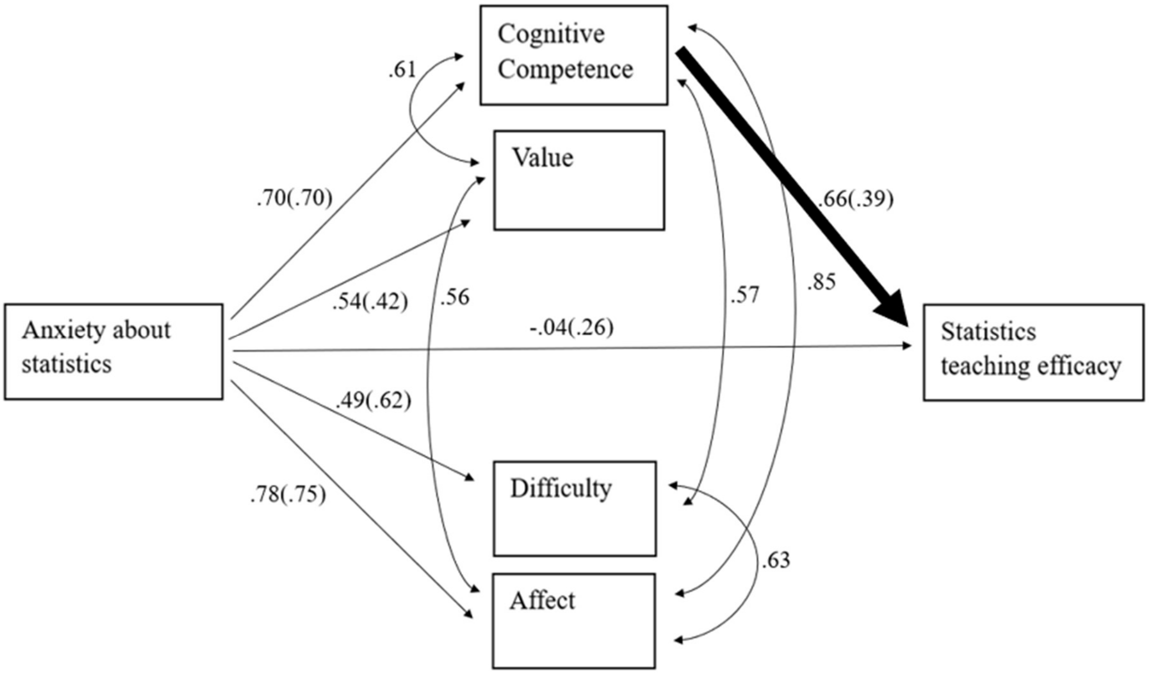
QQ-plot of Difficulty after Handling Outliers



Appendix F:

Cognitive Competence on Statistics Teaching Efficacy

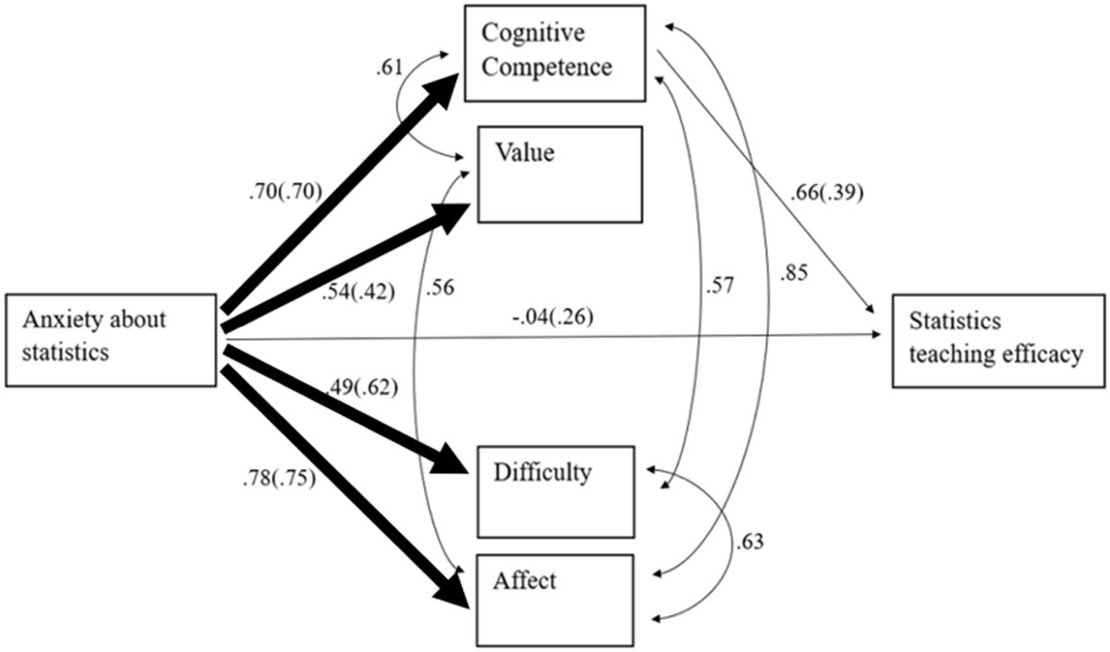
Cognitive competence had a significant positive path coefficient in its prediction on statistics teaching efficacy ($\beta = .66$, $Z = 4.47$, $p < .001$). The interpretation: Primary PSMTs' positive attitudes toward knowledge and intellectual skill in using statistical knowledge (cognitive competence) significantly predicted high confidence in teaching statistics. In the figure below, the dark highlighted directional path represents this finding.



Appendix G:

Anxiety on Attitudes Toward Statistics

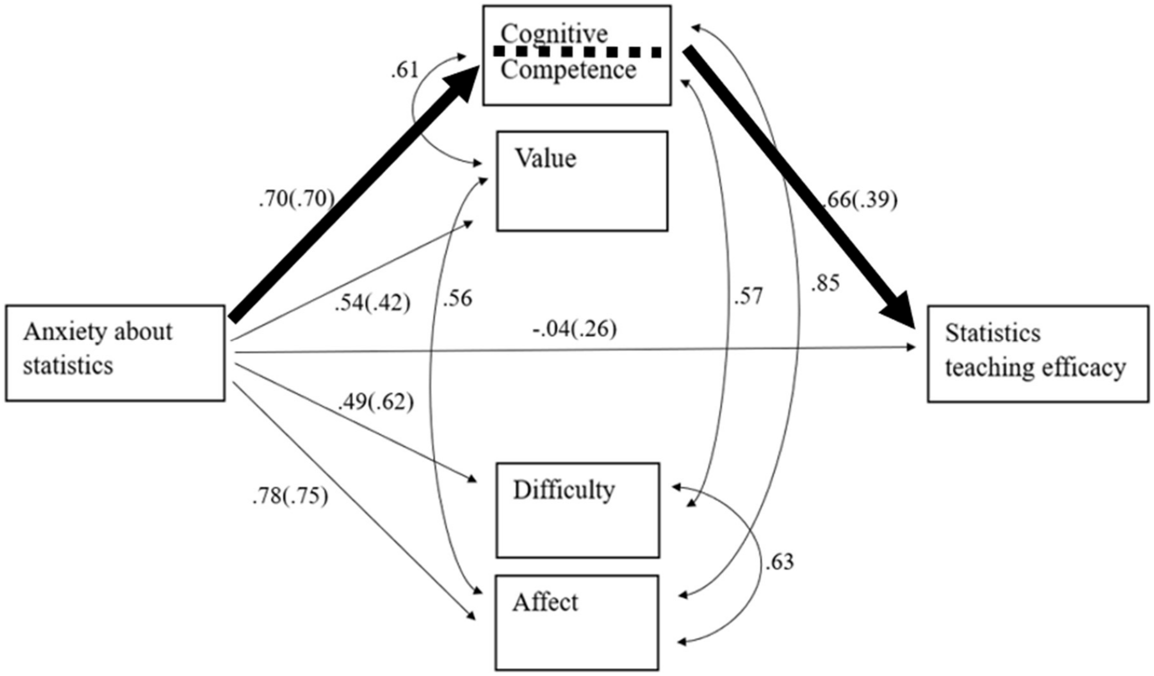
Anxiety about statistics had significant positive path coefficients in its' prediction of affect ($\beta = .78$, $Z = 17.14$, $p < .001$), cognitive competence ($\beta = .70$, $Z = 14.90$, $p < .001$), difficulty ($\beta = .49$, $Z = 11.48$, $p < .001$), and value ($\beta = .54$, $Z = 7.40$, $p < .001$). The interpretation: Primary PSMTs' who possessed low anxiety about statistics significantly predicted positive attitudes toward statistics through affect, cognitive competence, difficulty, and value. In the figure below, the dark highlighted directional paths represent this finding.



Appendix H:

Anxiety on Statistics Teaching Efficacy through Mediator

Anxiety about statistics was significantly predictive of statistics teaching efficacy through mediator cognitive competence ($\beta = .46$, $Z = 4.15$, $p < .001$). The interpretation: PSMTs who had low anxiety about statistics significantly predicted positive attitudes toward knowledge and intellectual skill in using statistical knowledge (cognitive competence), which, in turn, significantly predicted high confidence in teaching statistics. In the figure below, the dark highlighted directional path represent this finding.



Appendix I:

R Code Source File

```
#Libraries utilized
library(psych); library(dplyr); library(data.table);
library(tidyverse); library(naniar); library(DescTools); library(car);
library(MVN); library(lavaan)

#Import raw data
df <- read_excel("C:/Users/Surface Pro 4/Desktop/DataFrame.xlsx")

#Determine missing data percentage across the whole data frame
sum(is.na(df))/prod(dim(df))

#Determine missing data percentage for each column
apply(df, 2, function(col)sum(is.na(col))/length(col))

#Delete rows with 50% of more missing data.
#Rows represented individual responses to the survey.
keep <- rowMeans(is.na(df[,c(1:49)])) <= 0.5
keep
df<-df[keep,]

#For each subscale, delete row with 50% of more missing data
  #Anxiety subscale
  keep <- rowMeans(is.na(df[,c(40:49])) <= 0.5
  keep
  df<-df[keep,]

  #Efficacy subscale
  keep <- rowMeans(is.na(df[,c(1:11)) <= 0.5
  keep
  df<-df[keep,]

  #Affect subscale
  keep <- rowMeans(is.na(df[,c(40:49])) <= 0.5
  keep
  df<-df[keep,]

  #Cognitive subscale
  keep <- rowMeans(is.na(df[,c(14,20,31,34,35,38])) <= 0.5
  keep
  df<-df[keep,]

  #Difficulty subscale
  keep <- rowMeans(is.na(df[,c(15,17,28,29,33,37,39])) <= 0.5
  keep
  df<-df[keep,]

  #Value subscale
  keep <- rowMeans(is.na(df[,c(16,18,19,21,23,24,27,30,36])) <= 0.5
```

```
  keep
  df<-df[keep,]
  view(df)

#Use Little’s Missing Completely at Random Test to test if data is
missing completely at random
mcar_test(df)

#Replace missing data with median value of the corresponding column
dt <- data.table(df)
setkey(df, df[,c(1:53)])
dt[,value := ifelse(is.na(value), median(value, na.rm=TRUE), value)]
df<-dt
view(df)

#Descriptive statistics of items and demographics from the new data set
describe(df)

#Fequency and relative frequency count of items and demographic
variables
apply((df),2,table)

#Handle negatively worded items by reverse coding.
  #Anxiety
  #define columns to reverse code
  reverse_cols = c("item45","item46","item47","item48","item49")
  #reverse code columns
  df[ , reverse_cols] = 6 - df[ , reverse_cols]
  #view updated data frame
  view(df)

  #Affect
  #define columns to reverse code
  reverse_cols = c("item13","item22","item25","item32")
  #reverse code columns
  df[ , reverse_cols] = 8 - df[ , reverse_cols]
  #view updated data frame
  view(df)

  #Cognitive
  #define columns to reverse code
  reverse_cols = c("item14","item20","item31","item38")
  #reverse code columns
  df[ , reverse_cols] = 8 - df[ , reverse_cols]
  #view updated data frame
  view(df)
```

```
  #Difficulty
  #define columns to reverse code
  reverse_cols = c("item17","item29","item33","item37","item39")
  #reverse code columns
  df[ , reverse_cols] = 8 - df[ , reverse_cols]
  #view updated data frame
  view(df)

  #Value
  #define columns to reverse code
  reverse_cols =
  c("item16",”item21","item23","item27","item30","item36")
  #reverse code columns
  df[ , reverse_cols] = 8 - df[ , reverse_cols]
  #view updated data frame
  view(df)

#Calculate Cronbach's alpha for each scale
  #Anxiety
  Anxiety <- select(df, c("item40", "item41", "item42", "item43",
  "item44","item45","item46","item47","item48","item49"))
  psych :: alpha(Anxiety)

  #Efficacy
  Efficacy <- select(df, c(item1", "item2", "item3", "item4", "item5",
  "item6","item7","item8","item9","item10","item11"))
  psych ::alpha(Efficacy)

  #Cognitive
  Cognitive <- select(df, c("item14", "item20",
  "item31","item34","item35","item38"))
  psych :: alpha(Cognitive)

  #Difficulty
  Difficulty <- select(df, c("item15", "item17", "item28", "item29",
  "item33","item37","item39”))
  psych :: alpha(Difficulty)

  #Value
  Value <- select(df, c("item16", "item18", "item19", "item21",
  "item23","item24","item27","item30","item36”))
  psych :: alpha(Value)

  #Affect
  Affect <- select(df, c("item12", "item13", "item22", "item25",
  "item26","item32"))
  psych :: alpha(Affect)
```

```
#Create columns for the variables total score, which is the sum of
their items.
  #Anxiety
  cols_to_sum <- names(df) %in% c("item40", "item41",
  "item42","item43","item44","item45","item46","item47","item48","item4
  9")
  Anxiety <- rowSums(df[, cols_to_sum])

  #Efficacy
  cols_to_sum <- names(df) %in% c("item1", "item2",
  "item3","item4","item5","item6","item7","item8","item9","item10","ite
  m11")
  Efficacy <- rowSums(df[, cols_to_sum])

  #Affect
  cols_to_sum <- names(df) %in% c("item12", "item13",
  "item22","item25","item26","item32")
  Affect <- rowSums(df[, cols_to_sum])

  #Cognitive
  cols_to_sum <- names(df) %in% c("item14", "item20",
  "item31","item34","item35","item38")
  Cognitive <- rowSums(df[, cols_to_sum])

  #Difficulty
  cols_to_sum <- names(df) %in% c("item15", "item17",
  "item28","item29","item33","item37","item39”)
  Difficulty <- rowSums(df[, cols_to_sum])

  #Value
  cols_to_sum <- names(df) %in% c("item16", "item18",
  "item19","item21","item23","item24","item27","item30","item36”)
  Value <- rowSums(df[, cols_to_sum])

#Descriptive statistics of scales and demographics
psych::describe(df)

#Correlations: Pearson’s, polychoric, tetrachoric. View correlation
coefficient with p-values. P-values were calculate for polychoric and
tetrachoric using polychor()
psych:: mixedCor(df, c=1:6, d=7, p=8:10)
corrplot <- corr_coef(df)
plot(corrplot)
a <- polychor(df$Math, df$Program, std.err = TRUE)
print(a)
b <- polychor(df$Math, df$Stat, std.err = TRUE)
print(b)
c <- polychor(df$Math, df$Gender, std.err = TRUE)
print(c)
d <- polychor(df$Program, df$Stat, std.err = TRUE)
```

```
print(d)
e <- polychor(df$Program, df$Gender, std.err = TRUE)
print(e)
f <- polychor(df$Stat, df$Gender, std.err = TRUE)
print(f)

#Normality inspection using Jarque Bera Test from library DescTools
JarqueBeraTest(df$Anxiety, robust=TRUE,method=c("chisq","mc"), N=0,
na.rm=FALSE)
JarqueBeraTest(df$Efficacy, robust=TRUE,method=c("chisq","mc"), N=0,
na.rm=FALSE)
JarqueBeraTest(df$Cognitive, robust=TRUE,method=c("chisq","mc"), N=0,
na.rm=FALSE)
JarqueBeraTest(df$Difficulty, robust=TRUE,method=c("chisq","mc"), N=0,
na.rm=FALSE)
JarqueBeraTest(df$Value, robust=TRUE,method=c("chisq","mc"), N=0,
na.rm=FALSE)
JarqueBeraTest(df$Affect, robust=TRUE,method=c("chisq","mc"), N=0,
na.rm=FALSE

#View boxplots
boxplot(Anxiety)
boxplot(Efficacy)
boxplot(Cognitive)
boxplot(Difficulty)
boxplot(Value)
boxplot(Affect)

#View Histograms
hist(Anxiety)
hist(Efficacy)
hist(Cognitive)
hist(Difficulty)
hist(Value)
hist(Affect)

#View QQplots
qqPlot(Anxiety)
qqPlot(Efficacy)
qqPlot(Cognitive)
qqPlot(Difficulty)
qqPlot(Value)
qqPlot(Affect)

#Calculate z-scores for each scale
  #Anxiety
  z_anxiety <- scale(Anxiety, center=TRUE, scale = TRUE)
  z_anxiety
  summary(z_anxiety)
  psych :: describe(z_anxiety)
```

```
  #Efficacy
  z_efficacy <- scale(df$Efficacy, center=TRUE, scale = TRUE)
  z_efficacy
  summary(z_efficacy)
  psych :: describe(z_efficacy)

  #Cognitive
  z_cognitive <- scale(df$Cognitive, center=TRUE, scale = TRUE)
  z_cognitive
  summary(z_cognitive)
  psych :: describe(z_cognitive)

  #Difficulty
  z_difficulty <- scale(df$Difficulty, center=TRUE, scale = TRUE)
  z_difficulty
  summary(z_difficulty)
  psych :: describe(z_difficulty)

  #Value
  z_value <- scale(df$Value, center=TRUE, scale = TRUE)
  z_value
  summary(z_value)
  psych :: describe(z_value)

  #Affect
  z_affect <- scale(df$Affect, center=TRUE, scale = TRUE)
  z_affect
  summary(z_affect)
  psych :: describe(z_affect)

#Replace outlier 9 in the Value subscale with 15
conditions <- c("9")
replacement_values <- c("15")
df$Value <- replace(df$Value, df$Value %in% conditions,
replacement_values)
view(df)

#Replace outlier 43 in the Difficulty subscale with 37
conditions <- c("43")
replacement_values <- c("37")
df$Difficulty <- replace(df$Difficulty, df$Difficulty %in% conditions,
replacement_values)
view(df)

#Descriptive statistics of scales and demographics
psych::describe(df)
```

```
#Correlations: Pearson’s, polychoric, tetrachoric. View correlation
coefficient with p-values. P-values were calculate for polychoric and
tetrachoric using polychor()
psych:: mixedCor(df, c=1:6, d=7, p=8:10)
corrplot <- corr_coef(df)
plot(corrplot)
a <- polychor(df$Math, df$Program, std.err = TRUE)
print(a)
b <- polychor(df$Math, df$Stat, std.err = TRUE)
print(b)
c <- polychor(df$Math, df$Gender, std.err = TRUE)
print(c)
d <- polychor(df$Program, df$Stat, std.err = TRUE)
print(d)
e <- polychor(df$Program, df$Gender, std.err = TRUE)
print(e)
f <- polychor(df$Stat, df$Gender, std.err = TRUE)
print(f)

#Normality inspection using Jarque Bera Test from DescTools
JarqueBeraTest(Anxiety, robust=TRUE,method=c("chisq","mc"), N=0,
na.rm=FALSE)
JarqueBeraTest(Efficacy, robust=TRUE,method=c("chisq","mc"), N=0,
na.rm=FALSE)
JarqueBeraTest(Cognitive, robust=TRUE,method=c("chisq","mc"), N=0,
na.rm=FALSE)
JarqueBeraTest(Difficulty, robust=TRUE,method=c("chisq","mc"), N=0,
na.rm=FALSE)
JarqueBeraTest(Value, robust=TRUE,method=c("chisq","mc"), N=0,
na.rm=FALSE)
JarqueBeraTest(Affect, robust=TRUE,method=c("chisq","mc"), N=0,
na.rm=FALSE

#View boxplots
boxplot(Anxiety)
boxplot(Efficacy)
boxplot(Cognitive)
boxplot(Difficulty)
boxplot(Value)
boxplot(Affect)

#View Histograms
hist(Anxiety)
hist(Efficacy)
hist(Cognitive)
hist(Difficulty)
hist(Value)
hist(Affect)
```

```
#View QQplots
qqPlot(Anxiety)
qqPlot(Efficacy)
qqPlot(Cognitive)
qqPlot(Difficulty)
qqPlot(Value)
qqPlot(Affect)

#Mahalanobis’ Distance
#Delete the item values and demographic variables. This will leave the
subscale total scores.
data.frame = subset(df,select=-c(1:53))

#Mahalanobis Distance
maha <- mahalanobis(data.frame[,c(1,2,3,4,5,6)],
                                colMeans(data.frame[,c(1,2,3,4,5,6)]),
                                cov(data.frame[,c(1,2,3,4,5,6)]))

#Include MD in data frame
data.frame$maha<-round(maha, 3)
head(data.frame)
#Identify rows with CV>12.59
data.frame[data.frame$maha>12.59,]
#Remove the MD column
df_maha=subset(data.frame, select = -c(7))
head(df_maha)

#Mardias test
mvn(df_maha, mvnTest = "mardia")

#####################
####Path Analysis####
#####################

#Model 1
  #Hypothetical model

  model1 <-'
  Affect ~ Efficacy
  Cognitive ~ Efficacy
  Value ~ Efficacy
  Difficulty ~ Efficacy
  Anixety ~ Affect + Cognitive + Difficulty + Value + Efficacy

  Efficacy ~~ Efficacy
  Cognitive ~~ Cognitive
  Value ~~ Value
  Difficulty ~~ Difficulty
  Affect ~~ Affect
  Anxiety ~~ Anxiety
```

```
  '

  fit1<-lavaan(model1, data=dataframe, estimator = "MLM")
  summary(fit1, fit.measures = TRUE, standardized = TRUE, rsquare =
  TRUE)
  parameterEstimates(fit1)
  modindices(fit1, minimum.value = 10, sort = TRUE)

#Model 2
  #Reverse direction of model 1; guided by modindices()

  Model2 <-'
  Affect ~ Anxiety
  Cognitive ~ Anxiety
  Value ~ Anxiety
  Difficulty ~ Anxiety
  Efficacy ~ Affect + Cognitive + Difficulty + Value + Anxiety

  Efficacy ~~ Efficacy
  Cognitive ~~ Cognitive
  Value ~~ Value
  Difficulty ~~ Difficulty
  Affect ~~ Affect
  Anxiety ~~ Anxiety
  '

  fit2<-lavaan(model2, data=dataframe, estimator = "MLM")
  summary(fit2, fit.measures = TRUE, standardized = TRUE, rsquare =
  TRUE)
  parameterEstimates(fit1)
  modindices(fit2, minimum.value = 10, sort = TRUE)

#Model 3
  #Added covariances among the four subscales of attitudes toward
  statistics
  #This model fit had df=0 implying just-identified model

  model3 <-'
  Affect ~ Anxiety
  Cognitive ~ Anxiety
  Value ~ Anxiety
  Difficulty ~ Anxiety
  Efficacy ~ Affect + Cognitive + Difficulty + Value + Anxiety

  Affect ~~ Difficulty
  Affect ~~ Value
  Affect ~~ Cognitive
  Difficulty ~~ Value
  Difficulty ~~ Cognitive
  Value ~~ Cognitive
```

```
  Efficacy ~~ Efficacy
  Cognitive ~~ Cognitive
  Value ~~ Value
  Difficulty ~~ Difficulty
  Affect ~~ Affect
  Anxiety ~~ Anxiety
  '

  fit3<-lavaan(model3, data=dataframe, estimator = "MLM")
  summary(fit3, fit.measures = TRUE, standardized = TRUE, rsquare =
  TRUE)
  parameterEstimates(fit3)
  modindices(fit3, minimum.value = 10, sort = TRUE)

#Model 4
  #Covariance between Value and Difficulty was not significant, it was
  removed.
  #Difficulty, value, anxiety, and affect are not significant
  predictors of efficacy. Theses paths were removed
  #The model fit the data well

  model4 <-'
  Affect ~ Anxiety
  Cognitive ~ Anxiety
  Value ~ Anxiety
  Difficulty ~ Anxiety
  Efficacy ~  Cognitive + Anxiety

  Affect ~~ Difficulty
  Affect ~~ Value
  Cognitive ~~ Affect
  Cognitive ~~ Value
  Cognitive ~~ Difficulty

  Efficacy ~~ Efficacy
  Cognitive ~~ Cognitive
  Value ~~ Value
  Difficulty ~~ Difficulty
  Affect ~~ Affect
  Anxiety ~~ Anxiety
  '

  fit4<-lavaan(model4, data=dataframe, estimator = "MLM")
  summary(fit4, fit.measures = TRUE, standardized = TRUE, rsquare =
  TRUE)
  parameterEstimates(fit4)
  modindices(fit4, minimum.value = 10, sort = TRUE)

  #bootstraps = 5000 was ran on Model 4
```

```
model4b <-'
Affect ~ Anxiety
Cognitive ~ Anxiety
Value ~ Anxiety
Difficulty ~ Anxiety
Efficacy ~  Cognitive + Anxiety

Affect ~~ Difficulty
Affect ~~ Value
Cognitive ~~ Affect
Cognitive ~~ Value
Cognitive ~~ Difficulty

Efficacy ~~ Efficacy
Cognitive ~~ Cognitive
Value ~~ Value
Difficulty ~~ Difficulty
Affect ~~ Affect
Anxiety ~~ Anxiety
'

fit4b<-lavaan(model4b, data=dataframe, estimator = "MLM", test =
"bootstrap", bootstrap = 5000)
summary(fit4b, fit.measures = TRUE, standardized = TRUE, rsquare =
TRUE)

#Direct, Indirect and Total effects
model_effects <-'
Affect ~ Anxiety
Cognitive ~ a*Anxiety
Value ~ Anxiety
Difficulty ~ Anxiety
Efficacy ~ b*Cognitive + c*Anxiety

Affect ~~ Difficulty
Affect ~~ Value
Cognitive ~~ Affect
Cognitive ~~ Value
Cognitive ~~ Difficulty

Efficacy ~~ Efficacy
Cognitive ~~ Cognitive
Value ~~ Value
Difficulty ~~ Difficulty
Affect ~~ Affect
Anxiety ~~ Anxiety

#indirect effects
IE1:=a*b
```

```
  #total indirect effects
  TIE := IE1 + c
  '

  fit_effects<-lavaan(model_effects, data=dataframe, estimator = "MLM")
  summary(fit_effects, fit.measures = TRUE, standardized = TRUE,
  rsquare = TRUE)
  parameterEstimates(fit_effects)

#############################
####Multigroup Invariance####
#############################

dataframe <- read_excel("C:/Users/Surface Pro
4/Desktop/DataFrame_Adjusted.xlsx")

# adjusted data frame includes demographics
# Gender: Group 1 = male; Group 2 = female
# Number of Math courses Completed: Group 1 = 1 to 3 math course, Group
2 = 4 to 6 math courses
# Number of Statistics courses Completed: Group 1 = 1 to 2 stat course,
Group 2 = 3 to 6 stat courses
# Program years completed: Group 1 = 1, Group 2 = 2, Group 3 = 3, Group
4 = 4 = 4 and 5

#Multigroup Invariance on Gender
#Upsampling on gender to create equal size groups
#new data set contains 270 male and 270 female
df_upsample<-upsample(dataframe,cat_col = "Gender")
view(df_upsample)

model_effects <-'
Affect ~ Anxiety
Cognitive ~ Anxiety
Value ~ Anxiety
Difficulty ~ Anxiety
Efficacy ~  Cognitive + Anxiety

Cognitive ~~ Affect
Affect ~~ Difficulty
Affect ~~ Value
Cognitive ~~ Value
Cognitive ~~ Difficulty

Efficacy ~~ Efficacy
Cognitive ~~ Cognitive
Value ~~ Value
Difficulty ~~ Difficulty
Affect ~~ Affect
Anxiety ~~ Anxiety
```

```
'
measurementInvariance(model = model_effects, data=df_upsample, group=
"Gender", strict=TRUE)

#configural invariance
fit1<-lavaan(model_effects, data=df_upsample, estimator = "MLM", group
= "Gender")
#weak invaraince
fit2<-lavaan(model_effects, data=df_upsample, estimator = "MLM", group
= "Gender", group.equal = "loadings")
#strong invariance
fit3<-lavaan(model_effects, data=df_upsample, estimator = "MLM",  group
= "Gender", group.equal = c("intercepts","loadings"))
#strict invariance
fit4<-lavaan(model_effects, data=df_upsample, estimator = "MLM", group
= "Gender", group.equal = c("residuals", "intercepts","loadings"))

#model comparison test
lavTestLRT(fit1, fit2, fit3, fit4)

#Fit measures
fitMeasures(fit1, c("chisq", "df", "pvalue",
"cfi","tli","rmsea","srmr"))
fitMeasures(fit2, c("chisq", "df", "pvalue",
"cfi","tli","rmsea","srmr"))
fitMeasures(fit3, c("chisq", "df", "pvalue",
"cfi","tli","rmsea","srmr"))
fitMeasures(fit4, c("chisq", "df", "pvalue",
"cfi","tli","rmsea","srmr"))

#Try to establish partial multigroup invariance as there is a lack of
strong invariance (scalar model)

#lavtestScore() was used to view which fixed parameters in the model
should be released to improve the fit for the scalar model
lavTestScore(fit3)

#parTable() is used to print out the list of all parameters in the
model
parTable(fit3)

#By reviewing the p-value column, I identified the parameters that are
expected to have a significant impact on model fit (p<.05).
#lavTestScore() in addition with parTable() suggested to release the
constraints for the following parameters between the goups to establish
partial MI
#Efficacy ~ 1, Cognitive ~ 1, Value ~ 1, and Anxiety ~ 1
#Note: Difficulty ~ 1 did not work
```

```
fit3a<-lavaan(model_effects, data=df_upsample, estimator = "MLM",
group = "Gender", group.equal = c("intercepts","loadings"),
group.partial = c("Efficacy ~ 1", "Cognitive ~ 1", "Value ~ 1",
"Anxiety ~ 1"))

lavTestLRT(fit1, fit2, fit3a)

#Partial MI on strict invaraince
lavTestLRT(fit3a, fit4)

#Did not work. Proceed with partial MI as before
lavTestScore(fit4)
parTable(fit4)

#Results suggested to release the constraints for the following
parameters between the groups to see if it establishes partial MI
#Efficacy ~ 1, Cogntiive ~1, Difficulty ~ 1, and  Cognitive~~Cognitive
#Results still not work
#I decided to release the constraints for the following parameters
between the groups to establish partial MI. This is the same for strong
invariance. The results were good.
#Efficacy ~ 1, Cogntiive ~ 1, Value ~ 1, and Anxiety ~ 1
#Strict invariance

fit4a<-lavaan(model_effects, data=df_upsample, estimator = "MLM", group
= "Gender", group.equal = c("residuals", "intercepts","loadings"),
group.partial = c("Efficacy ~ 1", "Cognitive ~ 1", "Value ~ 1",
"Anxiety ~ 1"))

lavTestLRT(fit3a, fit4a)

#Multigroup invariance on number of math courses complete.
#Strict invariance existed. Strict invariance indicates the two groups
did not differ on the construct of the model. The groups can be
compared.

model_effects <-'
Affect ~ Anxiety
Cognitive ~ Anxiety
Value ~ Anxiety
Difficulty ~ Anxiety
Efficacy ~  Cognitive + Anxiety

Cognitive ~~ Affect
Affect ~~ Difficulty
Affect ~~ Value
Cognitive ~~ Value
Cognitive ~~ Difficulty

Efficacy ~~ Efficacy
```

```
Cognitive ~~ Cognitive
Value ~~ Value
Difficulty ~~ Difficulty
Affect ~~ Affect
Anxiety ~~ Anxiety
'
measurementInvariance(model = model_effects, data=dataframe, group=
"Math", strict=TRUE)

#configural invariance
fit1<-lavaan(model_effects, data=dataframe, group = "Math")
#weak invaraince
fit2<-lavaan(model_effects, data=dataframe, group = "Math", group.equal
= "loadings")
#strong invariance
fit3<-lavaan(model_effects, data=dataframe, group = "Math", group.equal
= c("intercepts","loadings"))
#strict invariance
fit4<-lavaan(model_effects, data=dataframe, group = "Math", group.equal
= c("residuals", "intercepts","loadings"))

#model comparison test
lavTestLRT(fit1, fit2, fit3, fit4)

#Fit Measures
fitMeasures(fit1, c("chisq", "df", "pvalue",
"cfi","tli","rmsea","srmr"))
fitMeasures(fit2, c("chisq", "df", "pvalue",
"cfi","tli","rmsea","srmr"))
fitMeasures(fit3, c("chisq", "df", "pvalue",
"cfi","tli","rmsea","srmr"))
fitMeasures(fit4, c("chisq", "df", "pvalue",
"cfi","tli","rmsea","srmr"))

#Multigroup invariance on number statistics courses completed
#Strict invariance existed. Strict invariance indicates the two groups
did not differ on the construct of the model.

model_effects <-'
Affect ~ Anxiety
Cognitive ~ Anxiety
Value ~ Anxiety
Difficulty ~ Anxiety
Efficacy ~  Cognitive + Anxiety

Cognitive ~~ Affect
Affect ~~ Difficulty
Affect ~~ Value
Cognitive ~~ Value
Cognitive ~~ Difficulty
```

```
Efficacy ~~ Efficacy
Cognitive ~~ Cognitive
Value ~~ Value
Difficulty ~~ Difficulty
Affect ~~ Affect
Anxiety ~~ Anxiety
'
measurementInvariance(model = model_effects, data=dataframe, group=
"Stat",

#configural invariance
fit1<-lavaan(model_effects, data=dataframe, group = "Stat")
#weak invaraince
fit2<-lavaan(model_effects, data=dataframe, group = "Stat", group.equal
= "loadings")
#strong invariance
fit3<-lavaan(model_effects, data=dataframe, group = "Stat", group.equal
= c("intercepts","loadings"))
#strict invariance
fit4<-lavaan(model_effects, data=dataframe, group = "Stat", group.equal
= c("residuals", "intercepts","loadings"))

#model comparison test
lavTestLRT(fit1, fit2, fit3, fit4)

#Fit measures
fitMeasures(fit1, c("chisq", "df", "pvalue",
"cfi","tli","rmsea","srmr"))
fitMeasures(fit2, c("chisq", "df", "pvalue",
"cfi","tli","rmsea","srmr"))
fitMeasures(fit3, c("chisq", "df", "pvalue",
"cfi","tli","rmsea","srmr"))
fitMeasures(fit4, c("chisq", "df", "pvalue",
"cfi","tli","rmsea","srmr"))

#Multigroup invariance on number of program years completed.

#Strict invariance existed. Strict invariance indicates the four groups
did not differ on the construct of the model.
```

```
model_effects <-'
Affect ~ Anxiety
Cognitive ~ Anxiety
Value ~ Anxiety
Difficulty ~ Anxiety
Efficacy ~  Cognitive + Anxiety

Cognitive ~~ Affect
Affect ~~ Difficulty
Affect ~~ Value
Cognitive ~~ Value
Cognitive ~~ Difficulty

Efficacy ~~ Efficacy
Cognitive ~~ Cognitive
Value ~~ Value
Difficulty ~~ Difficulty
Affect ~~ Affect
Anxiety ~~ Anxiety
'
measurementInvariance(model = model_effects, data=dataframe, group=
"Program", strict=TRUE)

#configural invariance
fit1<-lavaan(model_effects, data=dataframe, group = "Program")
#weak invaraince
fit2<-lavaan(model_effects, data=dataframe, group = "Program",
group.equal = "loadings")
#strong invariance
fit3<-lavaan(model_effects, data=dataframe, group = "Program",
group.equal = c("intercepts","loadings"))
#strict invariance
fit4<-lavaan(model_effects, data=dataframe, group = "Program",
group.equal = c("residuals", "intercepts","loadings"))

#model comparison test
lavTestLRT(fit1, fit2, fit3, fit4)

#Fit Measures
fitMeasures(fit1, c("chisq", "df", "pvalue",
"cfi","tli","rmsea","srmr"))
fitMeasures(fit2, c("chisq", "df", "pvalue",
"cfi","tli","rmsea","srmr"))
fitMeasures(fit3, c("chisq", "df", "pvalue",
"cfi","tli","rmsea","srmr"))
fitMeasures(fit4, c("chisq", "df", "pvalue",
"cfi","tli","rmsea","srmr"))
```