ABSTRACT

AZMY, CHRISTINA. Secondary Mathematics Preservice Teacher Development of Statistics Teaching Self-Efficacy. (Under the direction of Dr. Hollylynne Lee).

The purpose of this study was to examine the statistics teaching self-efficacy (TSE) beliefs of secondary mathematics preservice teachers (PSTs) and the impact of experiences on those beliefs. In order to research experiences and beliefs, I designed a two-phase qualitative collective case study within a situated context with thirty-four participants from two institutions in Phase I, and a subset of seven of those participants in Phase II. Phase I involved participants' submitted work in a two-week online module expertly designed for preparing mathematics PSTs to teach statistics. Submitted work included a survey about confidence to teach statistics, records of discussion forums, reflection assignments, written statistical investigations, and screencast statistical investigations. Phase II involved an online autobiographical survey documenting participants' past experiences with learning and teaching statistics, and transcripts of a recorded semi-structured interview to gain insight into how those experiences impacted their statistics TSE (STSE). All data was coded, guided by a theoretical framework, and descriptions and themes were developed. Main results from the study are communicated in two manuscripts.

The first manuscript's focus is the beliefs secondary mathematics PSTs hold and express about their self-efficacy for teaching statistics. Because of the unique nature of statistics, it is important to characterize both personal teaching efficacy beliefs and general statistics education beliefs when describing STSE beliefs. Both types of beliefs expressed within the online module and during interviews are described, with implications for teacher education. Specifically, the types of general statistics education beliefs, in addition to the types of personal efficacy beliefs, that participants hold that were impacted by participants' engagement in the module and that are

a result of prior experiences that were impacted by participants' engagement in the module are described.

This particular collective case engaged in a two-week online module, and so the situated context provided a set of similar experiences that could be examined, which is the focus of the second manuscript. Opportunities the module provided for impacting mathematics PSTs' STSE are described. In addition, evidence that those experiences had an impact, from submitted work from the module and post-use interviews, is presented.

Results from the study indicate that the examined two-week intervention, aimed at impacting secondary mathematics PSTs' general statistics education beliefs, was effective to a certain extent. In addition, although participants displayed an overall incomplete personal belief in teaching statistics that has been reported in past research, a closer look reveals specific types of personal confidence. Results on the effectiveness of the online materials indicate success in developing certain factors that we know to be important in statistics teaching self-efficacy (STSE) – PSTs' view of statistics, statistical knowledge, pedagogical knowledge, and use of technology.

© Copyright 2019 by Christina Azmy

All Rights Reserved

Secondary Mathematics Preservice Teacher Development of Statistics Teaching Self-Efficacy

by
Christina Azmy

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Mathematics Education

Raleigh, North Carolina
2019

APPROVED BY:

Dr. Hollylynne Lee Committee Chair	Dr. Gemma Mojica
Dr. Herle McGowan	Dr. Kevin Oliver

DEDICATION

This dissertation is dedicated to the two pillars in my life, my parents, Dr. Yousry Azmy and Dr. Inas Messiha, who have always pointed me to the Strong Tower, to whom I can run and be safe.

BIOGRAPHY

Christina Azmy was born in Knoxville, TN on December 28, 1986, to two immigrant parents, Dr. Yousry Azmy and Dr. Inas Messiha, who moved from Egypt to America to seek educational and career opportunities only this wonderful country could provide. Fifteen months later, her brother and best friend, Joseph Azmy, was born. In 2002, the whole family moved to State College, Pennsylvania during Christina's sophomore year of high school, and she stayed there to earn her bachelor's degree in Mathematics at the Pennsylvania State University, as part of the Schreyer Honors College.

After graduation and an amazing year teaching preschool at a small, Christian school, she was accepted into Duke Law School in Durham, North Carolina in 2010. She decided after one year of law school that it was not the path for her, and she enrolled in NC State University's Master of Arts in Teaching graduate program. She then taught at Clayton High School for 2.5 years, where she loved her students very much and saw their joy in excelling at the challenging subject of statistics. During this time, Joseph married her other best friend, Dr. Christine Azmy, joining two beautiful families.

Seeing her high school students graduate and excited for their futures inspired Christina to return to school, and so she joined the doctoral program at NC State University's College of Education in the STEM Education department, where she was fortunate enough to be mentored by Dr. Hollylynne Lee. Christina will graduate in summer of 2019, and hopes to help prepare the next generation of mathematics teachers, especially when it comes to their self-confidence to teach statistics!

ACKNOWLEDGMENTS

I would like to first thank my advisor, Dr. Hollylynne Lee, for her mentorship throughout my studies. She has challenged me and taught me in so many ways. I have grown so much from her leadership, and never knew I was capable of all the things she encouraged me to do. This dissertation was especially strengthened by her commitment to providing detailed feedback and ideas for revisions.

I would also like to thank my other mentor, Dr. Gemma Mojica, for her support during the last three years. Personally and academically, she helped me in more ways than she will ever know. She was my biggest cheerleader, even in her own rough times. I thank her for always pointing out what I was doing well, and pushing me to do better… and for coming up with a spreadsheet solution to any problem!

I would like to thank my committee members Dr. Herle McGowan and Dr. Kevin Oliver. Their feedback throughout the process, from initial ideas on research, to ways of improving my proposed dissertation study and providing suggested readings. Their willingness to help and their flexibility in meeting was always very much appreciated.

I would like to thank my fellow-minions, Heather Barker and Taylor Harrison, who kept me sane through their laughter and care. Also, Gregory Downing's fortuitous timing of writing his dissertation was heaven-sent. We both ended up writing our dissertations at the same time, and going through the writing process together was the motivation I really needed to get it done. I am grateful for his constant encouragement. I hope the four us continue to be HQties for life!

Finally, and most importantly, I would like to thank my family and friends who are like family for their prayers and love, without which I would have given up a million times.

The work in this dissertation was partially supported by the Enhancing Statistics Teacher Education with E-Modules [ESTEEM] project, funded by the National Science Foundation with a grant to NC State University (under Grant No. DUE 1625713). The opinions, findings, and conclusions or recommendations in this dissertation are my own, and do not necessarily reflect the views of the National Science Foundation.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1: Introduction

Recent changes in mathematics curriculum in the United States reflect the increasing importance of data analysis and statistical thinking in today's society. Standards for mathematics include an emphasis on data and probability (National Governors Association Center for Best Practice & Council of Chief State School Officers, 2010; National Council of Teachers of Mathematics (NCTM), 2000). In addition, the American Statistical Association has released two documents with a framework for assessment and instruction in statistics education—at the preK-12 level (Franklin et al., 2007) and the undergraduate level (American Statistical Association, 2005; American Statistical Association, 2016). These frameworks and standards emphasize the importance of student understanding of statistical concepts, like distribution, variation, and inferential reasoning and an active technology-supported approach to learning and teaching statistics.

With the addition of these statistics standards and professional documents, and the emphasis placed on conceptual understanding of statistics topics, there is a stronger need for teacher preparation in statistics education. In fact, there is evidence that many preservice mathematics teachers feel underprepared to teach statistics topics (Lovett & Lee, 2017; 2018; Harrell-Williams, Sorto, Pierce, Lesser, & Murphy, 2014). This may be partially due to the lack of confidence many teachers report in their own understanding of the statistics concepts they are expected to teach (Lovett & Lee, 2017; Begg & Edwards, 1999; Hannigan, Gill, & Leavy, 2013). This lack of confidence is indeed important as it relates to teaching self-efficacy (TSE). TSE is important because it has been found to be related to student outcomes (Tschannen-Moran & Hoy, 2001).

In an effort to end the cycle of teachers and students who are uncomfortable with statistics, it is especially important for teacher educators in teacher preparation programs and for teachers of statistics to be aware of the implications of the experiences they provide their students. For teacher educators, it is important to provide opportunities for preservice teachers (PSTs) to reflect on their past and current experiences with statistics, in terms of their content knowledge, pedagogical content knowledge, use of technological tools, and view of statistics compared to mathematics, as these all contribute to their perceived preparedness to teach statistics (Lovett & Lee, 2017). In addition, new experiences should be provided for preservice and inservice teachers in order to impact current statistics TSE (STSE) levels. The additional fact that education in general, and particularly teacher education, is increasingly being offered online means that preparing teachers to teach statistics in online settings is also a consideration. Online contexts provide unique affordances, and also constraints, that should be considered in providing teachers experiences that are beneficial to their STSE.

For teachers of statistics, it is important to be aware of their impact on future teachers in regards to the experiences they provide their students. If future teachers are to gain a sense from their statistics teachers of what statistics *is* and how it is *effectively* taught, it is vital that statistics teachers have appropriate ideas about statistics and statistics education, and enact those ideas in their classrooms. Because of this, it is important for teacher educators, preservice teachers, inservice teachers, and statistics teachers to be aware of the importance of STSE, in addition to the factors that contribute to STSE.

Significance of the Study and Research Questions

Because statistics is commonly included within mathematics departments and curricula, there is a tendency to treat statistics similarly to mathematics with regards to teaching and learning, even though the two topics have important inherent differences (e.g., Rossman, Chance, & Medina, 2006). For this reason, quantitative methods that have traditionally been used to measure beliefs and attitudes towards statistics may actually be confounding perceptions of statistics and of mathematics. For example, high levels of STSE or positive attitudes towards statistics combined with a lack of conceptual understanding of statistics may actually reflect the undesirable perspective that statistics and mathematics are the same, and should be taught and learned in the same way (Hannigan et al., 2013; Lovett & Lee, 2017).

The development of statistics teaching self-efficacy (TSE) of preservice secondary mathematics teachers is an area that has yet to be fully-explored. I am interested in the different pathways that mathematics preservice teachers (PSTs) may experience in developing their own STSE in general, and more specifically within the context of a specific intervention aimed at better preparing preservice teachers to teach statistics. The intervention consists of a two-part online module that includes opportunities for reading about teaching statistics, watching expert educators discuss issues in teaching statistics, engaging in statistical investigations, watching video cases of classrooms, being introduced to a guiding framework on teaching statistics, and discussing/reflecting with classmates on statistical teaching and learning experiences.

The module incorporates best practices of effective teacher education, like the focus on student thinking through things like the use of video cases of real students and the focus on content-based investigative tasks through things like written and screencast statistical investigations of real data sets (Hiebert, Morris, & Glass, 2003). In addition, the module incorporates best practices of effective online teacher education, like opportunities to view different types and models of instructional practice, to increase pedagogical content knowledge, to provide resources and example tasks for curriculum, and to offer spaces for reflection and dialogue (Burns, 2011). However, the online module is just one set of experiences within a larger scheme of a life's worth of experiences, and so to investigate the larger picture of PSTs' statistics teaching self-efficacy, the two research questions guiding this study are:

1) "What are the general statistics education beliefs and personal teaching efficacy beliefs that comprise preservice mathematics teachers' statistics teaching self-efficacy, and what experiences and factors most impact development of those beliefs?" and
2) "How do statistics pedagogy learning experiences impact preservice mathematics teachers' statistics teaching self-efficacy?"

I hypothesized that mathematics PSTs who engage with resources in the online module will identify those as impactful, potentially in a positive or negative way, on their TSE development. While the second research question is focused on the experiences within the online module, for the first question, I am also interested in prior and later experiences that mathematics PSTs identify as impactful on their TSE development.

Overview of Methodological Approach

In order to answer my research questions, I used a two-phase qualitative collective case study to understand experiences mathematics PSTs identify as impactful on their STSE. For the first phase, data in the form of assignments and discussion forums from the online intervention were used in order to identify efficacy beliefs expressed by participants, and experiences they described in relation to those beliefs. The second phase asked a subset of the first set of participants, chosen on voluntary basis, to complete an autobiographical survey and a an interview that delved deeply into past experiences with statistics in addition to experiences in the online module, and how those experiences impacted current STSE. Data analysis was done by several rounds of coding each data source for types of experiences that were important in STSE development and also for expressed STSE beliefs. Coding was both data-driven and based on a framework developed from a compilation of existing TSE frameworks, discussed below. After coding was complete, codes were reorganized and consolidated into themes.

Important Definitions

Teaching self-efficacy (TSE) – the perceived ability to enact behaviors that are effective for student learning

Secondary mathematics preservice teacher (PST) – an undergraduate or graduate student who is in the preparation phase to become a licensed mathematics teacher for middle and high school students

Hybrid – for the purposes of this research, a hybrid course is one that includes both synchronous and asynchronous online components and/or face-to-face meetings

Former experiences – any experience a secondary mathematics preservice teacher may have engaged in prior to a particular moment in time, even if it is immediately prior to that moment in time

Personal Teaching Efficacy (PTE) beliefs – beliefs about one's self and one's own abilities to teach

General Statistics Education (GSE) beliefs – beliefs about statistics education, including beliefs about the nature of statistics and how statistics should be taught

Chapter 2: Literature Review and Conceptual Framework

In this section, I lay the theoretical and empirical foundation for my proposed research. First, I give an overview of statistics education. Next, I synthesize general practices and goals of mathematics teacher education generally, followed by sections on statistics teacher education and online teacher education specifically. I then shift to describing teaching self-efficacy (TSE) generally, followed by research on statistics teaching self-efficacy. Then, I describe the framework used to inform data collection and analysis to describe how mathematics PSTs develop their STSE, which is based on a compilation of general TSE frameworks proposed by various researchers.

Statistics Education Overview

Beginning in the 1990's, a shift began to occur in statistics education towards a reformed curriculum and pedagogy. The reforms included allowing students to participate in the full investigative cycle, giving students real data in context to investigate, a focus on concepts, rather than computations, and the relative frequency approach to probability, rather than a theoretical model. The reform methods explicitly align with constructivist ideas, allowing for students to develop higher cognitive thought through problem-solving (Moore, 1997).

The reform movement in statistics education insists on more student *doing*, rather than teacher *telling*, so that students are able to construct their own knowledge, rather than passively receive information (Moore, 1997). Among the important statistics concepts about which students should develop a deep understanding is that of distributions (Wild, 2006). Research done on students and teachers actively engaging in investigations with real data has found that students are better able to view distribution as an aggregate (Konold, Higgins, Russell, & Khalil, 2015), and even to compare distributions by attending to variation (Makar & Confrey, 2004) in those contexts. Reform methods of teaching have shown that students can build on their case value perspectives (attending to individual points) of a distribution to work up to an aggregate perspective (viewing the distribution as a whole) (Konold et al., 2015). In addition, using real data that students have collected themselves allows them to construct knowledge about critically inspecting and cleaning data. For example, collecting data makes it easier to notice outliers that occur because of measurement error (Connor, Davies, & Holmes, 2006).

Importantly, both the NCTM *Principles and Standards for School Mathematics* (2000) and Common Core State Standards in Mathematics (2010) added statistics standards to the younger years, although only the NCTM standards have included them as early as elementary school. The inclusion of statistics standards have reflected the more general reform movement in teaching and learning. For example, the Common Core standards include eight mathematical practices that emphasize the importance of actively constructing knowledge of concepts, rather than memorizing formulas or procedures (e.g. "Model with mathematics," and "Look for and express regularity in repeated reasoning").

In their foundational paper on statistical thinking, Wild and Pfannkuch (1999) set up a framework for statistical thinking using inquiry which includes four dimensions, the investigative cycle, types of thinking, the interrogative cycle, and dispositions. Building on this work, two influential documents that have recently informed statistics education are the *Guidelines for Assessment and Instruction in Statistics Education (GAISE)* reports, one written for preK-12 statistics education (Franklin et al., 2007), and the other written for undergraduate statistics (American Statistical Association, 2005; American Statistical Association, 2016). The *GAISE* report for preK-12 develops a framework based on the NCTM standards, in which students are actively engaged in developing knowledge of statistics concepts from an early age.

The framework combines the investigative cycle and the different levels of student understanding during each phase of the cycle (Franklin et al., 2007, p. 14–15). For example, for the first phase, Formulate Question, the framework describes student thinking at Levels A, B, and C, with Level C being the most sophisticated type of statistical thinking. The levels are not associated with an age or grade level, but rather can be used to assess students' current level of thinking.

In a similar document, the ASA has recently released a report for the education of statistics *teachers* (Franklin et al., 2015). The report's recommendations for teachers at all grade levels is linked to the *GAISE* framework in that the statistics investigative cycle is emphasized, as well as big principles like analyzing real data and using statistical software to complete investigations. In summary, recommendations for both student and teacher statistics education have been informed by reform principles.

With the increasing importance of statistics in the world and in education, it is necessary for all citizens to be statistically literate (Kwasny, 2015). Thus, it is even more important for mathematics teachers to be well-prepared to teach statistics, from middle school to the university level. In public schools, statistical concepts are expected to be taught in mathematics classrooms. Common Core Statistics standards begin in Grade 6 and go all the way until Grade 12 (National Governors Association Center for Best Practice & Council of Chief State School Officers, 2010). In addition, public schools often offer an Advanced Placement Statistics course that can earn high school students college credit for an introductory statistics course. Some of the challenges for both teachers and students are a lack of adequate teacher preparation, a lack of student time to delve into statistics, and a lack of outreach from statisticians to students (Kwasny, 2015).

Goals and Practices of Mathematics Teacher Education

The unique and complex skill of teaching is not an inherent or natural gift; rather, teachers learn skills and ideas through different learning opportunities in different settings. However, for *effective* teaching, there are principles and research-based practices that teachers must learn. PSTs must learn a variety of skills and knowledge bases in order to be prepared for teaching in their specific contexts. For example, they should have an appropriate amount of content knowledge in order to teach content, but they also should build specialized knowledge required for teaching, called pedagogical content knowledge (PCK) (Shulman, 1986). Pedagogical content knowledge, or mathematical knowledge for teaching, includes knowledge of content and students, knowledge of content and teaching, and knowledge of curriculum (Ball, Thames, & Phelps, 2008). PSTs build this knowledge in a variety of settings; specifically in their undergraduate or graduate preparation, this most often occurs during coursework experiences that include both content courses and methods for teaching courses and during fieldwork experiences that include real classroom observations and teaching assignments.

Those responsible for the complex task of preparing future teachers to teach mathematics are called mathematics teacher educators (MTEs). However, because contexts, programs, requirements, and audiences differ so greatly, it may be difficult for MTEs to access research on best practices (Cochran-Smith & Villegas, 2015). Goals of MTEs differ depending on context, especially because there is no common curriculum or set of resources upon which MTEs can draw. Some researchers (Hiebert, Morris, & Glass, 2003) argue it is not feasible for mathematics PSTs to learn how to be an effective teacher without yet being a teacher. Hiebert et al. propose a model for math teacher preparation which focuses on two goals of preparing teachers to teach: (1) become proficient in mathematical pedagogical content knowledge, and (2) develop knowledge, competencies, and dispositions to learn to teach from teaching. These goals could be met by treating lessons as experiments, which means that PSTs should learn to pose research questions about their lessons, collect data from their lessons, and analyze and interpret that data in order to improve future practice. So, it is important for MTEs to include opportunities for mathematics PSTs to shift their focus from teachers' performance to student thinking and to reconsider mathematical ideas in new ways (Hiebert, Morris, Berk, & Jansen, 2007). This can be done using video cases of real classrooms, by creation or engagement in mathematical tasks, and reflection on mathematical pedagogical opportunities of those experiences (Hiebert et al., 2007).

There is evidence to say that PSTs who are exposed to videos of student-teacher interactions during coursework and asked to critically reflect on those videos include more student input in their own teaching and attend more to student thinking when reflecting on their own teaching, as opposed to focusing on themselves (Santagata & Yeh, 2014). PSTs who do not have those video case experiences during coursework tend to conduct more teacher-centered teaching, as well as focus on themselves in their teaching analysis (Santagata & Yeh, 2014). For these PSTs, only having fieldwork experience was not enough to create the desired shift in teaching practice and analysis. Overall, video cases of real classrooms in teacher preparation is increasing in use in all contexts and are potentially powerful for increasing levels of reform education, but effectiveness depends on the facilitation of analysis (Gaudin & Chaliès, 2015). In order to successfully implement the viewing of video cases, mathematics teacher educators (MTE's) must be conscious of implementation decisions-what videos they choose, how they decide to orchestrate discussion about the video, and how they react in the moment to types of comments made by teachers, requiring a "heightened listening" (Coles, 2014).

In addition to the positive impact that critical viewing, reflection, and discussion of video cases of real classrooms can have, PSTs' engagement in or creation of mathematical tasks can also have a positive impact on PCK learning.

> ...unlike tasks for students, a mathematical task for teachers rarely deals with just the mathematics. It can be seen as an opportunity to generalize from it to a large class of tasks, and to deal with many other aspects of teaching mathematics as well. (Zaslavsky, 2008, p. 112)

For preparation of mathematics teachers to teach statistics, MTEs have the unique challenge of especially perturbing existing ideas about statistics and lack of sufficient content knowledge. Thus, it is even more vital for MTEs to incorporate practices like those above for statistical pedagogical content knowledge.

Statistics Teacher Education

In order to inform different groups of educators, the American Statistical Association has released several documents for assessment and instruction in statistics education—one for the preK-12 level (Franklin et al., 2007), one for the undergraduate level (American Statistical Association, 2005; 2016), and one for teacher preparation programs (Franklin et al., 2015). These frameworks emphasize the importance of student understanding of statistical concepts, such as distribution, variation, and informal inference. An extension of the *GAISE* framework, the Framework for Supporting Students' Approaches to Statistical Investigations (SASI), identifies specific statistical concepts to which students should attend at different phases of the investigative cycle (Lee & Tran, 2015).

A major part of statistical thinking involves thinking how a statistician would (Wild & Pfannkuch, 1999). One instructional approach teachers often choose to support students' statistical thinking is by introducing students to statistical concepts in the context of engaging in a statistical investigation. Tasks that require students to find solutions to problems by going through the entire investigative cycle accustom them to the approach statisticians use to solve “messy” data analysis problems (Wild & Pfannkuch, 1999).

A statistics course is *not* equivalent to a mathematics course; statistics can be viewed as a science that heavily uses mathematics (Rossman et al., 2006). The fact that statistics is taught within mathematics classrooms can be problematic, since statistics and mathematics are very different subjects (Scheaffer, 2006; Rossman, Chance, & Medina, 2006). Mathematics is deterministic, based on proof, and involves deductive reasoning. On the other hand, statistics is probabilistic, based on data, and involves inductive reasoning (Scheaffer, 2006; Rossman et al., 2006). Furthermore, statistics cannot be stripped of context the way mathematics can be (Rossman et al., 2006). This difference may not be widely recognized by mathematics teachers. For example, in one particular study, one group of preservice mathematics teachers had positive attitudes towards statistics that did not correlate with a positive conceptual understanding of statistics (Hannigan, Gill, & Leavy, 2013). The teachers in this study seemed to be conflating mathematics and statistics, which led to the inconsistency in attitudes and knowledge. The beginning of knowledge for effective statistics teachers is the knowledge that statistics is inherently different from mathematics (Hannigan et al., 2013). To successfully teach statistics, educators should be aware of how the approach to teaching, modeling, and applying statistical concepts is different from mathematics. Thus, statistics instructors should use and create tasks to present concepts of statistics by methods other than “number crunching,” which is the mathematical approach to statistics.

Because of these issues, it is important for teacher educators to take on the unique task of preparing mathematics teachers to teach statistics (Franklin et al., 2015). Two successful programs implemented in Portugal connected teacher education in statistics with practice, one with pre-service teachers and one with in-service teachers, by providing participants with opportunities of (1) collaborating with fellow teachers to create effective tasks, (2) implementing those tasks, and (3) reflecting on how they went (Ponte, 2011). Similarly, a workshop focused on the topic of variation increased and deepened teacher knowledge of the types of sources of variation in data for many of the teachers (Arnold, 2008). The reasons teachers gave behind the effectiveness of the professional development were that they actually did the activities and that they did them within a learning community.

In addition to the characteristics of effective professional development for statistics teachers in the examples above, much research has been done on the positive effects of using technology to teach statistics to teachers and students (e.g. Hammerman & Rubin, 2004; Lee et al., 2014; Prodromou & Pratt, 2006). Educational statistical tools are a specific type of software that allow students to actively engage in statistical content by providing students the ability to analyze data, visualize data, problem-solve, and focus on concepts, as opposed to tools for doing statistics which focus on performing procedures and calculations (McNamara, 2015; Chance, Ben-Zvi, Garfield, Medina, 2007). For example, Hammerman and Rubin (2004) conducted a case study and found teachers and students were able to use strategies unique to technology to reason about distributions using binning and proportional reasoning options afforded to them by educational statistical tools. Similarly, teachers using two widely-known educational statistical tools, *TinkerPlots* and *Fathom*, were able to solve data analysis tasks using strategies uniquely-related to technology, like linking representations, and augmenting data by adding dividers or reference lines (Lee et al., 2014). Therefore, the evidence points to the fact that one way to increase preservice secondary math teachers' knowledge of statistics is by using educational statistical tools to teach them statistics concepts. Unfortunately, preservice high school teachers have reported a lack of exposure to using such software in learning or teaching statistics (Harrell-Williams, et al., 2014).

Related to the use of technology in teaching, technological pedagogical statistical knowledge (TPSK) is important in teacher education (Lee & Hollebrands, 2011). The TPSK framework describes the different components of knowledge that a statistics teacher should have. The base knowledge teachers should possess is statistical knowledge; situated within that statistical knowledge, they should also know how to use technology to perform statistics-related tasks, like analyzing data. The inner-most level of knowledge points to the very specific knowledge a teacher should have that combines the statistical content, the use of technology, and knowledge of effective implementation in a classroom, like strategies for implementing technology to increase student understanding. If PSTs possess a strong combination of the four categories that contribute to teaching statistics self-efficacy, they will have strong TPSK. The importance of TPSK in teacher preparation indicates that teachers should be engaging in statistics, using technology, and considering student thinking about statistics.

By combining those ideas of effective statistics teacher education, the relationship between conceptual learning, reform teaching methods, and teacher knowledge is a strong one. Wilson and Cooney (2002) found that:

> Students learn statistics more effectively in settings where collaboration is encouraged, and where progressive teaching methods such as discovery learning and problem solving are the focus. To successfully implement such strategies,

> teachers must possess the necessary attitude structures as well as good knowledge of mathematical and didactic aspects of school learning of statistics. (as cited in Estrada et al., 2011, p. 12)

Thus, because reform methods of teaching, attitude, content knowledge, and TPSK are so inextricably linked, I propose that preservice secondary mathematics teachers will increase self-efficacy to teach statistics by learning statistics topics, along with pedagogical content, using technology. I hypothesize that engaging opportunities that allow for increasing knowledge of statistics, and specifically TPSK, will lead to an improvement in teaching statistics self-efficacy.

Online Mathematics Teacher Education

Another aspect to consider in the preparation of mathematics and statistics teacher is the context in which they are learning. In fall of 2015, almost 30% of all higher education students were taking at least one online course (Allen & Seaman, 2017). Online learning has increased in prevalence in many different educational settings, and some of the attraction to online learning can be attributed to the accessibility it provides students to learning opportunities they may not have had otherwise, cost-efficiency, and unique opportunities for collaboration (Burns, 2011; Means, Toyama, Murphy, & Baki, 2013). Unlike previous waves of distance education, current capabilities of technology tools that can be used in online learning allow for rich student experiences, like live synchronous online sessions, interactive discussion forums, real-time collaborative word processing documents, and dynamic, free, online content-based software. As online education has become more common, teacher preparation programs are increasingly incorporating online aspects . Thus, recently, researchers have studied the effectiveness of online teacher preparation.

Online learning can be classified on a continuum of fully-online approaches to blended approaches that incorporate both online and face-to-face components. Researchers tend to use "hybrid" or "blended" learning interchangeably to refer to online learning that also incorporates face-to-face experiences, although some specify a proportion of content to be delivered online in order to satisfy the label (Burns, 2011). Research done on online or blended learning incorporate different theoretical frameworks due to the unique nature of learning online. In fact, "much of the literature on how to implement online or blended learning (e.g., Bersin, 2004; Martyn, 2003) is based either on interpretations drawn from theories of learning or on common practice rather than on empirical evidence" (Means, et al., 2013, p. 37). In the already complex skill of teaching, which involves so many moments of decision-making on the part of the instructor, the added complexity of making instructional decisions that are effective in an online space is an area yet to be fully explored.

One theoretical model that has been used in online learning is the Community of Inquiry model, which posits that teacher and social presence in an online space impact cognitive presence, or learning outcomes (Shea & Bidjerano, 2010). Some studies have found that learners also have a role in impacting learning outcomes, especially when it comes to characteristics like student self-efficacy (e.g. Noesgaard, & Ørngreen, 2015; Shea & Bidjerano, 2010). A fourth presence, learner presence, is sometimes included in the Community of Inquiry model. Thus, even in online environments, self-efficacy beliefs are an important construct that impact learning.

When comparing online learning to face-to-face in terms of effectiveness for student outcomes, meta-analyses have found online learning at least to be *as* effective, and in the case of blended instruction, *more* effective than solely face-to-face-instruction (e.g. Means et al., 2013; Noesgaard & Ørngreen, 2015). However, researchers are cautious to attribute this positive impact only due to the fact that instruction is conducted online; rather, they interpret results to be a result of many factors related to online learning, like learning time, types of learning activities, and collaboration opportunities. Thus, researchers emphasize that design of instruction and learning opportunities offered to students are vital for effectiveness.

Mirroring the general increase in online instruction being offered in educational settings, teacher education is also increasingly being offered online. Specifically in teacher education, most recent web-based models of online learning are attractive for providing continuous learning to a wider audience, opportunities to view different types of instructional practice, solutions to teachers' content knowledge needs, access to curriculum supports, spaces for reflection and dialogue, and models for good instructional practice (Burns, 2011). Thus, research that informs best practices is needed in order to aid in designing quality online instruction in general, and more specifically, in teacher education.

In one design-based research study, the authors implemented and revised online mathematics and science teacher preparation courses at their university over a long span of time based on data from the courses and interviews in order to increase the quality of the social, teaching, and cognitive presences of the courses (Niess & Gillow-Wiles, 2013). Their research resulted in a learning trajectory for teachers' knowledge for teaching with technology which finds that in an online asynchronous environment, a supportive community structure, purposeful instructor actions, technology, and shared content knowledge are essential. Some best practices identified include purposefully scaffolded content, collaborative work, inquiry-based activities, and opportunities for reflection. Thus, while best practices for teaching online may resemble those for teaching face-to-face, it is important for instructors to design learning opportunities that incorporate those practices differently, given constraints of an online space. Constraints of an online course are unique and provide unique challenges for MTEs in terms of conveying "listening" to students, building rapport with participants, and being constrained by the limitations of the technology tool (Kastberg, Lynch-Davis, & D'Ambrosio, 2014).

However, despite constraints of online learning, using the internet to teach also affords unique opportunities in terms of collaboration, multimodality, and performance (Borba, Clarkson, & Gadanidis, 2013). Synchronous collaboration allows for teachers and students to work on mathematical tasks together, interacting with each other and with mathematical software. In one study in online synchronous secondary mathematics teacher education, trends in interactions included small group discussion through the use of breakout rooms, being followed by whole group discussion, electronic responses from participants in the form of emojis or checkmarks, pauses for individual work or because of a break in discussion, choral responses in the chat, and the use of whiteboards, among others (Starling & Lee, 2015). Robust conferencing software now allows for these types of interactions to take place online.

Asynchronous collaboration requires that *all* students participate in order to be counted as present (e.g., completing a quiz after a video, submitting assignments, posting in a discussion board), which is not the case in face-to-face settings (Borba & Llinares, 2012). In addition, asynchronous collaboration encourages a less teacher-centered approach since students will likely be responding to each other more often, or more quickly, than an instructor would. Specifically in the case of asynchronous discussion forums, these unique spaces allow mathematics PSTs to thoughtfully read, comprehend, interpret, plan a response, and communicate a response to peers, which is not always possible in face-to-face synchronous situations (Borba & Llinares, 2012).

Related to unique ways of collaboration is the opportunities for multimodality; online mathematics PSTs have access to multiple ways of engaging with content, through text, audio, video, images, graphs, applets, and interactive PDFs (Borba et al., 2013). Interacting with content in such a variety of ways allows mathematics PSTs to make connections that might not have been possible without such online learning opportunities. Related to collaboration and multimodality is the opportunity for online spaces to be spaces for performance, in which mathematics PSTs can develop ways of expressing ideas in innovative, creative, and artistic ways (Borba et al., 2013). Allowing mathematics PSTs to "perform" by recording videos and podcasts, communicating pedagogical ideas in a discussion forum post, or sharing a mathematical idea using technology, requires them to build communication and presentation skills, while allowing for practice and refinement that is not available in face-to-face synchronous environments.

Best practices in online statistics teacher education are starting to emerge in various settings. In one online course focused on teaching mathematics with technology, with six weeks of statistics pedagogy-focused modules, participants displayed perspectives that had been emphasized during the course, on the nature of statistics, the features of a good statistical task, learning statistics, the practice of teaching statistics, and the role of technology (Harrison, Azmy, & Lee, 2018). Similarly, research on online professional development for statistics educators has also showed promising results for impacting perspectives and beliefs. For example, in one massive online open course designed for educators (MOOC-Ed), the *Teaching Statistics Through Data Investigations* (TSDI), participants reported feeling more confident to teach statistics and expressed changes in beliefs and perspectives regarding teaching statistics (Mojica, Lee, Lovett, & Azmy, 2018). New perspectives included engaging in statistics should involve exploring data, is enhanced by the use of dynamic technology, and should include investigative cycles and habits of mind, among others.

Research on Teaching Self-Efficacy

Historically, research on mathematics teachers has focused on the importance of teacher knowledge and teacher characteristics like teaching experience (e.g. Fetler, 1999; Hill, Rowan, & Ball, 2005; Wayne & Youngs, 2003). However, affective constructs have also become the focus of some research (Zan, Brown, Evans, & Hannula, 2006). In his chapter on affect in the first *Handbook of Research on Mathematics Teaching and Learning*, McLeod (1992) defines affect as “a wide range of beliefs, feelings, and moods that are generally regarded as going beyond the domain of cognition” (p. 576). He identified three main affective constructs to be addressed by researchers: (a) *beliefs*; (b) *attitudes*; and (c) *emotions*. *Values* is another main affective dimension included by some researchers (Zan et al., 2006). In addition to these four, many other affective constructs have been the focus of much research, including: confidence, self-concept, mathematics anxiety, learned helplessness, autonomy, intuition, metacognition, social context (Philipp, 2007), motivation, mood, and interest (Zan et al., 2006).

Definitions of the three main affective constructs—beliefs, attitudes, and emotions—usually involve a description of the degree of cognition that is associated with it (Estrada, Batanero, & Lancaster, 2011; Gal, Ginsburg, & Schau, 1997; Philipp, 2007). While there are varying definitions for these constructs, most definitions for each construct include some commonalities. In addition, while each construct is distinct, researchers admit that there is overlap between them (Gal et al., 1997). *Emotions* are the least cognitive of the three and involve positive or negative feelings or responses to experiences (Estrada et al., 2011). *Attitudes* are defined as the average of feelings, positive or negative, experienced as a learner and are viewed as less cognitive than beliefs but more cognitive than emotions (Gal et al., 1997). Similarly, attitudes are more stable than emotions but less stable than beliefs. Finally, *beliefs* are understandings of the world that are thought to be true (Philipp, 2007). They are more cognitive and stable, and thus harder to change, than attitudes and emotions.

Although definitions of affective constructs in education are not always clearly or consistently defined, synthesizing works have attempted to build foundational definitions based on past research. For example, one view is that beliefs are a subset of conceptions, and that *conceptions* are made up of beliefs and knowledge (Thompson, 1992). Teacher conceptions can be about math in general or about teaching and learning math (Philipp, 2007). In the context of research on affect, four types of teacher beliefs have been identified: (a) beliefs about mathematics; (b) beliefs about self; (c) beliefs about mathematics teaching; (d) beliefs about the social context (McLeod, 1992). Generally, in the context of statistics education, beliefs are defined to be about the nature of statistics, statistics classroom culture, and your own ability or need for statistics (Gal et al., 1997). For example, in an Australian survey, preservice and inservice teachers indicated they value statistics, and believe technology is helpful to teach and learn statistics, but they felt less knowledgeable about statistics content and expressed neutral emotions towards it (Marshman, Dunn, McDougall, & Wiegand, 2015). Not much research has been done on teachers' beliefs about statistics, especially compared to the vast amount of research done on teachers' beliefs about mathematics (Pierce & Chick, 2011).

Some research on teacher affect has measured teachers' *conceptions* (beliefs and knowledge) and *practices* to see if and how the two align (Philipp, 2007). Although some researchers have found inconsistencies in conceptions and practice of teachers they have studied, there has been a push to reinterpret those apparent inconsistencies as a lack of information on the part of researchers (Philipp, 2007). For example, inconsistencies can be explained by the fact that teacher beliefs are constructed and differ based on context (Hoyles, 1992).

The importance of teachers' affective characteristics are that they are thought to influence instructional practice, which influences student outcomes (Wilkins, 2008). In Wilkins' framework, based on Ernest's conceptual model (1989), teacher characteristics indirectly affect instructional practice through teachers' content knowledge, attitudes, and instructional beliefs. Instructional beliefs are also influenced by content knowledge and attitudes. A large-scale quantitative study found this model to be viable (Wilkins, 2008).

Measuring Attitudes and Beliefs

In this section, I focus on *attitudes* and *beliefs*, the two affective characteristics included in Wilkins' framework (2008). I first discuss common ways of measuring these constructs, the connection between research done on students and our assumptions about teachers, and the importance of developing ways to measure these constructs for mathematics education.

As with much affective research done to date, measuring attitude is mostly done using surveys with Likert-scale statements. For research on statistics' attitudes, there are several instruments, like the Survey of Attitudes Towards Statistics (Cashin & Elmore, 2005). However, these quantitative instruments do not provide information on *sources* of existing attitudes in students, so researchers have suggested qualitatively measuring for these sources (Gal et al., 1997). Within the context of statistics education, sources of *student* attitudes have been found to be: (a) previous experience with statistics in school and out of school; (b) previous experience with math; and (c) a belief that statistics is the same as math (Gal et al., 1997).

In mathematics education research, most research on beliefs has been done using two methods: case study and surveys (Philipp, 2007). The case study methodology, which has been used more predominantly in mathematics education, incorporates data collection techniques like interviews, observations, stimulated recall, and vignettes to gather a whole picture of teachers' beliefs. The survey method usually incorporates Likert-scale items to gain teacher-reported ideas of teachers' beliefs. Because surveys measuring beliefs leave a lot of information for researchers to assume and interpret (discussed in more detail below), some researchers have implemented alternative surveys, which require open-ended response reactions to video cases (Philipp, 2007).

The study of beliefs has blossomed to include different types of beliefs, like math-specific beliefs and self-efficacy beliefs, and beliefs of both teachers and students (Zan et al., 2006). Teaching self-efficacy is different from self-efficacy that one has a learner—but they are related (Philipp, 2007). "...the existing research shows that the feelings teachers experienced as learners carry forward to their adult lives, and these feelings are important factors in the ways teachers interpret their mathematical worlds" (Philipp, 2007, p. 258). For example, in considering attitudes and beliefs in statistics education, research has been done on students (e.g. Gal et al., 1997), and some argue that since teachers carry with them their experiences as students, affective research on students regarding beliefs and attitudes also applies to teachers (Estrada et al., 2011; Philipp, 2007; Pierce & Chick, 2011).

Learning about attitudes and beliefs is important for knowing how well prepared teachers are to teach mathematics, particularly statistics. In one study on the relationship between knowledge, attitudes, beliefs, and practices for elementary teachers, it was found that beliefs had the biggest impact on practices (Wilkins, 2008). In a mathematics education context, Cooney, Shealy, and Arvold (1998) describe the importance of teacher educators assessing their preservice teachers' beliefs in order to purposefully choose activities that might develop their beliefs. Gal et al. (1997) consider three ways that attitudes and beliefs about statistics particularly matter for students: (a) process considerations; (b) outcomes considerations; and (c) access considerations. In other words, attitudes and beliefs influence the teaching and learning process, how students emerge from courses thinking about statistics, and students' future decisions about careers that require statistics. Finding ways to assess student and teacher beliefs and attitudes within the context of statistics education is a worthwhile albeit complex endeavor.

Defining and Measuring Self-Efficacy Beliefs

As stated earlier, beliefs can be about content, self, teaching, and social context. These beliefs are related, but I focus on beliefs about self. Specifically, much research has been done on teaching self-efficacy [TSE] beliefs. Below, I address the theoretical foundations for self-efficacy research, research on sources of self-efficacy, and defining and measuring self-efficacy.

Research done on self-efficacy has traditionally been based on one of two theories—Bandura's social cognitive theory (1977) or Rotter's locus of control theory (1966). Rotter's locus of control theory distinguishes between the belief that outcomes are a result of one's own actions, internal control, or a result of external factors outside of one's control. Bandura's theory built on Rotter's by adding a dimension of beliefs in one's own capabilities, not just beliefs about humans' general ability to control circumstances (1977).

It is generally now accepted that measurements based on Bandura's theory accurately measure self-efficacy, rather than measurements based on Rotter's theory, which tend to measure general teaching efficacy rather than personal teaching efficacy. Bandura's (1977) original definition of self-efficacy is "beliefs in one's capabilities to organize and execute the courses of action required to produce given attainments" (p. 3). Bandura's general approach to self-efficacy was reinterpreted in the teaching context and much research has been done within that framework (Tschannen-Moran, Woolfolk Hoy, & Hoy, 1998). However, there have been many issues with past research on measuring TSE. One issue is that the very earliest studies (e.g. Armor et al., 1976) that paved the way for claiming the importance and effects of teaching self-efficacy were based on Rotter's theory. Another issue is that instruments used in later studies based on Bandura's theory were found to have measurement issues (Henson, Kogan, & Vacha-Haase, 2001).

Part of Bandura's social cognitive theory specifies four sources of self-efficacy: enactive mastery experiences, vicarious experiences, social persuasions, and physiological and affective states (Bandura, 1977). Morris, Usher, and Chen (2017) describe in more detail what these four sources mean in the teaching context. Enactive mastery experiences are experiences teachers have with actual teaching tasks, like developing and implementing lessons that result in student outcomes, positive or negative. Vicarious experiences occur when a teacher experiences another teacher's mastery experience; teachers must be able to relate to other teachers in some way for vicarious experiences to be influential. Social persuasions come in the form of encouragement, feedback, or assessments from peers or supervisors about one's teaching abilities. Finally, physiological and affective states are physical and emotional reactions during teaching experiences, like a fast heart-rate, sweaty hands, or anticipation.

Bandura claimed mastery experiences to be the most important of the four sources, which for teachers means successful experiences teaching students (Bandura, 1997). Empirical evidence has supported this claim (Tschannen-Moran & Hoy, 2007; Tschannen-Moran & McMaster, 2009). In addition, sources of TSE are not weighted evenly for novice teachers compared to experienced teachers; novice teachers are more impacted by vicarious experiences and contextual differences than are career teachers, presumably because of their lack of mastery experiences (Tschannen-Moran & Hoy, 2007).

It is important to note that these sources affect self-efficacy indirectly; teachers' interpretations of their experiences or affective states are what in turn affect their self-efficacy (Morris, Usher, & Chen, 2017). For example, two teachers may experience teaching vicariously by observing the same lesson by the same teacher, but may internalize very different reactions. Also important to note is that self-efficacy is a measure of *perception* of one's ability to produce certain behaviors or outcomes; not a measure of one's actual ability.

In addition, note the lack of "knowledge" as a source of self-efficacy according to Bandura. Although many researchers have claimed knowledge—content, pedagogical, and technological—to be an important source for teachers' self-efficacy, Bandura viewed it as a *result* of the four sources he described. As Morris et al. describe, "Knowledge is not, in itself, a source of self-efficacy; it is necessarily derived from previous experiences" (2017, p. 805). Other researchers also argue that knowledge is a form of mastery experience (Morris et al., 2017). The relationship is not certain. One example that may support Bandura's view is Wilkins' (perhaps) unexpected finding that increased mathematics content knowledge negatively relates with inquiry-based instructional practices for elementary teachers (Wilkins, 2008). So, while teachers may have succeeded in traditional mathematics classes and thus have increased content knowledge, their interpretation of that experience is what gives knowledge its direction of influence.

In defining TSE, there has been much confusion and debate among scholars. A widely used definition is from a seminal piece that says, "Teacher efficacy is the teacher's belief in his or her capability to organize and execute courses of action required to successfully accomplish a specific teaching task in a particular context" (Tschannen-Moran et al., 1998, p. 233). Recently, researchers have analyzed commonly used TSE definitions like this one to see if they portray the necessary complexity of TSE beliefs (e.g. Wheatley, 2005; Wyatt 2014). For example, many definitions describe agent-ends beliefs, rather than separating out agent-means and means-ends beliefs (Wyatt, 2014). Agent-means beliefs center around the idea of people's' beliefs that they can effectively enact some type of action or behavior. This is different from means-end beliefs, which refers to beliefs that certain actions elicit certain outcomes. For example, teachers can believe a certain teaching method is effective for student learning (means-end), but that they are unable to successfully use that teaching method (agent-means).

Agent-ends beliefs leave out information about the middle step, the means, that might impact an outcome; they are beliefs that one can effectively be the cause of a desired outcome. It is helpful to read TSE definitions with these types of beliefs in mind. For example, in a later paper, Tschannen-Moran & Woolfolk-Hoy construct a new definition, "A teacher's efficacy belief is a judgment of his or her capabilities to bring about desired outcomes of student engagement and learning, even among those students who may be difficult or unmotivated" (2001, p. 783), which incorporates only agent-ends beliefs rather than their former definition, which incorporates agent-means and means-end beliefs (Wyatt, 2014).

Differentiating between these three views ends up being very important when attempting to measure self-efficacy. In a landmark piece, Smith identifies the mathematics reform movement as a challenge to TSE (and vice-versa) because it is so much easier to attribute means or ends to agents (teachers) when teachers are doing the work of simply passing on content to students (1996). Smith discusses the reform movement in mathematics and how it fundamentally changes ideas about mathematics content, teaching, and learning. How do teachers assess their own self-efficacy when their role shifts from "telling" to providing opportunities for and guiding students to develop their own understandings based on their own actions with mathematical objects? The traditional view of teaching as telling provides teachers a direct link to their self-efficacy beliefs because they can very directly attribute the cause of student learning to their own actions, which involve presenting information in a clear way for students to passively receive knowledge that teachers have been successful in attaining themselves.

However, because the reform movement pushes teachers to reconsider how students learn from passive reception to active construction of knowledge, teachers must now view their role in terms of providing opportunities for students to mathematically act, by choosing tasks and leading classroom discourse. Thus, the link between teacher actions and student learning in the reform movement involves a complex teaching model, which does not provide teachers the direct causal link between actions and outcomes.

Past research on TSE generally has mostly been quantitative, and measurement has been done using surveys with Likert-scale responses or something similar. Affective constructs in general have an assortment of well-known surveys that have been used in research, but there is much debate as to the validity and reliability of those instruments. Originally, the most popular instrument for measuring TSE was the Teacher Efficacy Scale (TES) (Gibson & Dembo, 1984), but because of many measurement issues (Morris et al., 2017), it was modified into the Teachers' Sense of Efficacy Scale (TSES) (Tschannen-Moran & Woolfolk-Hoy, 2001). Quantitative instruments that have been used to measure TSE *sources* have been criticized for not being psychometrically strong and for having construct validity issues (Morris et al., 2017).

Most researchers agree that in measuring self-efficacy, it is best that measurements be context- and task-specific (e.g. Bandura, 1977; Smith, 1996; Wyatt, 2014). However, the widely used Teachers' Sense of Efficacy Scale (TSES), which measures TSE generally, does not include questions specifically about teaching in a specific context, like a mathematics classroom

(Tschannen-Moran & Hoy, 2007). Some researchers have chosen to use the TSES for mathematics teachers by adding mathematics teaching questions (e.g. Ross & Bruce, 2007). On the other hand, some instruments have been created which are subject-specific. One example of a math-specific survey is the Mathematics Teaching Efficacy Belief Instrument (MTEBI) (Enochs, Smith, and Huinker, 2000), which was modified from a science instrument (Bates, Latham, & Kim, 2011). An example of a statistics-specific survey is the Self-Efficacy to Teach Statistics (SETS) survey (Harrell-Williams, Sorto, Pierce, Lesser, & Murphy, 2014; Harrell-Williams et al., 2019).

There have been critiques about not only the measurement of TSE, but also the assumption of a simplistic model of benefits of high TSE. The traditional model posits that high self-efficacy results in better teaching, which results in desired outcomes, which strengthens TSE (Tschannen-Moran et al., 1998). Similarly, it is believed that unsuccessful experiences lower TSE, which lowers quality of teaching.

However, there has been some push-back on this simple cyclic model because self-efficacy beliefs may actually over- or under-estimate one's actual capabilities. For example, some researchers have pointed out that high TSE that does not actually match the level of teaching practice, like in the case of preservice or novice teachers, might not always be beneficial (Wheatley, 2005; Wyatt, 2014). In that case, some doubts about efficacy may push teachers to learn or develop new teaching strategies that are ultimately more effective. As Wheatley claims, "Doubt is crucial for reflection" (2005, p. 756). For example, in professional development models, it has been shown that some formats that do not include an authentic mastery experience actually lead to a drop in TSE for many teachers (Tschannen-Moran & McMaster, 2009), This may be because teachers were given knowledge about a new teaching method that then made them reassess their teaching task, but without a substantive mastery experience they did not feel like they could attain success with it.

Thus, there have been issues consistently defining TSE, measuring TSE, and interpreting TSE effects. For all of these reasons, there is a general consensus that future TSE research should include qualitative components (e.g. Ross & Bruce, 2007; Tschannen-Moran & Hoy, 2007; Tschannen-Moran et al., 1998; Wheatley, 2005; Wyatt, 2014; Wyatt, 2015). In fact, those in mathematics education have been advocating for these methods for over 20 years, in order to gain insight into what contributes to TSE and how it changes over time through the use of case studies (Smith, 1996).

Research on Statistics Teaching Self-Efficacy

The increasing use and availability of data in the modern world has led to a new focus on statistics education and the inclusion of statistics standards in curricula. Much of the research on statistics education for teachers has focused on teacher knowledge, which has generally found a lack of teacher knowledge sufficient to teach statistics (e.g. Lovett & Lee, 2017; 2018; Begg, & Edwards, 1999; Hannigan et al., 2013). Although teacher self-efficacy in general has an impact on student success (Tschannen-Moran & Hoy, 2001), not much research has been done on teachers' confidence or attitudes towards teaching statistics; the little that has been done specifically on teacher confidence is inconclusive (Estrada, Batanero, & Lancaster, 2011).

Some research has been done that indicates that advances in self-efficacy to teach statistics are needed. In a study of 15 primary and 28 secondary teachers, it was found that primary teachers are less confident to teach statistics than are secondary math teachers (Watson, 2001). Using the SETS instrument, middle school teachers have been found to be confident in teaching Level A topics, according to the GAISE framework levels of understanding, but not confident in teaching Level B topics (Harrell-Williams et al., 2014). Preservice high school teachers were found to be unconfident in teaching certain statistics topics. This was because of their lack of recent exposure to statistics and also because of the unique nature of probabilistic statistics reasoning compared to the deterministic nature of mathematics (Harrell-Williams et al., 2014). Similarly, preservice mathematics teachers in Ireland rated themselves less confident to learn statistics than to learn mathematics on a survey of attitudes towards statistics (SATS) (Hannigan et al., 2013). Although perceived ability to do statistics has been found to be correlated with perceived ability to teach statistics, that does not necessarily imply a correlation with actual ability to do statistics; in fact, the opposite has been shown (Harrell-Williams et al., 2019).

In a follow-up study Fitzmaurice, Leavy, and Hannigan used interviews to interpret why these preservice teachers in Ireland viewed statistics as difficult (2014). They found 5 common themes among prospective teachers: (a) influence of secondary school experience and teacher; (b) avoidance of statistics as a focus of study; (c) teacher knowledge (their own); (d) perception of difficulty; and (e) influence of teaching practice. Unfortunately, researchers found the impact on teaching is these PSTs would avoid teaching statistics if they could. One notable finding is that the preservice teachers who had taught statistics in their field-based teaching experience had positive reactions to their experience and stated they would like to teach it again. The authors draw a conclusion that teaching statistics led to higher self-efficacy, which they view as valuable because of their belief in the simplistic cyclical nature of TSE. While these results corroborate what we thought to be true about mastery experiences being an important source of high TSE, there should be more said about *how* these PSTs taught statistics and what they view statistics and statistics education to be.

There are some examples that exemplify this need. Begg and Edwards (1999) found that inservice and preservice primary teachers were mostly confident when teaching statistics. However, this same study found that the majority of these teachers viewed statistics narrowly, as charts and numbers, rather than as reasoning and investigating. In addition, a significant number of them were not familiar with a newly implemented statistics curriculum in New Zealand. This indicates that even when self-efficacy for teaching statistics is high, it does not necessarily correlate with modern views of statistics and how it is different from mathematics. In line with these findings, Irakleous and Panaoura (2015) found in their case study that the PSTs they interviewed had high STSE, even though they had insufficient statistics content knowledge. These authors were specifically focused on personal teaching efficacy (PTE), and they were surprised to find that despite PSTs lack of any mastery experience, one source of their high statistics PTE was their belief that primary school statistics content is easy to learn.

One factor often associated with a negative attitude towards statistics is a lack of knowledge of the topic (Estrada & Batanero, 2008). In Estrada and Batanero (2008) define "attitude" as "learned predispositions to respond positively or negatively to given objects, situations, concepts, or persons which includes self-efficacy," which is based on McLeod's (1992) classic definition. While Hannigan et al. found prospective math teachers' positive attitudes towards statistics not to be strongly correlated with conceptual understanding of statistics, the authors of the study admit that the participants had very little exposure to statistics in secondary school and might have confounded attitudes towards statistics with attitudes towards mathematics in general (2013). This actually corroborates earlier findings about the difficulty of teaching and using probabilistic statistics reasoning compared to the deterministic reasoning prevalent in mathematics.

Most recently, a mixed methods study of mathematics PSTs from a purposeful sample of eighteen U. S. institutions explored PSTs' STSE compared to other mathematics topics and factors that influence their TSE (Lovett & Lee, 2017). The authors also used a content assessment to measure PSTs' statistical knowledge, and found that it was generally lacking, with mathematics PSTs scoring a mean of 69% and struggling with such important concepts like p-values, sampling distributions, and variability. Although results from the SETS instrument (Harrell-Williams et al., 2014) showed PSTs to be generally confident to teach statistics (about a 4 on a 6-point scale), they ranked it lower than the other topics in secondary school such as calculus, algebra, geometry. Further, based on open-ended responses to the SETS instrument and interviews, the authors described factors that contribute to PSTs' perceived preparedness to teach statistics: (a) role of statistics knowledge; (b) role of pedagogical knowledge; (c) impact of using technology; and (d) view of statistics.

Statistics knowledge was cited as a factor by PSTs' to explain both why they felt confident or unconfident to teach a particular topic in statistics, and came up much more often than the other three factors, which were mostly cited as reasons why PSTs were confident (Lovett & Lee, 2017). For pedagogical knowledge, PSTs discussed the importance of having knowledge of strategies to teach a particular topic and common student conceptions about a statistics topic; when they lacked this knowledge for certain topics, they indicated not being highly confident to teach them. With regards to technology, they expressed comfort with particular computation tools and wrote about experiences with technology that led to deeper understanding of concepts. PSTs expressed confidence with particular topics by expressing a view that certain topics were procedures or by relating topics to algebra, which most of them ranked higher in terms of TSE. For all three of these factors, PSTs were more often citing them as reasons they were most confident to teach a statistics topic, but sometimes citing them as reasons they were least confident to teach a statistics topic.

Despite the lack of a full body of conclusive research on teacher self-efficacy for teaching statistics, the research that has been done indicates room for improvement. It also indicates that while mastery experiences and vicarious experiences can help to improve self-efficacy, weighted differently for inservice versus preservice teachers (Tschannen-Moran & Hoy, 2007; Tschannen-Moran & McMaster, 2009), it is important for researchers to gain an understanding of participants' perspectives on statistics education, since quantitative methods do not give vital information on the latter (Wheatley, 2005). In the grand picture of statistics education, this is tremendously important because teacher self-efficacy has been shown to have a positive impact on a large variety of student outcomes (Tschannen-Moran & Hoy, 2001).

Development of a Statistics Teaching Self-Efficacy Framework

In order to study the development of statistics teaching self-efficacy in a teacher preparation program, and specifically within a specially targeted online learning experience, I synthesized existing research to develop a statistics teaching self-efficacy framework. I began with Tschannen-Moran et al.'s influential work (1998), in which they propose a framework that attempts to reconcile research done using Rotter's locus of control as a theoretical basis and research done using Bandura's social cognitive theory as a theoretical basis (Figure 1). The model refers to general teaching efficacy belief as "analysis of teaching task" and personal teaching efficacy belief as "assessment of personal teaching competence." It is the cognitive processing of efficacy sources that form these two components of a teacher's overall self-efficacy. It is important to note that they describe the process as cyclical. Thus, what results from teachers' self-efficacy influences performance, which is then incorporated as a new source of efficacy information and then cognitively processed into analysis of teaching tasks and assessment of personal teaching competence.

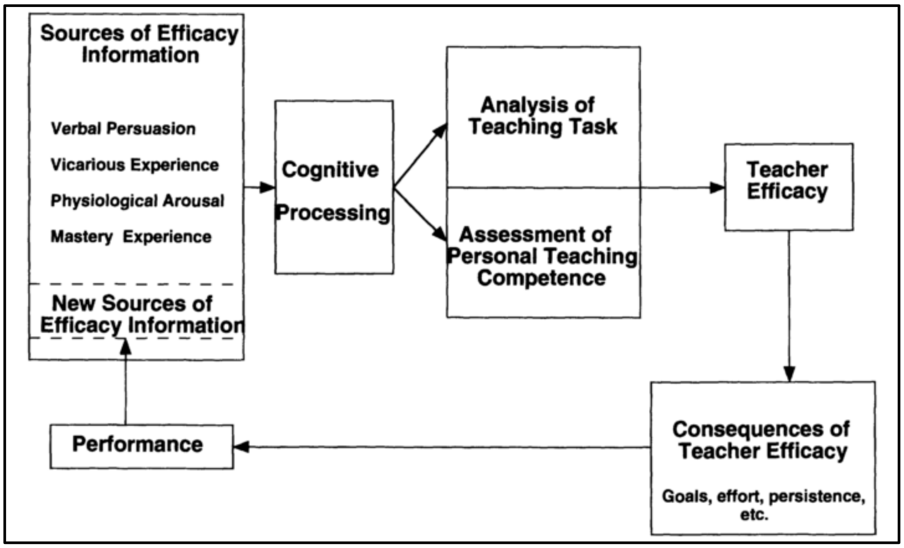


Figure 1. Tschannen-Moran et al.'s framework that describes the cyclic nature of teacher efficacy (1998).

In another framework that models teacher change (Figure 2), built based off a qualitative study of an intervention implementing teacher self-assessments, the development and impacts of TSE are also viewed as cyclical (Ross & Bruce, 2007). At the center of the model is the focus of the researchers, which is teacher self-assessment, impacted by student achievement, peer input, and professional development presenters. The impacts of TSE are seen as indirect; rather than directly impacting student achievement, the effect is mediated by goal setting and effort, which impact instructional practice, which then directly impacts student achievement. It is interesting to note the direct link between practice and student achievement, indicating a strong belief in means-end effects. Although this framework does not explicitly link to Bandura's four sources of self-efficacy (1986), I can map them onto the categories included; for example, student achievement could be considered a part of mastery experiences, peer and change agent input could be social persuasion, and the observation part of the professional development (PD) could be a vicarious experience. Although physiological and emotional states are not addressed in the framework, the authors claim that this dimension was addressed in the PD.

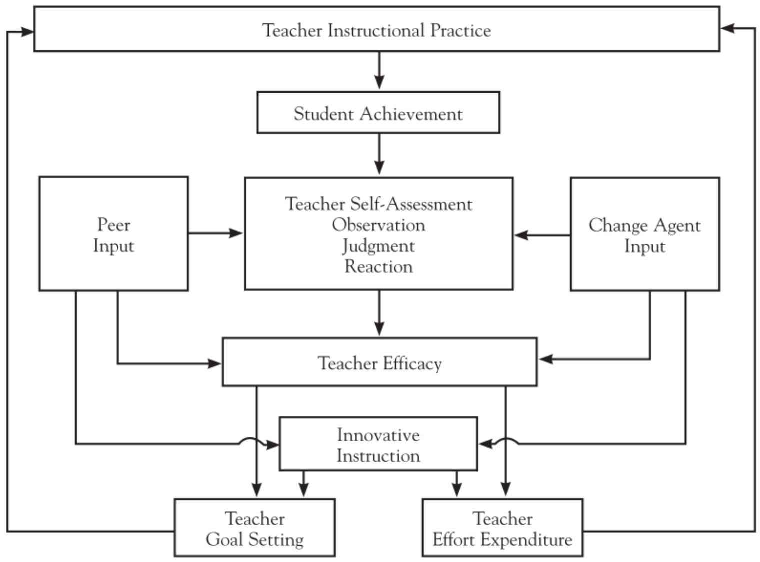


Figure 2. Model of teacher self-assessment as a mechanism for teacher change (Ross & Bruce, 2007).

In yet another cyclic model of TSE, Zee and Koomen's model (2016) proposes specific classroom processes that explain the relationship between TSE and both student and teacher outcomes (Figure 3). In this model, TSE is seen as influencing and being influenced by students' academic adjustment and teachers' well-being. In addition, the quality of classroom processes is a mediating variable.

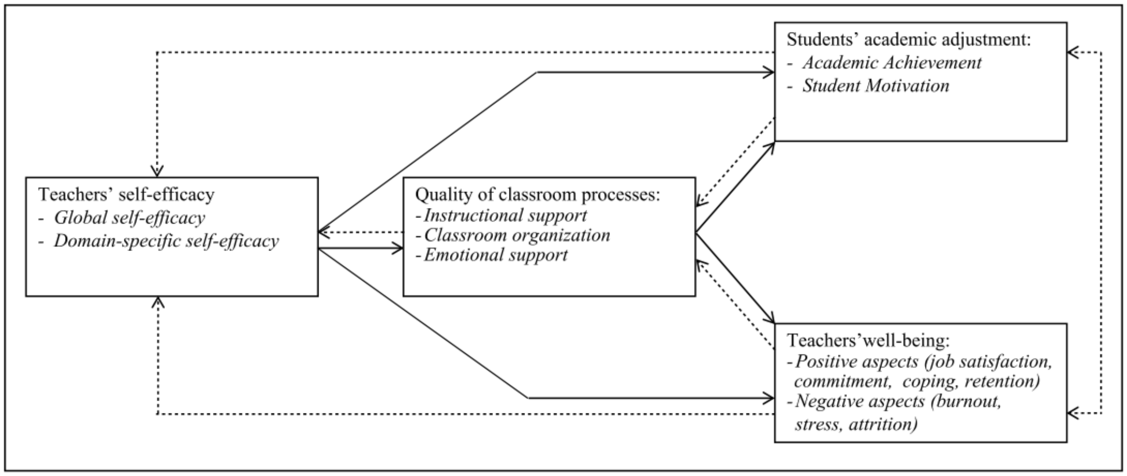


Figure 3. Another cyclic model of teacher self-efficacy (Zee & Koomen, 2016).

Although there are many more frameworks and models that incorporate TSE, I use these three as representative of important aspects that are most commonly included. First, it is important to mention that all three are viewed as cyclic. Second, TSE is commonly seen as indirectly impacting student outcomes by impacting a variety of teaching practices. Finally, most models are careful to emphasize the fact that it is a teacher's perception or processing of an experience that influence his or her TSE—not the experience itself.

Conceptual Framework for this Study

Although teacher affect includes many constructs, with various definitions, and is viewed by different researchers with different perspectives, I will use McLeod's distinction of three main branches: beliefs, attitudes, and emotions (McLeod, 1992). These three affective constructs are related with some overlap, but two main characteristics that distinguish them are their stability and the amount of cognition related to each. For example, the amount of cognition associated with beliefs, attitudes, and emotions goes from highest to lowest. The same goes for their stability.

Another common construct in the literature is the idea of teacher conceptions; conceptions are also defined differently by different researchers, but I am defining teacher conceptions to consist of beliefs and knowledge (Thompson, 1992). Thus, the four teacher characteristics that I focus on, acknowledging that there exist many more, are: knowledge, beliefs, attitudes, and emotions, where knowledge and beliefs form a “conception,” and where knowledge is the only construct not considered to be affective (Figure 4).

Within the affective construct of beliefs, there are different types of beliefs. McLeod specified four types of teacher beliefs: (a) beliefs about mathematics; (b) beliefs about self; (c) beliefs about mathematics teaching; (d) beliefs about the social context (1992). Within the statistics education context, teacher beliefs include, but are not limited to, beliefs about mathematics, beliefs about statistics, personal teaching statistics self-efficacy beliefs, and general teaching statistics self-efficacy beliefs (Estrada et al., 2011; Gal et al., 1997).

The categories that make up teacher characteristics are not completely distinct, and they affect each other. For example, former experiences affect all four characteristics (Estrada et al., 2011), but attitudes and emotions also affect beliefs. In my framework (Figure 4), dotted lines represent the fact that some characteristics are indirectly affected; they pass through a teacher’s perception of an experience, rather than the experience itself. Similarly, student outcomes are incorporated into self-efficacy information after being reflected on by teachers. Student achievement is also not directly impacted by teacher characteristics, but indirectly through other aspects of teacher practice like instructional support, classroom organization, and emotional support (Ross & Bruce, 2007; Zee & Koomen, 2016).

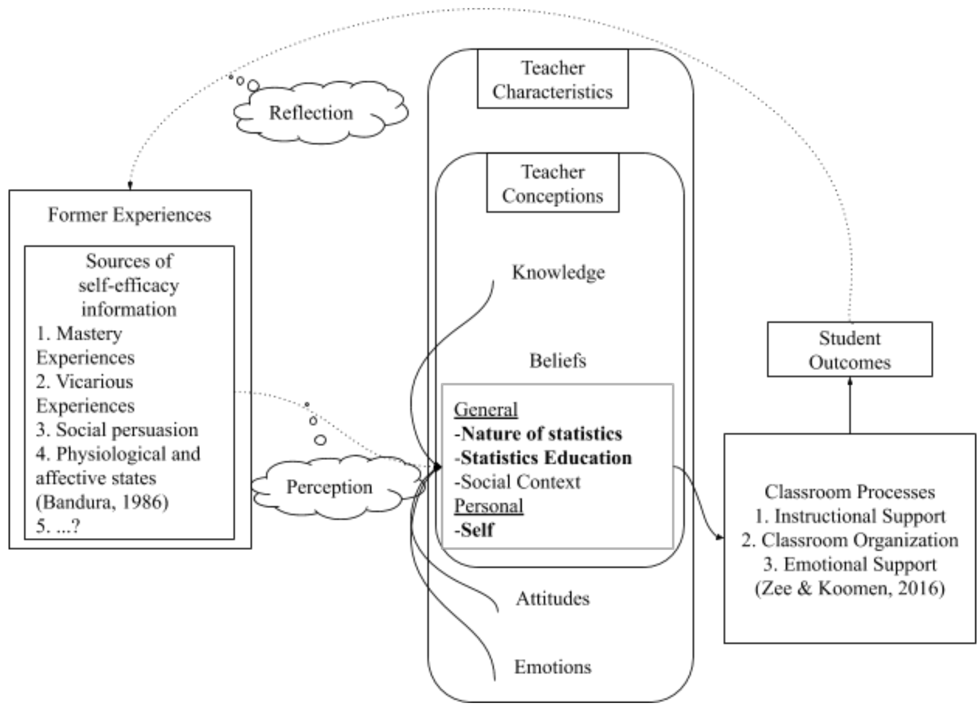


Figure 4. Framework on teaching self-efficacy, with former experiences indirectly affecting teacher characteristics through teacher perception and student outcomes being indirectly affected by teacher practice.

The focus of my research is on three types of teacher belief—belief about the nature of statistics, belief about statistics education, and belief about the self. Belief about the self includes personal self-efficacy to teach statistics. Bandura theorized four sources of this type of belief: mastery experiences, vicarious experiences, social persuasion, and physiological and emotional states. Again, it is a teacher's reflection on these four sources, or their perception, that forms a teacher's self-efficacy (Ross & Bruce, 2007). In my framework, I am allowing for the possibility for more than just the four that Bandura theorized. In addition, there is overlap between constructs in the framework. For example, attitudes and emotions can also be characterized as the last source of self-efficacy beliefs— physiological and emotional states. This is also reflected in literature that has found knowledge and attitudes to influence beliefs (Wilkins, 2008).

Self-efficacy includes two components—personal teaching self-efficacy and general teaching self-efficacy. The first, personal teaching efficacy (PTE) beliefs, concerns teachers' beliefs in their own abilities to perform certain behaviors or actions or to achieve certain outcomes, depending on the definition of self-efficacy you are using. The second is about teachers' beliefs in the role of teaching generally; for example, a teacher might not believe that *any* teacher can be effective in a certain context or with a certain population of students. With respect to statistics specifically, general statistics teaching self-efficacy could include a belief that teaching statistics is most effectively done by allowing students to experience the full investigative cycle. Because of the unique nature of statistics, these general statistics education (GSE) beliefs are especially relevant in how they relate to mathematics PSTs' PTE beliefs. For example, if PSTs display extreme confidence in teaching statistics, it is important to understand *what* they consider statistics is and *how* they think it should be taught.

Finally, it is important to note that I chose to include the cyclic aspect incorporated by most other researchers in their models of change in TSE. However, I have tried to address critiques in recent works of the simplistic model that assumes that successes or failures lead to higher or lower TSE respectively, which leads to further successes or failures. By explicitly highlighting teachers' reflection on student outcomes, I have attempted to incorporate some indication that teachers' doubts and perturbations to their current TSE state caused by what they deem unsuccessful experiences can also lead to positive teacher change.

The framework in Figure 4 not only situates the specific topic of statistics teaching self-efficacy within the larger domain of affect, but also synthesizes common characteristics of frameworks that have been used in TSE research. It displays relationships between different constructs and theorized directions of influence, based on past theory and research. My hope is to use the framework to ground my research on the development of STSE, focusing specifically on the left hand side of the diagram, namely *Former Experiences* and how perceptions of those experiences form *Teacher Beliefs*, because I am interested in preservice teacher STSE development. For this phase in a teacher's career, they do not have a chance to experience the full cycle on a consistent basis, where their teaching practice and student outcomes affect their TSE development. In order to focus on the relevant portion of the TSE development cycle, I present a zoomed-in look at that portion, which provides more detail on how *Former Experiences* are interpreted to form STSE (Figure 5).

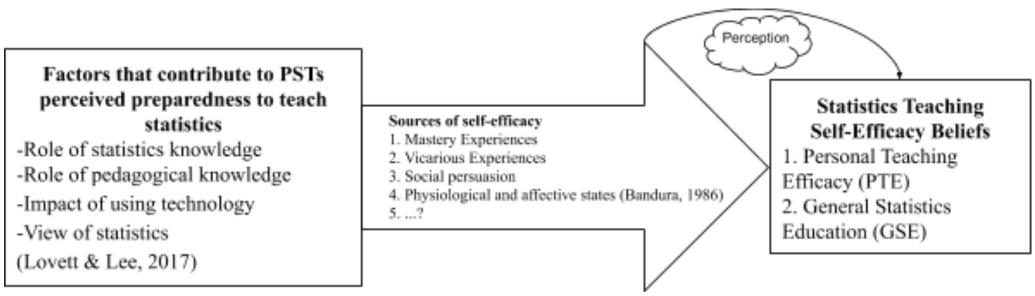


Figure 5. Zoomed-in view of larger framework representing the statistics TSE development cycle.

In this view of the framework, I insert an additional piece of information that cannot be seen in the larger framework. Here we see that teachers (both preservice and inservice) identify *factors* that contribute to STSE through *experiences* that can be considered sources of TSE. For example, teachers identify statistics knowledge as important in the development of their perceived preparedness to teach statistics. This statistics knowledge impacts STSE belief as it happens through an experience that occurs, for example, in a formal statistics course. Another example could be that a teacher identifies her comfort with using a graphing calculator to create linear regression models as an important factor in her STSE, but that comfort developed as a result of a professional development workshop she went to where the facilitator demonstrated how to do that skill. Thus the factor of using technology impacts her STSE *through* a former experience. During and after teachers' experiences, their interpretation of whether experiences positively or negatively impacted their knowledge, for example, leads to a change in STSE, which includes both beliefs about self (PTE) and beliefs about statistics education (GSE).

The addition of this upfront piece is important because PSTs do not view their former experiences in Bandura's language of the four sources. Rather, they view them through their interpretation of how those experiences impacted factors they find relevant to their teaching ability. In my research of discovering mathematics PSTs' pathways to STSE, I connected mathematics PSTs' identified impactful factors with Bandura's four sources of efficacy.

Mathematics PSTs may encounter a multitude of experiences in their preparation programs that contribute to factors they perceive as impactful to their STSE. Those experiences can then be categorized into Bandura's four sources of self-efficacy. For example, mathematics PSTs may experience being a statistics student during their coursework requirements. Content knowledge gained during that mastery experience of doing statistics may impact STSE; however pedagogical content knowledge may also be gained during that vicarious experience of viewing another teacher teaching statistics. In courses specifically focused on preparing mathematics teachers, all of the factors identified that most impact STSE could occur during experiences like lesson plan assignments, video case analyses, classroom discussions, peer teaching assignments, and engagement in statistical tasks. Each of these experiences could be considered a mastery experience for teaching or doing statistics or a vicarious experience. Finally, in fieldwork assignments, both mastery and vicarious experiences may impact STSE if mathematics PSTs have the opportunity to observe or help teach statistical content. Throughout all of these examples, mathematics PSTs' physiological and affective states doing assignments may impact their STSE, in addition to feedback they may receive from instructors, mentor teachers, peers, or students.

Conclusion

In this literature review, I have presented a summary of research on various relevant areas – statistics education, mathematics teacher education, statistics teacher education, online teacher education, teaching self-efficacy, and statistics teaching self-efficacy. I also presented a guiding conceptual framework and how it was developed, in addition to a zoomed-in perspective, on which to base this research study. Most researchers now agree that qualitative research on TSE and its sources is needed in order to understand how best to prepare teachers to teach mathematics and statistics. Thus, although there has been some research done on STSE specifically, it is sometimes inconclusive or conflicting, and has not fully answered the question of how mathematics PSTs develop their own STSE.

Chapter 3: Methods

The two research questions guiding this qualitative collective case study are:

1) "What are the general statistics education beliefs and personal teaching efficacy beliefs that comprise preservice mathematics teachers' statistics teaching self-efficacy, and what experiences and factors most impact development of those beliefs?" and
2) "How do statistics pedagogy learning experiences impact preservice mathematics teachers' statistics teaching self-efficacy?"

In order to answer these questions, I collected a variety of qualitative data related to different aspects of my theoretical framework from participants who were part of a situated context. Below I describe the study design, the situated context, all data sources, and data analysis.

Study Design

Qualitative research is effective in answering open-ended questions that require a deep understanding of a complex issue (Creswell, 2013). Specifically in researching TSE, where measurement issues have long been around, I complied to the overwhelming call for qualitative research (Ross & Bruce, 2007; Tschannen-Moran & Hoy, 2007; Tschannen-Moran et al., 1998; Wheatley, 2005; Wyatt, 2014; Wyatt, 2015). To answer my research questions, I used a collective case study method with two phases. Phase I included 34 preservice secondary mathematics teachers enrolled in courses using the same online material, and Phase II then dove deeper with 7 of those participants. Focusing on a few individuals' self-efficacy to teach statistics allowed me to delve into the "how" and "why" self-efficacy develops over the course of mathematics PSTs' educational careers.

As Smith called for over 20 years ago (1996), case study methodology is appropriate because it allowed me to gather multiple types of data in order to provide an in-depth description of PSTs' experiences with learning and teaching statistics (Creswell, 2013). It also allowed me to develop themes of self-efficacy development (Creswell, 2013), and to answer the questions of what are common sources of TSE and how does it change over time (Smith, 1996)?

Teacher Education Context

In this section, I will describe the context of this research study and how aspects of the context connect to my theoretical framework. Participants were from two institutions that offer a hybrid teaching mathematics with technology course, Institution A and Institution B (pseudonyms). These courses are field test sites for the Enhancing Statistics Teacher Education with E-modules (ESTEEM) project (NSF-funded grant Due 1625713). ESTEEM aims to design online statistics pedagogy modules to be used in preparatory courses for preservice mathematics teachers. The modules are flexible for use a purely online, hybrid, or face-to-face setting.

Because I view the ESTEEM module as a critical aspect of mathematics PSTs' opportunity to development of STSE, it provides a situated context, which all analysis and interpretations must take into consideration. I purposely chose the ESTEEM context to study mathematics PSTs' STSE development because it provided common learning opportunities which I know all the PSTs would have experienced. Mathematics PSTs would have had a variety of other experiences in terms of primary, secondary and tertiary statistics courses as a learner, statistics teaching experiences, and informal experiences. I did not systematically collect information from all participants about former statistics experiences, except for a subset of participants during an interview after their participation in the course. Thus, I am making the assumption for participants I did not interview that their experiences are varied, even if they are at the same institution. My assumption is based on the fact that even students who experienced the same college coursework trajectory may have taken courses at different times, with different instructors, and they may have had varying high school experiences, like AP statistics.

However, all participants in this study would have experienced the resources in ESTEEM's foundation module during their online course, which provided guaranteed opportunities for me to probe into if, how and why those experiences affected mathematics PSTs' STSE, in relation to the whole corpus of other experiences they have encountered throughout their life.

The ESTEEM foundational module provides mathematics PSTs important experiences with regards to specific aspects of my framework (Figure 5). Table 1 describes how the module is organized with materials to read and watch, dataset investigations, and discussion forums to synthesize and apply participants' learning, a total of 16 to 18 hours worth of material (Table 1).

Table 1. Overview of ESTEEM online foundational module and connections to literature on teacher education and STSE.

Section	**ESTEEM Learning Opportunity**	**Type of Resource**	**Teacher Education Connection**	**Hypothesized Impact on STSE**
ESTEEM Foundation Module Part 1				
Read and Watch Essential Materials	1.1a. How is statistics different from mathematics?	Online page reading	-Fostering awareness to similarities and differences -Identifying and overcoming barriers to students' learning -Sharing and revealing self, peer, and student dispositions	-Mastery experience
	1.1.b. Statistical investigations and habits of mind	Online page reading with instructional video	-Fostering awareness to similarities and differences -Identifying and overcoming barriers to students' learning -Sharing and revealing self, peer, and student dispositions	-Mastery experience
	1.1.c. Considering the importance of teaching statistics	Online page reading with expert panel video	-Fostering awareness to similarities and differences -Identifying and overcoming barriers to students' learning -Sharing and revealing self, peer, and student dispositions	-Mastery experience

Table 1 (continued)

	1.1.d. Quiz on Read and Watch material	Quiz	-Fostering awareness to similarities and differences -Identifying and overcoming barriers to students' learning -Sharing and revealing self, peer, and student dispositions	-Mastery experience -Social Persuasion -Physiological and affective state
Read and Watch Learn from Practice Videos	1.1.e Teaching statistics in the mathematics curriculum	Video of expert teacher in classroom	-learning from the study of practice -selecting and using (appropriate) tools and resources for teaching	-Vicarious experience
	1.1.f. Statistical investigation cycle in a classroom	Video of classroom	-learning from the study of practice -selecting and using (appropriate) tools and resources for teaching	-Vicarious experience
Engage with Data	1.1.g. Investigating older roller coasters in the US	Statistical investigation assignment with video of classroom	-developing adaptability -fostering awareness to similarities and differences -coping with conflicts, dilemmas and problem situations -selecting and using (appropriate) tools and resources for teaching -Sharing and revealing self, peer, and student dispositions	-Mastery experience -Vicarious experience -Physiological and affective state
Synthesize and Apply	1.1.h. Discuss learning statistics through investigations with real data	Discussion forum	-fostering awareness to similarities and differences -coping with conflicts, dilemmas and problem situations -identifying and overcoming barriers to students' learning -sharing and revealing self, peer, and student dispositions	-Vicarious experience -Social Persuasion -Physiological and affective state
	1.1.i. Compare and contrast online data analysis tools	Discussion forum	-fostering awareness to similarities and differences -coping with conflicts, dilemmas and problem situations -identifying and overcoming barriers to students' learning -sharing and revealing self, peer, and student dispositions	-Vicarious experience -Social Persuasion -Physiological and affective state
ESTEEM Foundation Module Part 2				
Read and Watch Essential Materials	1.2.a. Supports for Learning to Do Statistical Investigations	Online page reading with instructional videos	-Fostering awareness to similarities and differences -Identifying and overcoming barriers to students' learning -Sharing and revealing self, peer, and student dispositions	-Mastery experience
	1.2.b. A Guiding Framework for Teaching Statistics	Online page reading with instructional video and expert panel video	-Fostering awareness to similarities and differences -Identifying and overcoming barriers to students' learning -Sharing and revealing self, peer, and student dispositions	-Mastery experience

Table 1 (continued)

	1.2.c. Tasks as Opportunities for Statistical Learning	Online page reading with expert panel video	-Fostering awareness to similarities and differences -Identifying and overcoming barriers to students' learning -Sharing and revealing self, peer, and student dispositions	-Mastery experience
	1.2.d. Read & Watch quiz	Quiz	-Fostering awareness to similarities and differences -Identifying and overcoming barriers to students' learning -Sharing and revealing self, peer, and student dispositions	-Mastery experience -Social Persuasion -Physiological and affective state
Read and Watch Learn from Practice Videos	1.2.e. Expert Teacher Interview on Tools & Resources	Video of expert teachers	-Fostering awareness to similarities and differences -Identifying and overcoming barriers to students' learning -Sharing and revealing self, peer, and student dispositions	-Mastery experience
	1.2.f. Teaching Statistics Using Multiple Technologies	Video of classroom	-learning from the study of practice -selecting and using (appropriate) tools and resources for teaching	-Vicarious experience
Engage with Data	1.2.g. Investigating More Roller Coasters	Statistical investigation assignment	-developing adaptability -fostering awareness to similarities and differences -coping with conflicts, dilemmas and problem situations -selecting and using (appropriate) tools and resources for teaching -Sharing and revealing self, peer, and student dispositions	-Mastery experience -Vicarious experience -Physiological and affective state
	1.2.h. Examining Students' Work on the Roller Coaster Task	Discussion forum with video of classroom	-learning from the study of practice -fostering awareness to similarities and differences -coping with conflicts, dilemmas and problem situations -identifying and overcoming barriers to students' learning -sharing and revealing self, peer, and student dispositions	-Vicarious experience -Social Persuasion -Physiological and affective state
Synthesize and Apply	1.2.i. Supporting Statistical Discourse with the Roller Coaster Task	Reflection assignment with video of classroom	-learning from the study of practice -fostering awareness to similarities and differences -coping with conflicts, dilemmas and problem situations -identifying and overcoming barriers to students' learning -sharing and revealing self, peer, and student dispositions	-Vicarious experience

Table 1 (continued)

	1.2.j. Analyze Tasks and Discuss	Discussion forum with example tasks	-fostering awareness to similarities and differences -coping with conflicts, dilemmas and problem situations -identifying and overcoming barriers to students' learning -sharing and revealing self, peer, and student dispositions	-Mastery experience -Vicarious experience -Social Persuasion
Optional Assignment	Screencast Assignment	Statistical investigation assignment	-developing adaptability -fostering awareness to similarities and differences -coping with conflicts, dilemmas and problem situations -selecting and using (appropriate) tools and resources for teaching -Sharing and revealing self, peer, and student dispositions	-Mastery experience -Physiological and affective state

There are learning opportunities within the ESTEEM module that map onto Bandura's four sources of self-efficacy, and I hypothesized that they would be critical points in participants' STSE development (see Table 1). For example, with regards to mastery experiences, participants engage in two statistical investigations. When considering how to implement these types of tasks in their own classrooms, it could also be viewed as a vicarious experience for mathematics PSTs. With regards to vicarious experiences, there are opportunities for participants to watch video cases of teachers implementing tasks and teaching methods in real classrooms. In addition, discussion forum questions that ask participants to envision themselves enacting a task or teaching method provide an opportunity for training students to implement these types of tasks in their own classroom (Gerges, 2001). Answering reflective questions allows for the vicarious experience of considering how they would implement a task like this in their own classroom or how they would modify an existing task to better it (Morris et al., 2017). Social persuasion may have come in the form of feedback from instructors or appraisal/critique from classmates. Physical and emotional states may have been felt as a result of the materials or assignments.

While I recognized, and wanted to discover, the many other former experiences participants feel have affected their STSE development, the ESTEEM module provided a context which I knew would be informative in terms of mapping the process of development. In addition to two weeks of experiences within the ESTEEM module, participants from Institution A had additional learning experiences related to teaching statistics, specifically, four more weeks of statistics pedagogical content focused on center and variation, bivariate and multivariate concepts, inferential reasoning, and designing technology-based statistics tasks. At Institution B, three weeks were used for the ESTEEM materials, with no other statistics content during the course. In both cases, the optional screencast assignment of a statistical investigation was assigned and submitted after completion of the module. Institution A also included an additional statistics lesson plan assignment which could be completed in pairs, which may have further impacted participants' development STSE through a mastery experience of planning for instruction, emotional affective responses to the experience, and social persuasion as they work with a peer and receive feedback from the instructor.

Participants

After an all-day training workshop, instructors from an initial twelve different universities agreed to implement the ESTEEM project's foundation module, which facilitates an introduction to mathematics PSTs' learning of how to teach statistics. Participants for this research study come from the pool of mathematics PSTs from two of these institutions who have experienced ESTEEM's foundation module. Participants for this study were chosen using a convenience sample. They are PSTs enrolled in undergraduate and graduate programs at two institutions in the United States aimed at preparing mathematics teachers, Institution A and Institution B. These two institutions were chosen because they have opted to participate in the

ESTEEM project and they both offer a similar hybrid course focused on teaching mathematics with technology. For Phase I of the study, all students who taken a course using the ESTEEM foundational module during the Fall 2017 and Spring 2018 semesters and also agreed to participate in research were included. Thus, Phase I included 34 preservice secondary mathematics teachers. Demographic information was not available or collected

For Phase II, all eligible preservice teachers from the two institutions (34 participants) were invited to participate by completing an online autobiographical survey and participating in an interview. Encouraging a large number of participants t allowed for the possibility to gain a better understanding of themes and differences in critical experiences in the development of STSE. For example, participants had varied classroom experience; some mathematics PSTs who had engaged in ESTEEM's foundation module had not yet student taught, and some students who had engaged in ESTEEM's foundation module were completing or had completed student teaching.

My goal was to maximize the size of the collective case during Phase II in order to track pathways of how the development of self-efficacy among students is similar or different, especially within the context of the ESTEEM module. In order to encourage participation, I offered a $40 Amazon gift card incentive for anyone who agreed to and actually completed all required portions of Phase II data collection (i.e. autobiographical survey and interview). Although several reminder emails were sent, only 7 preservice secondary mathematics teachers participated fully in Phase II, 6 from Institution A and 1 from Institution B.

I was an instructor and teaching assistant of two of the courses included. This role made me more intimately knowledgeable about the data, participant experiences in their course, and the general context. In order to prevent bias, all data were blinded prior to analysis. During interviews, my presence may have biased responses, especially while asking participants about experiences they had during my course. I tried to mitigate this issue by reminding participants that their responses had no impact on their performance in the course since it had long been over and assuring them their responses, negative or positive, were going to be used as helpful feedback on improving the education of preservice teachers.

Data Sources

To answer my research questions, I gathered data from each participant in order to develop an understanding of mathematics PSTs' current STSE beliefs and the types of experiences that may have impacted those beliefs. The major data sources were pre- and post-self-efficacy for teaching statsitics survey responses, ESTEEM discussion forums/reflections and assignments, in addition to an autobiographical survey completed online, and a semi-structured interview conducted in person or virtually for Phase II participants. See Table 2 for a summary of data collected.

Phase I Data Sources

Self-Efficacy to Teach Statistics survey. All of the mathematics PSTs who engaged in ESTEEM's foundational module were asked to complete a pre- and post-SETS survey in order for researchers to gain an overall understanding of mathematics PSTs' STSE before and after engaging in the module. These surveys were given before PSTs completed any ESTEEM materials and again as soon after mathematics PSTs completed the materials as possible. The high school version of the instrument (Appendix B) which was used includes 44 Likert-scale items on a scale of 1 through 6 and asks participants to rate their confidence to teach specific statistical topics, with some questions aimed at Level A topics (11 questions), some at Level B (15 questions), and some at Level C (18 questions), according to the GAISE framework

(Harrell-Williams et al., 2014; Harrell-Williams et al., 2019). This version was found to be reliable and valid, after a study with 290 secondary mathematics PSTs from 20 universities across the United States found evidence for the three subscales (Harrell-Williams et al., 2019).

In addition to demographic questions and 44 Likert-scale responses, the SETS survey that was administered in the ESTEEM module also included open-ended responses, asking participants to choose a topic they indicated feeling least and most confident about teaching and to explain why they feel that way. There are a some discrepancies in two of the versions that were administered (both pre and post). The first discrepancy is that one of the versions asks for much more demographic information than the other. The second discrepancy is that one of the versions asks two separate open-ended questions; one for participants to identify the topic they feel least confident to teach and to describe why, and one for participants to identify the topic they feel most confident to teach and to describe why. That is opposed to the other version which asks in general for participants to consider topics most and least confident. The third discrepancy is that the wording of each item in one is less specific than the other. Note the difference between: “Please rate your confidence in teaching students the skills necessary to complete the following tasks successfully...” and “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 high school students the skills necessary to complete the following tasks successfully...” None of these discrepancies should have a significant impact on the results of the survey, with the exception that the qualitative data may be richer for the one that asked separately about topics identified as most and least confident.

Assignments and discussion forums/reflections descriptions. In order to gain insight into mathematics PSTs' STSE and experiences that may have impacted it, several data sources available from the ESTEEM foundational module were analyzed. Assignments and discussion forum and/reflections were submitted by participants' instructors (Appendices C, D, & E). Data submitted from the module included three discussion forums/reflections, one written statistical investigation, and one electronic screencast statistical investigation.

Discussion forum: 1.1.h. Discuss learning statistics through investigations with real data. In this discussion forum, participants are asked to reflect on what they learned in Part 1 of the module about teaching and learning statistics (Appendix C). They are also asked to compare this experience with their prior experiences with statistics. In addition, some instructors may ask participants to respond to peers' discussion forum posts.

Written statistical investigation: 1.2.g. Investigating more roller coasters. In this written assignment, participants investigate a set of 157 roller coasters in the United States (Appendix D). They use the free, web-based software CODAP to explore data and answer investigative questions by comparing different groups of roller coasters and exploring the relationship between different quantitative and qualitative attributes. Participants are asked to respond to questions, insert screenshots, and reflect on their feelings about implementing similar tasks in their own classrooms. While the investigation questions were common across both institutions, Institution B implemented a different method of submission from Institution A. Participants at Institution A submitted a written (e.g. Word or PDF) document, while most (all except one) participants at Institution B submitted a link to a CODAP document.

Discussion forum: 1.2.h. Examining students' work on the roller coaster task. In this discussion forum, participants are asked to watch a video case of different age groups of students exploring the same data set of US roller coasters that they investigated in the module (Appendix C). Participants are asked to reflect on the video by choosing among different prompts that encourage PSTs to compare and contrast the statistical reasoning of the pairs of students in the video, explain how the technology used supported or hindered students' reasoning, and describe how the student-teacher interactions supported student thinking.

Reflection: 1.2.i. Supporting statistical discourse with the roller coaster task. In this reflection, participants are again asked to watch a video case of students exploring the US roller coasters data set, but this time the focus is on how teachers support statistical discourse (Appendix C). Participants are asked to reflect by answering prompts that ask them to provide examples of evidence that the teacher supported students' use of statistical habits of mind, explain how the student-teacher interactions built on student thinking, explain how the teacher sequenced the student presentations, and explain how the teacher used student ideas to help students make connections between the various presented student work.

Screencast statistical investigation. In this electronic screencast assignment (Appendix E) participants choose to investigate a statistical question from the previous written statistical investigation on US roller coasters that they did *not* do already, or they choose to explore an entirely different data set with given prompts. They record themselves using CODAP to investigate the question for a maximum of 10 minutes and share with the instructor the URL to the video hosted on an online platform. At Institution B, it was due at the end of the ESTEEM module and at Institution A, it was due 2.5 weeks after the end of the ESTEEM module.

Phase II Data Sources

Autobiographical survey. In order to gain insight into participants' timeline of their statistics career, critical experiences with statistics, and perception of what statistics is, participants from Phase I who agreed to participate in Phase II of my study, completed an online survey (n=9), consisting of autobiographical open-ended prompts (Appendix F). Questions in this survey ask participants to describe early experiences with statistics as a student in elementary, middle, high, and tertiary school. In addition, participants are asked to describe any pedagogical experiences with statistics. Finally, participants are asked about any informal experiences with statistics.

The purpose of the autobiographical survey was both for me to obtain information on participants' past experiences and also for participants to remember and reflect on past experiences. My experience with one pilot interview informed my understanding that it is hard for PSTs to remember experiences on the spot, so answering questions prior to an interview in an online survey may help them jog their memory in order to be better prepared for face-to-face interviews (Bukor, 2011). In addition, I referenced critical moments that were identified in the autobiographical survey during face-to-face interviews.

Semi-structured interview. To gain direct insight into mathematics PSTs' thoughts on their STSE, I conducted one semi-structured interview per participant, from the participants who agreed to be a part of Phase II of my study and completed the online autobiographical survey (Appendix G). Although 9 participants took the online autobiographical survey, 7 of the 9 responded to multiples emails sent to schedule an interview. A semi-structured interview was chosen to be a rich source of data to understand how and why preservice teachers feel confident or unconfident about teaching statistics (Creswell, 2013). Because we know that high TSE is not *necessarily* a good thing if it overestimates actual ability (Irakleous & Panaoura, 2015), I used the interviews to dive deeper into participants' perspective on assignments and discussion forums/reflections and their expressed TSE in interviews. Thus, I was able to see whether certain types of experiences are related to certain expressed levels of TSE, with the qualitative advantage that PSTs were able to express their perception of experiences. This was true both for experiences in ESTEEM that may have introduced mathematics PSTs to new understandings of statistics pedagogy and STSE, but also for outside experiences. Interviews were either face-to-face (n=1) or conducted virtually with video conferencing software (n=6), and were video recorded and transcribed. They were approximately forty-five minute to an hour in length.

Table 2. Summary of data collection including data sources, collection dates, data analysis, and intended connections to the framework.

Data Source	**n**	**Collection Date**	**Data Analysis**	**Intended Framework Connection**
Phase I				
SETS Quantitative Responses	33	Fall 2017-Spring 2018	Exploratory data analysis and Wilcoxon Signed Rank Sum Test	-Teacher Beliefs (PTE)
SETS Open-Ended Survey Responses	33	Fall 2017-Spring 2018	Checking and modifying themes	-Former Experiences -Factors Affecting Perceived Preparedness -Teacher Beliefs (PTE & GSE)

Table 2 (continued)

Discussion forum: 1.1.h, Discussion forum: 1.2.h., Reflection: 1.2.i.	25-31	Fall 2017-Spring 2018	Re-reading, open coding, several rounds of coding, collapsing into themes	-Former Experiences -Factors Affecting Perceived Preparedness -Teacher Beliefs (PTE & GSE) -During interview: Former Experiences, Factors Affecting Perceived Preparedness, Teacher Beliefs (PTE & GSE)
Written Statistical Investigation: 1.2.g	32	Fall 2017-Spring 2018	Re-reading, open coding, several rounds of coding, collapsing into themes	-Former Experiences -During interview: Former Experiences, Factors Affecting Perceived Preparedness, Teacher Beliefs (PTE & GSE)
Optional: Screencast Statistical Investigation	31	Fall 2017-Spring 2018	Re-reading, open coding, several rounds of coding, collapsing into themes	-Former Experiences -During interview: Former Experiences, Factors Affecting Perceived Preparedness, Teacher Beliefs (PTE & GSE)
Phase II				
Autobiographical Survey	9	Fall 2018-Spring 2019	Re-reading, open coding, several rounds of coding, collapsing into themes	-Former Experiences -Factors Affecting Perceived Preparedness
Semi-structured Interview	7	Fall 2018-Spring 2019	Video recording, transcription, coding, re-reading	-Former Experiences -Factors Affecting Perceived Preparedness -Teacher Beliefs (PTE & GSE)

Data Analysis

Data analysis followed that which is typical for a collective case study and resulted in a description and themes found within the case, as well as important differences found within the case, and finally with an interpretation of what the data means (Creswell, 2013). To analyze data, files were organized so that pre- and post-SETS open-ended responses, discussion forums/reflections, written statistical investigations, links to screencast URLs, autobiographical survey responses, and interview transcripts were chunked together. Each piece of data was blinded, read or watched multiple times, coded multiple times with themes emerging, and a description written on each theme.

After open-coding was completed for large sections of data, a codebook was created using a combination of theory-driven and data-driven codes (DeCuir-Gunby, Marshall, & McCulloch, 2011). Coding was done in sections by "level of meaning" (DeCuir-Gunby et al., 2011, p. 145). Preliminary open-coding was done in order to inform a systematic and rigorous process of several rounds of coding after that. After the preliminary open-coding was complete, the first round of systematic coding was done in order to identify statements expressing beliefs and statements describing experiences. A second round of coding was done to categorize those beliefs and experiences, based on categories found during the open-coding. A third round of coding was done to further categorize statements, and a final round of coding was done to specify the nature of types of beliefs and experiences. For example, after identifying statements determined to be beliefs, I used codes based on the bolded areas in my framework that relate to STSE (Figure 4)—beliefs about the self, beliefs about statistics education, and beliefs about statistics—as broad categories. Thus, for a section that addresses participants' beliefs, on the second round of coding, I labeled that section with a code for beliefs about the self (PTE) or beliefs about statistics education (GSE). On the third round of coding, I used the broad categories determined by preliminary open-coding within each of those larger categories to further describe the types of beliefs expressed (e.g. beliefs about how statistics should be taught, and beliefs about ability to do statistics). On the fourth and final round of coding, I developed in vivo descriptive codes for the specific nature of those beliefs.

Take for example, this discussion forum post on former statistics experiences compared to experiences in the module taken from Discussion Forum: 1.1.h.:

> I did not take statistics before coming to college. It was only offered online (line 1)
> at my high school. Because of this, I completely agree with what the expert (line 2)
> panel was saying about introducing it earlier in schools. It is important for (line 3)
> students to have that mindset of statistics. It is different that math but still very (line 4)
> important. I like to think of it as "real-life numbers" because it is much less (line 5)
> abstract than say integrals. In the second video we watched, the teacher stated (line 6)
> that the students seemed to actually enjoy doing statistics whereas they don't (line 7)
> always enjoy math. I think introducing students to statistics helps them see that (line 8)
> even if they don't like higher level math, math and numbers in general can be (line 9)
> interesting and useful. Also, data and technology are so important these days (line 10)
> that it is impossible to not be dealing with data in some way. I want to make (line 11)
> sure my students have the capacity to deal with that data instead of being afraid (line 12)
> of it. (line 13)

This reflection includes both references to former experiences, as well as beliefs. So, on the first round of coding, it was coded for both beliefs and former experiences. Lines 3-6 express a belief about statistics and how this participant views it. Similarly Lines 8-13 also include some expressed beliefs about statistics teaching. On the second round of coding, focusing on beliefs, it was coded for GSE beliefs; there are no PTE beliefs expressed. On the third round of coding, it was coded for “beliefs about the nature of statistics” and for “beliefs about how statistics should be taught” because this participants expressed beliefs within both of these categories. On the final round of coding, the specific nature of those beliefs were coded; specifically, it was coded as “Because statistics and math are different, they should be taught differently” and “Mathematics and statistics are different.”

I also used the framework to describe *sources* of TSE for statements that related to participants’ former experiences. Codes for former experiences were theory-driven and data-driven. Theory-driven codes included Lee and Lovett’s factors that contribute to perceived preparedness to teach statistics (2017). The process of coding sources or former experiences was similar to that of coding for beliefs. First, a preliminary open-coding method was used in order to gain a sense of the data and the categories of former sources expressed within it. When the formal coding process began, again, the first round of coding was done to identify statements that described former experiences. The second round of coding specified broad categories that had been determined by the open-coding process and from Lee and Lovett’s factors (e.g. experiences as a statistics student, experiences as an informal learner). The third round of coding identified the participants’ judgment of that experience, whether they expressed a feeling that it was a positive experience, a negative experience, or not enough of an experience to make an impact. Finally, the fourth round of coding used in vivo descriptive codes for the specific nature of those experiences. Data-driven codes allowed for other sources of self-efficacy to be expressed by participants that may not have been considered by prior research.

For the quoted example above, the first round of coding identified that former experiences were being described. Lines 1-2 describe a lack of former experiences. In addition, Lines 2-3 and lines 6 -8 refer to experiences the participant just engaged with in the ESTEEM online module. The second round of coding identified two types of former experiences being described: experiences as a statistics students, and experiences within the module. The third round of coding identified the participants’ judgment of those experiences; the module experiences were described positively, whereas the experience as a statistics student was judged as being too few. Finally, the specific nature of those experiences were coded as “Participant has very few former statistics experiences” and “Reference to a video in the module is described.”

In order to validate the coding process, I met with an experienced researcher periodically, who reviewed artifacts and codes to ensure validity of constructs being coded. Once the rigorous coding process was complete, I collapsed and consolidated them in order to develop themes that emerged. Themes were built based off of frequencies and comparing between Phase I and Phase II data. Codes that appeared more than 10 times in module data were then compared to codes in similar categories in interview data. Themes were written to describe a comparison of those most frequently observed codes and how they appeared in the two phases of the study. I provided a rich description of themes in order to describe how mathematics PSTs develop STSE.

Assignments and Discussion Forums/Reflections Analysis

In order to analyze responses from assignments and discussion forums/reflections, I read all responses multiple times in order to gain a general picture of participants’ responses. Since participants are not always explicitly asked about their self-efficacy in these prompts, I was not necessarily able to code teacher beliefs, particular PTE beliefs, unless participants *chose* to write about their STSE. However, they did write about prior experiences, for example, in Discussion Forum/Reflection #1. These experiences were coded first with a preliminary open-coding pass-through, and then according to the rigorous process described above.

In discussion forums/reflections where participants are asked to reflect on video cases, again, they are not explicitly asked about their self-efficacy in these prompts. However, if they chose to discuss the impact of these video cases on their own STSE, those were coded as PTE beliefs. There were very few instances of that. More often, a discussion of the video cases and how they relate to their prior experiences were described, and these experiences were coded according to the coding process described above. Although PTE beliefs were not prevalent in the discussion forums/reflections, GSE beliefs were often expressed. In addition, since these video cases can be considered a form of *vicarious experience* (one of Bandura's 4 sources), they were referenced in the semi-structured interview.

Because the written statistical investigation and the electronic screencast statistical investigation are focused on statistical content, with the exception of a reflection question at the end of the written statistical investigation, they were not coded according to the STSE framework. These two artifacts were useful in that I viewed them as former experiences that could have a significant impact on STSE. Although participants did not discuss beliefs on these assignments, I could infer information about their beliefs from the way they engaged in the assignments. Analysis on the screencasts was done using an entire screencast as a unit of meaning. I watched each screencast and took notes on what features of CODAP were used and to which statistical habits of mind participants were attending. At Institution A, part of the assignment was to identify Common Core mathematical standards that might be addressed by the investigation, and so I also noted whether or not the standards mentioned matched the content of the screencast. After viewing a random selection of three of the screencasts, I open-coded these based on commonalities and differences I deemed were important aspects of the experience. I wrote a description of each screencast and developed a rubric based on that first round of open coding (Appendix H). I then watched all of the screencasts, and after taking notes on each one, I determined a rubric score for both evidence of STSE beliefs and evidence of the extent to which the screencast could be considered a STSE source. A total of 31 screencasts were available for analysis; participants who did not have a screencast either did not submit one for class or removed the video from YouTube before it was downloaded. Of the 31 screencasts, 2 of them could not be fully coded because the audio quality made it impossible to hear the majority of the investigation. Thus, a total of 29 screencasts were fully analyzed.

In order to analyze written statistical investigations (Appendix D), I focused on trying to answer the question: What pedagogical strategies can mathematics PSTs learn from doing their own statistical investigations? This was helpful for evaluating the effectiveness of the ESTEEM materials in providing opportunities for learning pedagogical content knowledge. Although the statistical investigations can also give insight into participants' content knowledge, I was not using this source of data for that purpose; the screencast statistical investigations provided me with a glimpse of that information. Rather, I was interested in analyzing whether the pedagogical purpose behind specific design choices matched what mathematics PSTs actually did. If it did, it provided evidence that they had the opportunity to learn specific pedagogical strategies that might impact their beliefs.

The process of analysis for questions 1 through 4 of the written statistical investigation began by reading through some responses to gain a sense of the data. Because question 5 was of a different nature, a reflection on participant experiences, I treated it separately and coded it along with the discussion forums/reflections data. Questions 1 through 4 were really content focused, whereas question 5 was a reflection asking mathematics PSTs to write about their own feelings about implementing a task like what they did in questions 1 through 4. Based on initial readings of responses to questions 1 through 4, I developed a list of things for which to code for each question on the task. I then coded all participant investigations from Institution A first and created a summary of Institution A participants' responses. Next, I coded all participants investigations from Institution B and modified the code categories for these responses based on the different submission requirements described above. Finally, I created a summary of Institution B participants' responses. While coding was taking place, I also took notes on important characteristics or features I noticed about the data.

Autobiographical Surveys and Interviews Analysis

In order to analyze responses from autobiographical survey responses and interviews, I read all interview transcriptions multiple times in order to gain a general picture of participants' responses. Autobiographical survey responses were embedded verbatim into interview protocols, so analysis of interview transcriptions necessarily included autobiographical responses. I coded each response according to the codebook that had been developed from the analysis of the ESTEEM module data, with the same rigorous four-phase process. During the fourth phase of coding specific beliefs and experiences, I added codes in vivo that had not been documented prior.

One of the purposes of interviews was to see if and how specific experiences in ESTEEM's foundational module impacted both: (a) participants' perceptions of statistics education and the need for implementing tasks similar to those found in the module; and (b) participants' belief in their ability to implement tasks like those. Thus, the interviews were used to gain a sense of participants' PTE beliefs, GSE beliefs, effectiveness of the ESTEEM module, and impact of other experiences. After several rounds of coding, codes from across the ESTEEM data (Phase I) and from the interviews (Phase II) were reorganized, collapsed, and consolidated into themes.

SETS Analysis

Of the 34 participants, 33 had both a pre- and a post-SETS survey response; however, not every response included both a pre- and post-response to the open-ended question(s). For the purposes of this qualitative study, statistical analysis of quantitative responses was used in order to gain insight into mathematics PSTs' STSE before and after engagement in the module. This was done using exploratory data analysis and then formal inferential significance testing. In order to analyze responses from open-ended SETS responses, I read all responses multiple times in order to gain a general picture of participants' responses. Then, I coded each response according to the major themes previously developed from ESTEEM module data and interview data. Codes were confirmed or modified based on the constant comparison method (Glaser, 1965).

Ethical Considerations

This study received Internal Review Board (IRB) approval from North Carolina State University (IRB Protocol 12680). The data collected from the ESTEEM project was also collected under IRB permission (IRB Protocol 7970). Participants from Phase I of this study had already signed a consent form in which they agreed to be a participant in research, which informed them that they may be contacted in the future for an interview. Thus, students who did not consent to participate were not included in the sample and were not invited to participate in Phase II. In order to protect participant identities, all data was stored on a secure password-protected online drive. Code names were used in all data, analysis, and reports, and were created to code institution name, instructor name, course number, semester, and participant name. For example, a student at Institution A taking a course that implemented ESTEEM materials with faculty member 01 for the first section of a course in fall of 2018 might be assigned a code name of InAFac01C01F18St08.

Chapter 4: Secondary Mathematics Preservice Teacher Development of Statistics Teaching Self-Efficacy

Journal

The focus of the *Statistics Education Research Journal* (*SERJ*) is research on the teaching and learning of statistics and probability at all levels, and this article serves to add to that body of literature related to statistics education. This article is written as a "report of original empirical research", which should range from 6,000 to 8,000 words, with a maximum of 10,000 words, not including elements outside the main text. Research that is qualitative in nature is accepted. The *SERJ* audience is broad, including statistics educators, statistics teacher educators, and researchers. The following research article is relevant to all contexts. For statistics educators, research on future teachers' self-efficacy to teach statistics may impact their own teaching practice and the experiences they provide their students. For statistics teacher educators, the same is true, with the additional impact of providing new experiences of the quality and quantity needed to potentially augment or readjust pedagogical beliefs based on their students' past experiences. For researchers, the results of this collective case study fill in a research gap on what we know qualitatively about how statistics teaching self-efficacy is developed and can be reexamined with broader and varied populations.

Abstract

In this study, statistics teaching self-efficacy (TSE) beliefs of secondary mathematics preservice teachers were identified, in addition to factors that impact those beliefs. Student work from a statistics pedagogy module implemented during a hybrid teaching mathematics with technology course was qualitatively analyzed in this collective case study. In addition to data from the course, a subset of voluntary participants completed an autobiographical survey and an

interview in order to provide insight into a lifetime of statistics experiences and their impact on statistics TSE (STSE). Analysis on secondary mathematics preservice teachers' expressed statistics teaching beliefs and factors that impacted those beliefs led to several themes focused on two large categories: beliefs about statistics education and beliefs about personal ability to teach statistics. Findings indicate that secondary mathematics PSTs within this situated context express some desirable general statistics education and personal teaching efficacy beliefs, but that there is still some room for improvement. In addition, it was found that the statistics pedagogy module positively impacted STSE beliefs, but a lack of positive experiences as a statistics student negatively impacted those beliefs.

Introduction

Recent changes in mathematics curriculum in the United States reflect the increasing importance of data analysis and statistical thinking in today's society. With the addition of statistics standards, and the emphasis placed on conceptual understanding of statistics topics, there is a need for teacher preparation in statistics education. However, there is evidence that preservice secondary mathematics teachers feel underprepared to teach statistics topics (Harrell-Williams, Sorto, Pierce, Lesser, & Murphy, 2014; Lovett & Lee, 2017; 2018). This may be partially due to the lack of confidence many teachers report in their own understanding of statistics concepts they are expected to teach (Lovett & Lee, 2017). This lack of confidence is important as it relates to teaching self-efficacy (TSE), which has been found to be related to student outcomes (Tschannen-Moran & Hoy, 2001).

Most researchers now agree that qualitative research on TSE and its sources is needed in order to understand how best to prepare teachers to teach mathematics and statistics. In researching TSE, where measurement issues have long been around, I am complying to the overwhelming call for qualitative research (e.g. Ross & Bruce, 2007; Tschannen-Moran & Hoy, 2007; Wheatley, 2005; Wyatt, 2015). Because statistics is commonly included within mathematics departments and curricula, there is a tendency to treat statistics similar to mathematics with regards to teaching and learning, even though the two topics have important inherent differences (e.g., Rossman, Chance, & Medina, 2006). For this reason, quantitative methods that have traditionally been used to measure beliefs and attitudes towards statistics may actually be confounding perceptions of statistics and of mathematics. For example, high levels of STSE or positive attitudes towards statistics combined with a lack of conceptual understanding of statistics may actually reflect the undesirable perspective that statistics and mathematics are the same, and should be taught and learned in the same way (Hannigan et al., 2013; Lovett & Lee, 2017).

In an effort to end the cycle of teachers and students who are uncomfortable with statistics, it is especially important for teacher educators and for teachers of statistics to be aware of the implications of experiences they provide their students. For teacher educators, it is important to provide opportunities for preservice teachers (PSTs) to reflect on their past experiences with statistics, in terms of their content knowledge, pedagogical content knowledge, use of technological tools, and view of statistics compared to mathematics (Lovett & Lee, 2017), as these all contribute to their perceived preparedness to teach statistics. The development of statistics teaching self-efficacy (STSE) of preservice secondary mathematics teachers is an area that has yet to be fully-explored. I am interested in the different pathways that mathematics preservice teachers (PSTs) may experience in developing their own STSE in general, and more specifically within the context of a specific online intervention aimed at better preparing preservice teachers to teach statistics. However, the online module is just one set of experiences within a larger scheme of a life's worth of experiences. To investigate the larger picture of PSTs' statistics teaching self-efficacy, the research question guiding this study is: What are the general statistics education beliefs and personal teaching efficacy beliefs that comprise preservice mathematics teachers' statistics teaching self-efficacy, and what experiences and factors most impact development of those beliefs?

Background

The increasing use and availability of data in the modern world has led to a new focus on statistics education and the inclusion of statistics standards in curricula. Much of the research on statistics education for teachers has focused on teacher knowledge, which has generally found a lack of teacher knowledge sufficient to teach statistics (e.g. Lovett & Lee, 2017; 2018; Begg, & Edwards, 1999; Hannigan et al., 2013). Although teacher self-efficacy in general has an impact on student success (Tschannen-Moran & Hoy, 2001), research on teachers' confidence or attitudes towards teaching statistics is only beginning to pick up steam (Estrada, Batanero, & Lancaster, 2011).

Some research indicates that advances in self-efficacy to teach statistics are needed. In general, preservice mathematics teachers are less confident to do and to teach statistics than other mathematics topics, like calculus, algebra, and geometry (Hannigan et al., 2013; Lovett & Lee, 2017). Fitzmaurice, Leavy, and Hannigan found 5 reasons why preservice teachers in Ireland viewed statistics as difficult: (a) influence of secondary school experience and teacher; (b) avoidance of statistics as a focus of study; (c) teacher knowledge (their own); (d) perception of difficulty; and (e) influence of teaching practice (2014). They also found that preservice teachers who had taught statistics in their field-based teaching experience had positive reactions to their experience and stated they would like to teach it again. While these results corroborate what is thought to be true about mastery experiences being an important source of high TSE, there should be more said about *how* these PSTs taught statistics and what they view statistics and statistics education to be.

There are some examples that exemplify this need because even when self-efficacy for teaching statistics is high, it does not necessarily correlate with modern views of statistics and how it is different from mathematics (Begg & Edwards, 1999; Irakleous & Panaoura, 2015). For example, inservice and preservice primary teachers who were mostly confident when teaching statistics viewed statistics narrowly, as charts and numbers, rather than as reasoning and investigating (Begg & Edwards, 1999). Most recently, a mixed methods study of mathematics PSTs from a purposeful sample of eighteen U. S. institutions found PSTs' statistical knowledge generally lacking. Mathematics PSTs scored a mean of 69% on a content assessment and struggled with important concepts like p-values, sampling distributions, and variability, despite generally expressing confidence in teaching statistics (Lovett & Lee, 2017).

Despite the lack of a full body of conclusive research on teacher self-efficacy for teaching statistics, the research that has been done indicates room for improvement. It also indicates that while mastery experiences and vicarious experiences can help to improve self-efficacy, weighted differently for inservice versus preservice teachers (Tschannen-Moran & Hoy, 2007; Tschannen-Moran & McMaster, 2009), it is important for researchers to gain an understanding of participants' perspectives on statistics education, since quantitative methods do not give vital information on the latter (Wheatley, 2005). There is evidence that teachers' participation in a course that focuses on statistics education foundational ideas is effective in positively impacting teachers' STSE (Thrasher, Starling, Lovett, Doerr, Lee, 2015). This is tremendously important because TSE has positive impacts on a variety of student outcomes (Tschannen-Moran & Hoy, 2001).

Theoretical Foundation

Teacher affect includes beliefs, attitudes, and emotions (McLeod, 1992). These three affective constructs are related with some overlap, but two main characteristics that distinguish them are their stability and the amount of cognition related to each. McLeod specified four types of teacher beliefs: (a) beliefs about mathematics; (b) beliefs about self; (c) beliefs about mathematics teaching; (d) beliefs about the social context (1992). Within the statistics education context, teacher beliefs include, but are not limited to, beliefs about mathematics, beliefs about statistics, personal teaching statistics self-efficacy beliefs, and general teaching statistics beliefs (Estrada et al., 2011; Gal et al., 1997).

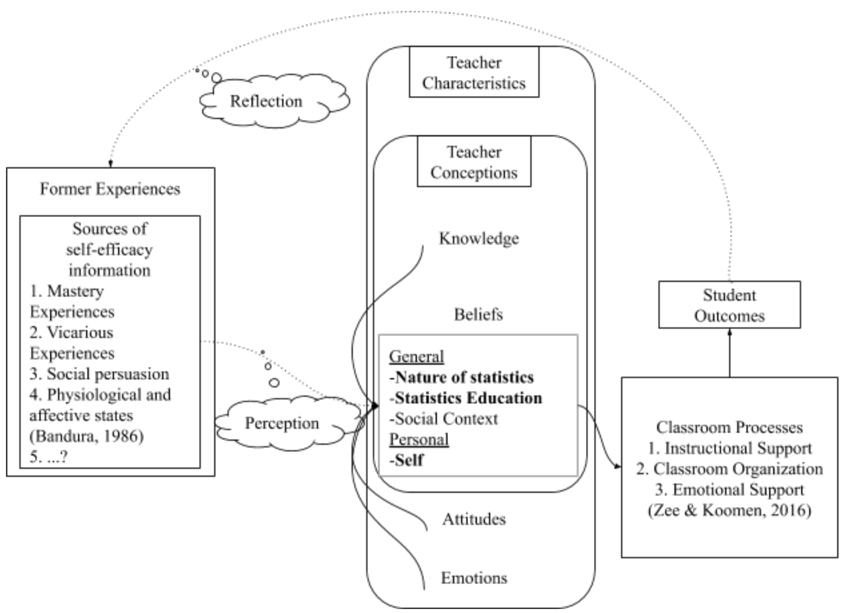


Figure 6. Framework on teacher self-efficacy with a focus on beliefs related to statistics teaching.

The focus of this research is on three types of teacher belief—beliefs about the nature of statistics, the self, and about statistics education. Belief about the self includes personal self-efficacy to teach statistics. Bandura theorized four sources of this type of belief: mastery experiences, vicarious experiences, social persuasion, and physiological and emotional states (1997). It is a teacher's reflection on these four sources, or their perception, that forms a teacher's self-efficacy (Ross & Bruce, 2007).

Because I am interested in preservice teacher STSE development, I focus specifically on the left hand side of the diagram in Figure 6, namely Former Experiences and how perceptions of those experiences form Teacher Beliefs,. For this phase in a teacher's career, they do not have a chance to experience the full cycle on a consistent basis, where their teaching practice and student outcomes affect their TSE development. I have addressed critiques in recent works of the simplistic model that assumes that successes or failures lead to higher or lower TSE respectively, which leads to further successes or failures by explicitly highlighting teachers' reflection on student outcomes. The addition of reflection indicates that perturbations to one's current TSE states caused by, what may be deemed, unsuccessful experiences, can also lead to positive change.

Research Question and Methods

This collective case study is part of a larger study examining the development of mathematics PSTs' STSE and the effectiveness of specifically designed statistics pedagogy online materials.

Context and Participants

This study is situated within the context of a specific intervention aimed at better preparing preservice teachers to teach statistics, which consists of a two-part online module created for the NSF funded grant Enhancing Statistics Teacher Education with E-modules (ESTEEM) project (Due 1625713). It includes opportunities for reading about teaching statistics, engaging in statistical investigations, watching video cases of classrooms, and discussing/reflecting with classmates on statistical learning experiences. However, my focus is not limited to experiences within the module. I am also interested in other experiences mathematics PSTs identify as impactful on TSE development.

I used a qualitative collective case study to understand the experiences mathematics PSTs identify as impactful on their STSE and how those experiences impacted beliefs. Participants for this study came from a convenience sample of PSTs enrolled in undergraduate and graduate programs aimed at preparing mathematics teachers at two institutions in the United States and who took a hybrid course on teaching mathematics with technology that implemented the online module. I define a hybrid course as one that includes both synchronous and asynchronous online components and/or face-to-face meetings. Phase I participants are all students enrolled in two sections from Institution A and one from Institution B, during the 2017-18 year, and who consented to participate in research (n=34). Phase II participants are students who completed a follow-up autobiographical survey and interview after completion of the course (n=7). Even though a $40 gift card incentive was offered for anyone who fully completed the survey and interview, recruitment for Phase II resulted in a small sample. Demographic information was not available or collected.

Because I view the ESTEEM module as a critical aspect of mathematics PSTs' development of STSE, it provides a situated context, which all analysis and interpretations must take into consideration. I purposely chose the ESTEEM context to study mathematics PSTs' STSE development because it provided common learning opportunities which I know all the PSTs would have experienced. Mathematics PSTs would have had a variety of other experiences in terms of primary, secondary and tertiary statistics courses as a learner, statistics teaching experiences, and informal experiences. A subset of participants were interviewed about 8-13 months after the course to inquire about former statistics experiences that may have impacted their STSE. I am making the assumption for all participants that their experiences are varied, even if they are at the same institution.

The ESTEEM foundational module provides mathematics PSTs important experiences with regards to specific aspects of my framework (Figure 6). The module is organized into two major sections, each one consisting of materials to read and watch, a dataset investigation, and discussion forums to synthesize and apply participants' learning, a total of 16 to 18 hours of material. Particularly, mathematics PSTs in the online module are introduced to the Common Online Data Analysis Platform (CODAP) software (https://codap.concord.org), which is a free, web-based educational statistical tool that allows users to explore and visualize data sets.

There are learning opportunities within the ESTEEM module that map onto Bandura's four sources of self-efficacy, and I hypothesized that they would be critical points in participants' STSE development. For example, with regards to mastery experiences, participants engage in two statistical investigations. When considering how to implement these types of tasks in their own classrooms, they could also be viewed as vicarious experiences for mathematics PSTs. With regards to vicarious experiences, there are opportunities for participants to watch video cases of teachers implementing tasks and teaching methods in real classrooms. In addition, discussion forum questions that ask participants to envision themselves enacting a task or teaching method provide an opportunity for training students to implement these types of tasks in their own classroom (Gerges, 2001). Answering reflective questions allows for the vicarious experience of considering how they would implement a task like this in their own classroom or how they would modify an existing task to better it (Morris et al., 2017). Social persuasion may have come in the form of feedback from instructors or appraisal or critique from classmates. Physical and emotional states may have been felt as a result of participants' experiences with the learning materials or assignments.

In addition to two weeks of experiences within the ESTEEM module, participants from Institution A had additional learning experiences related to teaching statistics, specifically, four more weeks of statistics pedagogical content focused on center and variation, bivariate and multivariate concepts, inferential reasoning, and designing technology-based statistics tasks. At Institution B, three weeks were used for the ESTEEM materials, with no other statistics content during the course. Institution A also included a statistics lesson plan assignment, which may have impacted participants' STSE.

Data Sources

Phase I of data collection included selected student work from the online module. For this study, data used for analysis from Phase I consisted of a pre- and post-quantitative survey measuring statistics teaching self-efficacy (n=33), open-ended responses from that survey (n=33), and 3 discussion forums/reflection assignments (n=25-31). For Phase II, data consisted of autobiographical survey responses (n=9) and transcripts of the interviews (n=7).

SETS survey. All of the mathematics PSTs who engaged in ESTEEM's foundational module were asked to complete a pre- and post-Self-Efficacy to Teach Statistics (SETS) survey (Harrell-Williams, Sorto, Pierce, Lesser, & Murphy, 2014; Harrell-Williams et al., 2019) in order to gain an overall understanding of mathematics PSTs' STSE before and after engaging in the module. These surveys were given before PSTs completed any ESTEEM materials and again as soon after PSTs completed the materials as possible. The high school version of the instrument which was used includes 44 Likert-scale items on a scale of 1 through 6. The SETS instrument (Harrell-Williams et al., 2014; Harrell-Williams et al., 2019) asks participants to rate their confidence to teach specific statistical topics, with questions aimed at 3 different levels of difficulty (Level A: 11 items, Level B: 15 items, and Level C: 18 items). The SETS survey also included open-ended responses, asking participants to choose a topic they indicated feeling least and most confident about teaching and to explain why they feel that way.

Discussion forums/reflections. Several times during their engagement in the module, participants were asked to reflect either in discussion forums or in submitted assignments on their experiences or react to their learning from the materials. For example, they were asked to reflect on what they learned in Part 1 of the module about teaching and learning statistics and to compare this experience with their prior experiences with statistics. As part of a statistical investigation on roller coaster data they were asked to complete, participants reflected on what makes them nervous or excited to implement tasks like the one they just did. In another example, participants were asked to watch a video case of different age groups of students exploring the same data set of US roller coasters that they investigated in the module and to reflect on the video by choosing among different prompts that encourage mathematics PSTs to compare and contrast the statistical reasoning of the pairs of students in the video, explain how the technology used supported or hindered students' reasoning, and describe how the student-teacher interactions supported student thinking. Finally, participants were again asked to watch a video case of students exploring the US roller coasters data set, but this time the focus of the assignment was on how a teacher can support statistical discourse and orchestrate classroom-level discussions of students' work.

Autobiographical survey and interview. In order to gain insight into participants' timeline of their statistics career, critical experiences with statistics, and perception of what statistics is, participants from Phase I who agreed to participate in Phase II of my study, completed an online survey (n=9), consisting of autobiographical open-ended prompts. Questions in this survey ask participants to describe early experiences with statistics as a student in elementary, middle, high, and tertiary school. In addition, participants were asked to describe any pedagogical experiences with statistics. Finally, participants were asked about any informal experiences with statistics. The interviews aimed to dive deeper into participants' perspective on assignments and discussion forums/reflections and their expressed TSE in interviews. Interviews were either face-to-face (n=1) or conducted virtually (n=6), and were video recorded and transcribed. They were approximately forty-five minute to an hour in length.

Quantitative Data Analysis

In order to analyze quantitative data from the pre- and post-SETS surveys, total scores were computed by summing the ratings given for each item across all 3 levels and dividing by the number of items. Sub-scores were computed by summing the ratings within each of the 3 levels and dividing by the number of items within each of the 3 levels. Total scores and sub-scores from both the pre- and post-SETS survey results were matched and mean gain scores computed for each participant, followed by an informal analysis of the differences, and finally a Wilcoxon Signed Rank Sum Test was performed to test for differences.

Qualitative Data Analysis

In order to analyze qualitative data, all data was read multiple times, with notes taken on researcher noticings. Then, systematic coding was done in several rounds, and a codebook was created using a combination of theory-driven and data-driven codes (DeCuir-Gunby, Marshall, & McCulloch, 2011). Coding was done in sections by "level of meaning" (DeCuir-Gunby et al., 2011, p. 145).

On the first round of coding of qualitative data, beliefs expressed and experiences described were identified and coded as such. Several rounds of coding followed in order to specify particular types of beliefs and experiences. For example, a second round of coding beliefs that had been coded as such further identified whether they were personal teaching efficacy beliefs (PTE), about the self, or about statistics education in general (GSE).Then, a third round of coding beliefs specified major categories within those two broad categories. For example, GSE beliefs might have been categorized as beliefs about the nature of statistics or as beliefs about how statistics should be taught. On the fourth and final round, I developed in vivo descriptive codes for the specific nature of those categorized beliefs, for example: "Because statistics and math are different, they should be taught differently". A similar method was used for coding participants' former experiences

Codes were collapsed and consolidated in order to develop themes. This collapsing and consolidating process was done by comparing most frequently expressed beliefs and most frequently described experiences in the modules to those expressed and described in the interviews. I checked the validity of these themes by using them to code open-ended responses from the pre- and post-SETS surveys, in order to see if the themes were consistent across multiple data sources. Finally, I provided a rich description of each theme in order to describe how mathematics PSTs develop STSE.

Results

In examining the types of experiences critical to the development of statistics teaching self-efficacy, I first analyzed quantitative data from the SETS survey that participants took prior to and immediately after engaging in the two-week module. I then qualitatively explored their beliefs during and much after the online learning experience and was able to observe if the beliefs they expressed in the modules lasted past their involvement in the course. Thus, I examined not only participants' statistics teaching self-efficacy and what had been instrumental in impacting it, but also more specifically what they were referring to when they spoke about statistics, statistics education, and good teaching in general.

Changes in Self-Efficacy to Teach Statistics Within a Course

In viewing a distribution of participants' overall mean gain scores on the SETS survey, we see that the majority (77%) of participants either had no change (n=1) or a positive change (n=22) in STSE (Figure 7). However, 23% (n=7) of the participants reported a lower mean score after their engagement in the online module.

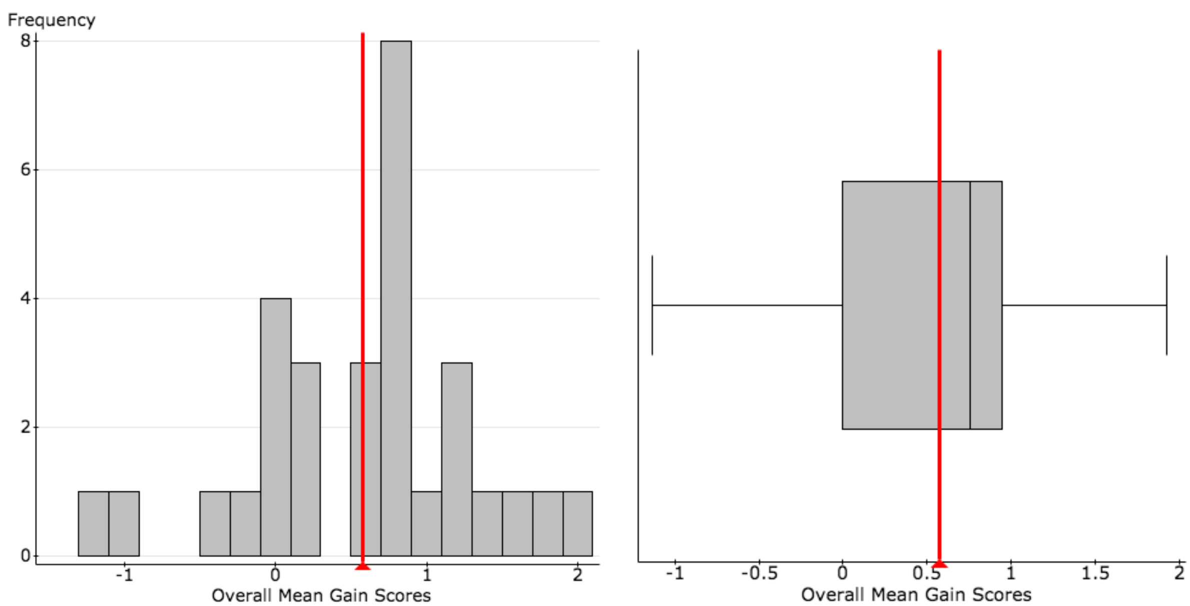


Figure 7. Distribution of mean overall gain scores from the SETS survey.

Table 3 contains statistical measures for overall SETS scores and each subscale. The average gain score was 0.58 with a standard deviation of 0.74. The middle 50% of the gain scores fell between 0.03 and 0.92. Based on this informal analysis, engagement in the module increased overall STSE scores between around 0.5 and 0.9 point on average, on a 6 point scale, but there *were* some individuals who experienced a decrease in average STSE. Formal nonparametric hypothesis testing confirm these results. A Wilcoxon Signed Rank Sum Test was used to test whether the median of the overall gain scores is greater than zero. Results confirmed a statistically significant trend towards positive gain in self-efficacy to teach statistics overall (Table 4).

For each level A, B, and C, we saw similar changes to those seen in overall STSE (Table 3). Again, there were some cases at each level where after participation in the module, lower average SETS scores were reported, but on average, participants increased in their confidence for each sub-scale. These results are confirmed by using a Wilcoxon Signed Rank Sum Test for each sub-score (Table 4).

Table 3. Summary of SETS Statistics

		Pre	Post	Gain
Overall	N	32	31	30
	Mean	3.95	4.63	0.58
	Median	3.86	4.75	0.76
	Standard Deviation	1.07	0.99	0.74
Level A	N	33	32	33
	Mean	4.28	4.96	0.53
	Median	4.20	5.00	0.45
	Standard Deviation	1.05	0.85	1.19
Level B	N	32	33	31
	Mean	4.01	4.66	0.62
	Median	4.03	4.67	0.73
	Standard Deviation	1.12	1.00	0.92
Level C	N	33	32	32
	Mean	3.60	4.35	0.74
	Median	3.56	4.50	0.72
	Standard Deviation	1.21	1.14	0.95

Table 4. Summary of Wilcoxon Signed Rank Sum Test Results.

	Test Statistic	p-Value
Overall	372.00	0.0004
Level A	383.00	0.0010
Level B	363.50	0.0008
Level C	410.50	0.0001

Quantitative results from the SETS survey show overall improvements in STSE, up to a full point increase in self-efficacy scores in some cases. This is true for overall scores, in addition to scores at each of the three increasing levels of statistical thinking.

The responses from the open-ended questions on the SETS survey were used to confirm or modify themes found within the other qualitative data sources. In order to delve more deeply into how and why self-efficacy to teach statistics may have increased or decreased, qualitative results are described below.

Expressed General Beliefs about Statistics Education

Beliefs expressed by participants about general statistics education, included the following:

- Beliefs about good teaching will persist over time, but beliefs about good statistics teaching are difficult to shift in one course;
- Mathematics PSTs believe in the usefulness of using educational statistical tools in teaching statistics, despite fears expressed about the possible drawbacks of using them in a classroom;
- Mathematics PSTs are not quite sure what statistics is, but they do believe it is distinct from mathematics.

Each of these themes are more fully described below.

Beliefs about good teaching will persist over time, but beliefs about good statistics teaching are difficult to shift in one course. In general, mathematics PSTs hold the belief that teaching practices compatible with the general reform movement, like using tasks and discourse for discovery learning, are effective. These beliefs were expressed in the course context, and also persisted till the time of the interviews. Within the module, one very frequently expressed idea was that the role of a teacher is one of a facilitator and that the role of a student is to explore, guided by the teacher. This belief was expressed, for example, when participants were reacting to videos of a classroom and commenting positively on the types of guiding questions the teacher in the video asked of students. They also responded positively to the open-ended nature of the task students in the video were engaging in and also to the open-ended nature of a similar task in which they themselves engaged in the module. A representative example of this belief was as follows:

> The role of the students in the video was to explore different statistics related to roller coasters and try to discover interesting relationships between the variables. The teacher's role was support the students exploration with probing questions and suggestions to further the understanding of CODAP and the different aspects of roller coasters. The interactions between the students and the teacher helped the students gain a better statistical understanding of what they were looking at, when they came up against some feature of the tool they didn't understand or some information they didn't quite understand, the teacher was there to help them over the hump and move them along the cycle of statistical reasoning.

In interviews, four of the 7 participants continued to express this belief of the teacher role as facilitator. However, even more detailed and nuanced general beliefs about teaching were also expressed in the interviews. For example, four interviewees expressed a desire to stay away from traditional practices like a focus on memorizing and procedures, direct instructions, definitions, and notes in their own teaching. Four of them also expressed a belief in using task-based instruction using discovery activities focused on concepts. The idea of good teaching being relevant to students or fun was expressed by several, but it also came out in module discussions and reflections. Finally, several interviewees also expressed a belief in focusing on student thinking. In responding to a question on the ways by which teachers impact students' learning in mathematics classrooms, Interview participant 2 responded:

> Every way. I think that recently there's been a big shift in teachers transferring knowledge to students versus helping facilitate student discovery of knowledge. And I think that's great because that almost does take some of the role off of us because it's more like we're teaching them techniques to learn, instead of just giving them the information, and I think that's really great. So I think that in some ways we are becoming less responsible for their learning by helping them learn how to do it on their own, helping them critically think on their own, and kind of facilitating. This is what we're trying to do at least. Give them the tools to learn on their own, especially in like high school math.

While student-centered, task-based, concepts-focused instruction is certainly applicable in statistics education, there are also reform practices that are specific to effective statistics instruction. Mathematics PSTs very frequently expressed some of these key ideas specific to statistics education in the context of the online module, like the ideas that (a) statistics and mathematics should be taught differently, (b) it is important to encourage statistical reasoning in students by building levels of understanding, (c) using or collecting real, large, messy interesting data is vital, (e) engaging students in the statistical investigative cycle is important, and (e) students should develop statistical habits of mind.

These ideas were present in the learning material in the module in a variety of ways. The first, that statistics and mathematics should be taught differently, mainly appeared in the module's very first resource, an online reading page on the differences between statistics and mathematics. However, the idea persisted throughout, with readings, videos, and a framework that expounded upon what makes statistics teaching unique. In the second part of the module, an online reading page presents a framework for teaching statistics, which incorporates the statistical investigative cycle, levels of students understanding at each phase of the cycle, and statistical habits of mind to attend to at each phase. Before this formal framework, PSTs were introduced to the cycle and habits of mind through a reading and an instructional video, in addition to seeing the cycle in action in an actual classroom and hearing an expert panel of professors discussing it in a video. This expert panel also discussed the importance of having students use or collect real and relevant, messy data. While some beliefs that mathematics PSTs expressed in the course context that were compatible with the statistics education reform movement persisted till the time of the interviews, many key ideas did not.

The major idea that persisted in the interviews was the idea that, in teaching statistics, teachers should ask students to explore and dive deep into understanding data. Again, 4 of the 7 expressed this idea in interviews. For example, Participant 1 expresses:

> So when I've gotten to make my ideal statistics lesson plans, I loved it. because I was going through and coming up with data that most students would probably be excited about. So either roller coasters or letting them choose data, those were my big two data sets, or big two data ideas that I went with normally so...And that was really fun because it was them noticing things of interest and finding patterns within the data.

Participants in interviews also mentioned the importance of choosing real and relevant data (3/7). What is more notable is the fact that only one participant mentioned the idea of developing students' statistical habits of mind and none of them explicitly mentioned involving students in the statistical investigative cycle. A few interviewees did implicitly invoke one of the statistical habits of mind by expressing the belief that it is important to encourage students to be skeptical of statistics. Thus, there is evidence that suggests the ideas of data being at the center of learning and teaching statistics and being skeptical about how data is collected or presented persisted over time, while explicit beliefs about how statistics and mathematics should be taught differently, focusing on the statistical investigative cycle, building up levels of understanding, and developing statistical habits of mind like uncertainty and variability, did not.

Mathematics PSTs believe in the usefulness of using educational statistical tools in teaching statistics, despite fears expressed about the possible drawbacks of using them in a classroom. Because participants were situated in a course context that was focused on teaching mathematics with technology, beliefs about technology in the module were frequently expressed and the focus of many prompts and forum posts. Overall, within the module, participants' views of teaching statistics using educational statistical tools like CODAP were overwhelmingly positive. And, although some hesitations were expressed in the module about implementing technology into a statistics lesson, positive views on the software's usefulness persisted in the interviews. There was only 1 negative reference to personally using technology in the module, compared to 26 positive references, with participants being excited not only to use the tool themselves, but also to use it with students in the future:

> One thing that excites me about implementing a similar investigation in my own classroom is that students can work with data sets that are large, real, and on topics that interest them (like roller coasters). Also, the features of CODAP allow students to pose and analyze their own questions, as they can quickly use tables and graphs to examine the relationships between multiple variables.

As demonstrated by this example, in particular, participants were most excited about CODAP's unique ability to provide easy-to-create interactive graphical displays, which they viewed as helpful for developing students' statistical reasoning. They also greatly appreciated the ability to graph multiple variables, which they also viewed as helpful for developing students' statistical reasoning.

Some fears that participants expressed in the online module regarding the use of technology to teach statistics were that it would take students too much time to learn the tool or that it would be too hard for them to learn. However, these fears were only expressed by one PST during the interviews. Five of the interviewees generally expressed their belief that it is beneficial to use technology when teaching statistics, and they also sometimes specified instances in which technology might be especially useful, like in taking random samples, in doing linear regressions, in playing around with data, and in visualizing or organizing big data. In addition, five of them specifically mentioned that learning how to use CODAP in the online module was useful.

The fact that mathematics PSTs' positive experiences endured time is especially impressive given that it was rare for them to see technology like CODAP in statistics classes; only 1 of the 7 interviewed saw it in in their own learning in a college-level statistics classroom, and in that case it was used by the teacher as a demonstration. Technology identified by participants as being used in their statistics courses included Statcrunch, Excel, online applets, SAS, and SPSS. Some participants viewed these technologies as useful (4/7), but the rest did not. Participant 13 says about learning how to use Excel: "I feel confident in that now I had not previously used Excel for anything except for like doing rosters and stuff like that. So being able to use it to do math was really cool." Contrast that to what Participant 14 says responding to whether or not she learned anything about teaching statistics from the tools her statistics instructors used to teach:

> ...she would use I think it was SPSS and she would demonstrate how like, oh this is what's happening in this concept, like look at it on, with this data set. This is what we're doing. But I only remember that a few times. I thought it was, I thought it was like more applicable to like how you would actually use statistics in like maybe a job. But yeah I mean, I don't, I don't think so.

In addition to conflicting ideas about the technology tools used in their own statistics classes, 3 of the mathematics PSTs adamantly made the point that their undergraduate statistics programming class was not helpful at all in terms of their learning of statistics content or how to teach it:

> I was able to kind of like use my programming skills of looking up how to use different code to do the statistics and so that became more of the focus rather than building statistics knowledge. I just approached it as I need to code this, what is, what is the code, what is the key word that I need to use. (Participant 11)

Thus, despite the lack of continued exposure to educational statistical tools in their statistics courses, the major idea of technology being useful when teaching statistics stuck with participants.

Mathematics PSTs are not quite sure what statistics is, but they do believe it is distinct from mathematics. As discussed earlier, participants in the module frequently mentioned the fact that statistics should be taught differently than mathematics, and that was coupled with a frequently expressed belief that mathematics and statistics are different. This expressed belief does not come as a surprise, since one of the major overarching ideas of the online module is that statistics and mathematics are different in their essence. The idea that statistics is different than math or that it is a more interesting or different type of math was expressed by five of the participants in interviews, so that belief persisted.

However, what came out from explicit questioning during the interviews are two additional beliefs that did not emerge in the module. The first is a clear distinction between statistics and probability, mentioned by 6 of the participants. The second is a perceived lack of clarity about what statistics is and what statistics content is covered in K-12 curriculum.

Although probability underlies many statistics concepts, mathematics PSTs differentiated the two topics, and some expressed a preference or comfort with probability over statistics (4/7). They were able to identify a convincing reason for this preference; that learning probability is much like learning mathematics, which they are comfortable with. In addition, they had much longer experiences they identified to be about probability that they could reference, including informal experiences with board games, and fun early childhood classroom activities like skittle counts.

When participants in interviews were asked how they define statistics, 5 mentioned it to be the study of real world data. They often also included in their descriptions specific statistical terms and concepts they associate with statistics like mean, median, mode, distributions, boxplots, z-score, p-values, residuals, confidence intervals, tests, and standard deviation. However, these descriptions were also coupled with statements about how they are not quite sure what statistics is (5/7) or what is covered by curriculum standards they would soon be teaching (6/7). In fact, when asked about what they would change about their teacher preparation programs, 4 explicitly mentioned including more time to cover statistics standards, understanding they mean, and how to teach them. This lack of confidence in knowing exactly what statistics is was coupled frequently with an explanation that they had been exposed to other mathematics topics for the majority of their lifetime, compared with very late and infrequent exposure to statistics. As Participant 11 shared:

> I'm still hazy about statistics like I still just, I have not been taught it enough I don't feel like. I don't feel very prepared to teach it in the classroom, like I think I could maybe do one lesson on statistics if it's the topic that I know, which is basically like, I don't know the CODAP things with like calculating probabilities and things like that. But outside of that I'm just like, I don't really get it still. So I think that's partially because I, I only had statistics in college so I never learned it in high school at all. And then in college it was like a class of a bunch of people, and I just didn't really understand what was going on. And then all my education classes they like, there's a section about statistics like teaching statistics but it's not like the whole thing is about teaching statistics so…

This sense of "haziness" due to lack of sufficient experiences was shared by some of the other participants. However, as we will see later, these general statistics beliefs conflict with some of the personal STSE beliefs and experiences they describe related to their participation in the online module.

Expressed Personal Beliefs about Statistics Teaching Self-Efficacy

In the second category, beliefs about personal statistics teaching self-efficacy included beliefs that participants did not feel comfortable teaching statistics yet, but also that they were confident in certain aspects of teaching statistics. In addition to personal teaching self-efficacy beliefs, participants also described and attributed certain experiences that impacted current self-efficacy levels; for example, participants frequently attributed a lack of confidence to a lack of enough statistics experiences, mostly vicarious, but also mastery. However, participants also attribute aspects of high self-confidence to their positive experiences with the online learning experiences, which included mastery, vicarious, socially persuasive, and affective experiences.

The two major themes related to personal beliefs about statistics teaching self-efficacy include:

- Overall, experiences in the module were helpful for positively impacting mathematics PSTs STSE; however, they were not viewed as *enough* because of a lack of quality experiences with statistics as a student and as a teacher.
- Although mathematics PSTs generally express a lack of confidence in their ability to teach statistics overall, they do express confidence in key aspects of teaching statistics.

Each theme is elaborated upon in what follows.

Overall, experiences in the module were helpful for positively impacting mathematics PSTs STSE; however, they were not viewed as *enough* because of a lack of quality experiences with statistics as a student and as a teacher. References to the online ESTEEM module were generally positive in terms of impacting STSE, both in the module and during the interviews. In fact, in the discussion forums and reflections analyzed, there was only one negative experience described in relation to using technology. All other references to experiences within the module were positive, including references to the module in general, references to specific articles, references to the roller coaster investigation, references to using technology, and references to videos within the module. References to experiences within the module overwhelmingly occurred in a discussion forum in the first half of the module, in which PSTs reflected on their past experiences with statistics compared to their experiences in the module. Of those references, they overwhelmingly spoke about videos within the module, and specifically focused on one video where an expert panel of four statistics educators discussed various topics, like what it means to do statistics and why it is important to teach statistics. One particularly strong response to the video read: “After completing this Module, I feel like my opinion regarding Statistics and introducing Statistics to my students has changed. I strongly believe that the expert panel video took my negative view of Statistics and changed it into something positive and useful.” This is also an example of one of the very few times personal self-efficacy beliefs were expressed in the module; when these personal beliefs were expressed, the majority were expressing confidence or developing confidence due to the module.

When participants were interviewed many months after their online course experience, they still had positive views of the statistics pedagogy module. In fact, all 7 expressed at least one positive experience from the module, whether it was about when they recorded their computer screens while they were doing a statistical investigation (n=4), learned how to use CODAP (n=5), completed a written statistical investigation on a data set of US roller coasters (n=5), or watched videos of students in real classrooms (n=5). Participants not only identified those experiences as positive, but also 6 of the 7 directly expressed a positive impact on confidence due to those experiences. For example, Participant 15 expressed:

> I remember in the [teaching mathematics with technology] class we did roller coaster data so that was really cool, finding things that are, is readily available, car data, I did a lot of car data stuff, stuff like that I think was definitely very helpful and made me feel better that I'd be able to find stuff that's relevant in a high school classroom.

In fact, in all the interviews, there were only two negative references about the roller coaster written investigation. One participant did not view the content as advanced enough to be taught in high school. Another thought the assignment page had too much information and was intimidating, but still found the experience to be helpful and positive overall.

Despite the fact that mathematics PSTs were positively impacted by the online module, the experiences within the module were not viewed as enough for them to feel completely confident to teach statistics. Again, participants view their experiences with statistics in K-12 settings, in college courses, and in teacher preparation as so brief compared to their extensive entire life of experiences with mathematics that even the experiences they do have are not viewed as enough. Participant 11 expresses:

> When I think about how statistics teaching has been introduced in those classes, it almost seems like kind of like a, like a fleeting idea where if like, because we talk a lot about linear equations and it's... a lot of the content that we’ve been taught how to teach is a lot of like Math 1 [the first in a 3-part series of integrated high school-level math courses] and so, but even though statistics is in Math 1, I feel like we haven't really like, talked about how to teach statistics in a more detailed way compared to like other content that we've talked about.

In general, mathematics PSTs expressed a lack of confidence in doing and teaching statistics; only one of the 7 interviewed participants expressed confidence in her ability to teach statistics. Five of the seven interviewees expressed that they are uncomfortable with statistics content and do not know it well enough to teach it. However, when a lack of confidence in doing and teaching statistics was expressed, it was most frequently attributed to the lack of experiences with statistics as a student or as a teacher. Even when mathematics PSTs did have statistics experiences upon which to draw, they were either viewed as negative experiences, or as not enough, especially compared to the extensive mathematics experiences they have had, as discussed earlier in relation to the uncertainty over what statistics actually entails.

In reflections and discussions during the module, the majority of references to experiences prior to the module were described either as negative or as being very few. Most often, negative experiences as a statistics student were described as being too similar to a traditional mathematics course that was focused on memorizing and procedures. For example, one student connected with a classmate about a negative past experience:

> I agree with your past experiences in statistics. The class I took in college was strictly lecture and did not involve any real data. I think that giving students data without context, takes away some of the more enjoyable tasks, such as going and collecting data, which is what statistics is all about and what makes statistics different than other mathematics topics. Can we really say we are learning/teaching statistics without real data?

We see here that the negative experience was due to the teaching style and the lack of real data used. The lack of using real data or focusing on context was commonly described. The negative experiences were often blamed for lack of knowledge of concepts, for example, "I remember just accepting the data at face value and performing procedures like excluding outliers without any idea of why this was done."

Due to the nature of the reflection questions in the module, negative experiences discussed in the module were most often compared with experiences in the module. Thus, descriptions of former experiences as a statistics student often included references to pedagogical ideas discussed in the module, like context, real data, technology, and the statistical investigative cycle. However, when former experiences as a statistics student were discussed in interviews, they did not necessarily include those same ideas. Negative experiences described in interviews by five of the participants were related to general teaching principles not specific to statistics, like a lack of understanding concepts and just getting through the class without really learning anything due to traditional direct instruction teaching methods.

Of the five interviewees who did not take a statistics course in high school, all five of them identified the lack of a statistics course as impacting their lack of confidence in their ability to do and to teach statistics. Of the two who *did* take a statistics course in high school, one viewed the experience as negative due to the teacher. Five of the participants attributed the lack of seeing it being taught in their fieldwork observations; four of the participants also attributed their lack of experience teaching it. This lack of experience in an actual high school classroom, either in a mastery or vicarious type of context, negatively impacted STSE because they did not know “what it should look like”:

> ...because I'm sort of in that transition level of I didn't really learn statistics all throughout high school so I don't know what that looks like as a teacher, and I don't know what, like, bad teaching statistics in high school looks like. I don't know what good teaching statistics in high school looks like. Like, I just don't really have any idea. Like with my other math courses in high school I, I had a lot of bad math teachers, so I kind of got this knowledge of, like, this is what I don't want to do whenever I'm teaching this class. But with statistics I just, I haven’t really seen much. (Interview Participant 4)

The lack of vicarious experience, either as a statistics student in high school or as an observer or guest teacher greatly negatively impacted mathematics PSTs’ STSE. This need to see statistics being taught in a secondary classroom is also related to their belief that they do not quite know what statistics is or what it should look like in a secondary classroom.

In addition to identifying lack of high school experiences as a statistics student or as a preservice teacher fieldwork visitor as negatively impacting their STSE, mathematics PSTs also identified a lack of positive college experiences as a statistics student. When mathematics PSTs referred to “few” experiences in college, it ranged from one to three courses, but even the mathematics PST who took three courses viewed that as not enough, confirming the trend that a lack of vicarious teaching experiences or mastery statistics experiences were majorly impactful on STSE.

Although mathematics PSTs generally express a lack of confidence in their ability to teach statistics overall, they do express confidence in key aspects of teaching statistics. As described, above, mathematics PSTs generally express a lack of confidence in their ability to teach statistics. However, qualitative analysis allowed for a more nuanced perspective on their STSE; rather than a yes or no answer to the question of STSE, my research found that mathematics PSTs are confident in *certain aspects* of teaching statistics. Specifically, they express confidence in their ability (a) to teach particular statistics topics, (b) to take the time to learn particular statistics topics that they would then be able to teach, and (c) to draw on experiences and contexts that make statistics content relevant to their students.

When interviewed participants were asked about their confidence to teach a particular Common Core high school statistics standard on interpreting differences in shape, center, spread, and accounting for outliers, 6 of the 7 expressed confidence to teach that standard. Although participants expressed a lack of confidence in their knowledge of what the statistics standards are, and how to teach them, discussed earlier, they expressed confidence when they were presented with a *specific* standard that addresses topics that were a major focus of their experiences in the online module. PSTs had opportunities in the module to analyze and interpret distributions of different groups in a written statistical investigation on roller coasters, and in a recorded screencast investigation on either roller coasters or another data set, in addition to watching a video of an expert describe distributions of children per women by comparing data for different countries. Participants did not express the same amount of confidence when presented with another statistics standard that was about using data from a random sample to draw inferences and generating multiple samples to gauge variation in estimates. For participants at Institution A, they were exposed to simulated samples during other pedagogical learning experiences, but evidently those experiences were not effective in positively impacting their confidence in teaching that standard.

In post-SETS open-ended responses, participants also expressed confidence in statistics related to specific topics; many expressed being most confident in teaching topics most related to algebra or calculations, like linear regressions, reasoning that they were exposed to those ideas for the longest amount of time. However, many participants' post-SETS open-ended responses (14/31) identified statistical topics they were most confident to teach that were explicitly emphasized in the module, like interpreting data, using technology to choose and display appropriate graphical representations of data, and engage in the investigative cycle, and oftentimes cited the module as the reason for that confidence. For example:

> Using the cycle of statistics (Pose question, collect data, analyze, interpret) and using statistical habits of mind (sampling, context, uncertainty, etc.) to find real, meaningful answers from data with reasonable confidence. These were major themes of the Esteem unit that I just completed.

This response is an example of participants highlighting aspects from the module that impacted what they were most confident to teach.

When asked in general about their confidence to teach statistics or about particular standards, all 7 interview participants expressed a confidence in their ability to learn material they could then teach:

> I do not feel like I'm very good at statistics. I don't know if that's true, like I've taken a statistics class and I did fine, but like I don't know if I could go up to someone right now and explain them certain topics in statistics, especially I don't really know what you learn in high school with statistics necessarily. So I'm like my confidence is mid to like lower but I know that I could teach myself. I'm confident that I could reteach myself like certain skills I'm going to teach someone else. (Participant 14)

This belief was also evident in some of the open-ended SETS responses, with two participants expressing it in their pre-survey, and 4 students expressing it in their post-survey. In their explanations of specific topics they indicated having low confidence to teach, these participants said they had forgotten those topics, and that they would have to relearn them before teaching them. Thus, even in their expressions of a lack of confidence, there was an accompanying *different* type of confidence, which may be beneficial or necessary for future teachers to have in order to cope with the overwhelming amount of knowledge they need to have before beginning to teach – the personal confidence in being able to learn content as needed in a way appropriate for being able to teach it.

Finally, 5 of the 7 expressed confidence in being able to use prior informal experiences with real-world contexts, like board games and daily risk assessments, in their statistics lessons to make them relevant to students. For example, Participant 1 cited her familiarity with roller coasters from originally wanting to be a roller coaster engineer as a source of confidence in being able to use extensive knowledge and personal experiences with roller coasters in order to implement an activity like a roller coaster investigation. Participant 15 described an experience where her child was moving around and eating, and choked on her food – she identified situations like that where risk is involved and certain actions lead to certain events being more likely as a source of confidence in being able to draw on those contexts in teaching statistics. Participant 11's comfort with board games and video games helps her feel confident in using ideas of probability that are built into those games in her own understanding of probability, but also a context "I can reference if I get anxious about teaching a probability lesson."

Summary and Conclusion

In summary, mathematics PSTs within this situated context at two institutions, were positively impacted by their experiences within a two-week online module focused on preparing them to teach statistics. This was evidenced by quantitative results from the SETS survey, and qualitative results from discussion forums, reflections, open-ended SETS responses, and interviews. However, the impacts of their self-efficacy were not comprehensive enough to combat a general expressed lack of confidence to teach statistics. Quantitative results indicate that self-efficacy improved directly after engagement with the online module, and qualitative analysis supports the impact on certain topics within statistics. For example, mathematics PSTs indicated confidence to teach a specific standard in the Common Core curriculum standards, which was a major focus of the online module, much after their engagement in the module. In addition, their confidence in certain topics was more fully described by probing their general beliefs about statistics education – for example, that statistics is different from mathematics and that it should be taught differently. However, these positive results are tempered by long-lasting uneasiness due to uncertainty about what statistics is and lack of mastery and vicarious experiences related to teaching statistics to build on.

This research study has described some important aspects of the development of secondary mathematics PSTs' STSE, confirming prior theory and research, in addition to providing evidence for new recommendations for teacher preparation. In terms of the experiences that are most important in impacting teaching self-efficacy, Bandura theorized that mastery experiences would be the most impactful (Bandura, 1997). Empirical evidence has supported this claim (Tschannen-Moran & Hoy, 2007; Tschannen-Moran & McMaster, 2009), and the mathematics PSTs in this study reconfirm their lack of experiences teaching statistics as one factor negatively impacting their STSE. However, researchers have found that vicarious experiences impact novice teachers more (Tschannen-Moran & Hoy, 2007) and the results of this study has provided some insight into the effects of those experiences. The mathematics PSTs in this study describe their lack of vicarious experiences, having rarely seen statistics taught in a high school, as majorly impacting their lack of knowledge of what statistics is and what it should look like in a classroom. In addition, videos of classrooms where statistics was being taught was frequently cited as an experience that was useful in positively impacting STSE. The importance of vicarious and mastery experiences is related to prior research that identified pedagogical knowledge as an important factor in mathematics PSTs' perceived preparedness to teach statistics (Lovett & Lee, 2017).

To reflect the expressed importance of various types of experiences on mathematics PSTs' STSE, I present a modified framework that provides more details related to the development of statistics teaching self-efficacy (Figure 8). In this framework, the types of experiences are ordered in importance and are described in more detail as to the different ways PSTs might have these experiences.

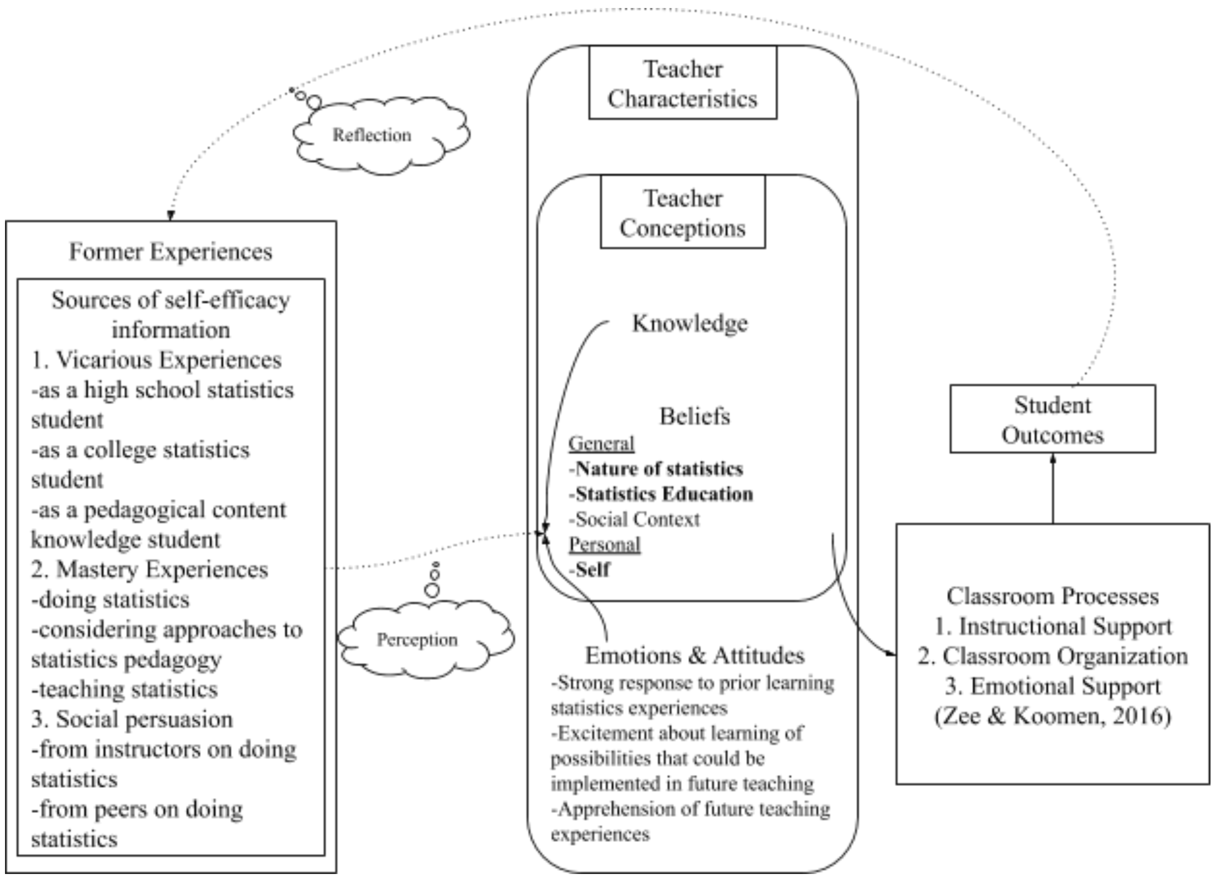


Figure 8. Modified framework on teaching self-efficacy, reflecting importance of impact as expressed by mathematics PSTs.

The mathematics PSTs in this study also both confirm and disconfirm Smith's hypothesis that the mathematics reform movement makes it harder for teachers to develop self-efficacy when they are no longer expected to directly pass on information to students through traditional methods (1996). PSTs in fact did express a general discomfort with teaching statistics, which they attributed to their lack of being able to see what good statistics teaching should look like; that might mirror PSTs' desire to see general reform methods in action, especially if they have not seen such methods in their own experiences as a student (Smith, 1996). However, PSTs did express confidence in teaching certain topics that they experienced in their teacher preparation program (e.g., interpreting differences in shape, center, spread, and accounting for outliers).

Statistics knowledge has previously been identified as a factor that contributes to mathematics PSTs' perceived ability to teach statistics (Lovett & Lee, 2017). They also expressed confidence in their ability to figure out how to teach topics. This ability to learn content in order to teach it has been seen in prior research (Irakleous &Panaoura, 2015).

These results are also an important complement to existing empirical work that provides evidence that mastery experiences are impactful for STSE; researchers have called for qualitative research on mathematics PSTs' views on statistics and statistics education in conjunction to their personal self-efficacy beliefs (Fitzmaurice et al., 2014, Wheatley, 2005). What the mathematics PSTs in this study demonstrated is that some course experiences can positively impact views on statistics and statistics education; however, some of those important beliefs may not last long if they are not reinforced throughout teacher preparation programs. These PSTs demonstrated some lasting desirable beliefs, like the belief that using educational statistical tools in teaching is useful, the belief that statistics and mathematics are different, and the belief that is important to use real, relevant data and approach data with a certain amount of skepticism.

In terms of the framework, participants expressed both general beliefs about statistics education and personal statistics teaching efficacy beliefs, and very frequently attributed certain sources to those beliefs. Sources that were most commonly explicitly referenced were vicarious experiences, as a high school statistics student, as a college statistics student, and as a pedagogical content knowledge student. Mastery experiences were also commonly referenced, like experiences of doing statistics, considering approaches to statistics pedagogy, and teaching statistics. Social persuasion experiences were explicitly brought up during interviews, in the context of receiving feedback from instructors on statistics lesson plans; participants did not remember feedback or think it was impactful. However, social persuasion experiences may have also occurred implicitly within the online module context when participants were discussing with each other, and possibility disagreeing or offering different perspectives. In addition, physiological or affective states were most often discussed *as a result of* a mastery or vicarious experience, rather than as a source of self-efficacy. For example, participants may have expressed a hatred or a love of statistics due to an experience in a college-level statistics course. To reflect those conclusions, the modified framework (Figure 8) lists social persuasion last in importance, and also specifies some types of social persuasion that do not include feedback from instructors on activities like lesson planning. More research should be done on various types of social persuasion and their impact on mathematics PSTs' STSE, like feedback from students on statistics teaching. In addition, the modified framework does not list physiological or affective states as a former experience, but rather, those can be found within "attitudes" and "emotions" in the resulting teacher characteristics.

I initially included room in my initial framework for potential other sources of self-efficacy. Researchers in the past have disagreed on where to place "knowledge"; some argue it is a result of mastery experiences, while some argue it should be considered a source. In the online module context that participants experienced, there were learning experiences in the form of readings, a framework, and instructional videos that could lead to mastery experiences, but cannot be considered a mastery experience of doing statistics or teaching statistics, or a vicarious experience of teaching statistics. One might argue that readings and instructional videos written or recorded by expert statistics educators could be categorized as social persuasion; the experience of learning about statistics education from experts is impactful because of the position of the author. However, in teaching self-efficacy research, social persuasion usually refers to personal feedback about one's own performance, not learning general pedagogical information. As evidenced by participant reflections, readings on the statistical investigative cycle and videos of an expert panel were particularly impactful on beliefs. Because the ideas found in those experiences did not necessarily match with participants' prior mastery experiences of doing statistics or vicarious experiences of teaching statistics, they provided an opportunity for PSTs to pause and reflect, and possibly, change their beliefs, either general or personal. However, because reactions to those types of learning experiences were rarely expressions of personal efficacy beliefs, in contrast to experiences like doing a statistical investigation or watching a video of a real classroom, they may in fact be more impactful on general statistics education beliefs than they are on personal teaching efficacy beliefs. The newly revised framework is useful in the field of statistics teacher education in that it highlights specific types of experiences that are most impactful for preservice teachers, implying suggestions for possible ways of adapting teacher education programs to provides those types of experiences.

Some recommendations for teacher preparation programs and statistics programs, partly based on the desires expressed by mathematics PSTs in this study, follow. There is a need for dispersed statistical pedagogical experiences so that desirable ideas about teaching and learning statistics persist like they do for desirable general reform teaching practices. However, even if there is limited time in a teacher preparation program for including statistics pedagogical content, there is evidence that even a two-week online module is positively impactful on mathematics PSTs' STSE, both general and personal. Specifically, the focus on data, mathematics and statistics being different, being skeptical about how data is collected or presented, and using educational technology to teach statistics were points of improvement. As we know from the participants in this study, much of what makes PSTs uncomfortable with statistics is their lack of familiarity with it. Teacher preparation programs should require mathematics PSTs to observe statistics lessons being taught in real secondary classrooms, and provide them with tools and knowledge to reflect on the effectiveness of those lessons they observe. In addition, programs should focus on providing mathematics PSTs with opportunities to use educational statistical tools to explore data, and to reflect on those experiences, like was done in the online module, because there is evidence that those types of opportunities were particularly impactful.

In addition, there is a need in statistics courses to implement teaching practices with the benefit of both modeling teaching for future teachers, and also making statistical concepts clear to learners (Franklin et al., 2015). For mathematics PSTs who have very few pedagogical experiences with statistics, much of their knowledge of what statistics is and how to teach statistics comes from the very few statistics courses they take, either in high school or in college. Statistics courses that model teaching using the investigative cycle, building up student levels of understanding, and emphasizing statistical habits of mind will reinforce desirable ideas about statistics education that mathematics PSTs encounter in their teacher preparation program. Otherwise, those ideas will become muddled and further contribute to the hazy lack of clarity that exists among PSTs on what statistics is.

Although the mathematics PSTs in this study were positively impacted by the online module in general, there are opportunities for improvement. For example, personal efficacy beliefs were rarely expressed in the context of discussion forums and reflections. In order to invoke more expressions, and to possibly change, personal efficacy beliefs, discussion prompts could be revised to explicitly ask PSTs to build their confidence for teaching statistics. For example, when viewing video cases of real students, prompts could ask PSTs to pinpoint aspects of the video that might be particularly helpful to them in developing their own teaching strategies for teaching statistics.

Chapter 5: Examining the Effectiveness of Online Materials Designed for Impacting Statistics Teaching Self-Efficacy

Journal

The focus of the *Journal of Mathematics Teacher Education* (*JMTE*) journal is research on the improvement of the education of mathematics teachers throughout all stages. This article addresses the use of online materials in preparing future mathematics teachers to teach statistics. It is written as a research paper, which can be about initiatives and programs to be used in teacher education. The word limit should be 10,000 words, not including references. Research that is qualitative in nature is accepted. The *JMTE* audience is broadly teacher educators and education researchers. The following research article is relevant to mathematics teacher educators who may not know the benefits of or have access to resources to help prepare mathematics teachers to teach statistics. For researchers, the results of this study report the effectiveness of expertly designed free, online materials that can be accessed and used on a wide-scale with teachers.

Abstract

In this collective case study, the effectiveness of online teacher education materials expertly designed as an introduction to pedagogical issues in teaching secondary level statistics is examined. Specific activities in the materials are examined for the opportunities they provide mathematics preservice teachers (PSTs) in developing factors that impact their perceived preparedness to teach statistics, particularly, their view of statistics, statistics knowledge, pedagogical knowledge, and use of technology. Student data from two institutions that implemented the online module were analyzed for evidence of the impact of those opportunities. Students were preservice secondary mathematics teachers, and their experiences in the module

included engaging in statistical investigations and analyzing and reflecting on teacher practice and student thinking through video cases. Findings indicate that due to the impact of the learning opportunities, participants displayed the beginning of a shift towards viewing statistics as distinct from mathematics, increasing knowledge of pedagogical strategies for teaching statistics, and developing knowledge of using educational technology tools to teach statistics.
Keywords: statistics, teacher education, online

Introduction

Teacher education has the unique opportunity and responsibility of preparing mathematics teachers who are responsible for creating the next generation of statistically literate citizens (Franklin et al., 2015). Our world is saturated with data, and it is increasingly important for professionals in all fields to be able to make data-based decisions, in addition to everyday citizens to be able to critically process information from media and peers (Kwasny, 2015). Thus, students need to be provided experiences for building up statistical thinking, which means that mathematics teachers need to have access to resources and the ability to implement those types of experiences. It is critically important for teacher educators, then, to address gaps in statistical pedagogical content knowledge and to effectively prepare future mathematics teachers to be competent and confident to teach statistics (Franklin et al., 2015).

In most K-12 settings, statistical concepts are expected to be taught in mathematics classrooms. Common Core Statistics standards begin in Grade 6 and go all the way until Grade 12 (National Governors Association Center for Best Practice & Council of Chief State School Officers, 2010). In addition, high schools in the U.S. often offer an Advanced Placement Statistics course that can earn high school students college credit for an introductory statistics course. Throughout these contexts, some of the challenges for both teachers and students are a lack of adequate teacher preparation, a lack of student time to delve into statistics, and a lack of outreach from statisticians to students (Kwasny, 2015).

Statistics is uniquely situated in teacher education because it is a distinct subject from mathematics, but usually taught within mathematics departments, where it can be treated like mathematics (Horton, 2015). And so mathematics teacher educators have added obstacles in fully preparing mathematics PSTs to teach it, since PSTs' experiences and ideas with statistics are usually limited and often mathematical in nature, rather than statistical. PSTs' statistics teaching self-efficacy, or their beliefs in their ability to teach statistics, in addition to their beliefs about what statistics is and how it should be taught, are likely very much related to the ways they will teach their students (Ernest, 1989; Wilkins, 2008; Eichler, 2008). Thus, it is important for teacher preparation programs to develop favorable ideas about statistics education and to provide PSTs experiences that positively impact their perspectives on their ability to teach statistics.

In this collective case study, one effort at providing teacher educators with a comprehensive introduction to teaching statistics is described and examined. This effort consists of an expertly designed two-part online module that is freely accessible to teacher education faculty and instructors, created by an NSF funded grant Enhancing Statistics Teacher Education with E-modules (ESTEEM) project (Due 1625713). By using student work data from two universities that implemented the online module, I answer the research question: How do statistics pedagogy learning experiences impact preservice mathematics teachers' statistics teaching self-efficacy?

Background and Theoretical Perspective

In this section, I will summarize research and principles of teacher education generally, and then for statistics teacher education and online mathematics teacher education specifically. I will then present the teaching self-efficacy framework used to guide this study.

Goals and Practices of Mathematics Teacher Education

PSTs must learn a variety of skills and knowledge bases in order to be prepared for teaching in their specific contexts. For example, they should have an appropriate amount of content knowledge in order to teach content, but they also should build specialized knowledge required for teaching, called pedagogical content knowledge (PCK) (Shulman, 1986). Pedagogical content knowledge, or mathematical knowledge for teaching, includes knowledge of content and students, knowledge of content and teaching, and knowledge of curriculum (Ball, Thames, & Phelps, 2008). PSTs build this knowledge in a variety of settings; specifically in their undergraduate or graduate preparation, this most often occurs through experiences that include both content courses and methods for teaching courses, and during fieldwork experiences that include real classroom observations and teaching assignments.

Those responsible for the complex task of preparing future teachers to teach mathematics are mathematics teacher educators (MTE). However, because contexts, programs, requirements, and audiences differ so greatly, it may be difficult for MTEs to access research on best practices (Cochran-Smith & Villegas, 2015). Goals of MTEs differ depending on context, especially because there is no common curriculum or set of resources upon which MTEs can draw. Some researchers (Hiebert, Morris, & Glass, 2003) argue it is not feasible for mathematics PSTs to learn how to be an effective teacher without yet being a teacher. Hiebert et al. propose a model for math teacher preparation which focuses on two goals of preparing teachers to teach: (1)

become proficient in mathematical pedagogical content knowledge, and (2) develop knowledge, competencies, and dispositions to learn to teach from teaching. These goals could be met by treating lessons as experiments, which means that PSTs should learn to pose research questions about their lessons, collect data from their lessons, and analyze and interpret that data in order to improve future practice.mSo, it is important for MTEs to include opportunities for mathematics PSTs to shift their focus from teachers' performance to student thinking and to reconsider mathematical ideas in new ways (Hiebert, Morris, Berk, & Jansen, 2007). Ideas on how this can be done include using video cases of real classrooms, by creation or engagement in mathematical tasks, and reflection on mathematical pedagogical opportunities of those experiences (Hiebert et al., 2007; Hiebert et al., 2003; Zaslavsky, 2008).

Overall, video cases of real classrooms in teacher preparation is increasing in use in all contexts and are potentially powerful for increasing levels of reform education, but effectiveness depends on the facilitation of analysis (Gaudin & Chaliès, 2015). In order to successfully implement the viewing of video cases, MTEs must be conscious of implementation decisions- what videos they choose, how they decide to orchestrate discussion about the video, and how they react in the moment to types of comments made by teachers, requiring a "heightened listening" (Coles, 2014). There is evidence that PSTs who are exposed to videos of student-teacher interactions during coursework and asked to critically reflect on those videos include more student input in their own teaching and attend more to student thinking when reflecting on their own teaching, as opposed to focusing on themselves (Santagata & Yeh, 2014). PSTs who do not have those videocase experiences during coursework tend to conduct more teacher-centered teaching, as well as focus on themselves in their teaching analysis (Santagata & Yeh, 2014). A framework by Wilson, Lee, and Hollebrands (2011) identifies the process preservice teachers use when attending to student thinking via video cases – describing, comparing, inferring, and restructuring, which can be helpful for MTEs in their efforts to guide preservice teachers to restructuring their own thinking based on analysis of video cases.

In addition to the positive impact that critical viewing, reflection, and discussion of video cases of real classrooms can have, PSTs' engagement in or creation of mathematical tasks can also have a positive impact on PCK learning. "...unlike tasks for students, a mathematical task for teachers rarely deals with just the mathematics. It can be seen as an opportunity to generalize from it to a large class of tasks, and to deal with many other aspects of teaching mathematics as well" (Zaslavsky, 2008, p. 112). Thus, the two practices of analyzing video cases and engaging in mathematical tasks, when carefully implemented by MTEs, have the potential to greatly impact PSTs' learning.

Statistics Teacher Education

For preparation of mathematics teachers to teach statistics, MTEs have the unique challenge of especially perturbing existing ideas about statistics and lack of sufficient content knowledge. Thus, it is even more vital for MTEs to incorporate practices like those above for statistical pedagogical content knowledge. The American Statistical Association has released several documents for assessment and instruction in statistics education—one for the preK-12 level (Franklin et al., 2007), one for the undergraduate level (American Statistical Association, 2005; 2016), and one for teacher preparation programs (Franklin et al., 2015). These frameworks emphasize the importance of student understanding of statistical concepts, such as distribution, variation, and informal inference. The *Guidelines for Assessment and Instruction in Statistics Education (GAISE)* report's framework defines three increasing levels of sophistication (Level A, Level B, and Level C) of statistics understanding for different phases of the investigative cycle (Franklin et al., 2007). The phases are 1) Formulate Question, 2) Collect Data, 3) Analyze Data, and 4) Interpret Results. The levels are not associated with an age or grade level, but rather assess students' current level of thinking. Related to this is an extension of the GAISE framework, the Framework for Supporting Students' Approaches to Statistical Investigations (SASI), which identifies specific statistical concepts and productive statistical habits of mind to which students should attend at different phases of the investigative cycle (Lee & Tran, 2015).

The fact that statistics is taught within mathematics classrooms can be problematic, since statistics and mathematics are very different subjects (Scheaffer, 2006; Rossman, Chance, & Medina, 2006). This difference may not be widely recognized by mathematics teachers, or mathematics teacher educators. The beginning of knowledge for effective statistics teaching is the knowledge that statistics is inherently different from mathematics (Hannigan, Gill, & Leavy, 2013). To successfully teach statistics, educators should be aware of how the approach to teaching, modeling, and applying statistical concepts is different from mathematics. Tasks that require students to find solutions to problems by going through the entire investigative cycle accustom them to the approach statisticians use to solve "messy" data analysis problems (Wild & Pfannkuch, 1999). Contrast that to a mathematical task, where deductive reasoning is used to answer a mathematical question; the task conclusion is expressed with certainty given the assumptions of the problem and is proven for every case using logic. Statistics (mathematics) teacher educators should use tasks that perturb potentially existing ideas about mathematical approaches to statistics.

Because of these issues, it is important for MTEs to take on the unique task of preparing mathematics teachers to teach statistics (Franklin et al., 2015). Two successful programs implemented in Portugal connected teacher education in statistics with practice, one with preservice teachers and one with in-service teachers, by providing participants with opportunities of (1) collaborating with fellow teachers to create effective tasks, (2) implementing those tasks, and (3) reflecting on how they went (Ponte, 2011). Similarly, a workshop focused on the topic of variation increased and deepened teacher knowledge of the types of sources of variation in data for many of the teachers (Arnold, 2008). The reasons teachers gave behind the effectiveness of the professional development were that they actually did the activities and that they did them within a learning community. In addition to the characteristics of effective professional development for statistics teachers in the examples above, much research has been done on the positive effects of using technology to teach statistics to teachers and students (e.g. Hammerman & Rubin, 2004; Lee et al., 2014; Prodromou & Pratt, 2006). For example, teachers using two widely-known educational statistical tools, *TinkerPlots* and *Fathom*, were able to solve data analysis tasks using strategies uniquely-related to technology, like linking representations, and augmenting data by adding dividers or reference lines (Lee et al., 2014). Unfortunately, some preservice high school teachers have reported a lack of exposure to using statistical software in learning or teaching statistics (Harrell-Williams, et al., 2014).

Online Mathematics Teacher Education

Mirroring the general increase in online instruction being offered in educational settings, teacher education is also increasingly being offered online. Specifically in teacher education, most recent web-based models of online learning are attractive for providing continuous learning to a wider audience, opportunities to view different types of instructional practice, solutions to teachers' content knowledge needs, access to curriculum supports, spaces for reflection and dialogue, and models for good instructional practice (Burns, 2011). In one design-based research study, the authors implemented and revised online mathematics and science teacher preparation courses at their university over a long span of time based on data from the courses and interviews in order to increase the quality of the social, teaching, and cognitive presences of the courses (Niess & Gillow-Wiles, 2013). Their research resulted in a learning trajectory for teachers' knowledge for teaching with technology which finds that in an online asynchronous environment, a supportive community structure, purposeful instructor actions, technology, and shared content knowledge are essential. Some best practices identified include purposefully scaffolded content, collaborative work, inquiry-based activities, and opportunities for reflection.

Thus, while best practices for teaching online may resemble those for teaching face-to-face, it is important for instructors to design learning opportunities that incorporate those practices differently, given constraints of an online space. Constraints of an online course are unique and provide unique challenges for MTEs in terms of conveying "listening" to students, building rapport with participants, and being constrained by the limitations of the technology tool (Kastberg, Lynch-Davis, & D'Ambrosio, 2014). However, despite constraints of online learning, using the internet to teach also affords unique opportunities in terms of collaboration, multimodality, and performance (Borba, Clarkson, & Gadanidis, 2013). For example, asynchronous collaboration requires that *all* students participate in order to be counted as present, which is not the case in face-to-face settings. In addition, asynchronous collaboration encourages a less teacher-centered approach since students will likely be responding to each other more often, or more quickly, than an instructor would. Related to unique ways of collaboration is the opportunities for multimodality; online students have access to multiple ways of viewing content, through text, audio, video, images, graphs, applets, and interactive PDFs (Borba et al., 2013). Related to collaboration and multimodality is the opportunity for online spaces to be spaces for performance, in which students can develop ways of expressing ideas in innovative, creative, and artistic ways (Borba et al., 2013).

I hypothesize that engaging online with expertly designed opportunities that allow for increasing knowledge of statistics and PCK for statistics and incorporate educational statistical tools will lead to an improvement in statistics teaching self-efficacy. The format of the learning experiences at the focus of this study was at least partially implemented online, but I do not specifically address the implications of the nature of the online experience.

Theoretical Framework

The increasing use and availability of data in the modern world has led to a new focus on statistics education and the inclusion of statistics standards in curricula. Much of the research on statistics education for teachers has focused on teacher knowledge, which has generally found a lack of teacher knowledge sufficient to teach statistics (e.g. Lovett & Lee, 2018; Begg, & Edwards, 1999; Hannigan et al., 2013). In addition to knowledge, teacher self-efficacy in general has an impact on a large variety of student outcomes (Tschannen-Moran & Hoy, 2001).

Most recently, a mixed methods study of a purposeful sample of mathematics PSTs from eighteen institutions around the United States explored PSTs' statistics teaching self-efficacy (TSE) compared to other mathematics topics and factors that influence their TSE (Lovett & Lee, 2017; 2018). The researchers also used a content assessment to measure PSTs' statistical knowledge, and found that it was generally lacking, with mathematics PSTs scoring a mean of 69% and struggling with such important concepts like p-values, sampling distributions, and variability. Although results from the Self-Efficacy to Teach Statistics (SETS) survey (Harrell-Williams, Sorto, Pierce, Lesser, & Murphy, 2014; Harrell-Williams et al., 2019) showed PSTs to be confident to teach statistics, they ranked it lower than other topics taught in high school

(Lovett & Lee, 2017). Further, based on open-ended responses to the SETS instrument and interviews, Lovett and Lee described factors that contribute to PSTs' perceived preparedness to teach statistics: (a) role of statistics knowledge; (b) role of pedagogical knowledge; (c) impact of using technology; and (d) view of statistics.

Statistics knowledge, or lack thereof, was cited as a factor by PSTs' to explain both why they felt confident or unconfident to teach a particular topic in statistics, and came up much more often than the other three factors, which were mostly cited as reasons why PSTs were confident (Lovett & Lee, 2017). For pedagogical knowledge, PSTs discussed the importance of having knowledge of strategies to teach a particular topic and common student conceptions about a statistics topic; when they lacked this knowledge for certain topics, they indicated not being highly confident to teach them. With regards to technology, they expressed comfort with particular computation tools and wrote about experiences with technology that led to deeper understanding of concepts. PSTs expressed confidence with particular topics by expressing a view that certain statistics topics were procedures or by relating topics to algebra, which most of them ranked higher in terms of TSE. They also discussed discomfort with teaching statistics because of its uncertainty. For all three of these factors, PSTs were more often citing them as reasons they were most confident to teach a statistics topic, but sometimes citing them as reasons they were least confident to teach a statistics topic.

These four factors identified by PSTs as impacting their confidence to teach statistics occur within experiences as a student and as a teacher. Bandura theorized four sources of self-efficacy beliefs: mastery experiences, vicarious experiences, social persuasion, and physiological and emotional states (1977). Morris, Usher, and Chen describe in more detail what these four sources mean in the teaching context (2017). Enactive mastery experiences are experiences teachers have with actual teaching tasks, like developing and implementing lessons that result in student outcomes, positive or negative. Vicarious experiences occur when a teacher experiences another teacher's mastery experience; teachers must be able to relate to other teachers in some way for vicarious experiences to be influential. Social persuasions come in the form of encouragement, feedback, or assessments from peers or supervisors about one's teaching abilities. Finally, physiological and affective states are physical and emotional reactions during teaching experiences, like a fast heart-rate, sweaty hands, or anticipation.

It is a teacher's reflection on these four sources, or their perception, that forms a teacher's self-efficacy (Ross & Bruce, 2007). Despite the lack of a full body of conclusive research on teacher self-efficacy for teaching statistics, the research that has been done indicates room for improvement. It also indicates that while mastery experiences and vicarious experiences can help to improve self-efficacy, weighted differently for inservice versus preservice teachers (Tschannen-Moran & Hoy, 2007; Tschannen-Moran & McMaster, 2009), it is important for researchers to gain a qualitative understanding of participants' perspectives on statistics education, since quantitative methods do not give vital information on the latter (Wheatley, 2005).

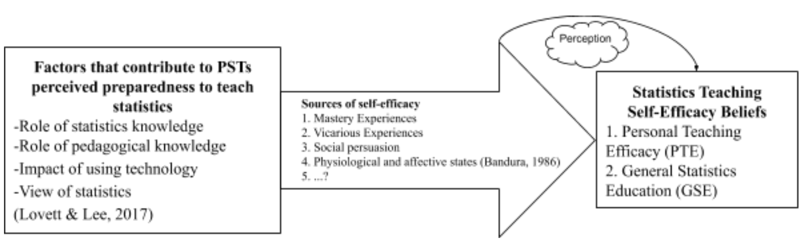


Figure 9. Part of the statistics TSE development cycle, in which teachers factors contribute to perceived preparedness through different types of experiences.

In the framework above (Figure 9), we see that teachers (both preservice and inservice) identify *factors* that contribute to statistics teaching self-efficacy (STSE) through *experiences* that can be considered sources of TSE. For example, teachers identify statistics knowledge, perhaps gained through mastery experiences, as important in the development of their perceived preparedness to teach statistics. This statistics knowledge impacts STSE belief as it happens through an experience that occurs, for example, in a formal statistics course. Another example could be that a teacher identifies her comfort with using a graphing calculator to create linear regression models as an important factor in her STSE, but that comfort developed as a result of a professional development workshop she went to where the facilitator demonstrated how to do that skill. Thus the factor of using technology impacts her STSE *through* a former experience. During and after teachers' experiences, their interpretation of whether experiences positively or negatively impacted their knowledge, for example, leads to a change in STSE. Note that STSE beliefs include both beliefs about the self (PTE) and beliefs about general statistics education (GSE). This distinction is important because when PSTs express a certain level of confidence in their ability to teach statistics, their view of what statistics is and how it should be taught underlies that belief. In courses specifically focused on preparing mathematics teachers, all of the factors identified that most impact STSE could occur during experiences like engagement in statistical tasks, video case analyses of students and teachers in classrooms, peer discussions, lesson planning assignments, and peer teaching assignments.

Online Learning Context

The ESTEEM project aims to fill a gap in materials for mathematics teacher educators to better prepare secondary preservice mathematics teachers to teach statistics. The project has created three online modules, each with two parts. The first module in the collection (the Foundational module) provides mathematics PSTs important experiences aimed at preparing them to teach statistics through opportunities to think and reflect on the learning and teaching of statistics in a way that may be unfamiliar. Indeed one of the major intents is to offer a view of statistics that emphasizes its uniqueness from mathematics and its importance in secondary curriculum. The module is organized into two major sections, each one consisting of materials to read and watch, a technology-enabled data investigation, and discussion forums to synthesize and apply participants' learning, for a total of 16 to 18 hours of instructional material (Table *5*). Participants engage in two statistical investigations of relatively large multivariate datasets on roller coasters in the United States, using the free, web-based software Common Online Data Analysis Platform (CODAP), which allows for various types of graphical representations, displays statistical measures, links multiple representations of data, among other features. They also watch video cases of teachers implementing tasks like the roller coaster investigation and teaching methods in real classrooms. In addition, participants interact in discussion forums and respond in assignments to reflective questions in order to analyze teacher-student interactions or to consider their own experiences and plans for implementing similar pedagogical techniques.

Design Principles for Teacher Learning Materials

During the design of the ESTEEM materials, specific design principles were adhered to in order to ensure effectiveness, to reflect a core set of values, and to achieve consistency. In their 2018 paper, Hudson et al. described the essential design elements for the online Foundational module for statistics teacher education. The design principles for content include: (a) use of free and accessible tools for data investigation, (b) investigation of engaging, real, larger, and messy data sets, (c) connection to classroom practice through videos (Hudson et al., 2018), and (d) use of a framework about the statistical investigative cycle, statistical habits of mind and levels of student understanding. These design principles are informed by prior research on teaching in an online context (Niess & Gillow-Wiles, 2013) and by recommendations for statistics teacher education (Franklin et al., 2015). The design principles for delivery include: (a) common organization and structure for units, (b) opportunities for meaningful interactions with an instructor and other secondary mathematics PSTs through discussion, (c) keeping readings and videos brief, and (d) readings, videos, activities, and discussions to vary the way secondary mathematics PSTs have an opportunity to think about statistics pedagogy (Hudson et al., 2018). These design principles echo literature on the unique affordances of online education (Borba et al., 2013).

Table 5. Outline of ESTEEM's Two-Part Online Module

Section	ESTEEM Learning Opportunity	Type of Resource	Type of Factor	Hypothesized Impact on STSE
ESTEEM Foundation Module Part 1				
Read and Watch Essential Materials	1.1a. How is statistics different from mathematics?	Online page reading	-View of statistics	-Mastery experience
	1.1.b. Statistical investigations and habits of mind	Online page reading with instructional video	-View of statistics -Statistics knowledge	-Mastery experience
	1.1.c. Considering the importance of teaching statistics	Online page reading with expert panel video	-View of statistics -Pedagogical knowledge	-Mastery experience
	1.1.d. Quiz on Read and Watch material	Quiz	-View of statistics -Pedagogical knowledge	-Mastery experience -Social Persuasion -Physiological and affective state
Read and Watch Learn from Practice Videos	1.1.e Teaching statistics in the mathematics curriculum	Video of expert teacher in classroom	-View of statistics -Pedagogical knowledge	-Vicarious experience
	1.1.f. Statistical investigation cycle in a classroom	Video of classroom	-View of statistics -Pedagogical knowledge	-Vicarious experience
Engage with Data	1.1.g. Investigating older roller coasters in the US	Statistical investigation assignment with video of classroom	-View of statistics -Statistics knowledge -Pedagogical knowledge -Using technology	-Mastery experience -Vicarious experience -Physiological and affective state
Synthesize and Apply	1.1.h. Discuss learning statistics through investigations with real data	Discussion forum	-View of statistics -Pedagogical knowledge	-Vicarious experience -Social Persuasion -Physiological and affective state
	1.1.i. Compare and contrast online data analysis tools	Discussion forum	-View of statistics -Pedagogical knowledge -Using technology	-Vicarious experience -Social Persuasion -Physiological and affective state

Table 5 (continued)

ESTEEM Foundation Module Part 2				
Read and Watch Essential Materials	1.2.a. Supports for Learning to Do Statistical Investigations	Online page reading with instructional videos	-View of statistics -Statistics knowledge	-Mastery experience
	1.2.b. A Guiding Framework for Teaching Statistics	Online page reading with instructional video and expert panel video	-View of statistics -Statistics knowledge -Pedagogical knowledge	-Mastery experience
	1.2.c. Tasks as Opportunities for Statistical Learning	Online page reading with expert panel video	-View of statistics -Statistics knowledge	-Mastery experience
	1.2.d. Read & Watch quiz	Quiz	-View of statistics -Statistics knowledge -Pedagogical knowledge	-Mastery experience -Social Persuasion -Physiological and affective state
Read and Watch Learn from Practice Videos	1.2.e. Expert Teacher Interview on Tools & Resources	Video of expert teachers	-View of statistics -Pedagogical knowledge -Using technology	-Mastery experience
	1.2.f. Teaching Statistics Using Multiple Technologies	Video of classroom	-View of statistics -Pedagogical knowledge -Using technology	-Vicarious experience
Engage with Data	1.2.g. Investigating More Roller Coasters	Statistical investigation assignment	-View of statistics -Statistics knowledge -Pedagogical knowledge -Using technology	-Mastery experience -Vicarious experience -Physiological and affective state
	1.2.h. Examining Students' Work on the Roller Coaster Task	Discussion forum with video of classroom	-View of statistics -Pedagogical knowledge -Using technology	-Vicarious experience -Social Persuasion -Physiological and affective state
Synthesize and Apply	1.2.i. Supporting Statistical Discourse with the Roller Coaster Task	Reflection assignment with video of classroom	-View of statistics -Pedagogical knowledge -Using technology	-Vicarious experience
	1.2.j. Analyze Tasks and Discuss	Discussion forum with example tasks	-View of statistics -Statistics knowledge -Pedagogical knowledge -Using technology	-Mastery experience -Vicarious experience -Social Persuasion
Post-Module Assignment	Screencast Assignment	Statistical investigation assignment	-View of statistics -Statistics knowledge -Pedagogical knowledge -Using technology	-Mastery experience -Physiological and affective state

Module Opportunities for Impacting Statistics Teaching Self-Efficacy

There are learning opportunities within the ESTEEM module that are hypothesized to be critical points in participants' STSE development. See Table 5 for ways each element of the module is hypothesized to be a potential factor that could impact STSE through a source experience. For example, with regard to mastery experiences of doing statistics, participants engage in two statistical investigations (see 1.1.g and 1.2.g in Table 5). When considering how to implement these types of tasks in their own classrooms, these mastery experiences could also be viewed as a vicarious experience of teaching statistics for mathematics PSTs. There are also vicarious opportunities for participants to watch video cases of teachers implementing tasks and teaching methods in real classrooms (see 1.1.f, 1.2.f, 1.2.h, 1.2.i in Table 5). In addition, discussion forum questions that ask participants to envision themselves enacting a task or teaching method provide an opportunity for preparing PSTs to implement these types of tasks in their own classroom (Gerges, 2001). Answering reflective questions allows for the vicarious experience of considering how they would implement a task like this in their own classroom or how they would modify an existing statistics task to better it (Morris et al., 2017). Aside for the two roller coaster data investigations, the PSTs analyze several tasks for their potential to promote statistical habits of mind (see 1.2.j in Table 5). Social persuasion about teaching statistics may have come in the form of feedback from instructors on assignments or in discussion forums or appraisal or critique about their ideas from peers in the forums. Physical and emotional states may have been felt as a result of the materials or assignments, and those states may impact how PSTs feel in terms of their confidence to teach statistics in the future.

Research Question and Methods

My research question was: How do statistics pedagogy learning experiences impact preservice mathematics teachers' statistics teaching self-efficacy? I used a collective case study methodology with two phases; Phase I involved data collection from courses using ESTEEM online materials, and Phase II involved additional data collection from a subset of the Phase I participants through a survey and interview. To answer my research question, I analyzed qualitative data from the collective case in order to develop an understanding of how the ESTEEM online materials may have impacted factors known to contribute to PSTs' perceived preparedness to teach statistics (Figure 9). Resources from the ESTEEM module hypothesized to be impactful were analyzed using evidence from ESTEEM assignments for Phase I participants, and a semi-structured interview conducted in person or virtually for Phase II participants.

Participants

Participants for this study were chosen using a convenience sample. They are PSTs who were enrolled in undergraduate and graduate mathematics teacher education programs at two institutions in the United States, Institution A (n=19) and Institution B (n=15). Both institutions opted to participate in the ESTEEM project and they both offer a similar hybrid course focused on teaching mathematics with technology in which the online Foundational module was used in its entirety. I define a hybrid course as one that includes both synchronous and asynchronous online components and/or face-to-face meetings. For Phase II, all eligible PSTs from the two institutions were invited and encouraged to participate, which involved completing an online autobiographical survey and participating in an interview. Only 7 preservice secondary mathematics teachers participated fully in Phase II, 6 from Institution A and 1 from Institution B. Demographic information was not available or collected for all of the participants.

In addition to two weeks of experiences within the ESTEEM module, participants from Institution A had other learning experiences related to teaching statistics, specifically, four more weeks of statistics pedagogical content focused on center and variation, bivariate and multivariate concepts, inferential reasoning, and designing technology-based statistics tasks. At Institution B, three weeks were used for the ESTEEM materials, with no other statistics-focused content during the course. In both cases, a screencast statistical investigation assignment, described further below, was assigned and submitted after completion of the module; at Institution B, it was due at the end of the ESTEEM module and at Institution A, it was due 2.5 weeks after the end of the ESTEEM module, while PSTs were still engaged in other statistics pedagogy learning material. Institution A also included an additional statistics lesson plan assignment due at the end of the statistics pedagogy unit. Participants at Institution B did not create a lesson plan for teaching a statistics topic within the course using the ESTEEM Foundational module.

Data Sources and Analysis

Student work completed during engagement with the Foundational module forms the main body of data used to analyze the effectiveness of the learning experience. Student work comes from discussion forums (1.1.h, 1.2.h), reflection assignments (1.2.g, 1.2.i), written statistical investigations (1.2.g), as well as the screencasts of statistical investigations. Interview transcripts are also used to gain insight into participants' perspectives on the module one to one and a half years after the online learning experiences.

In order to analyze discussion forums, reflections, written statistical investigations, and screencast statistical investigations, I read or watched all artifacts multiple times in order to gain a general picture of participants' responses. Coding was done in sections by "level of meaning"

(DeCuir-Gunby, Marshall, & McCulloch, 2011, p. 145). Discussion forums and reflections were coded during several rounds, with increasing levels of specificity at each round, guided by my framework (Figure 9) and notes from initial readings of the data. A codebook was created using a combination of theory-driven and data-driven codes (DeCuir-Gunby et al., 2011). Codes were confirmed or modified based on the constant comparison method (Glaser, 1965). Analysis on the screencasts was done using an entire screencast as a unit of meaning. I watched each screencast and took notes on what features of CODAP were used and to which statistical habits of mind participants were attending. After open-coding of screencasts, a rubric was developed and each screencast was watched and scored according to the rubric, in order to analyze for evidence of impact on STSE beliefs. A total of 29 screencasts were fully analyzed. In order to analyze written statistical investigations, based on initial readings, I developed a list of things for which to code for each question on the task. I then coded all participant investigations and created a summary of each institution.

Results and Discussion

Opportunities afforded by engagement in the ESTEEM module for impacting PSTs' statistics teaching self-efficacy can be categorized into one of the four factors that contribute to PSTs' preparedness to teach statistics (Figure 9):

- View of statistics
- Statistics knowledge
- Pedagogical knowledge
- Use of technology

Most of the time, specific ESTEEM materials were designed to impact more than one factor; in fact, some materials incorporated all four. For example, in the last activity, "1.2.j. Analyze tasks and discuss", mathematics PSTs engage in a pair of tasks, usually using technology, and then reflect on the experience and discuss with peers the opportunities for statistical thinking the two tasks provided. This activity could potentially impact PSTs' view of statistics, statistics knowledge, pedagogical knowledge, and use of technology. Below, selected opportunities in the module to impact each of the four factors are highlighted, in addition to evidence from student work that illustrate how those opportunities actually impacted participants.

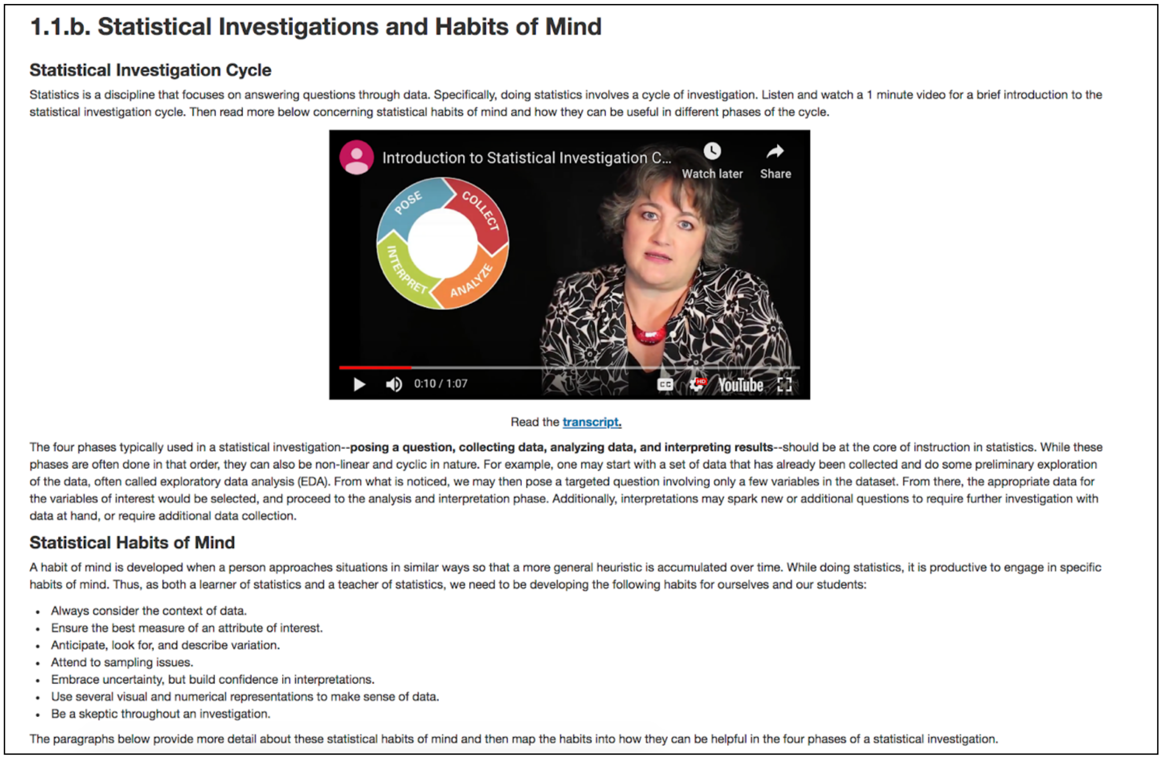
Opportunities for Impacting View of Statistics

I classified every single resource and material in the two-part module as an opportunity for impacting mathematics PSTs' view of statistics. This is because there was a deliberate design decision to present statistics throughout the module not as mathematical computations and procedures, but as the "art and science" (1.1.c., 1.1.d.) of investigating real data (Hudson et al., 2018). This view was emphasized throughout the entire module, from the readings, to the expert panel discussions, to the type of investigative questions that PSTs were asked to answer. I highlight below those resources I consider most impactful on PSTs' view of statistics, based on their positioning in the module and on evidence from PSTs' in their discussions, reflections, and interviews.

"1.1.a. How is statistics different from mathematics?" is the very first resource PSTs encounter in the module, and it is a very brief reading on an online page that highlights the key differences between statistics and mathematics. The main differences described are that statistics heavily relies on data and context, must consider measurements issues, and incorporates variability and uncertainty when making claims.

Immediately after this reading, in "1.1.b. Statistical investigations and habits of mind", PSTs are introduced to the statistical investigation cycle on an online page that also includes a very brief embedded instructional video of an expert educator describing the cycle (Figure 10).

The page also includes statistical habits of mind that should be attended to at certain phases of the cycle, with examples of each. The investigative cycle is emphasized in other parts throughout the module, and PSTs particularly mention the video in “1.1.f. Statistics investigation cycle in a classroom” as being impactful. In this video, they see an example of the investigative cycle being implemented in a classroom and a teacher’s reflection on how students are able to engage in it. This is again later emphasized in the second part of the module in “1.2.b. A guiding framework for teaching statistics”, where PSTs read about and see a visual of the Students’ Approaches to Statistics Investigations (SASI) framework that incorporates the statistical investigative cycle and statistical habits of mind, in addition to different levels of understanding at each phase of the cycle (Figure 11, Lee & Tran, 2015). Two instructional videos are embedded on this page that show expert educators describing examples of different levels of understanding of specific statistical concepts. All of these materials include opportunities for shifting mathematics PSTs’ view of statistics to one that involves collecting, analyzing, interpreting, and making conclusions about data.



1.1.b. Statistical Investigations and Habits of Mind

Statistical Investigation Cycle

Statistics is a discipline that focuses on answering questions through data. Specifically, doing statistics involves a cycle of investigation. Listen and watch a 1 minute video for a brief introduction to the statistical investigation cycle. Then read more below concerning statistical habits of mind and how they can be useful in different phases of the cycle.

Read the **transcript.**

The four phases typically used in a statistical investigation--**posing a question, collecting data, analyzing data, and interpreting results**--should be at the core of instruction in statistics. While these phases are often done in that order, they can also be non-linear and cyclic in nature. For example, one may start with a set of data that has already been collected and do some preliminary exploration of the data, often called exploratory data analysis (EDA). From what is noticed, we may then pose a targeted question involving only a few variables in the dataset. From there, the appropriate data for the variables of interest would be selected, and proceed to the analysis and interpretation phase. Additionally, interpretations may spark new or additional questions to require further investigation with data at hand, or require additional data collection.

Statistical Habits of Mind

A habit of mind is developed when a person approaches situations in similar ways so that a more general heuristic is accumulated over time. While doing statistics, it is productive to engage in specific habits of mind. Thus, as both a learner of statistics and a teacher of statistics, we need to be developing the following habits for ourselves and our students:

- Always consider the context of data.
- Ensure the best measure of an attribute of interest.
- Anticipate, look for, and describe variation.
- Attend to sampling issues.
- Embrace uncertainty, but build confidence in interpretations.
- Use several visual and numerical representations to make sense of data.
- Be a skeptic throughout an investigation.

The paragraphs below provide more detail about these statistical habits of mind and then map the habits into how they can be helpful in the four phases of a statistical investigation.

Figure 10. An online reading page that introduces PSTs to the statistical investigative cycle and habits of mind.

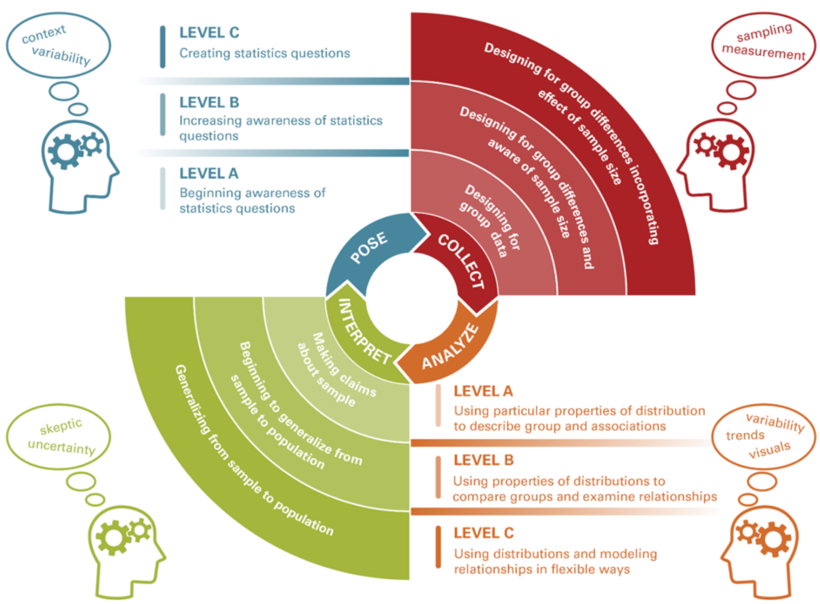


Figure 11. A visual of the SASI framework (original in Lee & Tran, 2015).

The most referenced resource in the first part of the module is "1.1.c. Considering the importance of teaching statistics" which includes two parts – a video of an expert panel discussing what statistics is and why we should teach it, and a very brief reading with links and reflection questions on the role statistics plays in various curriculum standards. Although there are many topics covered by the expert panel in 1.1.c., the very first question addressed by the panel is how they each define statistics. In addition, the experts give examples of practices they incorporate in their own classrooms, further emphasizing the shift to viewing statistics as data-focused, not computations-focused.

Finally, mathematics PSTs have several opportunities to complete their own investigations of data sets, for example in “1.1.g. Investigating older roller coasters in the US”, “1.2.g. Investigating more roller coasters”, and in a screencast statistical investigation assignment. In 1.1.g., participants watch a video of a teacher launching a similar investigation with middle school students, and then dive into the data themselves using CODAP to investigate older roller coasters’ attributes like maximum height and track length. In 1.2.g., participants investigate a larger data set of 157 roller coasters in the United States with 16 quantitative and qualitative variables, like height, park, speed, and type of material. They are asked to investigate individual cases, in addition to describing distributions, comparing groups, and analyzing relationships between variables (Figure 12 and Figure 13). In the screencast statistical investigation, PSTs use CODAP to investigate the roller coaster data set again or to choose from two other data sets, and record their voice and screens while they investigate. Throughout these experiences investigating data, again a shift towards viewing statistics as the art and science of analyzing data is emphasized implicitly through the types of tasks chosen for PSTs to complete.

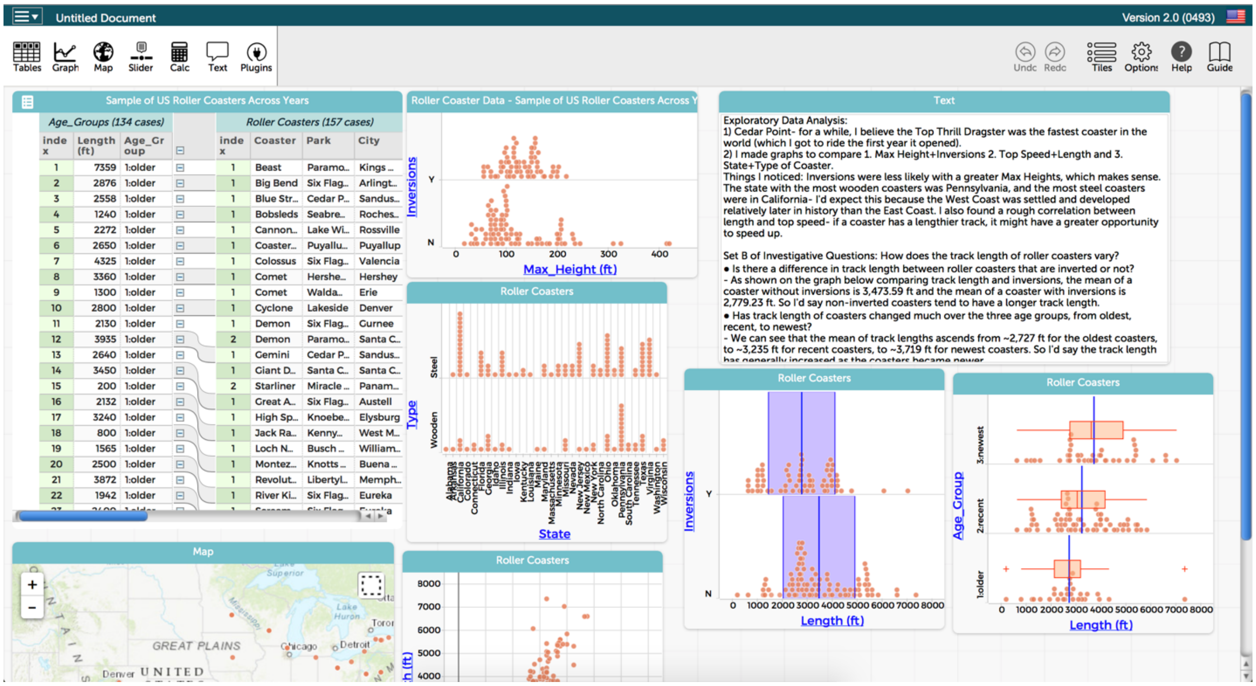


Figure 12. Example of a PSTs' work in CODAP to investigate a large set of US roller coasters.

Exploratory Data Analysis

1. Use the Map and Table to find a park that is close to your location or one you have visited. Describe some interesting facts about the roller coasters in that amusement park.

2. Open two or more Graph windows (click twice on the Graph icon on the top shelf) and explore attributes about coasters that interest you (e.g., track length, inversions). Create several different graphs to explore any patterns and trends for these attributes. Record some things you noticed and insert screenshots of graphs you created.

Comparing Groups

One of the essential ways to help students develop their ability to describe a distribution of data is to have them compare two or more groups. We can use categorical attributes to help do this. Students tend to naturally start asking questions, or wonder about, how something compares for different groups of data.

3. Choose **ONE** of the following sets of questions to investigate. Be sure to use different tools in the Graph menu to add measures to the graphs. Also use the Table to reorganize the data and compute statistical measures of interest. Insert screenshots and describe the work you did to answer the questions.

Set A of Investigative Questions: How fast do roller coasters tend to be?

- Are there any differences between the top speed of older, newer, or more recent roller coasters? Explain.
- Does the type of material make a difference in speed?
- If you were going to ride a coaster that was built before 1980, what would you expect for a typical top speed? What about if you were riding a newer roller coaster?

Set B of Investigative Questions: How does the track length of roller coasters vary?

- Is there a difference in track length between roller coasters that are inverted or not?
- Has track length of coasters changed much over the three age groups, from oldest, recent, to newest?
- If you were going to ride a coaster that was built before 1980, what would you expect for a typical track length? What about if you were riding a newer roller coaster?

Bivariate Association

Bivariate data is the term used to describe data that have two variables for each observation. The cases (observations) in our dataset have many variables (attributes). When examining two attributes in a data set, our attention is on how the values for those attributes co-vary. Thus for bivariate data, covariation involves correspondence of the variation possible in each variable. We could examine association between categorical variables or two quantitative variables.

4. Choose **ONE** of the following questions to investigate (and the follow-up question listed below it). Be sure to use different tools in the Graph menu to add measures to the Graphs. You could also use the Table to reorganize the data and compute statistical measures of interest. Insert screenshots and describe the work you did to answer the questions.

Set A Investigative Question: Is the type of material used to make a roller coaster related to whether or not the passengers get inverted during a ride?

- Are you more or less likely to get inverted on a Steel coaster? Explain why based on your data, and then explain why based on what you know, or can find out, about how coasters are built.

Set B Investigative Question: Is the length of the drop on a roller coaster related to its top speed?

- What other attributes in the design of a coaster may impact how drop and top speed are related?

Reflection (3-5 sentences)

What excites you and makes you nervous about implementing tasks like the roller coaster investigation in your own classroom? Make connections to what you experienced doing the investigation as well as what you have been learning about in this module related to supporting students' learning of statistics through investigations.

Figure 13. The 4 main questions that PSTs are asked to investigate in 1.2.g, ending with a final reflection question.

Evidence of impact on view of statistics. From their written statistical investigations, screencast statistical investigations, and interviews, there is evidence that the module had a positive impact on *developing* mathematics PSTs' view of statistics. Participants displayed a developing view of the nature of statistics as distinct from mathematics, by engaging in aspects of the statistical investigation cycle, attending to certain statistical habits of mind, and developing their use of statistical language. In semi-structured interviews participants were able to explicitly express if and how specific opportunities in the module impacted their STSE. When asked about the roller coaster investigations, (1.1.g and 1.2.g), 6 of the 7 interviewees indicated that it positively impacted their confidence to teach statistics because they learned that statistics could be fun and different from how they learned it.

PSTs started to shift from mathematical and procedural oriented approaches to investigations to statistical and exploratory investigations. Some desired views of statistics that were missing from their investigations were treating an investigation as being guided by a statistical question. Evidence of these views are discussed below.

Context. Participants were able to develop their understanding of statistics as data-centered by considering the context of data through explicit scaffolding that asked them to do so. For example, in the first question (Q1) of 1.2.g. (Figure 13), students became familiar with the data context through a case or multiple cases, which is a natural starting point for early learners (Konold, Higgins, Russell, & Khalil, 2015). There is evidence that PSTs were indeed able to use a case to start to explore the data. All but one student from both Institution A and B referenced data in responding to Q1, with some supplementing their responses with personal knowledge about the roller coaster context. For example, one participant wrote:

> Roller coaster number 74 in the chart, the Volcano at Kings Dominion in Doswell, Virginia. I can not tell you how many times I have rode this coaster. It starts off slow and then all of a sudden you blast off straight forward then up and out of the head of the volcano. By looking at the data I realized it it opened in the 1990s which was a shock to me. I also noticed that there is not a duration recorded for the coaster in the chart.

We see here this participant is supplementing personal knowledge from riding the coaster with information found in the table. Because the majority of participants did in fact base their responses on data from the table and/or a graph, rather than personal or outside knowledge, there is evidence that they were able to consider a case in order to start exploring data; an early starting point in the shift to viewing statistics as the study of data.

Visuals and trends. Statistical habits of mind introduced in 1.1.b included using multiple visuals to examine data and to look for and describe trends. Participants were able to explore trends by using a variety of representations, again shifting their view of statistics to one of exploring data, rather than computing single measures such as a mean or completing procedures for generating specific graphs. In the second question (Q2) of 1.2.g. (Figure 13), PSTs explored the roller coaster data in an open-ended way, choosing attributes that were interesting to them and that they might predict are associated with each other. A variety of different types of attributes were explored by PSTs in Q2, and the most popular type of trend explored at both Institutions was a bivariate quantitative scatterplot. In fact, 6 of the 17 participants from Institution A *only* explored bivariate quantitative associations in Q2. At Institution B, 4 of the 14 participants *only* explored bivariate quantitative associations in Q2; interestingly, 4 of the 14 participants at Institution B *only* explored univariate distributions in Q2.

Although most participants explored more than one different type of graphical representation, there was a large portion (14/31) that only looked at one type, particularly scatter plots and dotplots. This may lead to (or be a result of) a belief that scatter plots and dotplots are the only way to visualize data. Even though this may be the case, questions 3 and 4 (Figure 13) require *different* types of graphical representations than those two, so the scaffolding provided in choosing different types of attributes, forced the opportunity to see different ways of exploring data. Thus, there is evidence that participants were beginning to shift to a view that statistics involves using "several visual and numerical representations to make sense of data", which is one of the statistical habits of mind emphasized throughout the module. Again, this was seen in the variety of graphs participants chose to use to investigate data, incorporating bivariate categorical, bivariate quantitative, univariate, and multiple univariate representations in their investigations.

Some habits of mind related to making claims backed by evidence, and communicating those claims, are: "Coordinate graphs and statistical computations to reason about distributions in the aggregate", "Reason quantitatively and make arguments supported by data" and "Make a claim connected to the context of the questions" (Lee & Tran, 2015). For questions 2 through 4 of the written statistical investigation (1.2.g) participants supported claims or conclusions with graphical representations of data. Participants from Institution A included screenshots in their responses where they were required, but not when they were not, as in Q1. For Q3 and Q4 at Institution A, the majority included screenshots of individual objects, like of a graph or a table, rather than a screenshot of the entire CODAP window. While the investigation questions were common across both institutions, Institution B implemented a different method of submission from Institution A. Participants at Institution A submitted a written (e.g. Word or PDF) document, while most (all except one) participants at Institution B submitted a link to a CODAP document. Because the submissions at Institution B were links to CODAP documents rather than Word or PDF documents, it did not make sense for participants to include screenshots in their responses. Thus, I looked at whether or not they kept evidence of their responses, in the form of graphs or tables, in their CODAP document. At Institution B, almost all participants who responded to Q3 still retained evidence of their response in their CODAP window; all did for Q4. For Q1, the majority of participants who responded based on data kept evidence of their response in the saved document, even though a screenshot was not required.

Variability and uncertainty. In analyzing the screencast investigations for appropriate use of statistical habits of mind (Lee & Tran, 2015), most of the screencast investigations at least attended to uncertainty and variability in their language and made infrequent statistical language errors. However, while in many screencast investigations, participants attended to variability of data and expressing uncertainty in their conclusions, many lacked other important statistical habits of mind, like attending to issues of sampling and measurement. For example, in Figure 14, this participant states that the majority of data points for the South lie at or below the mean students per teacher ratio, while the majority of data points for the West lie at or above the mean, indicating that she is attending to the variability of the data. She also has the West states highlighted on the map, which she states are mostly a darker shade than the Southern states, expressing a certain amount of uncertainty in her conclusions. At Institution A, participants were asked to identify relevant standards that were addressed by their work in the screencasts. The fact that standards identified by participants sometimes did not match the content of screencast investigations and that statistical language errors occurred indicates insecurity in statistical content that may be related to comfort with some statistical habits of mind over others.

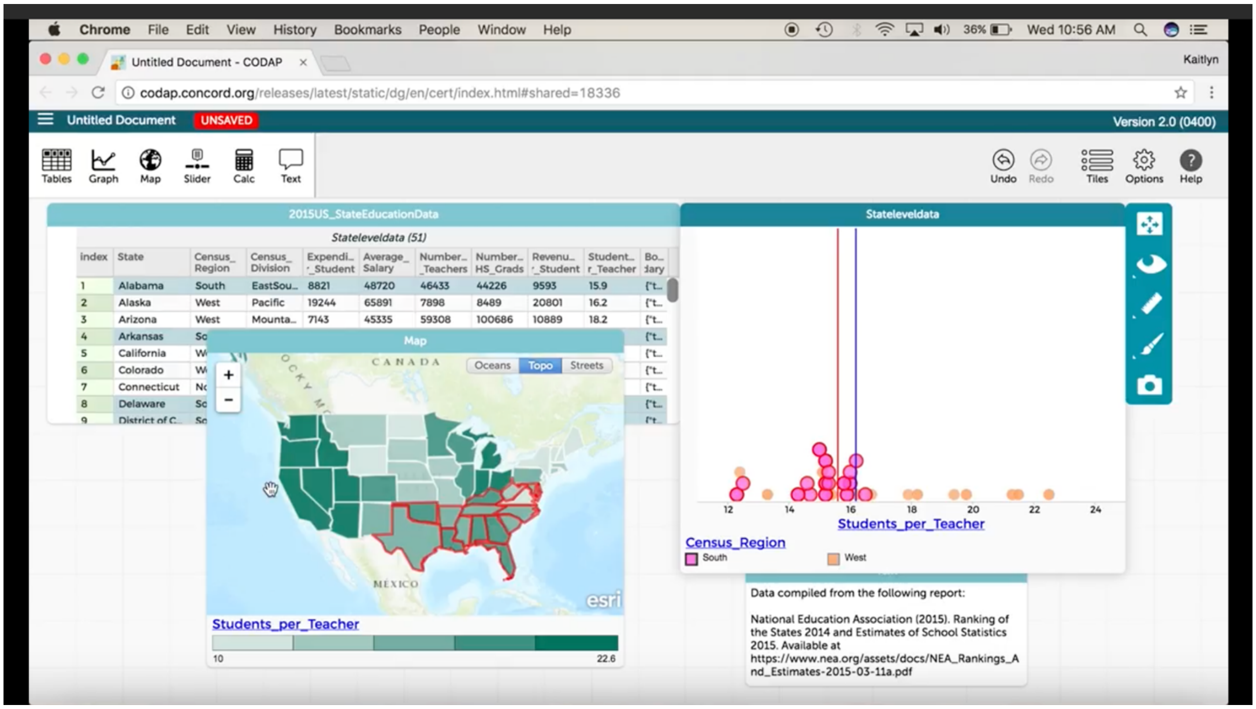


Figure 14. Screenshot of a PST's screencast statistical investigation.

In addition, although statistical habits of mind may have been present in portions of a screencast investigation, they were not always consistently expressed by a participant throughout the entirety of a screencast. For example, participants may have compared three univariate distributions separated by age group by comparing their measures of center and spread and using language that indicated trends, like "on average" and "generally", but then they may make a final conclusion that the three distributions are "definitely" increasing as time passes (Figure 15). Their definitive final conclusion lacks the attention to variability and uncertainty that they formerly expressed in their analysis.

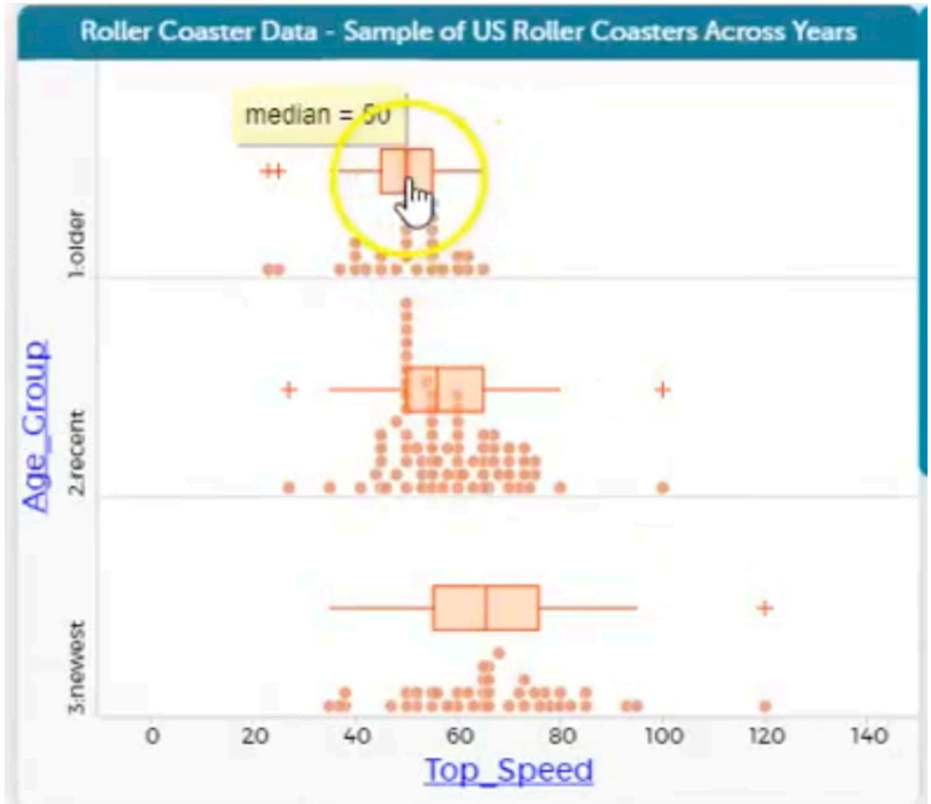


Figure 15. Screenshot of a PST's use of comparing univariate distributions.

Limited use of cycle of investigation. Some aspects of exploring a large, multivariate dataset that were encouraged throughout the module were not seen in PSTs work. Even though PSTs were exposed to the idea of investigations involving a statistical question, through a visual of the SASI framework and as part of a framework for analyzing statistical tasks, at Institution B, only 1 of the 14 participants included a statistical question they were attempting to answer in Q2. Although some participants at Institution A included a statistical question they were attempting to answer in Q2, the vast majority just listed their wonderings and noticings, without listing a question.

In addition, even though PSTs were exposed to the idea of investigations being a cycle in an online page that displayed and described the SASI framework as a visual cycle and in a classroom video of students engaging in the cycle, with conclusions leading to new questions, only 1 of the 14 PSTs at Institution B and none of the PSTs at Institution A built on their first exploration in Q2; the majority listed completely separate explorations in their responses. Since participants were asked to open multiple graphs and explore different attributes, it was a chance for them to build on their exploration. After noticing or finding out something interesting, they could have used that to prompt a related exploration, but that was not a skill or belief they seemed to gain from the module.

Mirroring some of the evidence seen in the written statistical investigation, the majority of participants had screencast investigations that were exploratory, but not completely rigorous statistical investigations. For example, a participant may have explored the differences among age groups, but their method of doing that may have been based solely on means and they may have repeated that method of analysis for a comparison among two different groups of data. While there were some very high quality exploratory statistical investigations at Institution A, there were none at Institution B and, on the other hand, there were some that lacked any type of exploration and took on a very rigid method of investigation. More participants from Institution A created a screencast investigation that attended to at least three habits of mind, made no statistical language errors, and matched their standards to their statistical content than those who may have only attended two habits of mind, made infrequent statistical language errors, and did not match their standards to statistical content.

Screencast investigations that were labeled as "developing" in the category of appropriate method of statistical investigation, attended to both measures of center and spread in their exploration of the data, however, conclusions were usually based solely off of a measure of center, sometimes with a brief mention of the differences in variability between two sets of data. This routinized method of investigating echoes prior work studying teachers' use of dynamic statistical software in statistical tasks (Lee et al., 2014). Participants who displayed evidence of an exploratory and rigorous statistical investigation relied on more than just means to conduct their analysis; they did things like compare chunks of data, investigate specific extreme cases to explore other attributes that made them different, and used multivariate thinking by considering more than two attributes at a time. These actions were only possible if participants felt comfortable taking advantage of the affordances of the technology tool, discussed in more detail below, and also if they felt the nature of a statistical investigation necessitated exploring further than simply a difference in means.

So, while participants were on their way to expressing statistical thinking in an advanced way, they were not quite there yet; thus, I inferred a developing view of statistics that is impacted by their experiences in the module. PSTs' experiences in the module tended to emphasize certain habits of mind over others, which reflected the ones they attended to in their screencast investigations.

Opportunities for Impacting Statistical and Pedagogical Knowledge

Another major focus of the module was to provide opportunities for mathematics PSTs to increase their statistics pedagogical knowledge, a factor we know contributes to their perceived preparedness to teach statistics (Figure 9). The statistical content used in all aspects of the module was purposely designed to be mainly equivalent to the statistical topics included in middle school and early high school content, or would be covered in the first part of a college level introductory statistics course. Thus, it was *not* an explicit goal of the module to increase PSTs' statistics content knowledge; however, sometimes statistics content knowledge was implicitly included, especially in resources that gave examples of pedagogical strategies, which is why I am combining these two factors for the purposes of this paper. The major resources highlighted here that included opportunities for developing PSTs' statistical and pedagogical knowledge are similar to the resources highlighted above for impacting their view of statistics, with the addition of two video cases of real classrooms which PSTs discussed or reflected on to consider the pedagogical strategies used.

The resources that emphasize the statistical investigation cycle and statistical habits of mind, like in "1.1.b. Statistical investigations and habits of mind", "1.1.f. Statistics investigation cycle in a classroom", "1.2.b. A guiding framework for teaching statistics", and "1.1.c. Considering the importance of teaching statistics", not only include opportunities for implicitly increasing statistical knowledge, but also explicitly include opportunities for increasing pedagogical knowledge for teaching statistics to others. The emphasis on using tasks and projects in classrooms where students engage in the entire cycle, investigate real and large data sets using technology, and attend to statistical habits of mind is emphasized throughout these materials. In addition, the two written investigations and the screencast investigation give PSTs examples of statistics tasks, data sets, and a technology tool that can be implemented in their future classroom. After completing the second written statistical investigation (1.2.g.), they were asked to reflect on what excites and makes them nervous about implementing similar tasks in their own classrooms, a reflection on their future practice.

In addition to these resources, "1.2.h. Examining students' work on the roller coaster task" and "1.2.i. Supporting statistical discourse with the roller coaster task" are also critical points in the module where PSTs can increase their pedagogical knowledge. In these activities, PSTs watch a video of a classroom where students are engaging in the roller coaster investigation. Then, they are asked to reflect or discuss explicit pedagogical concerns, like the affordances of the technology for statistical reasoning or how the teacher built on student thinking through her interactions.

Evidence of impact on pedagogical knowledge. Because the module was not specifically aimed at impacting statistics knowledge, I discuss only evidence of impacts on pedagogical knowledge. As evidenced in their recording of a screencast statistical investigation, reflecting on a written statistical investigation, and reflecting on videos cases of real classrooms, participants displayed *developing* pedagogical knowledge, including the beginning of a shift to focusing on student thinking and comfort in doing a statistics investigation, and possibly taking on a teacher role while completing one.

In 1.2.h. And 1.2.i., participants were asked to reflect on certain aspects of two video cases. Although two of the discussion forum prompts use language referencing the specific teacher and students in the video cases, the prompts in the second discussion forum also include language that can be interpreted as more general, and thus resulted in the expression of general beliefs, which I viewed as evidence of restructuring, rather than descriptions of the specific video case, which I viewed as describing and inferring. However, participants were at least able to begin the process of attending to student thinking by describing and inferring from what they saw in the videos. When asked about how watching videos of teachers and students impacted their self-confidence to teach statistics, 6 of the 7 interviewed participants said they were positively impacted because the videos provided examples of seeing how statistics should be taught in a classroom.

Reflections written about implementing tasks like 1.2.g expressed excitement about statistics teaching practices that mirrored those seen throughout the rest of the module, like that statistics should be taught differently from mathematics, using the investigative cycle, and education statistical technology. However, excitement expressed in reflections was often conditioned on perceived constraints regarding beliefs about the realities of a real classroom. For example, participants discussed constraints in the amount of time they would have to plan effective lessons, constraints in the amount of time they would have to implement effective lesson plans, and constraints in the amount of prior knowledge with which students would enter into their classroom. For example, one participant worried: "I am a little nervous that I will struggle with the creation of activities like these, because it does require a lot of planning, understanding of students' interests, and brain power." Another said: "I like, with CODAP, that I am able to investigate the data at hand, but I am nervous about implementing a tool that may be new, and having to add into my instruction, how to use the tool itself instead of being able to spend even more time on statistical learning", which brings up another constraint that was discussed – the fear that technology would either not be available or fail, or that it would take too much time for students to learn.

In addition to evidence that PSTs were increasing pedagogical knowledge by shifting their focus to student thinking, they also showed their ability to do, and possibly model for students, a statistical investigation. Because PSTs have limited teaching experience and have limited access to real classrooms and students, teacher educators can provide opportunities for teaching that are not truly authentic, but rather "approximations of practice" (Santagata & Yeh, 2014). For example, the screencast statistical investigation could have been used as an approximation of practice if it was implemented in a way where the mathematics PSTs were asked to take on a teacher, rather than student, role. Most participants did not take on the teacher role while recording themselves doing the statistical investigation. The ones who did, did so very clearly; they spoke to an imaginary class or imaginary students, put their investigation in the context of a larger unit, and explained how or why they were doing things to the viewer. It is more feasible to consider the screencast investigation assignment as a mastery teaching statistics experience for those participants who did these things. If participants viewed the screencast investigation assignment as an experience in teaching statistics, it has more of a chance of impacting their PTE because mastery experiences should involve the actual activity being mastered, which in this case is teaching, not doing, statistics. Those who did not take on the teacher role very much took the role of an investigator doing a statistical exploration themselves. It was treated more of a documentation of their own investigation, with little or no reference or acknowledgement of the viewer; in that case, mathematics PSTs still had the opportunity to display pedagogical knowledge of how students should engage in a statistical investigation, by modeling it themselves. None of the participants from Institution B showed evidence of rehearsal or took on the teacher role. Contrast that to Institution A, where half of the participants did both of those things. The difference in results among the two institutions can be attributed to the way the task was presented to participants. At Institution A, there were specific written instructions given that indicated participants should rehearse their screencast investigation, include advanced features of the technology, use appropriate statistical language, and incorporate statistical habits of mind. In addition, there were informal oral instructions given to encourage participants to treat the screencast investigation as if they had a student who was absent and this was their way of filling in that student. At Institution B, participants had completed a written statistical investigation on the same data set, which allowed for student choice in which statistical questions to answer; for the screencast investigation assignment, they were asked to record themselves answering the questions they had not previously answered.

Another aspect to consider in analyzing the screencasts is the amount of rehearsal done before recording. The vast majority of students in this collective case provided no evidence of rehearsal. In most cases, this was an example of failing to reject; I cannot say with certainty that they did not rehearse, but they do not provide any evidence that they did. That is contrasted with some students who clearly rehearsed and planned their investigation as evidenced by planned slides, prepared graphs, and websites that have already been pulled up. However, in some instances, there *is* very clear evidence that there was no rehearsal prior to the recording. This evidence partly comes in the form of rereading investigative questions and intonations in reading that convey an attempt to understand what is being asked. It also comes in the form of talking through decisions of which graphs to make, silent pauses to interpret a resulting graph, and intonations that may express surprise or unanticipated results.

In screencast investigations that were rehearsed and showed evidence of teacher role, the learning opportunity uniquely allowed for mathematics PSTs to consider pedagogical ideas, like where their screencast would be placed within a unit, what standards the screencast might address, and answering anticipated questions viewers might have about why they chose to make certain decisions with the technology tool. When they were not rehearsed and showed no evidence of teacher role, the learning opportunity allowed for mathematics PSTs to display their technology abilities and knowledge about completing a statistical investigation, while still providing them an example of a task that could be used with future students. Mathematics PSTs who rehearsed and took on the teacher role had the benefit of both types of learning opportunities.

For 4 of the 7 interviewed participants, the screencasts helped them feel confident in using that particular method of teaching, recording a screencast for students, but not necessarily for teaching statistics. Some of them (4/7) expressed similar feelings as the ones expressed about the written statistical investigation; that it was a new, fun experience with statistics and showed them another example of how to do and teach statistics with interesting data.

Opportunities for Impacting Using Technology

Finally, there were opportunities in the online module for mathematics PSTs to learn how to use and how to teach with education statistical tools, particularly CODAP. Not only did participants have the opportunity to use educational tools to investigate data sets, they also had the chance to reflect on the affordances and constraints of certain education technology tools, to watch CODAP being used in real classrooms, and to hear from expert educators on how and why they use technology with their students. Again, I highlight the written and screencast statistical investigations as major opportunities for impacting the use of technology, since they required PSTs to practice using the technology to explore data sets and to communicate conclusions based on explorations. Another critical point is a discussion forum in which participants compared and contrasted two educational statistics tools, CODAP and Tuva, another free, web-based data analysis software designed for educational settings ("1.1.i. Compare and contrast online data analysis tools"). Analysis of these two tools may have impacted their confidence in using and teaching with them. The two video cases of real classrooms discussed above ("1.2.h. Examining students' work on the roller coaster task" and "1.2.i. Supporting statistical discourse with the roller coaster task") were also moments when participants could consider how students use CODAP, and compare that to their own use. There were additional opportunities in the module to consider the use of technology in teaching and learning statistics, but these are the experiences I consider most impactful.

Evidence of impact on ability to use technology. Using mainly the screencast investigations as evidence of the module's impact on mathematics PSTs' ability to use education statistical tools, participants displayed a *developing* comfort to use CODAP. In analyzing screencast investigations for the appropriate use of technology, the most frequent uses of CODAP were to create a dotplot, split a dotplot by a categorical variable, create a boxplot, and overlay an attribute onto a plot. Uses of CODAP that were present and desirable, but not used as frequently, were linking between multiple representations and computing a measure in the table. Institution A had a fairly even split of participants in terms of their use of advanced features of CODAP, whereas Institution B had the majority of participants using a medium amount of features of CODAP.

The collective case displayed developing knowledge of the advantages that technology can afford in a statistical investigation. Again, participants completed this screencast investigation within the context of an online module, which introduced them to a specific educational statistical tool, CODAP, for the first time for many of them. The features of the tool they decided to include in their screencast investigation is expressive of what they believe the tool to be capable of providing in terms of benefits; since the screencast investigations were submitted as a graded assignment, participants at least had an external motivation to show off their technology skills. Despite this fact, many participants kept their usage fairly basic, rather than using some of the more advanced features of the tool, like creating a hierarchy in the table, linking multiple representations, computing measures, and using the map. These features were demonstrated in an instructional video on how to use CODAP in 1.1.g.Six of the 7 interviewees indicated that the roller coaster investigations positively impacted their confidence to teach statistics, with one reason being that it provided them an example of how they could use technology to teach statistics.

Discussion and Implications

Overall, the ESTEEM online materials were effective in offering opportunities for PSTs to increase their confidence to teach statistics. They provided opportunities for them to shift their view of statistics, increase their statistics and pedagogical knowledge, and have meaningful experiences with educational statistics technology. As noted in the literature, it is unreasonable to expect preservice teachers who are not yet teaching to master the art and skill of teaching; so, it is desirable to teach PSTs to learn from their teaching (Hiebert et al., 2003).

This can be done through a variety of practices, like analyzing videos of teachers and students and reflecting on mathematics tasks in order to start to learn how to attend to student thinking. It is hard to shift PSTs' perspectives from focusing on themselves (or on teachers) to focusing on students (Zaslavsky, 2008). Specifically for statistics education, it is important for PSTs to consider students' habits of mind and engagement in doing statistics (Lee & Tran, 2015). The module provided those opportunities, and there is evidence that the mathematics PSTs at both institutions were developing their statistics pedagogy by engaging in statistical investigations and discussing videos of students doing statistics. The preservice teachers in the discussion forums and reflections with student video cases did much describing, what Wilson et al. (2011) refer to as the most simplistic way to analyze students' work by simply restating what they noticed happening during a video. PSTs also did some inferring, making connections or assuming things beyond what they saw in the video (Wilson et al., 2011). However, they did not do as much restructuring, incorporating new ways of thinking into their own existing knowledge (Wilson et al., 2011). One possible implication of this is that asking preservice teachers to reflect on particular teachers/students may not be helpful in changing or solidifying or eliciting beliefs if they are not asked to generalize those reflections to their future students or students in general. Because educational reform involves shifting instruction from being teacher-focused to being student-focused, teachers must know more about students' understandings & restructure their own understandings based on exposure to various ways of thinking (Wilson et al., 2011).

Another pedagogical consideration expressed by participants was the idea of imagined constraints in a classroom, which corroborates prior research (Harrison, Azmy, & Lee, 2018). In that study, incongruence between PSTs' expressed beliefs and planned implementation in a written statistical lesson plan was due to imagined constraints of a real classroom. The belief that constraints in a classroom may hinder the implementation of valued, effective practices may be due to former experiences in the role of a student or of a teacher.

There was also evidence that the module impacted participants' views on the importance of using technology in teaching; however, mathematics PSTs still showed a developing ability to use advanced features of educational statistical tools. Since the most common uses of the tool were to create dotplots and boxplots, or several dotplots separated by categories, and to display measures like mean and median, I can infer from this a belief about the types of ways technology can aid in a statistical investigation. The ways that participants used the technology was as an amplifier (Hollebrands, 2016), asking the tool to perform tasks that they could have done by hand, but more efficiently with technology. However, participants less frequently used the tool as a reorganizer, accomplishing tasks that can only be done using technology, and changing the way they are able to interact with statistical concepts inherent in the task (Hollebrands, 2016). Even though PSTs were exposed to some of these ways of using the tool, it may not have been emphasized enough for them to believe that they were useful and necessary.

The implementation decisions at both universities also proved to be important in several cases. Lee et al. (2014) noted that when teachers submit work with screenshots in a document, rather than within a technology file, we lose some insight into some of the interactions they may have attended to. I argue that this may be equally true when asking PSTs to submit via a technology file. The investigations submitted via CODAP document provided a different perspective of how PSTs' communicate results; they have limited space (practically speaking) in which to include responses and graphical and tabular representations, rather than the unlimited space of a Word document that is structured in a way that pretty much forces participants to communicate results in a certain way. Another thing to note is that as an instructor it is easier to know which objects were used in response to a question if they are directly connected to the question, like in a Word document. In the CODAP document because of window space issues, it was not always clear, although it was usually possible to know based off of the attributes described in the question response.

Whether this is a drawback or an affordance depends on the goals of the instructor; do we want PSTs to include as much detail as possible and to specify relevant visual representations, in a format similar to a written report? Or do we want PSTs to include limited detail that will fit onto a window screen and to organize visual representations in a way that makes sense to them, in a format similar to a visual presentation like a poster or infographic? Either way, the different experiences these PSTs at these two different institutions had provides evidence that even the method of submission is an opportunity for learning a pedagogical strategy, however implicit.

Several implications for practice follow. In order to positively impact PSTs' perceived preparedness to teach statistics, teacher educators should provide opportunities that combat a lack of prior experiences with statistics, or negative prior experiences with statistics. Even within a short 2-3-week module, purposefully designed materials were effective in impacting certain factors we know from prior research to be important in developing STSE (Lovett & Lee, 2017). Specifically, allowing mathematics PSTs opportunities to engage in statistical investigations and to analyze and reflect on teacher practice and student thinking were critical activities that incorporated many of the important factors. For example, these types of activities can help to develop PSTs' beliefs in the importance of using large, relevant, multivariate data sets and allowing for student exploration, asking students to complete a statistical investigation, attending to statistical habits of mind like backing up claims with evidence and organizing and communicating results, and using technology to teach statistics in order to build up higher levels statistical thinking.

We saw evidence that differences in implementation led to differences in ways PSTs viewed and interacted with similar activities, and thus the learning goals that were met. For example, using the screencast statistical investigation as a statistical investigation rather than using it as a guide to someone else's statistical investigation impacted experience and results. Thus, the implication for teacher preparation is that if teacher educators consider it important to provide mastery experiences to impact PSTs' statistics *teaching* self-efficacy, they must incorporate aspects into tasks that explicitly ask PSTs to develop skills needed for teaching. In addition, they must give specific rubrics or instructions that indicate to PSTs the types of ideas they should be incorporating into their teaching and/or doing of statistics. For example, although PSTs may be reading articles and watching videos on the nature of statistics and statistical habits of mind, they may not incorporate those ideas into their own mastery experiences if it is not explicitly asked for.

Finally, we address the issue of the online context. Although both institutions implemented the module in a hybrid course, the hybrid contexts were different. At Institution A, all course activity was online, with most of it asynchronous, and some synchronous online meetings. At Institution B, most course activity was online, with some synchronous face to face meetings. The fact that both institutions provided opportunities for synchronous communication between instructors and students is a factor to take into consideration; these meetings may have impacted the way participants engaged in the asynchronous online aspects of the course. That is an issue for further study.

Chapter 6: Discussion

This chapter summarizes the research questions answered by this collective case study and the main findings, making connections with current literature. In addition, I describe some limitations of this study. I end with suggesting some implications of the research and some directions for future research. This collective case study is meant to supplement current statistics education research by offering qualitative explanations for what impacts statistics teaching self-efficacy and how.

Summary of Research Question 1 and Findings

What are the general statistics education beliefs and personal teaching efficacy beliefs that comprise preservice mathematics teachers' statistics teaching self-efficacy, and what experiences and factors most impact development of those beliefs?"

To answer the first research question, I was interested in the broad spectrum of experiences that impact statistics teaching beliefs, but chose a collective case within a situated context to examine how specific experiences fit within the landscape of a lifetime of statistics experiences. Through analysis of participants' expressed beliefs within the context of an online module focused on developing statistics pedagogical content knowledge and also reflection after the module on experiences from childhood until college, I identified a variety of themes related to participants' beliefs about statistics education and beliefs about their personal statistics teaching efficacy.

General Statistics Education Beliefs

I identified three themes expressed by participants throughout the module and in their interviews. The first was that beliefs about good teaching persist over time, but beliefs about good statistics teaching are difficult to shift in one course. Because beliefs are the least cognitive and most stable of beliefs, attitudes, and emotions, they are also the hardest to change (Philipp, 2007). My research confirms this idea, since mathematics PSTs experiences with statistics education were much shorter in duration than their experiences with mathematics education and general teaching practices, like using task-based instruction and viewing teachers as facilitators. So, although their beliefs about good mathematics teaching may have positively changed as a result of experiences within their teacher preparation program, beliefs about good statistics teaching did not always stand the test of time.

One belief that PSTs consistently expressed was the belief that using dynamic statistical software in teaching statistics is useful. The fact that participants had all taken a course focused on teaching mathematics with technology may have contributed to this fact, but prior research shows that even when learning technology software for the first time, teachers are able to use it to analyze data using unique software features (Lee et al., 2014). Participants mentioned that learning how to use CODAP, an online dynamic statistical software, was useful, and that it would be beneficial when teaching statistics. Participants' predictions of their future use of CODAP is supported by research on practicing teachers who learned about teaching and learning statistics with technology in their teacher preparation program. In that research, McCulloch, Hollebrands, Lee, Harrison, and Mutlu (2018) found that novice secondary mathematics teachers purposely used educational statistical tools like *TinkerPlots* and *Fathom* (which are similar in nature to CODAP), in addition to online applets, to help students understand statistical concepts.

Another belief that did stand the test of time was the belief that statistics and mathematics are distinct. The third theme expressed by participants was that they were unsure about what constitutes statistics, but that it is distinct from mathematics. This may actually explain incongruent prior research that sometimes shows only a weak correlation between statistics knowledge and statistics teaching self-efficacy (Lovett, 2016) and moderate to strong evidence of a correlation (Harrell-Williams et al., 2019). When PSTs show a positive correlation, it may be because their knowledge and efficacy are assessed together; thus their self-efficacy assessment is based on ideas of what statistics is found within the content assessment. If the efficacy assessment is not accompanied closely in time to a content assessment, PSTs' lack of knowledge of what constitutes statistics may be impacting their responses. They may be responding based on their mathematics teaching self-efficacy beliefs (Hannigan, 2013) or incomplete ideas of what statistics is.

Personal Statistics Teaching Efficacy Beliefs

I identified two themes expressed by participants about their own personal statistics teaching efficacy beliefs. The first was that mathematics PSTs identified themselves as being positively impacted by the context of the online module aimed at preparing them to teach statistics; however, the experience was not viewed as enough, even for participants who had an extended number of weeks in addition to the two week module. Because participants already had a lack of statistics experiences outside of their teacher preparation program, the short intervention in one of their methods classes was not viewed as enough to fully prepare them to teach statistics. However, PSTs identified specific aspects of the module as being helpful. The module provided PSTs with some vicarious experiences of teaching statistics and mastery experiences of doing statistics (Bandura, 1997), but PSTs wanted face-to-face vicarious experiences that were not provided within the module.

In addition, mathematics PSTs generally lack confidence in doing and teaching statistics, and that is attributed to a lack of quality experiences with statistics. Participants mostly talked about how they did not have enough statistics experiences upon which to draw, but even when they did, they were viewed negatively. Most statistics experiences were not until high school, and even then, they were lecture-based and procedural.

The results were not all negative – even though mathematics PSTs were generally not confident to teach statistics, they were confident in their ability to teach certain topics, to learn topics they would have to teach, and to use their own personal life experiences to make statistics relevant. For example, they expressed confidence in teaching topics that were a particular focus of the ESTEEM online module, like the Common Core standard on interpreting differences in shape, center, spread, and accounting for outliers. Although these results confirm prior research that shows PSTs' lack statistical knowledge (Lovett & Lee, 2018), the good news is that they recognize it and plan to address it when they teach statistics in the future.

Impactful Experiences

As discussed in Chapter 4, I revised my initial framework to reflect sources of beliefs that were identified by mathematics PSTs as being most important to the development of their beliefs. As novice teachers, their vicarious experiences, or lack thereof, were most impactful, confirming prior research that made this claim (Tschannen-Moran & Hoy, 2007). Those included vicarious experiences as a statistics student and as a "teacher" student developing their pedagogical content knowledge through experiences such as viewing a classroom from a teacher perspective, confirming prior research that identified pedagogical content knowledge as a major factor impacting STSE (Lovett & Lee, 2017). Mastery experiences were also impactful, although mathematics PSTs mostly had mastery experiences *doing*, rather than *teaching*, statistics. In addition, they all shared common mastery experiences of considering approaches to statistics pedagogy. For example, participants used an educational statistical tool that they could use in their future classroom; use of technology has also been identified as a major factor impacting STSE (Lovett & Lee, 2017).

Social persuasion experiences were not frequently identified as impactful when they were viewed as feedback from instructors on assignments like lesson plans. However, other types of social persuasion experiences, like feedback from instructors or peers on ability to do statistics, arose in participant responses during interviews, and are included in the framework as a place for possible future research. Finally, physiological and affective states, like strong emotional responses to prior statistics learning experiences and excitement about learning statistical pedagogical methods, were shifted from the list of impactful experiences to the list of teacher characteristics, as they were most frequently identified as a result of an experience, not an experience in and of itself (see Figure 8 in Chapter 4).

Summary of Research Question 2 and Findings

How do statistics pedagogy learning experiences impact preservice mathematics teachers' statistics teaching self-efficacy?

The second research question dives deeper into how experiences make a difference in PSTs' statistics TSE. To answer the second research question, I was interested in the online module's impact on PSTs' beliefs about statistics education and about their ability to teach statistics. I analyzed participants' work on statistical investigations and their reflections on teacher practice and student thinking. By making connections with analysis and the entirety of the module experiences, I found that the module provided numerous opportunities for mathematics PSTs to encounter factors known to contribute to their perceived preparedness to teach statistics - their view of statistics, their statistical knowledge, their pedagogical knowledge, and their experiences with technology (Lovett & Lee, 2017).

In categorizing materials in the online module for different types of learning opportunities, most materials have the potential to be impactful in more than one factor. For example, when mathematics PSTs engaged in statistical investigations, those experiences had the potential to impact their view of statistics, their statistical and pedagogical knowledge, and their comfort with using technology. In order to verify that those opportunities indeed did have an impact, student work submitted in the module was analyzed for evidence of those four factors. In general, participants displayed developing abilities or knowledge for three of the four factors - view of statistics, pedagogical knowledge, and ability to use technology. Increasing statistical knowledge was not a primary goal of the module, and therefore was not explicitly analyzed.

By reading short papers, watching videos of experts and of classrooms, being introduced to a framework, discussing and reflecting on ideas with peers, and participating in statistical investigations, mathematics PSTs began to develop their view of statistics. This was displayed by the fact that they attended to important habits of mind while completing statistical investigations. The habits of mind that were most frequently included in investigations by PSTs were expressing uncertainty, taking into account variability, coordinating graphs and computations to reason about distributions, and reasoning quantitatively to make data-based arguments (Lee & Tran, 2015). However, PSTs did not always consistently express these habits of mind, and some habits of mind were not as frequently included in investigations at all, like attending to issues of sampling and measurement.

In addition, there were some statistical habits of mind that were noticeably absent from their investigations, like attending to measurement and sampling issues. In addition, PSTs were able to conduct statistical investigations that considered relevant contexts, and were exploratory in nature by examining different types of relationships among attributes, but their investigations were sometimes procedural and only based on a single measure, reflecting prior research that mathematics PSTs' statistical content knowledge is lacking (Lovett & Lee, 2017). In addition, this reflects prior research that it is hard for students to shift their perspective of data from the case to a distribution (Konold et al., 2015), but the design of the online materials made the beginning of this shift possible for some of the participants.

Particularly through analysis of video cases, and through reflection on statistical pedagogical concepts, mathematics PSTs began to develop their knowledge of statistics pedagogical ideas. For example, in their opportunities to approximate practice and to reflect on practice and student thinking, participants displayed developing beliefs about their own role as a teacher and the importance of attending to student thinking. For example, some mathematics PSTs took on the role of a teacher when recording their statistical investigations, whereas some used it more of a record of their own statistical investigation. In the former case, it provides evidence that those participants felt confident in the skills of teaching students how to investigate data. In addition, participants displayed a developing ability to reflect on teacher practice and considering student thinking. When they viewed and analyzed video cases, they most frequently described and inferred what was happening, but less often internalized their analysis and applied it to themselves. This indicates that PSTs were at the beginning of the process of attending to students' thinking (Wilson et al., 2011). In Hiebert et al.'s model of teaching as experimentation, which describes a full cycle of attending to student thinking as planning to collect, collecting, analyzing, and interpreting student thinking (2003). I would argue that analysis of video cases lands in the third and fourth phases of the cycle. However, these participants did not display evidence that they were yet fully capable of analyzing student thinking, restructuring their own thinking based on that analysis, and planning for future teaching based on that analysis and restructuring.

Finally, through learning about and using technology tools, mathematics PSTs began to develop a belief in the affordances of using dynamic statistical software to investigate data. This was evidenced by the fact that almost all PSTs were able to take advantage of basic functions of the tool in order to complete their investigations; however, features that were more advanced were only sometimes displayed.

Evidence from the design of the ESTEEM online module, in addition to evidence from PSTs' submitted work, indicates that the module was effective in providing opportunities for experiences related to the four factors that impact STSE. These factors impact STSE *through* mastery, vicarious, social persuasion, and physiological or affective experiences. For example, when PSTs completed a written statistical investigation where they explored a large data set of roller coasters, they were exposed to a specific view of statistics, given the opportunity to increase statistics and pedagogical knowledge, and use technology *through* an experience that could be considered mastery for doing statistics or vicarious for teaching statistics. Based on post-module interviews, PSTs attributed the module generally, and experiences with the module specifically, as positively impacting their confidence to teach statistics, which reflects prior research on the effectiveness of mathematics PSTs' viewing video cases (Santagata & Yeh, 2014) and best practices of online teacher education online, like spaces for reflection and dialogue, and models for good instructional practice (Burns, 2011).

Limitations

One limitation of this study is that it is based on a convenience sample of PSTs whose instructors volunteered to field test the ESTEEM foundational module, and so results may not be generalizable to other secondary mathematics PSTs. In addition, the small sample size also limits the generalizability. However, the sample includes participants from two different institutions, and so results are not institution- or instructor-specific. In addition, statistics experiences of participants were not so uncommon as to differently impact statistics teaching self-efficacy from a typical secondary mathematics PST.

Another limitation of this study is that I, as a researcher on this study, was also part of a project team charged with developing the ESTEEM resources. Although my involvement with the creation of the resources may indicate that I am more likely to have a favorable view of them, I believe that my intimate knowledgeable about the design of the resources also gives a unique perspective that an outside researcher may not have. For example, I was able to examine whether purposeful design decisions were effective when analyzing participant work.

Implications

The results of this study provide insight into the beliefs of secondary preservice mathematics teachers about statistics education and their personal teaching self-efficacy, and how different types of experiences impact those beliefs. Based on the results, some implications follow both for teacher preparation, and for research.

First, this study provides evidence that an expertly designed two-week module aimed at preparing mathematics PSTs to teach statistics is impactful for both general statistics beliefs and personal teaching efficacy beliefs. Thus, teacher preparation programs wishing to improve statistics teaching self-efficacy, but that do not have vast amounts of time to dedicate to statistics, should consider using the module in teacher preparation.

However, the brief module was not enough to completely change undesirable beliefs or to instill desirable beliefs. For example, using the statistical investigative cycle to teach was not mentioned by participants in interviews after the module. Because the statistical investigative cycle mirrors ways of working in many fields of study, teacher preparation programs should consider incorporating the model throughout PST experiences.

The ESTEEM foundational module was designed to be fully implemented online, but also flexible enough to be implemented in a variety of contexts. In this study, two hybrid contexts were examined - both with mostly asynchronous online activity, but one with some synchronous online meetings, and the other with some synchronous face to face meetings. The module was impactful in both types of implementation.

In terms of the design of the module, some slight changes may prove helpful. For example, in order to develop mathematics PSTs pedagogical knowledge even more, prompts encouraging PSTs to move further in the progression of attending to student thinking via video cases can be included. In addition, specific advanced uses of statistical educational tools like CODAP can be modeled by instructors so that PSTs are exposed to their benefit and shift towards viewing technology as tools that can substantially change the way one does or thinks about mathematics or statistics (reorganizers), rather than as tools that simply make doing math or statistics quicker or more efficient (amplifiers) (Hollebrands, 2011).

One main implication for research is the usefulness of a teaching self-efficacy framework specifically created for statistics education. The framework included in this study has two views - one more general view that includes the entire cycle, and one zoomed-in view that provides more detail on the pre-teaching time period. The frameworks proved useful in guiding research focused on both beliefs and sources of those beliefs, specifically for statistics education. Improvements on the framework based on this research (see Figure 8 in Chapter 4) are two-fold. First, because mastery and vicarious experiences were those mostly described by participants, with social persuasion rarely emerging, and physiological and affective states mostly being described as impacts, not experiences, there should be some indication in the framework as to the importance of the four sources. Second, because there are some experiences, like online readings or instructional videos, that do not fall neatly within one of the four categories, further discussion of where those types of experiences belong is warranted. Those types of materials tended to impact general statistics education beliefs, more than personal teaching efficacy beliefs, and so that relationship should also be reflected in the framework.

Future Research Directions

For mathematics teacher educators, the need for preparing confident statistics teachers is great. However, providing online learning opportunities where PSTs can engage in statistical tasks using dynamic statistical software and consider pedagogical concerns has been shown to be effective in positively impacting self-efficacy and changing beliefs about statistics education. Results from this collective case study suggest future research to be done.

The first branch of future research to be done concerns the impact of the online context; how does the design of the module match with best practices of online teaching? How does the online component of the module impact its effectiveness? Research that considers the impact of the module in different contexts, like face-to-face or completely online, versus the hybrid model used here is also something to be considered, especially because many teacher preparation programs still use mostly face-to-face experiences. For example, incorporating synchronous collaboration via conferencing software allows for small group discussion through the use of breakout rooms, electronic responses from participants in the form of emojis or checkmarks, time for individual work, choral responses in chats, and the use of whiteboards (Starling & Lee, 2015). Future research can study the differences in outcomes of a face-to-face implementation and an online implementation that incorporates such affordances of virtual synchronous interactions.

In addition, the research done here should be considered for a larger group of mathematics PSTs from a wide range of institutions. The ESTEEM module has been implemented in institutions across the United States in a variety of contexts, so further research should be done on the generalizability of the current study's results in order to confirm the effectiveness of the materials in impacting mathematics PSTs' STSE. This research might mean differentiating between different grade levels and courses of implementation. It might also mean using different types of cases, based on PSTs' past experiences with statistics.

Finally, in order to examine the impact of statistics experiences in general, and the ESTEEM foundational module in particular, on teaching practice, it will be important to follow PSTs into classrooms to see impact on practice. In general, teaching self-efficacy is important in the way that it impacts a teacher's practice, which impacts student outcomes. Because beliefs may shift when PSTs enter a classroom, due to school and/or classroom context, examining the long-lasting impact of prior experiences, including the ESTEEM foundational module, on actual practices will give a more complete picture of the cycle of teaching self-efficacy development (Figure 8).

REFERENCES

Allen, I. E., & Seaman, J. (2017). Digital Learning Compass: Distance Education Enrollment Report 2017. Babson Survey Research Group. https://onlinelearningsurvey.com/reports/digtiallearningcompassenrollment2017.pdf

American Statistical Association. (2005). *Guidelines for Assessment and Instruction in Statistics Education: College Report.* Alexandria, VA: Author.

American Statistical Association. (2016). *Guidelines for Assessment and Instruction in Statistics Education College Report 2016*. Available at http://www.amstat.org/asa/files/pdfs/GAISE/GaiseCollege_Full.pdf

Appova, A., & Taylor, C. E. (2019). Expert mathematics teacher educators' purposes and practices for providing prospective teachers with opportunities to develop pedagogical content knowledge in content courses. *Journal of Mathematics Teacher Education, 22*, 179–204. https://doi.org/10.1007/s10857-017-9385-z

Armor, D., Conroy-Oseguera, P., Cox, M., King, N., McDonnell, L., Pascal, A., Pauly, E., & Zellman, G. (1976). *Analysis of the school preferred reading programs in selected Los Angeles minority schools* (Rep. No. R-2007-LAUSD). Santa Monica, CA: RAND. (ERIC Document Reproduction Service No. 130 243)

Arnold, P. (2008). Developing new statistical content knowledge with secondary school mathematics teachers. In Batanero, C., 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, ICMI and IASE, Monterrey. Online: www.stat.auckland.ac.nz/~iase/publications.

Ball, D., Thames, M. H., & Phelps, G. (2008). Content knowledge for teaching: What makes it special? *Journal of Teacher Education*, *59*(5), 389-407.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84*, 191–215.

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

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

Bates, A. B., Latham, N., & Kim, J. A. (2011). Linking preservice teachers' mathematics self-efficacy and mathematics teaching efficacy to their mathematical performance. *School Science and Mathematics*, *111*(7), 325—333.

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. Online: www.aare.edu.au/99pap/beg99082.htm.

Bersin, J. (2004). *The blended learning book: best practices, proven methodologies, and lessons learned.* San Francisco, CA: Wiley.

Borba, M. C., Clarkson, P., & Gadanidis, G. (2013). Learning with the use of the internet. In M. A. (Ken) C. et Al. (Ed.), *Third International Handbook of Mathematics Education* (pp. 691–720). New York: Springer Science+Business Media. https://doi.org/10.1007/978-1-4614-4684-2

Borba, M. C., & Llinares, S. (2012). Online mathematics teacher education: Overview of an emergent field of research. *ZDM Mathematics Education, 44*, 697–704. https://doi.org/10.1007/s11858-012-0457-3

Bukor, E. (2011). *Exploring teacher identity: Teachers' transformative experiences of re-constructing and re-connecting personal and professional selves* (Doctoral dissertation). Retrieved from http://hdl.handle.net/1807/31700.

Burns, M. (2011). Distance education for teacher training: Modes, models, and methods. Washington, DC: EDC. Retrieved from: http://tinyurl.com/h7umzy8

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

Chance, B., Ben-Zvi, D., Garfield, J., & Medina, E. (2007). The Role of technology in improving student learning of statistics. *Technology Innovations in Statistics Education*, 1(1). Retrieved from https://escholarship.org/uc/item/8sd2t4rr

Cochran-Smith, M., & Villegas, A. M. (2015). Framing teacher preparation research: An overview of the field, part 1. *Journal of Teacher Education, 66*(1), 7–20. https://doi.org/10.1177/0022487114549072

Coles, A. (2014). Mathematics teachers learning with video: The role, for the didactician, of a heightened listening. *ZDM Mathematics Education, 46*, 267–278. https://doi.org/10.1007/s11858-013-0541-3

Connor, D., Davies, N., & Holmes, P. (2006). Using real data and technology to develop statistical thinking. In G. Burrill (Ed.), *Thinking and reasoning with data and chance: Sixty-eighth yearbook* (pp. 185-194). Reston, VA: National Council of Teachers of Mathematics.

Cooney, T. J., Shealy, B. E., & Arvold, B. (1998). Conceptualizing belief structures of preservice secondary mathematics teachers. *Journal for Research in Mathematics Education, 29*, 306–333.

Creswell, J.W. (2013). Qualitative inquiry and research design: Choosing among five approaches. Los Angeles, CA: SAGE Publications, Inc.

DeCuir-Gunby, J. T., Marshall, P. L., & McCulloch, A. W. (2011). Developing and using a codebook for the analysis of interview data: An example from a professional development research project. *Field Methods*, *23*(2), 136–155. http://doi.org/10.1177/1525822X10388468

Eichler, A. (2008). Teachers' classroom practice and students' learning. 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*. Monterrey: ICMI and IASE.

Enochs, L. G., Smith, P. L., & Huinker, D. (2000). Establishing factorial validity of the mathematics teaching efficacy beliefs instrument. *School Science and Mathematics, 100*, 194–202.

Ernest, P. (1989). The knowledge, beliefs and attitudes of the mathematics teacher: A model. *Journal of Education for Teaching*, *15*(1), 13–33.

Estrada, A., & Batanero, C. (2008). Explaining teachers' attitudes towards statistics. 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.* Monterrey, Mexico: International Commission on Mathematical Instruction and

International Association for Statistical Education. Online: www.stat.auckland.ac.nz/~iase/publications.

Estrada, A., Batanero, C., & Lancaster, S. (2011). Teachers' attitudes towards statistics. In C. Batanero (Ed.). *Teaching statistics in school mathematics-challenges for teaching and teacher education* (pp. 163–174). Springer, Dordrecht.

Fetler, M. (1999). High school staff characteristics and mathematics test results. *Education Policy Analysis Archives*, *7*(9), 11.

Fitzmaurice, O., Leavy, A., & Hannigan, A. (2014). Why is statistics perceived as difficult and can practice during training change perceptions? Insights from a prospective mathematics teacher. *Teaching Mathematics and Its Applications, 33*(4), 230-248.

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

Franklin, C.A., Bargagliotti, A.E., Case, C.A., Kader, G.D., Scheaffer, R.L., & Spangler, D.A. (2015). *The statistical education of teachers (SET)*. Alexandria, VA: American Statistical Association. www.amstat.org/education/SET/

Gal, I., Ginsburg, L., & Schau, C. (1997). Monitoring attitudes and beliefs in statistics education. *The assessment challenge in statistics education*, *12*, 37–51.

Gaudin, C., & Chaliès, S. (2015). Video viewing in teacher education and professional development: A literature review. *Educational Research Review, 16*, 41–67. https://doi.org/10.1016/j.edurev.2015.06.001

Gerges, G. (2001). Factors influencing preservice teachers' variation in use of instructional methods: Why is teacher efficacy not a significant contributor? *Teacher Education Quarterly, 28*(4), 71–88.

Gibson, S., & Dembo, M. (1984). Teacher efficacy: a construct validation. *Journal of Educational Psychology, 76*, 569–582.

Glaser, B. G. (1965). The constant comparative method of qualitative analysis. *Social problems, 12*(4), 436-445.

Hammerman, J. K., & Rubin, A. (2004). Strategies for managing statistical complexity with new software tools. *Statistics Education Research Journal*, *3*(2), 17-41.

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.

Harrell-Williams, L., M., Sorto, M. A., Pierce, R., Lesser, L. M., & Murphy, T. J. (2014). Using the SETS instruments to investigate sources of variation in levels of pre-service teacher efficacy to teach statistics. In K. Makar, B. de Sousa, & R. Gould (Eds.), *Sustainability in statistics education. Proceedings of the Ninth International Conference on Teaching Statistics (ICOTS9, July 2014) Flagstaff, Arizona, USA*. Voorburg, The Netherlands: International Statistical Institute.

Harrell-Williams, L. M., Lovett, J. N., Lee, H. S., Pierce, R. L., Lesser, L. M., & Sorto, M. A. (2019). Validation of scores from the high school version of the Self-Efficacy to Teach Statistics instrument using preservice mathematics teachers. *Journal of Psychoeducational Assessment, 37*(2), 194-208.

Harrison, T., Azmy, C., & Lee, H. S. (2018). Online learning experiences and impact on statistics education perspectives. In Hodges, T.E., Roy, G. J., & Tyminski, A. M. (Eds.). (2018). *Proceedings of the 40th annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (pp 751 – 754). Greenville, SC: University of South Carolina & Clemson University.

Henson, R. K., Kogan, L. R., & Vacha-Haase, T. (2001). A reliability generalization study of the teacher efficacy scale and related instruments. *Educational and Psychological Measurement*, *61*(3), 404–420.

Hiebert, J., Morris, A. K., Berk, D., & Jansen, A. (2007). Preparing teachers to learn from teaching. *Journal of Teacher Education, 58*(1), 47–61. https://doi.org/10.1177/0022487106295726

Hiebert, J., Morris, A. K., & Glass, B. (2003). Learning to learn to teach: An "experiment" model for teaching and teacher preparation in mathematics. *Journal of Mathematics Teacher Education, 6*, 201–222.

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.

Hollebrands, K. (2016). Technology as an Amplifier or Reorganizer: Posing Questions to Capitalize on the Power of Technology. In Teaching mathematics with technology MOOC-Ed, Friday Institute for Educational Innovation: NC State University, Raleigh, NC. Retrieved from: http://fi-courses.s3.amazonaws.com/tmt/unit_2/Amplifier_Reorganizer.pdf

Horton, N.J. (2015). Challenges and opportunities for statistics and statistics education: Looking back, looking forward. *The American Statistician, 69*, 138-145.

Hoyles, C. (1992). Mathematics teaching and mathematics teachers: A meta-case study. *For the Learning of Mathematics, 12*(3), 32–44

Hudson, R., Lee, H. S., Casey, S., Finzer, B., Mojica, G. M., Azmy, C., & Eide, A. (2018). Designing e-modules to support preservice mathematics teachers' statistical thinking. In M. A. Sorto, A. White, & L. Guyot (Eds.), *Looking back, looking forward. Proceedings of the Tenth International Conference on Teaching Statistics* (ICOTS10, July, 2018), Kyoto, Japan. Voorburg, The Netherlands: International Statistical Institute. Online: https://iase-web.org/icots/10/proceedings/pdfs/ICOTS10_9A2.pdf.

Irakleous, P., & Panaoura, A. (2015). Personal teaching efficacy in teaching statistics: A case study. *Mathitudes, 1*(1), 1–18.

Kastberg, S., Lynch-Davis, K., & D'Ambrosio, B. (2014). Examining mathematics teacher educators' emerging practices in online environments. *Contemporary Issues in Technology and Teacher Education, 14*(4), 384–400. Retrieved from https://www.learntechlib.org/primary/p/131301/

Konold, C., Higgins, T., Russell, S. J., & Khalil, K. (2015). Data seen through different lenses. *Educational Studies in Mathematics, 88*, 305–325. http://doi.org/10.1007/s10649-013-9529-8

Kwasny, M. J. (2015). Statistics in K-12: Educators, students, and us: Challenges and resources for K-12 educators and students. *CHANCE*, *28*(4), 26–29. http://doi.org/10.1080/09332480.2015.1120124

Lee, H. S., Kersaint, G., Harper, S. R., Driskell, S. O., Jones, D. L., Leatham, K. R., ... & Adu-Gyamfi, K. (2014). Teachers' use of transnumeration in solving statistical tasks with dynamic statistical software. *Statistics Education Research Journal, 13*(1), 25-52.

Lee, H. S., & Tran, D. (2015). Statistical habits of mind. In *Teaching statistics through data investigations MOOC-Ed*, Friday Institute for Educational Innovation: NC State University, Raleigh, NC. Retrieved from https://s3.amazonaws.com/fi-courses/tsdi/unit_2/Essentials/Habitsofmind.pdf

Lovett, J. N. (2016). *The preparedness of preservice secondary mathematics teachers to teach statistics: A cross-institutional mixed methods study* (Unpublished doctoral dissertation). North Carolina State University, Raleigh. Available at: https://repository.lib.ncsu.edu/handle/1840.16/11008

Lovett, J. N., & Lee, H. S. (2017). New standards require teaching more statistics. *Journal of Teacher Education*, 1–13. https://doi.org/10.1177/0022487117697918

Lovett, J. N., & Lee, H. S. (2018). Preservice secondary mathematics teachers' statistical knowledge: A snapshot of strengths and weaknesses. *Journal of Statistics Education, 26*(3), 214–222. https://doi.org/10.1080/10691898.2018.1496806

Makar, K., & Confrey, J. (2004). Secondary teachers' reasoning about comparing two groups. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning, and thinking* (pp. 353–373). Dordrecht, the Netherlands: Kluwer Academic Publisher.

Martyn, M. (2003). The hybrid online model: Good practice. *Educause Quarterly*. Retrieved from http://www.educause.edu/ir/library/pdf/EQM0313.pdf

McCulloch, A.W., Hollebrands, K., Lee, H. S., Harrison, T., & Mutlu, A. (2018). Factors that influence secondary mathematics teachers' integration of technology in mathematics lessons. *Computers & Education, 123*, 26—40.

McLeod, D. B. (1992). Research on affect in mathematics education: A reconceptualization. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 575–596). New York: Macmillan

McNamara, A. (2015). *Bridging the gap between tools for learning and for doing statistics* (Doctoral dissertation). University of California, Los Angeles.

Means, B., Toyama, Y., Murphy, R., & Baki, M. 2013. The effectiveness of online and blended learning: A meta-analysis of the empirical literature. *Teachers College Record*, *115*(030303), 1–47.

Mojica, G. M., Lee, H. S., Lovett, J. N., & Azmy, C. A. (July 2018). Impacts of a Teaching Statistics MOOC on educators' perspectives and practice. In M. A. Sorto, A. White, & L. Guyot (Eds.), *Looking back, looking forward. Proceedings of the Tenth International Conference on Teaching Statistics* (ICOTS10, July, 2018), Kyoto, Japan. Voorburg, The Netherlands: International Statistical Institute. Online: https://icots.info/10/proceedings/pdfs/ICOTS10_C159.pdf

Moore, D. S. (1997). New pedagogy and new content: The case of statistics. *International Statistical Review / Revue Internationale de Statistique*, *65*(2), 123. http://doi.org/10.2307/1403333

Morris, D. B., Usher, E. L., & Chen, J. A. (2017). Reconceptualizing the sources of teaching self-efficacy: A critical review of emerging literature. *Educational Psychology Review*, *29*(4), 795–833.

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

National Governors Association Center for Best Practice & Council of Chief State School Officers. (2010). *Common core state standards for mathematics*. Washington D.C.: Author.

Niess, M., & Gillow-Wiles, H. (2013). Developing asynchronous online courses: Key instructional strategies in a social metacognitive constructivist learning trajectory. *The Journal of Distance Education*, *27*(1), 1–17.

Noesgaard, S. S., & Ørngreen, R. (2015). The effectiveness of e-learning: An explorative and integrative review of the definitions, methodologies and factors that promote e-Learning effectiveness. *Electronic Journal of E-Learning*, *13*(4), 278–290.

Philipp, R. (2007). Mathematics teachers' beliefs and affect. In F. Lester (Ed.), *Second handbook of research in mathematics teaching and learning*. New York: Information Age.

Pierce, R., & Chick, H. (2011). Teachers' beliefs about statistics education. In *Teaching statistics in school mathematics-Challenges for teaching and teacher education* (pp. 151–162). Springer, Dordrecht.

Ponte, J. (2011). Preparing teachers to meet the challenges of statistics education. In C. Batanero, G. Burrill, & C. Reading (Eds.), *Teaching Statistics in School Mathematics-Challenges for Teaching and Teacher Education A Joint ICMIIASE Study* (pp. 299–309). Springer Science+Business Media.

Prodromou, T., & Pratt, D. (2006). The role of causality in the co-ordination of two perspectives on distribution within a virtual simulation. *Statistics Education Research Journal*, *5*(2), 69-88.

Ross, J., & Bruce, C. (2007). Professional development effects on teacher efficacy: Results of randomized field trial. *The Journal of Educational research*, *101*(1), 50–60.

Rossman, A., Chance, B., & Medina, E. (2006). Some important comparisons between statistics and mathematics, and why teachers should care. In G. Burrill (Ed.), *Thinking and Reasoning with Data and Chance: Sixty-eighth Yearbook* (pp. 323-333). Reston, VA: National Council of Teachers of Mathematics.

Rotter, J. B. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs, 80*, 1–28.

Santagata, R., & Yeh, C. (2014). Learning to teach mathematics and to analyze teaching effectiveness: Evidence from a video- and practice-based approach*. Journal of Mathematics Teacher Education, 17*, 491–514. https://doi.org/10.1007/s10857-013-9263-2

Scheaffer, R. (2006). Statistics and mathematics: On making a happy marriage. In G. Burrill (Ed.), *Thinking and Reasoning with Data and Chance: Sixty-eighth Yearbook* (pp. 309-321). Reston, VA: National Council of Teachers of Mathematics.

Shea, P., & Bidjerano, T. (2010). Learning presence: Towards a theory of self-efficacy, self-regulation, and the development of a communities of inquiry in online and blended learning environments. *Computers and Education*, *55*, 1721–1731. https://doi.org/10.1016/j.compedu.2010.07.017

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

Smith III, J. P. (1996). Efficacy and teaching mathematics by telling: A challenge for reform. *Journal for Research in Mathematics Education*, 387–402.

Starling, T., & Lee, H. (2015). Synchronous online discourse in a technology methods course for middle and secondary prospective mathematics teachers. *Contemporary Issues in Technology and Teacher Education, 15*(2), 106–125.

Thompson, A. G. (1992). Teachers' beliefs and conceptions: A synthesis of the research. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 127–146). New York: Macmillan.

Tschannen-Moran, M., & McMaster, P. (2009). Sources of self-efficacy: Four professional development formats and their relationship to self-efficacy and implementation of a new teaching strategy. *The Elementary School Journal*, *110*(2), 228–245.

Tschannen-Moran, M., & Woolfolk Hoy, A. (2001). Teacher efficacy: capturing an elusive construct. *Teaching and Teacher Education, 17*, 783–805.

Tschannen-Moran, M., & Woolfolk Hoy, A. (2007). The differential antecedents of self-efficacy beliefs of novice and experienced teachers. *Teaching and Teacher Education*, *23*(6), 944–956.

Tschannen-Moran, M., Woolfolk Hoy, A., & Hoy, W. K. (1998). Teacher efficacy: Its meaning and measure. *Review of Educational Research*, *68*(2), 202–248.

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*, 305–337.

Wayne, A. J., & Youngs, P. (2003). Teacher characteristics and student achievement gains: A review. *Review of Educational Research*, *73*(1), 89–122.

Wheatley, K. F. (2005). The case for reconceptualizing teacher efficacy research. *Teaching and Teacher Education*, *21*, 747–766.

Wild, C. (2006). The concept of distribution. *Statistics Education Research Journal*, *5*(2), 10–26. http://www.stat.auckland.ac.nz/~iase/serj/SERJ5(2)_Wild.pdf

Wild, C. J., & Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. *International Statistical Review / Revue Internationale de Statistique*, *67*(3), 223–256. http://doi.org/10.2307/1403699

Wilkins, J. L. M. (2008). The relationship among elementary teachers' content knowledge, attitudes, beliefs, and practices. *Journal of Mathematics Teacher Education*, *11*(2), 139–164.

Wilson, M. S., & Cooney, T. J. (2002). Mathematics teacher change and development: The role of beliefs. In G. C. Leder, E. Pehkonen, & G. Torner (Eds.), *Beliefs: A hidden variable in mathematics education?* (Vol. 31, pp. 127-148). London: Kluwer Academic Publishers.

Wilson, P. H., Lee, H. S., & Hollebrands, K. F. (2011). Understanding prospective mathematics teachers' processes for making sense of students' work with technology. *Journal for Research in Mathematics Education, 24*(1), 39–64.

Wyatt, M. (2014). Towards a re-conceptualization of teachers' self-efficacy beliefs: Tackling enduring problems with the quantitative research and moving on. *International Journal of Research and Method in Education*, *37*(2), 166–189.

Wyatt, M. (2015). Using qualitative research methods to assess the degree of fit between teachers' reported self-efficacy beliefs and their practical knowledge during teacher education. *Australian Journal of Teacher Education*, *40*(1), 7.

Zan, R., Brown, L., Evans, J., & Hannula, M. S. (2006). Affect in mathematics education: An introduction. *Education Studies in Mathematics*, *63*(2), 113–121.

Zaslavsky, O. (2008). Meeting the challenges of mathematics teacher education through design and use of tasks that facilitate teacher learning. In B. Jaworski & T. Wood (Eds.), *The Mathematics Teacher Educator as a Developing Professional* (pp. 93–114). Sense Publishers.

Zee, M., & Koomen, H. M. (2016). Teacher self-efficacy and its effects on classroom processes, student academic adjustment, and teacher well-being: A synthesis of 40 years of research. *Review of Educational Research*, *86*(4), 981–1015.

APPENDICES

Appendix A - Informed Consent Form

North Carolina State University
INFORMED CONSENT FORM for RESEARCH (ESTEEM College Student Interviews)
This consent information is valid [date] through [date]

Title of Study: Enhancing Statistics Teacher Education with E-Modules [ESTEEM]

Principal Investigator: Hollylynne S. Lee   Faculty Sponsor (if applicable): N/A

What are some general things you should know about research studies?
You are being asked to take part in a research study. Your participation in this study is voluntary. You have the right to be a part of this study, to choose not to participate or to stop participating at any time without penalty. The purpose of research studies is to gain a better understanding of a certain topic or issue. You are not guaranteed any personal benefits from being in a study. Research studies also may pose risks to those that participate. In this consent form you will find specific details about the research in which you are being asked to participate. If you do not understand something in this form it is your right to ask the researcher for clarification or more information. A copy of this consent form will be provided to you. If at any time you have questions about your participation, do not hesitate to contact the researcher(s) named above.

What is the purpose of this study?
The ESTEEM project has three primary goals: 1) create online resources for statistics pre-service teacher education, 2) design modules and approaches for using these resources, and 3) implement resources and modules in undergraduate mathematics teacher education programs. The purpose of this study is to better understand the value and utility of resources that have been created. Additionally, the study seeks to better understand actual integration of resources into curriculum; barriers and support factors for integration of resources into curriculum; and the impact of ESTEEM resources on pre-service mathematics teachers' self-perception of preparedness to teach statistics.

What will happen if you take part in the study?
If you agree to participate in this component of the study, you may be invited to take part in personal interview. If you choose to participate, the research team will contact you to schedule an interview. Sessions will be audio-recorded. The recordings will be transcribed and used for analysis. A semi-structured interview protocol will be used to guide the interview. You will be asked about various issues, including your perceptions of the value and utility of ESTEEM resources. While the exact duration will vary, we anticipate that interviews will last approximately 45 minutes.

Risks
As with any discussion among professionals, participants could conceivably experience discomfort or uncertainty relating to topics or questions raised. This however is no larger a risk to any individual than any routine online or personal discussion such professionals could encounter in their daily professional lives, so does not represent any risk particular or unique to this project.

Benefits
No specific direct benefits are predicted from participation in the IRB-approved research. However, consistent with the intent of the ESTEEM Project, you may gain your own insight into particular statistical tasks, strategies, and knowledge that you would not have otherwise gained, may become more engaged with statistics as a topic of interest, and may broaden your understanding of student learning and statistical instruction, become more observant of statistical activity in the classroom, and/or become more confident in your educational practice.

Confidentiality
All interviews will be digitally recorded and immediately downloaded to password-protected computers. The recordings will be transcribed and the transcriptions will be stored securely on Friday Institute or NCSU-contracted servers. Data will be accessible to project staff and NCSU technical support staff only. All identifying information of individual participants will be removed from the transcriptions. What is learned from the research project will be included in research presentations, publications, and reports. To ensure privacy and confidentiality, all participants will be identified by pseudonyms. Publications will NOT contain any information that identifies you.

Compensation
You will not receive anything for participating.

What if you are a NCSU student?
Participation in this study is not a course requirement and your participation or lack thereof, will not affect your class standing or grades at NC State.

What if you are a NCSU employee?
Participation in this study is not a requirement of your employment at NCSU, and your participation or lack thereof, will not affect your job.

What if you have questions about this study?
If you have questions at any time about the study or the procedures, you may contact the director of the research, Dr. Hollylynne Lee (919.513.3544, hollylynne@ncsu.edu) at the Friday Institute for Educational Innovation, North Carolina State University, 1890 Main Campus Road, Campus Box 7249, Raleigh, NC 27606.

What if you have questions about your rights as a research participant?
If you feel you have not been treated according to the descriptions in this form, or your rights as a participant in research have been violated during the course of this project, you may contact Deb Paxton, Regulatory Compliance Administrator, Box 7514, NCSU Campus (919/515-4514).

Consent To Participate
I have read and understand the above information. I have received a copy of this form. I agree to participate in this study with the understanding that I may choose not to participate or to stop participating at any time without penalty or loss of benefits to which I am otherwise entitled.

____________________________ ____________________________
Signature   Name (please print)

____
Date

Appendix B - SETS Survey Instrument

NC STATE UNIVERSITY

Instructions: Rate your confidence in *teaching high school students* the skills necessary to complete successfully the task given by selecting your choice on the following scale: 1 = not at all confident, 2 = only a little confident, 3 = somewhat confident, 4 = confident, 5 = very confident, 6 = completely confident.

For example,

	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.	○	○	○	✓	○	○

For the open-ended questions, please include as much detail as you feel comfortable sharing.

Your responses are voluntary and confidential. You may simply skip any question you are unable or unwilling to answer, but we hope that you will answer as many questions as possible.

Copyright 2009-2013 Research Cluster Study, Consortium for the Advancement of Undergraduate Statistics Education (CAUSE). Not for reproduction or use without written consent of Research Cluster members. Item development partially supported by NSF grant DUE-0618790.

0% 100%

>>

2. 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 high school students* the skills necessary to complete the following tasks successfully:

	Not at all confident (1)	Only a little confident (2)	Somewhat confident (3)	Confident (4)	Very confident (5)	Completely confident (6)
1. Collect data to answer a posed statistical question in contexts of interest to high school students.	○	○	○	○	○	○
2. Recognize that there will be natural variability between observations for individuals.	○	○	○	○	○	○
3. Select appropriate graphical displays and numerical summaries to compare individuals to each other and an individual to a group.	○	○	○	○	○	○
4. Create dotplot, stem and leaf plot, and tables (using counts) for summarizing distributions.	○	○	○	○	○	○
5. Use dotplot, stem and leaf plot, and tables (using counts) for describing distributions.	○	○	○	○	○	○
6. Create boxplots for summarizing distributions.	○	○	○	○	○	○
7. Use boxplots, median, and range for describing distributions.	○	○	○	○	○	○
8. Identify the association between two variables from scatterplots.	○	○	○	○	○	○
9. Generalize a statistical result from a small group to a larger group such as the whole class.	○	○	○	○	○	○
10. Recognize that statistical results may be different in another class or group.	○	○	○	○	○	○
11. Recognize the limitation of making inference (i.e. generalization) from a classroom dataset to any population beyond the classroom.	○	○	○	○	○	○

3. 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 high school students* the skills necessary to complete the following tasks successfully:

	Not at all confident (1)	Only a little confident (2)	Somewhat confident (3)	Confident (4)	Very confident (5)	Completely confident (6)
12. Distinguish between a question based on data that vary and a question based on a deterministic model (for example, specific values of rate and time determines a particular value for distance in the model d = r x t).	○	○	○	○	○	○
13. Identify what variables to measure and how to measure them in order to address the question posed.	○	○	○	○	○	○
14. Describe numerically the variability between individuals within the same group.	○	○	○	○	○	○
15. Create histograms for summarizing distributions.	○	○	○	○	○	○
16. Use histograms for comparing distributions.	○	○	○	○	○	○
17. Compute interquartile range and five-number summaries for summarizing distributions.	○	○	○	○	○	○
18. Use interquartile range, five-number summaries, and boxplots for comparing distributions.	○	○	○	○	○	○
19. Recognize the role of sampling error when making conclusions based on a random sample taken from a population.	○	○	○	○	○	○
20. Describe numerically the strength of association between two variables using linear models.	○	○	○	○	○	○

21. Explain the differences between two or more groups with respect to center, spread (for example, variability), and shape.	○	○	○	○	○	○
22. Recognize that a sample may or may not be representative of a larger population.	○	○	○	○	○	○
23. Interpret measures of association.	○	○	○	○	○	○
24. Distinguish between an observational study and a designed experiment.	○	○	○	○	○	○
25. Distinguish between "association" and "cause and effect."	○	○	○	○	○	○
26. Recognize sampling variability in summary statistics such as the sample mean and the sample proportion.	○	○	○	○	○	○

4. 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 high school students* the skills necessary to complete the following tasks successfully:

	Not at all confident (1)	Only a little confident (2)	Somewhat confident (3)	Confident (4)	Very confident (5)	Completely confident (6)
27. Describe characteristics of a normal distribution, such as general shape of distribution, symmetry, how standard deviation influences shape, and area under the curve.	○	○	○	○	○	○
28. Estimate percentages via the empirical rule (i.e., percentage of observations within 1, 2, or 3 standard deviations from the mean) using the mean and standard deviation of a dataset which has an approximately bell-shaped distribution.	○	○	○	○	○	○

29. Estimate a specified area under the normal curve using technology or a statistical table.	○	○	○	○	○	○
30. Summarize categorical data using two-way tables (i.e., contingency tables, frequency tables).	○	○	○	○	○	○
31. Calculate and interpret relative frequencies using two-way tables (i.e., contingency tables, frequency tables).	○	○	○	○	○	○
32. Find conditional and marginal frequencies from two-way tables (i.e., contingency tables, frequency tables).	○	○	○	○	○	○
33. Fit an appropriate model (e.g., linear, quadratic, or exponential) using technology for a scatterplot of two quantitative variables.	○	○	○	○	○	○
34. Assess the fit of a particular model informally by plotting and analyzing its residuals.	○	○	○	○	○	○
35. Identify the slope and y-intercept coefficients of a linear model and interpret them in the context of data.	○	○	○	○	○	○

5. 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 high school students* the skills necessary to complete the following tasks successfully:

	Not at all confident (1)	Only a little confident (2)	Somewhat confident (3)	Confident (4)	Very confident (5)	Completely confident (6)
36. Calculate, using technology, the correlation coefficient between two quantitative variables.	○	○	○	○	○	○
37. Evaluate whether a specified model is consistent with data generated from a simulation.	○	○	○	○	○	○

38. Explain the role of randomization in surveys, experiments and observational studies.	○	○	○	○	○	○
39. Describing purposes and differences among surveys, experiments, and observational studies.	○	○	○	○	○	○
40. Evaluate how well the conclusions of a study are supported by the study design and the data collected.	○	○	○	○	○	○
41. Estimate a population mean or proportion using data from a sample survey.	○	○	○	○	○	○
42. Develop a margin of error for an estimate of a population mean or proportion using simulation models.	○	○	○	○	○	○
43. Compare two treatments from a randomized experiment by exploring numerical and graphical summaries of data.	○	○	○	○	○	○
44. Determine if the difference between two population means or proportions is statistically significant using simulations.	○	○	○	○	○	○

0% 100%

>>

Open-Ended Questions:

a) Choose one topic that you indicated feeling LEAST confident about teaching high school students. Think about the reason(s) you feel this way. Use the space below to identify the topic and explain your reason(s).

b) Choose one topic that you indicated feeling MOST confident about teaching high school students. Think about the reason(s) you feel this way. Use the space below to identify the topic and explain your reason(s).

0% 100%

>>

1. Gender:

Male ○ Female ○

2. Institution Name:

3. Instructor Name:

4. Which option below best describes the course you are currently enrolled in while completing this survey?

○ A course focused entirely on teaching methods for teaching statistics

○ A course focused on methods for teaching mathematics that contains units or lessons focused on teaching statistics

○ A statistics course that includes lessons or assignments focused on teaching methods for statistics

○ Other. Please describe:

5. Upon completion of your teacher preparation program, which grade levels will you be certified to teach. If you already hold a license, please indicate the level for which you are certified. (Choose ONE response)

Middle School	High School	Both Middle and High School	Elementary
○	○	○	○

6. From the following list, which level degree program are you currently enrolled?

Bachelor's	Master's	Both a Bachelor's and a Master's	Doctoral	Non-Degree Post-Bachelor's Studies
○	○	○	○	○

7. Indicate the courses you have taken in the past (not including the current course) in which you have learned instructional strategies for teaching statistics topics. Check all that apply:

- ☐ A course focused entirely on teaching methods for statistics
- ☐ A course focused on teaching methods for mathematics that contained units or lessons focused on teaching statistics
- ☐ A statistics course that had lessons or assignments focused on teaching methods for statistics
- ☐ None
- ☐ Other:

8. Did you take AP Statistics in High School?

Yes	No
○	○

9. Including this semester/term, how many classes have you taken at a college or university in statistics. (Choose ONE response)

None	1 Class	2 Classes	3 Classes	4-5 Classes	6-7 Classes	8 or More Classes
○	○	○	○	○	○	○

10. Please name the statistics courses that you have taken at a college or university.

11. How well prepared do you feel to teach high school statistics topics required in either the Common Core State Standards or your local state's standards?

1 (Completely Unprepared)	2	3	4 (Somewhat prepared)	5	6	7 (Sufficiently prepared)	8	9	10 (Completely prepared)
○	○	○	○	○	○	○	○	○	○

**12. As a secondary mathematics teacher you have to be prepared to teach a wide variety of subjects: algebra, geometry, pre-calc/advanced algebra, calculus, and statistics.
Given these five different topics, please rank these in order of how well you feel prepared to teach them from most (1) to least (5). Drag the topic up or down to change its rank position.**

- Algebra
- Geometry
- Pre-Calc/Advanced Algebra
- Statistics
- Calculus 5

Appendix C - Discussion Forum/Reflection Prompts

Discussion forum: 1.1.h. Discuss learning statistics through investigations with real data

Reflect on what you learned in this module about teaching and learning statistics. How does this compare with your own prior learning experiences with statistics (and/or teaching if you have previously taught statistics)? Make explicit connections to readings, videos, and your experience in the investigations.

You must write an initial post (2 pts) and then respond to at least one post from a peer (2 pts), for a total of 4 pts.

Reflection: 1.2.i. Supporting statistical discourse with the roller coaster task

Now that you have used CODAP to investigate differences, trends, and relationships about characteristics of a sample of 157 US roller coasters, consider how students reason about the same statistical investigation. Watch as students in 6th grade, 7th grade and high school AP Statistics explore this data set. This is the first time that students used CODAP to conduct a statistical investigation.

https://www.youtube.com/watch?v=RvzAxKlHr0E&feature=youtu.be (embedded link)

Post a paragraph reflection (3-5 sentences) that focuses on one of the questions below. Your response should reflect what you have been learning about in this module in relation to supporting students' learning of statistics through investigations. Make explicit connections to readings, videos, and your experience in the investigations.

1. Compare and contrast how the four pairs of students reasoned statistically in relationship to the following ideas. Be sure to support your claims with evidence that includes what students did and said.

    a. The ways the context supports or hinders students' statistical thinking as they engage in the investigation.

    b. The ways they engaged in posing a question of interest to them.

    c. The ways the students were analyzing and interpreting the data at different levels of sophistication. (Hint: The order students appear in the video is not by levels of sophistication.)

2. Students used several visual and numerical representations to make sense of the data. In what ways did features in CODAP support or hinder students' statistical reasoning? Explain.

3. Throughout the video, the teacher asked students various questions while they engaged in the statistical investigation. Describe the role of the teacher and students, and explain the extent to which these interactions supported productive statistical thinking.

Discussion forum: 1.1.h. Discuss learning statistics through investigations with real data

While selecting a statistically rich task that ties together the learning goal, data, context, and investigative cycle is foundational in providing students opportunities to develop more sophisticated statistical thinking, it is as important that teachers consider the implementation of the task and how that implementation might promote reasoning that builds on productive habits of mind. Teachers can support students in developing statistical thinking by encouraging them to communicate their own ideas about engaging with data and consider the thinking of others through discourse. Smith and Stein (2011) elaborate on a model for supporting discourse about students' work on tasks which involves: anticipating students' responses to a task; monitoring students' responses to a task; selecting specific students to present mathematical ideas; sequencing students' responses that will be publicly displayed; and connecting between student responses and to key ideas. To learn more about how teachers can use students' work with tasks as the launching point for orchestrating productive whole class discussions that advance important statistical ideas, read this two page paper.

Part I: Watch a Classroom Statistical Investigation

Watch the following video, where a teacher launches a statistical investigation about roller coasters in a seventh grade classroom, in which students use CODAP for the first time. The teacher monitors student work, selects and sequences several students' findings to discuss, and leads a whole class discussion connecting students' statistical ideas.

https://youtu.be/ETNF_542DvU (embedded link)

Part II: Reflect on Students' Statistical Reasoning

Write a 1-2 page reflection in response to the following questions. Your response should reflect what you have been learning about in this module in relation to supporting students' statistical reasoning, including designing tasks and making sense of students' work. Make explicit connections to readings, videos, and your experience in the investigations.

a. Is there evidence that the teacher supported students' use of statistical habits of mind (e.g., role of context, sampling, attending to variability, measurement, being skeptical, accounting for uncertainty)? Provide specific examples.

b. In what ways did the teacher's interactions with pairs or the whole class build on students' thinking and move that reasoning forward? Explain.

c. How did the teacher sequence the order of students' sharing of their work to account for different student approaches to analysis and interpretations? Explain your thinking.

d. Did the teacher use student ideas to assist the class in making connections between the statistical ideas that were reflected in the shared strategies and representations? If so, how?

References

Smith, M. S., & Stein, M. K. (2011). 5 practices for orchestrating productive mathematics discussions. Reston, VA: National Council of Teachers of Mathematics.

Appendix D - Written Statistical Investigation

1.2.g. Investigating more roller coasters

ENHANCING STATISTICS TEACHER EDUCATION WITH E-MODULES

Mod 1: Part 2

Investigating MORE Roller Coasters Developed Over the Past 100 Years!

CONTEXT: Amusement Parks are located in various locations across the United States, and many Americans include a trip to an amusement park as a favorite vacation option. These parks often have one or more roller coasters. While not all students have ridden roller coasters before, they have seen them in the media, and the internet is full of videos made by roller coaster enthusiasts. Some parks have older coasters that they have continually maintained over the years, like the Jack Rabbit at Kennywood park, built in 1921 and still in operation today. Many parks, like Cedar Point in Ohio, also continually try to build new coasters to attract new visitors and keep their existing visitors coming back, like the Millenium Force which opened in 2000, shown in picture. Advances in engineering over the years have certainly expanded how coasters are built.

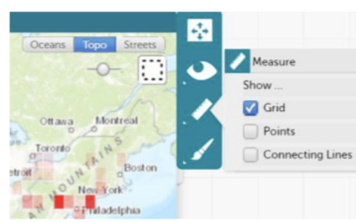
(image from: https://en.wikipedia.org/wiki/Millennium_Force)

In this investigation, you have an opportunity to use CODAP online to explore a sample of 157 roller coasters at parks across the United States.

Getting Familiar with the data set.

Open the data in CODAP. There are 157 cases, organized hierarchically based on when they were built (the attribute Age Group to the left of the table). Click on the different age groups and explore the table. You can make the Table window larger and also scroll to the right and down to see all 16 attributes and all 157 cases.

Review the attribute definitions by hovering over the name of an attribute in the column header. The definitions of all attributes are also explained more in the Appendix to this document.



Open the Map (click on Map on the top shelf). Each location of the amusement parks is indicated by a point. Click on the Map. In the Map menu (right-hand side), click on the ruler and change Point to Grid. The rectangular grids shown on the map are different shades of red, depending on how many coasters from that park are in the sample of data.

Click on various grids on the map for a location of an amusement park and then scroll through the table to see which roller coasters correspond.

Lee, H. S., & Mojica, G. F. (2017). Enhancing Statistics Teacher Education with E-Modules. Retrieved from: https://fi-esteem.s3.amazonaws.com/module1/part2/esteem_mod1_part2_roller_coaster_investigation.pdf

Exploratory Data Analysis

1. Use the Map and Table to find a park that is close to your location or one you have visited. Describe some interesting facts about the roller coasters in that amusement park.

2. Open two or more Graph windows (click twice on the Graph icon on the top shelf) and explore attributes about coasters that interest you (e.g., track length, inversions). Create several different graphs to explore any patterns and trends for these attributes. Record some things you noticed and insert screenshots of graphs you created.

Learning a Few Ways of Working in CODAP

Before moving on, take some time to learn how to work more with Tables and compute statistical measures for a parent-level (or top level) that organizes data in a hierarchy. You can change the hierarchy to be organized by a different attribute by dragging an attribute label in the column headers to the left (or right to remove a hierarchy). Even though the data is initially organized by Age Group as the first level of hierarchy, the table can be reorganized to suit a user's preference to help display cases in a meaningful way.

To remove a level of hierarchy in a table.

- In the column header, click and grab the name of the attribute and drag to the right until a black vertical bar appears between the two columns where you would like to move the attribute.
- When the mouse is released the cases will be reorganized accordingly.

To add a level of hierarchy in a table.

- To add a level of hierarchy, grab any categorical attribute name and drag it to the far left of the table within the yellow vertical box on the left. By dragging another categorical attribute to the left side of the table, additional hierarchies can be created.

On the next page, there are several more specific investigations to pursue.

Lee, H. S., & Mojica, G. F. (2017). Enhancing Statistics Teacher Education with E-Modules. Retrieved from: https://fi-esteem.s3.amazonaws.com/module1/part2/esteem_mod1_part2_roller_coaster_investigation.pdf

Comparing Groups

One of the essential ways to help students develop their ability to describe a distribution of data is to have them compare two or more groups. We can use categorical attributes to help do this. Students tend to naturally start asking questions, or wonder about, how something compares for different groups of data.

3. Choose **ONE** of the following sets of questions to investigate. Be sure to use different tools in the Graph menu to add measures to the graphs. Also use the Table to reorganize the data and compute statistical measures of interest. Insert screenshots and describe the work you did to answer the questions.

Set A of Investigative Questions: How fast do roller coasters tend to be?

- Are there any differences between the top speed of older, newer, or more recent roller coasters? Explain.
- Does the type of material make a difference in speed?
- If you were going to ride a coaster that was built before 1980, what would you expect for a typical top speed? What about if you were riding a newer roller coaster?

Set B of Investigative Questions: How does the track length of roller coasters vary?

- Is there a difference in track length between roller coasters that are inverted or not?
- Has track length of coasters changed much over the three age groups, from oldest, recent, to newest?
- If you were going to ride a coaster that was built before 1980, what would you expect for a typical track length? What about if you were riding a newer roller coaster?

Bivariate Association

Bivariate data is the term used to describe data that have two variables for each observation. The cases (observations) in our dataset have many variables (attributes). When examining two attributes in a data set, our attention is on how the values for those attributes co-vary. Thus for bivariate data, covariation involves correspondence of the variation possible in each variable. We could examine association between categorical variables or two quantitative variables.

4. Choose **ONE** of the following questions to investigate (and the follow-up question listed below it). Be sure to use different tools in the Graph menu to add measures to the Graphs. You could also use the Table to reorganize the data and compute statistical measures of interest. Insert screenshots and describe the work you did to answer the questions.

Set A Investigative Question: Is the type of material used to make a roller coaster related to whether or not the passengers get inverted during a ride?

- Are you more or less likely to get inverted on a Steel coaster? Explain why based on your data, and then explain why based on what you know, or can find out, about how coasters are built.

Set B Investigative Question: Is the length of the drop on a roller coaster related to its top speed?

- What other attributes in the design of a coaster may impact how drop and top speed are related?

Reflection (3-5 sentences)

What excites you and makes you nervous about implementing tasks like the roller coaster investigation in your own classroom? Make connections to what you experienced doing the investigation as well as what you have been learning about in this module related to supporting students' learning of statistics through investigations.

Lee, H. S., & Mojica, G. F. (2017). Enhancing Statistics Teacher Education with E-Modules. Retrieved from: https://fi-esteem.s3.amazonaws.com/module1/part2/esteem_mod1_part2_roller_coaster_investigation.pdf

Roller Coaster Data: Attributes and Definitions

#	Attribute Name	Description	Units
1	**Coaster**	Name of the roller coaster	
2	**Park**	Name of the park where the roller coaster is located	
3	**City**	City where the roller coaster is located	
4	**State**	State where the roller coaster is located	
5	**Type**	Material of track (Steel or Wooden)	
6	**Design**	How a passenger is positioned in the roller coaster	
	Design Types:	Bobsled - designed like a bobsled run -- without a fixed track. The train travels freely through a trough.	
		Flying - a roller coaster ridden while parallel with the track.	
		Inverted - a roller coaster which uses trains traveling beneath, rather than on top of, the track. Unlike a suspended roller coaster, an inverted roller coaster's trains are rigidly attached to the track.	
		Pipeline - a coaster where riders are positioned between rails instead of above or below.	
		Sit Down - a traditional roller coaster ridden while sitting down.	
		Stand Up - a coaster ridden while standing up instead of sitting down.	
		Suspended - a roller coaster using trains which travel beneath the track and pivot on a swinging arm from side to side, exaggerating the track's banks and turns.	
		Wing - a coaster where pairs of riders sit on either side of a roller coaster track in which nothing is above or below the riders.	
7	**Opened**	Year when roller coaster opened	
8	**Top Speed**	Maximum speed of roller coaster	mph
9	**Age Group:**	1:Older (Built between 1900-1979)	
		2:Recent (1980-1999)	
		3:Newest (2000-current)	
10	**Max Height**	Highest point of roller coaster	ft
11	**Drop**	Length of largest gap between high and low points of roller coaster	ft
12	**Length**	Length of roller coaster track	ft
13	**Duration**	Time length of roller coaster ride	seconds
14	**Inversions?**	Whether or not roller coaster flips passengers at any point (Yes or No)	
15	**# of Inversions**	Number of times roller coaster flips passengers	

Data Sources: rcdb.com wikipedia.com ultimaterollercoaster.com

Lee, H. S., & Mojica, G. F. (2017). Enhancing Statistics Teacher Education with E-Modules. Retrieved from: https://fi-esteem.s3.amazonaws.com/module1/part2/esteem_mod1_part2_roller_coaster_investigation.pdf

Appendix E - Screencast Statistical Investigation

Screencast Using Technology Tool: Statistics

It is important for teachers to go through the mathematical tasks before using them in their classroom. It is also important for teachers to develop their own learning investigating mathematical concepts using technological tools that they will be using in their classroom. This assignment will allow you to illustrate your learning of technology used during this course and how you can apply that to illustrate using a technology tool to investigate a statistics task.

Assignment:
Your assignment is to create a screencast of you using a technology tool while investigating your choice of one of the tasks on the **provided list below. The video should be less than 10 minutes.**

The purpose of the video is to illustrate how to use a technology tool to investigate a statistics question. You want to show off your ability to use **advanced and powerful features of a tool to go in-depth into an analysis**. *The language you use during the video should be statistically accurate and you should be modelling enacting statistical habits of mind, such as considering variation, understanding measurement, connecting to the context, using multiple representations and measures to display data, and being uncertain in your claims.*

Tasks to choose from are at the bottom of this document.
You can pick the screencast software of your choice. Suggested ones include:

Free	Soapbox (browser extension) Screencastomatic (web-based recorder and storage) QuickTime (already installed on your MAC!) Icecream recorder Apowershot (online recorder) but also have download versions for Mac and PC
Cheap	Snagit ($30 for students) with free trial for 30 days

Other options for Screencast software.

Your Screencast video should be housed on YouTube or other website hosting service (some of the above web-based screen recording tools host videos for free on their site). Then you need to provide a URL to link to the video, or ideally, embed the video on a page on your website. For example, here is a How To for WordPress.
You get to choose the format/layout of the page. Keep in mind your screencast will be public to your classmates.

Rubric for Screencast

<table>
<tr><th>Category</th><th>Points Possible</th><th>Points Earned</th></tr>
<tr><td>Mathematical Content Standards; provide the following on the webpage:
Identify A FEW Common Core mathematical standards that could be addressed by using the chosen task.</td><td>2</td><td></td></tr>
<tr><td>Technology use:
The ways technology is used should illustrate best practices, and advanced skills as learned through different course tasks and resources. (e.g., creating more than one graph, linking between multiple representations, using appropriate graphs and adding measures to a graph, computing measures in a table, creating new attributes based on a formula, arranging data hierarchically, viewing a map if appropriate)</td><td>10</td><td></td></tr>
<tr><td>Problem solving and language
The language you use during the video should be statistically accurate and you should be modelling enacting statistical habits of mind, such as considering variation, understanding measurement, connecting to the context, using multiple representations and measures to display data, and being uncertain in your claims</td><td>10</td><td></td></tr>
<tr><td>Quality and length of video
Video should be high quality (practice a few times), and under 10 minutes.
Video should be posted on YouTube or other videohosting service provided by a screencast recorder website so that it can be embedded on your website and a direct URL link provided.</td><td>3</td><td></td></tr>
<tr><td>TOTAL</td><td>25</td><td></td></tr>
</table>

Choose one of the following THREE tasks to create a screencast. You should PRACTICE doing the task several times before you record the screencast.

Choice 1: Open the sample of 157 roller coasters in CODAP
http://tinyurl.com/157UScoasters
Investigate the following questions related to how fast roller coasters tend to be:

- Are there any differences between the top speed of older, newer, or more recent roller coasters? Explain.
- Does the type of material make a difference in speed?
- If you were going to ride a coaster that was built before 1980, what would you expect for a typical top speed? What about if you were riding a newer roller coaster?

Go the following website https://rcdb.com/3762.htm and enter all the information you can find about this roller coaster in a new row in the data table. If you do not know a value for a given attribute, leave that cell empty.

- How does this new case fit within the analysis you just did? Does this new case change any claims or inferences you would make based on your analysis?

Choice 2: Open the 2015 US State Education data in CODAP
http://tinyurl.com/2015StateEducationData

The South and West seem like a great place to live because of the nice weather. But, how do the regions compare in things like average teacher salaries or student per teacher ratio? To investigate, let's do some comparisons of the school data in the South and West. Create a plot of the attribute Average_Salary. Since we want to compare the West and South, we can separate the cases in the plot to display four distributions, separated by the qualitative attribute Census_Region
Tech Tip: If you want to remove the Northeast and Midwest data points, select those cases and choose "Hide selected cases" from the Eye menu.

1. Based on the average teacher salaries for states in the South and the West, where would you prefer to teach and why? Use the various tools (under the Ruler in the menu) in the graph window as well as the map to investigate and visualize how these regions compare in average teacher salary. Be sure to illustrate appropriate technique and language for comparing distributions.

2. Explore the other quantitative attributes in this data set and compare the distributions for the South and West using graphs and the map. Based on the data you have examined, in which region would you prefer to teach and why? Provide a detailed description of your comparisons.

Choice 3:

Open **http://tinyurl.com/2015VehiclesSample** file in CODAP

In this problem you will be investigating relationships between Engine_Displacement and City_MPG. (read more about engine displacement if you need to understand more about how this is measured)

1. First determine if there is a strong linear relationship between Engine displacement and City mpg. Consider which attribute should be the predictor variable and which should be the response variable. Build a linear model that can be used for predictions. Explain your results statistically, and related to the context.
2. Investigate the relationship between Engine displacement and City mpg using a third attribute, Transmission type, and describe what additional information about the data and relationship this attribute provides.
3. Remove the Transmission type attribute from the scatterplot. Construct a plot of Residuals City mpg versus Engine displacement. Describe this plot and what it tells you about the relationship between City mpg and Engine displacement.

Appendix F - Autobiographical Survey Protocol

Thank you for participating in this study aimed at understanding preservice teachers' self-efficacy to teach statistics, with the ultimate goal of finding ways to make teachers feel more comfortable with this important topic in mathematics. The purpose of this survey is to gain insight into a timeline of your statistics career, critical experiences you may have had with statistics, and perception of what statistics is. Please remember, there are no right or wrong responses; these questions are to gain a better understanding of your perspectives. Thus, I ask that you please be as thorough as you can be in the detail of your responses.

1. How many and which statistics-related courses did you take during your undergraduate (or graduate, if applicable) studies? Please give me a little bit of information on the content of each course and the teaching methods used by the instructor.
2. How many and which mathematics or statistics *teaching* courses did you take during your undergraduate (or graduate, if applicable) studies?
3. Did you have the opportunity to teach statistical concepts during any of your filed experiences in high school or middle school classrooms? Please describe these experiences.
4. What are some of the *informal* experiences in which you have experienced statistics used in the media or in conversations with others in any non-school related setting (e.g. home, club, religious)?
5. Please describe your most memorable teenage experiences or experiences as a student in high school that were statistics related.
6. Please describe your most memorable early childhood experiences or experiences as a student in elementary or middle school that were statistics related.

Appendix G - Interview Protocol

Hello. Thank you for participating in this study aimed at understanding preservice teachers' self-efficacy to teach statistics, with the ultimate goal of finding ways to make teachers feel more comfortable with this important topic in mathematics. If you would like to stop at any time, please let me know. Please remember, there are no right or wrong responses. I am asking these questions to gain a better understanding of your perspectives.

1. I know you are preparing to be a high school or middle school math teacher. Where are you in your program?
2. What types of classes or field experiences do you have this semester?
3. When will you graduate?
4. In what ways do you believe teachers impact students' learning in math classrooms?
5. How confident are you in your abilities to impact what students learn?

Now we are going to focus more specifically about teaching statistics, which is only one of the areas of the mathematics curriculum you will be teaching.

6. What is your general feeling about statistics?
7. What characteristics do you think a good statistics teacher has?
8. Think back to times when you have been preparing a lesson to teach a statistics topic. How did you feel while you were preparing this lesson? This could have been in a course or while you were in a field experience.
9. How would you describe your general feeling about your ability to teach statistics to high school students?
10. In your survey, you stated ## early experiences with teaching statistics. How did these experiences impact your confidence to teach statistics?

11. In your survey, you stated that you have ## experience with teaching statistics. How did this experience impact your confidence to teach statistics?
12. Have you received feedback from others about your teaching of statistics? What types of feedback have you received from others about your teaching? Who, if anyone, has provided you encouragement/and or strategies for overcoming obstacles in teaching statistics?

The next several questions are going to focus on your experiences in statistics classes in college.

13. In your survey, you stated that you took ## courses in statistics. Did those courses contribute to you feeling prepared to teach statistics? If so, how? If not, why not?
14. Did you learn anything about teaching statistics from the way your instructors taught statistics? How do you think what you learned could be impacting how you feel about your ability to teach high school statistics?
15. Did you learn anything about using tools to teach statistics from the tools your instructors used to teach statistics? How do you think what you learned could be impacting how you feel about your ability to teach high school statistics?

In the next several questions, we are going to focus on your education courses that prepare you to teach mathematics.

16. Was there anything you experienced in your ***methods*** courses that helped you learn how to teach statistics?
    1. Possible follow-up question: What was it about _________ that was impactful/that left you confused?
17. [Note: Researcher will identify places with perceived key shifts from other data sources and use stimulated recall for the interviewee to reflect upon the experience. For example:]

You recently took an online class on teaching math with technology. Think back to the roller coaster investigation you did using CODAP that class (show screenshot) . In that investigation there were 157 cases of roller coasters and you were asked to investigate certain questions about different attributes using CODAP. How did this experience impact your confidence to teach statistics?

18. In that same class, you were able to view videos of students completing a similar roller coaster investigation and a teacher orchestrating a discussion around the same investigation (show screenshot of page). How did this experience impact your confidence to teach statistics?
19. In that same class, you recorded a screencast of yourself completing a statistical investigation (show screenshot of assignment instructions). How did this experience impact your confidence to teach statistics?

For the next set of questions, I am going to ask about your overall confidence in teaching statistics at this point in your teacher preparation.

20. On a scale of 1-10, how well prepared do you feel to teach high school statistics topics required in either the Common Core State Standards or your local state's standards? 1 means completely unprepared, 4 means somewhat prepared, 7 means sufficiently prepared, and 10 means completely prepared.
21. I am going to show you two different standards that are in the secondary curriculum. For each one, for each one rate how well prepared to you feel to teach this particular Common Core State Standard, On a scale of 1-10, and why?:

    CCSS.MATH.CONTENT.HSS.ID.A.3

    Interpret differences in shape, center, and spread in the context of the data sets, accounting for possible effects of extreme data points (outliers).

22. CCSS.MATH.CONTENT.7.SP.A.2 Use data from a random sample to draw inferences about a population with an unknown characteristic of interest. Generate multiple samples (or simulated samples) of the same size to gauge the variation in estimates or predictions.
23. What experiences do you think you would need to feel more prepared to teach statistics? How could teacher education programs do a better job at this?
24. Is there anything else you would like to add or comment on that has not been addressed?

Thank you again for your time and participation.

Appendix H – Rubric for Screencast Analysis

Participant pseudonym:

Rubric 1: Evidence of STSE Beliefs

Characteristic	Yes (2)	Developing (1)	No (0)	Score
Appropriate use of statistical language	Attends to many (at least 3) habits of mind, like context, variability, uncertainty, skepticism. No statistical language errors. Standards addressed match content.	Attends to some (at least one) habits of mind and makes infrequent (no more than 2) statistical language errors. Standards addressed do not match content.	Does not attend to habits of mind and/or makes frequent statistical language errors (more than 2)	
Appropriate use of technology	Takes advantage of many CODAP features (at least 8)	Takes advantage of some CODAP features (at least 5 and not more than 7)	Takes advantage of only a few, basic CODAP features (less than 5)	
Appropriate method of statistical investigation	Method is very exploratory and statistical, rather than rigid and mathematical.	Method is exploratory, but lacks rigor of a statistical investigation.	Method is very rigid and mathematical, rather than exploratory and statistical.	
Total				/6

Rubric 2: Evidence of STSE Source

Characteristic	Yes? (1)	No? (0)	Score
Evidence of rehearsal	Evidence includes prepared items like a Powerpoint, graphs, text boxes, or pulled up websites	Evidence includes referring back to questions frequently and taking time to do things that could have been done before, like pull up a website	
Evidence of teacher role	Participant is doing an investigation, but expresses that it is being done with an audience (the viewer) in mind	Participant is doing an investigation, with no indication that it is being done with an audience in mind	
Total			/2