ABSTRACT

HARRISON, TAYLOR RAY. Decision-Making of Secondary Statistics Teachers. (Under the direction of Dr. Hollylynne Lee).

An increasing number of high school students are enrolling in statistics courses. Hence, an increasing number of teachers are needed to plan and implement these courses. Yet, many teachers report feeling unprepared to teach statistics. When planning and implementing these courses, teachers draw upon their knowledge and beliefs when making decisions about their instruction. These decisions can have a significant impact on the learning of the students in those courses. However, little is known about how and why teachers of statistics courses make the decisions they do. This study investigates the decisions that secondary statistics teachers make, both when they are planning statistics instruction and when they are implementing statistics instruction. The study attempts to identify the knowledge and beliefs that teachers are drawing upon when making decisions, and contextual factors that may be inhibiting these knowledge and beliefs from being put into practice.

To investigate secondary statistics teachers' decision-making, an instrumental collective case study was performed. Seven high school statistics teachers participated in a series of interviews and observations designed to assess their decision-making processes. Participants' decision-making was examined first during their planning of instruction, and then during their implementation. Regarding participants' planning of instruction, the study identified five areas of beliefs and two areas of knowledge that participants drew upon. Contextual factors ranging from short class periods to limited planning time often inhibited four of these areas of beliefs and both areas of knowledge from being put into practice. Regarding implementation of lessons, the study identified several types of events that resulted in participants making decisions to modify their initial instructional plan by altering a student task, a lecture, or a whole-class discussion. These modifications ranged from altering the instructions for a task to adding new topics to a discussion. When making these decisions to modify lessons, participants drew upon four areas of beliefs and two areas of knowledge. Contextual factors often mediated whether and how participants used their knowledge and beliefs in their decision-making.

Results suggest that statistics teacher educators should help prepare teachers to make decisions, both planned and in response to unanticipated events. Results also suggest that teachers should be aware of the variety of contextual factors that they will face in the reality of the classroom. Recommendations are presented for preparing statistics teachers to make decisions in the classroom.

Decision-Making of Secondary Statistics Teachers

by
Taylor Ray Harrison

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

Learning and Teaching in STEM

Raleigh, North Carolina

2020

APPROVED BY:

Hollylynne Lee
Committee Chair

Margaret Blanchard

Karen Hollebrands

Eric Wiebe

BIOGRAPHY

Taylor Harrison was raised by parents Teresa and Jeff in Mesic, NC, where he lived until the age of 16. A first-generation college student, he attended the University of North Carolina, where he obtained a Bachelor of Science degree in Computer Science. After working in the private sector for five years after graduation, he decided to enter the field of teaching to make a lasting impact on students' lives. He obtained a Master of Science degree in Mathematics Education from North Carolina State University, studying the use of technology in the mathematics classroom. He then taught mathematics and statistics at ECPI University and Davidson County Community College before returning to North Carolina State University to enroll in the Learning and Teaching in STEM doctoral program, with a program area of study in Mathematics and Statistics Education.

Taylor will graduate from the program in 2020, after which he plans to continue his work in mathematics and statistics teacher education. His research interests include statistics teacher education, technology use, online teacher education, and statistics classroom practice.

ACKNOWLEDGMENTS

I would like to acknowledge my parents Jeff and Teresa, who instilled in me a passion for learning and a desire to help others. I would like to thank my advisor and chair of my committee, Dr. Hollylynne Lee, for pushing me to achieve my best and for making possible multiple opportunities for me to further my knowledge and passions. I would like to thank the various members of the research teams I have been involved with, for valuing my contributions and helping me develop my research skills. I would like to thank the other members of my committee, Dr. Margaret Blanchard, Dr. Karen Hollebrands, Dr. Eric Wiebe, and Temple Walkowiak, whose valuable feedback and suggestions have helped shaped not only this study, but my perspectives and outlook on research and teaching. I would like to thank the Friday Institute for Educational Innovation, whose support through the Graduate Student Fellows Award helped make this study possible.

I would like to thank Christina Azmy and Heather Barker, who have worked tirelessly alongside me, providing encouragement, care, joy, and laughter, and have made this journey much more bearable. I would also like to thank the others who have undertaken this doctoral journey along with me, providing inspiration, encouragement, and companionship. These include, but are not limited to, Asli, Blain, Brittney, Brooke, David, Elyse, Emily, Greg, Hamid, James, Jessica, Jonathan, Kristi, Latoya, Meetal, Michele, Mike, Mona, Nate, and Zach.

Lastly, I would like to thank the teachers who participated in this study, who were, without exception, welcoming, accommodating, enthusiastic, and generous with their time and energy.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

Background

In today's increasingly data-driven world, there is a need for its citizens to be statistically literate (Wild et al., 2018). The job of developing this statistical literacy largely falls to mathematics teachers in K-12, though aspects of data literacy are taught through subjects such as science and the social sciences. Though statistics content is often integrated into other mathematics courses across various grade levels, an increasing number of students are taking standalone statistics courses in high school. As the number of these courses increases, so does the number of teachers required to teach these courses. However, many preservice and inservice mathematics teachers report being unprepared and feeling a lack of confidence to teach statistical topics (Banilower et al., 2018; Lovett & Lee, 2017).

How teachers plan and enact a lesson has a significant impact on the content that students learn (Eichler, 2011; Hiebert & Grouws, 2007). Three aspects typically form the foundation upon which a teacher plans a lesson: specific goals for the lesson, a hypothesized learning path that students can take to get to those goals, and a series of activities to help direct or guide students down that path (Simon, 1995). Any of these three aspects can change over the course of a lesson, and a number of factors can influence what a teacher plans to do and ultimately does in a lesson. The goal of this study is to investigate these factors for high school statistics teachers, paying particular attention to how the context in which teachers teach interacts with their knowledge and beliefs to shape and inform the instructional decision they make.

Significance of the Study

Statistics is central to many aspects of our lives. Every day, we are confronted with media reports containing statistics about topics including crime, sports, the spread of viruses, the economy, public opinion, and various other topics. We make decisions about what cars to buy, where to live, where to attend school – decisions that have immense impact on our lives, and decisions that can be aided by statistical literacy. Now, more than ever, data governs the world. Thus, there is an increasing need in schools to develop statistically literate citizens (Wild et al., 2018). Because of this, many have called for a reform in the way in which statistics is taught (e.g., Franklin et al., 2007; Pfannkuch, 2018).

Indeed, statistics education in the United States does look markedly different than it did 20 years ago, particularly at the secondary level (Zieffler et al., 2018). Over half of all high schools in the United States now offer a standalone course in statistics, more than double the number in 2000 (Banilower et al., 2018; Weiss et al., 2001). In many of these courses, students are learning statistical topics that were once reserved for the post-secondary level. In 2019, nearly 220,000 students took the Advanced Placement (AP) Statistics Exam, twice as many as in 2008 (College Board, 2008; 2019). Classrooms have increased access to technology that impacts not only *how* statistics is taught, but *what* statistical topics are taught (e.g., topics such as linear regression were less accessible and not in the curriculum before access to graphing calculators and computers were widespread). Pedagogical innovations often found in other science, technology, engineering, and mathematics (STEM) classrooms, such as discovery learning, flipped classrooms, collaborative learning, and project-based learning are also being adopted by some teachers when they are teaching statistics (Zieffler et al., 2018).

Because of these rapid changes, research in statistics education must continue to expand and evolve. Though research in statistics education has become much more prominent in the last two decades, most of this research has focused on two main areas: 1) curricular issues, including assessment, standards and guidelines, and technology, and 2) students' (and sometimes teachers') knowledge of statistics, sometimes described as statistical literacy or statistical reasoning (Petocz, Reid, Gal, 2018; Watson, 2016). Students' and teachers' affect and beliefs toward statistics have also received some attention (e.g., Lee et al., 2017; Zieffler et al., 2012). What has received considerably less attention is what is actually happening on a day-to-day basis in classrooms, particularly the actions of the teacher. In a review of literature on statistics teachers' classroom practice, Eichler (2011) found that:

> One of the most striking results of the overview of research described in this chapter is the minor status of research on statistics teachers' intended and enacted curricula and their influence on students' learning. If we accept that a potentially successful way to change teachers' central beliefs is through teachers' assimilation of new ideas in contrast to accommodation (Pajares, 1992) it seems worthwhile to increase the research addressing the understanding of statistics teachers' central beliefs, and to understand the relationships among teachers' central beliefs, their classroom practice, and students' learning. (p. 184)

In addition, according to Pearl et al. (2012), *teacher practice* is one of six areas in statistics education that should be a priority for future research. As will be revealed in the Chapter 2, few researchers have answered this call for research on teachers' practices in statistics teaching, particularly at the high school level.

Purpose and Focus of the Study

In this study, teacher practice was examined through the decisions teachers made during planning and during implementation of statistics lessons, using a conceptual framework building primarily on the theories of the *mathematics teaching cycle* and *teaching-in-context*. In the mathematics teaching cycle (Simon, 1995), teachers draw upon their knowledge and beliefs to plan a lesson. Upon implementing the lesson, teachers' assessments of student learning impact their knowledge and beliefs, resulting in modifications to the teachers' instructional plans, either in real time or for future lessons. In teaching-in-context (Schoenfeld, 1998), the specific knowledge and beliefs that are impacting teachers' decisions at any given time are determined by the current context of the classroom including any events that have just occurred, as well as by their past experiences.

This study built upon these theories while drawing upon other literature on instructional practice to examine the instructional decision making of high school statistics teachers. The study investigated the processes by which these teachers' instructional plans were created, and the factors that impacted these processes and influenced teachers' planned curriculum. Factors of interest included the teachers' knowledge and beliefs, the context in which they teach, and the interaction between this context and their knowledge and beliefs. Additionally, the study investigated how these instructional plans were modified during and after lessons. It examined events that occurred that had the potential to result in modified instructional plans, how teachers responded to these events, and whether and how teachers did in fact modify their instructional plans as a result of an event occurring.

The following questions guided this study:

1. For teachers of secondary statistics, how are instructional plans for teaching statistics lessons created?

   1a. What knowledge and beliefs do teachers draw upon when creating instructional plans?

   1b. What current contexts seem to influence which of these knowledge or beliefs take priority?

2. How are secondary statistics teachers' instructional plans for statistics lessons modified after creation?

   2a. What events cause teachers to modify their instructional plans for statistics lessons, either during or after implementation?

   2b. Which knowledge and beliefs take priority due to these events?

To answer these research questions, a collective instrumental case study approach (Stake, 2005; Yin, 2009) was used. Seven statistics teachers participated in a series of interviews and classroom observations designed to assess the creation and modification of their instructional plans.

Organization of the Study

Chapter 2 will present a review of the literature on the learning and teaching of statistics, influences on instructional practice and decision-making, theories of instructional planning, and on teachers' knowledge, beliefs, and attitudes regarding statistics. Chapter 3 will present the proposed methods for this study, including a guiding conceptual framework. Chapter 4 will present a journal-ready manuscript concerning decisions that secondary statistics teachers make when planning instruction. Chapter 5 will present a journal-ready manuscript concerning decisions that secondary statistics teachers make in real-time during implementation. Chapter 6 will present a summary and discussion of the findings, implications for the classroom and for teacher education, and suggestions for future research endeavors.

CHAPTER 2: LITERATURE REVIEW

Before we can investigate statistical instructional practice, we must unravel what it means to teach statistics, and what it is that we want students to learn in statistics. The American Statistical Association (ASA) describes statistics as "the science of learning from data, and of measuring, controlling and communicating uncertainty" (ASA, 2019). These two elements of *data* and *uncertainty* play a foundational role in all of statistics. In regard to *teaching* statistics, Wild et al. (2018) state that "the mission of statistics education is to provide conceptual frameworks (structured ways of thinking) and practical skills to better equip our students for their future lives in a fast-changing world" (p. 6). Of course, this is the goal of education in various subjects, but it is important to keep in mind that statistics teachers should not only be teaching statistical methods, but also how to *think statistically*.

Statistical Thinking

Though there is general consensus that developing *statistical thinking* is a primary goal of statistics education, defining statistical thinking can be difficult, even for statisticians (Wild & Pfannkuch, 1999). The goal of the current study is not to assess whether participants are effective at developing statistical thinking in students. However, having a model of statistical thinking can help determine if and how teachers are considering the development of statistical thinking in their students when making decisions when teaching statistics. An early attempt at creating a model of statistical thinking can be found in Wild and Pfannkuch (1999). Based on interviews with statisticians and statistics students, they created a framework of statistical thinking. Their framework consists of four dimensions, and has been highly influential in statistics education research in the past two decades.

Dimension one of Wild and Pfannkuch's (1999) framework is the PPDAC model of the investigative cycle (MacKay & Oldford, 2000), consisting of five phases: identifying and defining a *problem*, creating a *plan*, collecting and managing *data*, *analyzing* data, and making *conclusions*. This is similar to how others have envisioned the investigative cycle. For example, one common alternative model is the PCAI model: *posing* a question, *collecting* data, *analyzing* data, and *interpreting* data (Franklin et al., 2007; Graham, 1987). Porkess (2012) describes a *data handling cycle*, consisting of specifying the problem and plan, collecting data, processing and representing the data, and interpreting the data. It is important to note that each of these models are cyclical; that is, interpreting and/or making conclusions about data can lead to further questions that can be investigated or problems that need to be solved.

Dimension two of Wild and Pfannkuch's (1999) framework is the types of thinking employed when learning or doing statistics. Some types of thinking are applicable to problem-solving in general. These include strategic thinking, seeking explanations, modeling, and applying techniques. Other types are specific and fundamental to statistical thinking. These include the recognition of need for data, transnumeration of data (changing representations to engender understanding), consideration of variation, reasoning with statistical models, and integrating the statistical and contextual (p. 226). Wild et al. (2018) suggest adding inductive or inferential reasoning to this list. Chance (2002) differentiates between components of statistical thinking and those of statistical literacy. For statistical thinking, she includes seeing the investigative process as a whole and as iterative, understanding the meaning and relationship of variation in this process, exploring data in novel ways, and generating additional questions as a result of this process.

Dimension three of Wild and Pfannkuch's (1999) framework is the interrogative cycle. The interrogative cycle is a series of general thinking processes that a statistical thinker utilizes when solving problems, and the entire cycle can appear at any point in the investigative cycle. The interrogative cycle consists of five components: (a) *generating*, brainstorming possibilities for questions to ask, possible causes, ways to collect data, ways to analyze data, etc.; (b) *seeking* either internal knowledge or external information; (c) *interpreting*, processing the results of seeking; (d) *criticizing*, checking information against internal or external reference points; and (e) *judging*, deciding to keep or reject information, plans, or conclusions. Grolemund and Wickham (2014) encapsulate this cycle into a model of data analysis. In their model, the thinker continually compares data to a model, makes judgements on the differences, and then evaluates these judgements against personal schemas.

The final dimension of Wild and Pfannkuch's (1999) framework is dispositions that can promote statistical thinking. These dispositions include curiosity and awareness, engagement, imagination, skepticism, being logical, a propensity to seek deeper meaning, openness, and perseverance (pp. 233-234). Many of these are mirrored in Hahn and Doganaksoy's (2012) list of characteristics of a successful statistician, which also include flexibility, persistence, enthusiasm, a passion for lifelong learning, and a proactive mindset.

As this framework and other related works show, statistical thinking goes well beyond simply choosing and applying the correct statistical techniques. However, some teachers have a rather narrow view of statistics (Begg & Edwards, 1999; Zieffler et al., 2012) or are not aware of the ways in which their students can think and reason statistically (Cai & Gorowara, 2002; Chick & Pierce, 2008). As will be discussed later, these beliefs and lack of knowledge can have a substantial impact on instructional practices of teachers.

Differences Between Mathematics and Statistics

It should be evident from the previous section that there are several aspects of statistics and statistical thinking that are different than mathematics. However, it may be worthwhile to explicitly enumerate some of these differences, given that statistics is often taught by teachers who have been primarily prepared to teach mathematics. It is possible that teachers who are aware of these differences may make different decisions when teaching statistics than when teaching other mathematics courses. Thus, as researchers, having a clear understanding of these differences can help us identify when and how teachers are attending to these differences and whether having knowledge of these differences impacts teachers' instructional decision-making.

According to Franklin et al. (2007), the key feature that sets statistics apart from mathematics is the focus on variability. The source of variability can be natural (e.g., people are naturally different heights), or it can be induced (e.g., students who study more may receive better grades). It can be due to tools used to take measurements, or it can be due to randomness of a selected sample. Though mathematics can be used to *describe* this variability, it is the omnipresent role of thinking about and dealing with variability–in formulating statistical questions, in collecting data, in analyzing data, in interpreting data, and in drawing and presenting conclusions–that sets statistics apart from mathematics.

Another key difference in statistics and mathematics is the role of context. Context does, of course, play a role in mathematics; however, this role is quite different:

> Although mathematicians often rely on applied context both for motivation and as a source of problems for research, the ultimate focus in mathematical thinking is on abstract patterns: the context is part of the irrelevant detail that must be boiled off over the flame of abstraction in order to reveal the previously hidden crystal of pure structure.

> *In mathematics, context obscures structure.* Like mathematicians, data analysts also look for patterns, but ultimately, in data analysis, whether the patterns have meaning, and whether they have any value, depends on how the threads of those patterns interweave with the complementary threads of the story line. *In data analysis, context provides meaning* (Cobb & Moore, 1997, p. 803).

Context is the fundamental difference between statistical reasoning and mathematical reasoning, according to delMas (2004). When learning new concepts in the classroom, mathematics problems and statistics problems may both start with a context. Once students become familiar with the mathematical concept, the concept becomes abstract and does not require a context to understand. Two mathematics problems with similar structures, but different contexts, can be approached identically. In statistics problems, on the other hand, the context drives the selection of models and the approach to analysis. The context is a key part of the problem-solving process, rather than a shell that must be peeled away (delMas, 2004).

A final fundamental difference between statistics and mathematics is the issues surrounding measurement. In addition to considering issues of measurement variability previously discussed, one must consider *how* and *what* to measure. Particularly when measuring abstract concepts (e.g., intelligence, levels of depression), part of the problem-solving process involves making decisions about how to turn these concepts into data that can be analyzed (Rossman et al., 2006).

Issues for Teachers of Statistics

There are certain issues that teachers face that are specific to teaching statistics. Many statistics courses are taught by teachers who have had most of their preparation and experience in teaching mathematics. Because of the aforementioned differences in mathematics and statistics, additional issues can arise for these teachers.

Statistics education has rapidly evolved, particularly around the area of data science. There is a large number of choices when deciding what content and learning objectives are appropriate for a course, and what instructional methods are best to reach those objectives (Wild et al., 2018). This is amplified by the fact that more and more statistical content that was previously taught at the tertiary level has trickled down to the secondary level (Zieffler et al., 2018). Many teachers, even at the secondary level, have had relatively little coursework in statistics (Banilower et al., 2018; Lee & Harrison, 2020). These are perhaps some of the reasons why there is large variation when examining what teachers of statistics are doing in high school classrooms (Zieffler et al., 2018).

Many teachers rely on what they know about mathematics education, either through experience or professional development, to decide how to teach statistics (Zieffler et al., 2018). This has its advantages and disadvantages. Many of the recommended pedagogical strategies for mathematics can also be effective for teaching statistics. Using more student-centered approaches with less reliance on lectures, encouraging classroom discourse, using formative assessments, building on students' prior knowledge, and providing appropriate scaffolding for students' thinking are all strategies that can be as effective in teaching statistics as they are in teaching mathematics (Ben-Zvi et al., 2018; Watson et al., 2018). Approaches such as activity-based learning, flipped classrooms, and collaborative learning, all of which can promote student learning, are also being adopted by some teachers when they are teaching statistics (Zieffler et al., 2018). Of course, carrying over these practices to their teaching of statistics presupposes that all these teachers are already implementing these practices when teaching mathematics, which is assumedly not the case. On the other hand, taking a mathematical approach to teaching statistics can hinder the development of statistical thinking (Cobb & Moore, 1997; Pfannkuch, 2018). Unlike in mathematics, mathematical structures should be treated as a means by which to discover information, rather than an end to be abstracted out (Wild et al., 2018). Understanding mathematical theory, or even statistical theory, is not sufficient for developing statistical thinking (Cobb & Moore, 1997). Using off-the-cuff contrived examples will not adequately illustrate statistical concepts as it might mathematical concepts (Cobb & Moore, 1997). Students must be given firsthand experiences with data collection and exploration in order to develop components of statistical thinking that cannot be adequately developed with a mathematical approach (delMas, 2004). In short, “statistics should be taught as statistics” (Cobb & Moore, 1997, p. 814).

One catalyst for rapid change in statistics education is the evolution of technology. Technology makes statistics more accessible, and calculations quicker and more accurate (Zieffler et al., 2018). This allows more curricular time to be focused on interpretation of results and on conceptual learning of statistical content (Chance et al., 2008; Franklin et al., 2007). However, teachers must be able to adapt their instruction to the changes that technology brings. They should carefully structure explorations, observe and support students in their use of technology, and keep students’ focus on the statistical content being explored (Ben-Zvi et al., 2018; Chance et al., 2008). Successfully incorporating novel technologies into one’s teaching is not a trivial task. Several obstacles must be overcome, including potential lack of buy-in from colleagues and administrators, limited time and support to learn to use technologies, and teachers’ own resistance to change (Chance et al., 2008).

Because statistics education is rapidly evolving, researchers must continue to study what is happening in classrooms, and in particular, to the mechanisms that support successful statistics teaching (Zieffler et al., 2018).

Effectively Investigating and Explaining Teacher Practice

If we want to understand what teachers are doing in the classroom, the initial and perhaps most obvious method might be to simply go and observe teachers while they teach. But if our goal is to *understand* rather than *describe*, observations alone are not enough. According to Lampert (2004), "teaching practice is what teachers do, but it is more than how teachers behave with students or the actions of individual teachers; action is behavior with meaning, and practice is action informed by a particular organizational context" (p.2). Thus, teacher actions must be examined in the context of a larger lesson, unit, or year, along with the relationships between the teacher and the student and relationships between the teacher and the school environment. Behavior, meaning, and context cannot be disentangled, and are all part of teacher practice (Lampert, 2004). Similar sentiments are shared by other researchers, as well, as detailed below.

Koehler and Grouws (1992) proposed four different models of increasing sophistication for researching teacher practice. In their highest-level model, a teacher's behavior should be examined along with the influential factors of their knowledge (of content, pedagogy, and student learning), their beliefs (about teaching and about mathematics), and their attitudes, along with their students' behavior and characteristics. Based on a review of literature on teacher practice, van der Sandt (2007) expanded each of the three teacher characteristics to explicitly include curriculum knowledge, beliefs about students as learners, beliefs about the learning of mathematics, and teacher attitudes toward students, mathematics, and the teaching of mathematics. In addition, she added a factor of social context, claiming that the relationship between each of these factors is context-dependent, and that which of these factors are reflected in actual instruction is driven by context.

Through examining 14 secondary mathematics teachers, Artzt and Armour-Thomas (1998) developed the Teacher Metacognition Framework to describe the link between a teacher's metacognition and their instructional practice. Per this framework, in order to understand the reasons behind teachers' practice, researchers should examine teachers during three stages: preactive (planning), interactive (monitoring and regulating), and postactive (assessing and revising). During each of these stages, in addition to observing what a teacher *does*, we must try to ascertain the teacher's metacognition – their knowledge, beliefs, and goals – if we want to truly understand a teacher's practice.

These definitions, models, and frameworks provide us with a foundation of what must be considered when investigating teacher practice--much more than simply the observable actions of a teacher in a classroom. However, they do not prescribe methods for operationalizing teacher practice, or for collecting or analyzing data to inform us about teacher practice. For that, we turn to more recent studies that can help us determine how to best examine these multiple aspects of teacher practice.

Methods for Examining Teacher Practice

Three main methods have been used to measure instructional practice and influences on instructional practice–surveys/questionnaires, interviews, and observations–though other methods have been employed, as well. Many researchers use a combination of methods.

Surveys are sometimes used as a direct measure of instructional practice (e.g., Hamilton et al., 2003; Huffman et al., 2003). These types of surveys typically ask teachers how frequently they use certain practices, and typically give options to choose from such as *Once or twice a* *week*. Other times, surveys concern more general teaching orientations or styles (e.g., Askew et al., 1997; Eichler & Erens, 2015). These surveys are often used to classify teachers in order to make comparisons between groups. Both types of surveys are useful for quantitative analysis–for example, measuring association with students' test scores. Sometimes, questionnaires are used to gather information about instructional practice that is difficult to gather via observation, such as teachers' policies for evaluating student work and assigning grades or how teachers organize and plan lessons (e.g., Askew et al., 1997; Stipek et al., 2001). Rather than directly asking about instructional practice or teaching styles, questionnaires are also commonly used to ask about factors that may be influential on instructional practice, such as beliefs about teaching and learning (Askew et al., 1997; Muijs & Reynolds, 2002), self-efficacy (Baker et al., 2004; McCaffrey et al., 2001) or the teaching environment (Baker et al., 2004).

Interviews are common in studies examining instructional practice. Sometimes, interviews are used for some of the same purposes as surveys – e.g., to measure frequency of use of instructional practices (Baker et al., 2004) or to assess how lessons are planned (Artzt & Armour-Thomas, 1998; Chidziva, 2017). More often, however, interviews are used to gather more detailed information than surveys, or measure attributes that would be difficult to measure via surveys or observations. Interviews that are conducted before (or instead of) observations often ask about the teacher's beliefs about teaching and learning mathematics (Askew et al., 1997; Eichler & Erens, 2015; Sztajn 2003), or about their goals for the upcoming lesson (Chidziva, 2017; Eichler, 2008). After an observation, interviews are often used to probe teacher's decision-making, investigating the reasons behind teachers' practice (Artzt & Armour-Thomas, 1998; Thompson, 1984).

With the advent of audio- and video-recording devices, many studies investigating teacher practice have turned to the use of stimulated recall interviews following observations. In these interviews, teachers are typically shown recordings of portions of a lesson that they taught and are asked to discuss events in the recording, often focusing on the decisions the teacher made and the reasons behind those decisions. Given the constraints of the classroom, it is generally not feasible for teachers to verbalize their decision-making process in real-time. Thus, stimulated recall would seem to be a viable alternative to examining teacher decision-making. However, care must be taken when employing this method. When recalling decisions in a stimulated recall interview, interviewees tend to describe their decision-making process as more deliberate than it actually was in the moment (Lyle, 2003). That is, interviewees tend to describe steps they took and logical thoughts they had that led to a decision, when the decision was, in fact, a reactionary one. Additionally, interviewees tend to describe their decision-making process in less detail when compared to verbalizing their decision-making process while it is happening (Kuusela & Paul, 2000). Despite these limitations, however, stimulated recall remains one of the most effective ways to probe decision-making, particularly when the interview takes place shortly after the decision-making and when the interviewee is in a comfortable environment (Lyle, 2003).

Classroom observations allow the most direct measurement of a teacher's instructional practice, at least the part of practice that happens during a lesson (as opposed to before or after a lesson). The instructional practice that is observed, however, runs the risk of not being representative of a teacher's entire instructional practice, especially if the number or length of observations is small. Though some studies rely on an observation of a single class period to assess practice, others (e.g., Blanton & Kaput, 2005) observe dozens of class periods per teacher.

Observations are often video- or audio-recorded, both for analysis purposes and to be used in stimulated recall post-observation interviews with teachers. Observations that are not audio- or video-recorded generally either involve observation protocols with which the observers are attempting to describe or identify specific features of instruction (e.g., Baker et al., 2004; Muijs & Reynolds, 2002), or they are observations that take place over extended periods of time over many days (Blanton & Kaput, 2005). These observation protocols can range from simple organizational tools for field notes to research-based rubrics with detailed criteria and scoring systems. For example, Walkowiak et al. (2014) developed the Mathematics Scan (M-Scan) protocol to measure eight different dimensions of a teachers' instruction, such as the use of mathematical tools, or the cognitive demand of the lesson. Indicators are provided for each of the eight dimensions, and teachers are assigned values of one to seven depending on the extent each of the indicators is present in the teachers' instruction. Hill et al. (2008) used an observation protocol to measure the Mathematics Quality of Instruction (MQI) of teachers across six dimensions, including the presence of mathematical errors and how the teacher responded to students' questions and answers. These more detailed protocols are useful for large-scale observations by multiple observers, or if there is a need to quantify instructional practice or classify or sequence teachers based on certain aspects of their instruction. Classroom observations are often accompanied by field notes, especially when observations are not recorded or when detailed observation protocols are not used, in order to detail the instructional practices of the teacher and the occurrences of the classroom. Even when observations are recorded or observation protocols are used, field notes are often used to capture things that the recording or protocol may not capture or to make a record of particular moments of interest.

Though less common, other forms of data have also been used to assess teachers' instructional practice. Blanton and Kaput (2005) asked one teacher to provide written reflections after several unobserved lessons, providing descriptions of classroom activity, conversations that happened, her perceptions of how the lesson unfolded, and what she thought was noteworthy about student thinking. Copur-Gencturk et al. (2014) used student surveys (in addition to other instruments) to measure the frequency of certain instructional practices of the teacher. Artifacts of teaching such as lesson plans and classroom handouts have also been used to determine a teachers' intended instructional practice (e.g., Chidziva, 2017; Sztajn, 2003). One artifact-based method that has become more prominent over the last several years is the "scoop notebook," in which a teacher is asked to regularly select and annotate classroom artifacts, which are then combined with their answers to reflection-prompting questions to create a composite instrument that can be used to assess instructional practice (Borko et al., 2005; Martínez et al., 2012).

Each of the methods used has benefits and drawbacks. Observations require a large amount of time from the researcher, whereas interviews require teachers to devote time in addition to their normal instruction. Surveys are relatively convenient for researchers and teachers but are of questionable reliability and validity when examining instructional practice. Mayer (1999) administered two identical surveys to teachers four months apart, asking them about the frequency with which they use certain instructional practices. Although composite scores (i.e., *standards-based practice* and *traditional practice*) were fairly reliable, responses for individual practices were not. He also observed teachers to determine whether reported frequencies matched actual practice. Although the relative ordering of the frequency of practices were fairly valid, absolute frequencies reported for individual practices were generally not. He also found that the short descriptions of practices given on the surveys were not indicative of the quality of implementation of those practices--even when two teachers report using a practice for equal amounts of time, the resulting instruction can look significantly different (Mayer, 1999). The student surveys in Copur-Gencturk et al. (2014) also had similar validity issues, with responses not in accord with observations. More novel instruments, such as the scoop notebook, seem to seek to alleviate some of the validity issues of surveys, while only moderately increasing the amount of time and effort required by researchers and teachers. What follows is a discussion of the results of empirical studies using the aforementioned methods in order to examine influences on instructional practice.

Influences on Instructional Practice

Several empirical studies have identified factors that impact instructional practice. There seems to be broad consensus that characteristics of the teacher matter. Two teachers who are teaching the same content from the same curriculum in similar environments (Eichler, 2008) or even using the same task (Burgess, 2008; Chick & Pierce, 2008; Hill & Charalambous, 2012) can yield very different instructional practices. Two broad areas of teacher characteristics have received markedly more attention than others in the literature on influencing instructional practice: knowledge and beliefs.

Knowledge

A teacher's knowledge takes many forms, including content knowledge, pedagogical knowledge, knowledge of particular students, curricular knowledge, technological knowledge, and so on. Shulman (1986) presented a framework for the knowledge a teacher needs to be an effective teacher and posited that teachers make decisions based on their knowledge. In the decades since, researchers have attempted to examine teachers' knowledge and investigate the impact that this knowledge has on a teacher's instructional practice.

In a study of elementary school teachers, Askew et al. (1997) classified 18 grade 1-6 teachers into three broad teaching styles. They found that teachers with deeper, more conceptual knowledge of the mathematics being taught were more likely to use a *connectionist* or *discovery* teaching style, rather than a *transmission* style. Hill and Charalambous (2012) investigated how two seventh-grade teachers with different levels of mathematical knowledge for teaching (MKT) enacted an identical task. They found that the teacher with a higher level of MKT made more quality adaptations to the task, was better able to assess students' unusual but promising strategies, more often encouraged multiple solution strategies, and was more successful at keeping the focus of the lesson on the intended learning goal. Charalambous et al. (2012) found that seventh-grade mathematics teachers with higher levels of MKT were more likely to show more skillful use of representations, use more appropriate explanations during instruction, and were more able to capitalize on student contributions while directing the class toward the learning goal. They also found that teachers with higher levels of MKT were better able to compensate for curriculum materials that were lacking. Similarly, Lewis and Blunk (2012) found that eighth-grade teachers with higher levels of MKT explained concepts more completely and concisely, used multiple representations more fluently, and better incorporated students' ideas into the instruction. Conversely, Desimone et al. (2016) found that for beginning seventh- and eighth-grade mathematics teachers, there was no significant relationship between MKT and the rigor of tasks used, the cognitive demand expected of students, or the amount of time spent on basic instruction.

König and Pflanzl (2016) assessed vocational school teachers' general pedagogical knowledge using a standardized and validated assessment, and then compared it with student questionnaires regarding the instructional quality of those teachers. Students of teachers with higher levels of general pedagogical knowledge tended to rate the teachers higher in the areas of managing the classroom, the clarity of their instruction, and maintaining better relationships with the students.

Aubrey (1997) found that elementary school teachers with higher levels of content knowledge were better able to apply their pedagogical content knowledge, leading to shifts in instructional practice. For example, even if a teacher is aware of a common mistake that students make, and notices that a student is making that mistake, the assistance that the teacher provides may not be beneficial to the student's thinking if the teacher's content knowledge is lacking. In other words, "without clear subject content knowledge neither sophisticated theories concerning children's learning nor scaffolded approaches will necessarily lead to effective teaching" (p. ix). Ormrod and Cole (1996) found similar results, and in addition, found that not only did a teacher's knowledge influence their practice, but their practice impacted their knowledge, leading to an iterative cycle of shifting knowledge and practice.

Few studies have directly examined this connection between knowledge and practice for secondary mathematics teachers. Zaksis and Leiken (2010) used questionnaires and interviews to examine how 52 secondary mathematics teachers reported they used their advanced mathematical knowledge (content knowledge obtained in tertiary coursework) when teaching. Although the frequency that teachers reported using their advanced mathematical knowledge varied, the most common topics for which teachers did so were calculus and statistics. Most teachers, however, were unable to articulate specific examples of problems or tasks in which they used their advanced mathematical knowledge. Rather, the teachers indicated that their advanced mathematical knowledge helped them to use more precise language, make connections across the curriculum and to real-life situations, respond to student questions, and be more comfortable and confident in their teaching.

Knowledge, of course, is not fixed. Professional development intended to increase teachers' knowledge can have a substantial impact on instructional practice. Harland and Kinder (1997) found that providing teachers with knowledge and skills was one of the two factors (out of nine) in continuing professional development that had the highest impact on teacher's practice. Copur-Gencturk et al. (2014) saw shifts in the instructional practice of science teachers who completed a master's program designed to increase content knowledge and pedagogical content knowledge. The teachers' lessons became more inquiry-oriented, they more often engaged students in reasoning and meaning-making, and they used more high-quality tasks. They also were more likely to create classroom environments in which students were more engaged with others' ideas, and in which the needs and ideas of all students were attended to. Garet et al. (2016) conducted a rare experimental study in which 165 fourth-grade teachers were randomly assigned to a professional development or a control group. The core of the professional development was designed to increase teachers' mathematical content knowledge, with supplemental components of the professional development designed to help teachers enact this knowledge in the classroom. The measured increase in teachers' knowledge was significantly correlated with the richness of mathematics that they used in the classroom. Other observed changes in practice, such as the level of student engagement and the precision and amount of errors in instruction, were in the intended direction, but not at a statistically significant level.

Beliefs

Much research has also been done regarding teacher beliefs and their impact on instructional practice. The beliefs of a teacher entail many different things, including beliefs about mathematics, beliefs about effective teaching, beliefs about how students learn, and more. Studies investigating the link between a teacher's espoused beliefs and their practice have returned mixed results. Although there seems to be a consensus that beliefs matter when it comes to instructional practice, more specific questions such as how much they matter, when they matter, and which ones matter most are still questions being answered. In some theoretical models, such as Guskey's (1986) model of teacher change, it is suggested that it is a teacher's practice (and the resulting observable outcomes in students) that has an impact on their beliefs, rather than the other way around.

Examining preservice teachers, Abd-El-Khalick et al. (1998) discovered that the teachers' beliefs about the nature of science and about teaching the nature of science were, for the most part, not reflected in their instructional planning or student teaching. Participants articulated several reasons for this, including other beliefs about curriculum having priority, a lack of resources, a lack of planning time, and their own discomfort with the content being taught. Likewise, Chen (2008) found that secondary teachers' beliefs about teaching and learning were not reflected in the way they integrated technology into their instructional practice. That is, teachers who espoused constructivist-style teaching beliefs tended to not use technology in ways that would support students' learning under those systems of beliefs. Reasons for this discrepancy mirrored many of Abd-El-Khalick et al.'s (1998) findings, and included a lack of resources or access to resources, insufficient planning time, inadequate administrative or technical support, discomfort with technology, and other conflicting beliefs about assessment and curriculum.

However, several studies *have* observed links between teachers' beliefs and their instructional practice, particularly in mathematics. Through surveys and observations on 21 fourth–sixth-grade teachers, Stipek et al. (2001) found that teachers who exhibited more traditional beliefs in the surveys also tended to exhibit traditional practices in the classroom and relied on more traditional ways of evaluating their students. For example, teachers who viewed mathematics as a set of operations and procedures to be learned were more likely to emphasize performance and speed, and less likely to emphasize student understanding in their teaching and assessment. Thompson (1984) conducted a case study of three junior high school mathematics teachers, investigating whether their beliefs were in congruence with their practice, and whether differences in practice between the three teachers could be explained by differences in beliefs. She found that beliefs about mathematics played a significant but subtle role in impacting instructional practice, most visibly in the way the teachers presented mathematical content. Beliefs about things other than mathematics, such as the social and emotional makeup of the classroom, also had a significant effect, and for some teachers, these beliefs took priority over their beliefs about mathematics. Similarly, Sztajn (2003) found that beliefs about students' needs, which are in turn affected by the social setting of the school, can have a strong impact on teachers' practice. For example, when working with students from low socio-economic neighborhoods, teachers may believe that they can best serve their students by providing mathematics instruction that is highly structured with clearly organized goals and with deliberate practice. On the other hand, teachers working with higher socio-economic students may believe that they can best serve their students by building on what they have learned at home and by ensuring they are engaged and interested in mathematics instruction that is relevant to them.

Muijs and Reynolds (2002) found that the level of connectionist beliefs, measured through surveys, was the strongest predictors of *both* constructivist and traditional teaching practices, suggesting, like Thompson (1984), that the relationship between beliefs and practice, though significant, is a complex one. Even when policies and reform initiatives are instituted in an attempt to change teacher practice, teachers' beliefs can strongly impact how teachers adopt new policies and institute change in the classroom (Cohen & Ball, 1990; Johnson, 2006).

Beswick (2007) sought to understand the beliefs of teachers who tended to establish classroom environments that were consistent with constructivism. From her sample of 25 secondary mathematics teachers, she identified two that met this criterion. She found that for these two teachers, their teaching styles were very much grounded in their beliefs. However, when examining the beliefs that were central to each of the teachers' practice, the beliefs of the two teachers were of different natures. One teacher's central beliefs mainly involved the nature of mathematics and mathematics learning, whereas the other teacher's central beliefs all involved the role of the teacher. Thus, despite having different (though not necessarily contradictory) beliefs, the two teachers both created classroom environments consistent with constructivism. Beswick posits that the reasons that the two teachers' central beliefs were of a different nature may be due to contextual reasons such as previous professional learning experiences or their experiences as students.

Other Factors

Though teacher knowledge and beliefs have dominated the literature on factors influencing teacher practice, several other factors have received attention. A teacher's goals or objectives for a lesson seem to have a significant influence on their practice (Schoenfeld, 1998). Statistics teachers can have *central* goals and *peripheral* goals for a lesson; central goals seem to be much more attended to and be more likely to have an impact on a teacher's practice (Eichler, 2008; Eichler & Erens, 2015). Others have found that there are often discrepancies between what teachers intend to do and what actually happens. Paparistodemou et al. (2006) found that teachers' practice often lacked central aspects of their planning, and that events that occurred in the classroom often caused teachers to deviate from their plans.

The context and environment in which a teacher teaches can also influence instructional practice and present barriers to change. Teachers report that a lack of time--both instructional time and planning time--can prevent them from implementing changes and practices into their teaching (Johnson, 2006; Lee & Harrison, 2020). A lack of resources similarly impacts teachers' practice, as does the pressures of student assessments (Johnson, 2006; Lee & Harrison, 2020). The social norms of a school can also cause teachers to teach in ways that are not aligned with their beliefs (Skott, 2009). Other factors suggested to be influential for teacher practice include teacher attitudes (Koehler & Grouws, 1992; van der Sandt, 2007), teacher assessment of students' knowledge (Simon, 1995), curricular materials (Hill & Charalambous, 2012), student behavior (Koehler & Grouws, 1992; van der Sandt, 2007), and general student characteristics (Lee & Harrison, 2020).

Teachers' Knowledge, Beliefs, Attitudes Regarding Statistics

Teachers' knowledge, beliefs, and attitudes all impact their instructional practice when teaching statistics. The following sections will describe what we know, as researchers, about teachers' knowledge, beliefs, and attitudes regarding statistics. Although these three concepts are all intertwined and interdependent (Begg & Edwards, 1999), for the most part they will be discussed serially.

Knowledge

A teacher's knowledge of statistics and knowledge of teaching statistics can both have a significant impact on their instruction. Given that the teachers in this study were at various points in their curriculum and were teaching varied topics, a brief overview of research on teachers' knowledge (both content knowledge and pedagogical knowledge) on a selection of several relevant topics in statistics will be given. Studies involving inservice secondary teachers' knowledge about statistics and teaching statistics are relatively sparse, so research involving either preservice teachers or elementary teachers will be used to supplement this analysis when necessary.

Modeling. Lee and Mojica (2008) examined the instructional practice of nine middle school teachers planning and implementing a lesson involving a statistical investigation using simulations with hands-on or technology-assisted models. They found that these teachers often failed to make substantial connections between theoretical and empirical probabilities, and in general, failed to integrate ideas about probability into other statistical concepts such as variation, sample size, and inference. Teachers, perhaps because of their discomfort with randomness and variation, often turn statistical modeling activities into mathematical activities such as graph construction or mathematical computations (Lee & Mojica, 2008; Liu & Thompson, 2009). Biehler et al. (2018) found that when analyzing models, the elementary preservice teachers in their study tended to use their contextual knowledge about the phenomena being modeled much more than their statistical reasoning when deciding if a model was correct. Their contextual knowledge drove how the preservice teachers reasoned about the model, and their statistical knowledge was only briefly used as a validation of the model after their initial analysis was performed. Justice et al. (2018) investigated four inservice secondary teachers' conceptions about statistical modeling. Rather than viewing models and simulations as a way to model and examine variability, the teachers viewed models as a way to make a decision or inference about an experimental result and to faithfully replicate the data collection process.

Although the differences in these viewpoints are subtle, Justice et al. (2018) argue that taking the focus off variation can have negative consequences for these teachers' teaching of statistics.

Statistical Investigations. Heaton and Mickelson (2002) examined 44 preservice elementary teachers, first as they designed and carried out a statistical investigation of their choosing, and then as they implemented a unit in a classroom in which their students carried out a statistical investigation. In both parts of the study, the teachers struggled with posing good statistical questions. Questions typically had a single correct answer (and were thus more mathematical than statistical in nature), had little practical purpose, or were unable to be realistically answered with data that might be available. During the classroom unit, lessons often evolved away from ideas around statistical investigations, and toward instruction on less statistical concepts such as the technical aspects of graph creation.

Hannigan et al. (2013) administered the Comprehensive Assessment of Outcomes in Statistics (CAOS) instrument to 115 preservice secondary mathematics teachers, both undergraduate and graduate. The instrument assesses respondents' knowledge on material that any student completing an introductory statistics course should be expected to understand. All of the undergraduate preservice teachers were in the top 10% of mathematics achievement in secondary schools in Ireland, and all of the graduate preservice teachers had a degree with a strong mathematics component. The researchers found that these preservice teachers struggled with some statistical content. For example, only 9% correctly identified the purpose of randomization in an experiment, and only 22% correctly identified the factors of a sample that allowed conclusions to be generalized to the population.

Burgess (2002) gave a group of 30 elementary preservice teachers a small multivariate data set and asked them to produce a written report on what they discovered by analyzing the data in a way of their choosing. Nearly a third of them failed to produce any real conclusion about the data, but simply performed statistical procedures (e.g., creating a box plot) without any perceivable reason for doing so. Less than half of the preservice teachers attempted to make any type of generalizations beyond the data at hand. Lee et al. (2014) also provided (mostly preservice) teachers with a data set and asked them to generate a question and provide a written report. Approximately one-third of teachers proposed questions that were considered broad (i.e., open-ended with multiple solution paths). Nearly all of the teachers used appropriate graphical representations and statistical measures to answer their question, and 41% showed evidence that they linked multiple representations together to help them answer their question. The difference in these two studies may be explained by the affordances of the dynamic statistical software used in Lee et al. (2014), which allowed easier creation and modification of graphical representations and easier calculation of statistical measures.

Chick and Pierce (2008) examined preservice elementary teachers' pedagogical knowledge of statistics by providing them with a rich data set and asking them to plan a lesson for a sixth-grade class that uses that data set. Only three of the 13 lessons produced used the data set in a "deep and extended" way. Ten of the lessons were poorly articulated, and it was not clear that the concepts would be taught in a pedagogically appropriate or mathematically correct way. When algorithmic procedures were included (e.g., calculating the mean), no connections were made to how the numeric answer relates to the data or what it reveals about the data. Over half of the lessons produced, however, did include time in the lesson to let students become accustomed to the data. In each of these studies, the preservice teachers tended to lose sight of the bigger picture involving statistical investigations and the investigative cycle, and the "why" behind the procedures being performed.

However, there is evidence that teachers' pedagogical knowledge of statistical investigations can be improved. Casey et al. (2020) examined the characteristics of technology-enhanced statistical investigative tasks created by 75 preservice teachers after they had engaged with a set of curricular materials designed to prepare preservice teachers to teach statistics. Most of the questions in the tasks created by the preservice teachers asked students to coordinate multiple aspects of the data set or graphs, and nearly half of the tasks also included questions requiring students to generalize beyond the data set at hand. Most of the tasks continually situated the investigation within the context of the data set being investigated. Additionally, most tasks asked students to engage in multiple phases of the statistical investigation cycle.

Chance and Probability. Begg and Edwards (1999) investigated the statistical knowledge of 22 inservice and 12 preservice elementary teachers by giving them statistical tasks to complete and asking them to assess the written work of students. In the area of chance and probability, just over two-thirds of teachers showed a good understanding of the concept of equally likely events, about half understood randomness, and less than half showed a good understanding of independence. Canada (2006) gave 30 elementary preservice teachers a series of tasks involving flipping a coin 50 times. Prior to an intervention, very few of the teachers attended to variation in the number of heads expected in 50 flips. Nearly 30% did not exhibit any type of proportional reasoning when explaining their answer. When asked to predict the number of heads for multiple sets of 50 flips, 41% gave answers that were unlikely to occur, with ranges that were too large or small, or not centered around the expected value. After the intervention, which included activities involving probability and other topics, results on similar tasks showed general improvement in teachers' knowledge, though inaccurate or inarticulate responses remained. Sánchez and García (2008) found similar (mis)conceptions in inservice middle school teachers on a task involving rolling a die. In addition, they identified another inaccurate belief, with three of the six teachers believing that as the number of rolls increased, the *absolute* frequency difference from the expected value would be expected to decrease (rather than the *relative* frequency difference).

Statistical Representations. Teachers in the Begg and Edwards (1999) study mentioned above tended to view graphs as communicative tools, rather than tools to assist in analyzing and interpreting data. When working with graphs, they tended to overlook important details, and did not seem to consider data sets as a whole. When asked to choose an appropriate type of graph to represent data, most teachers' choices depended on the current representation of the data, rather than the structure and type of data itself. For example, when percentages were parenthetically added next to frequencies in a table with data, participants more often chose a pie graph to represent the data, even though the addition of percentages did not change the structure or content of the data. A lack of graphical fluency seems to be common among teachers, even at the secondary level (Eichler, 2011). Many teachers focus on the process of creating graphs, losing sight of the goal of the statistical investigation (Heaton & Mickelson, 2002). This focus on the graphing process can cause teachers to miss instructional opportunities to strengthen their students' understanding of statistics, as shown in Lee and Mojica (2008). In the assessment by Hannigan et al. (2013) discussed above, only 20% of participants showed a correct understanding that box plots cannot provide accurate estimations of percentiles other than quartiles. Only 23% were able to correctly interpret the median in the context of a box plot.

Measures of Center. Most of the studies examining teachers' knowledge of measures of center seem to be focused on elementary or middle school teachers. Jacobbe (2012) examined three "exemplary" (as identified by their supervisors) elementary teachers' knowledge of mean and median. He found that all three teachers did not possess a conceptual knowledge of mean and median. The knowledge that they did have was limited to algorithmic or procedural knowledge, and none of the three were able to successfully identify key differences between the mean and median or give reasons why one might be used over the other. Leavy and O'Loughlin (2006) found similar results with their sample of 263 preservice elementary teachers. Only about a quarter of their sample showed a conceptual understanding of the mean on at least two of the six tasks administered to assess their knowledge. Cai and Gorowara (2002) compared the knowledge of the mean of inexperienced and experienced middle school teachers. They found that although both groups had sufficient content knowledge about the mean, the experienced group exhibited much more knowledge about different ways students might approach solving problems using the mean and about possible misconceptions students might have. Begg and Edwards (1999) found that inservice teachers actually showed *less* conceptual understanding of measures of center than preservice teachers, with only 21% of inservice and 58% of preservice teachers correctly recognizing and interpreting all three of the concepts of mean, median, and mode when being asked to review student work. Importantly, Jacobbe and Carvalho (2011) found that misconceptions about measure of center held by elementary teachers are often mirrored in their students.

Variation. In one of the few studies that focused on inservice secondary mathematics teachers, da Silva and Coutinho (2008) classified teachers' knowledge of variation into five types of reasoning. Most teachers were classified as having verbal reasoning, where they understood some aspects of variation, but were unable to apply those understandings to correctly reason about actual situations. No teachers were classified as showing integrated process understanding, being able to relate understanding of multiple aspects of variation together. Less than half of the teachers in the study were correctly able to identify which of a selection of histograms represented a data set with the lowest variation from the mean. Makar and Confrey (2005) examined how preservice secondary math and science teachers talked about variation when they were asked to compare two groups using dot plots. They found that the teachers often used non-standard language when discussion variation, such as “mounds,” “chunks,” “spread out,” “clustered,” or the use of “range” to describe an interval. However, they argue that the use of this non-standard language still shows statistical thinking by the teacher, and that the use of these words has meaning for the teacher and may be more accessible to a wider variety of their potential students. Even on a multiple-choice assessment, however, Hannigan et al. (2013) found that 88% of preservice secondary mathematics teachers did not correctly answer a question intended to assess their understanding of properly describing a univariate distribution.

Peters (2009) examined the conceptions of variation and the robustness of understanding of variation possessed by 16 AP Statistics teachers. Half of the teachers viewed variation as an expected deviation, and as something that can be modeled and accounted for. Two of the teachers viewed variation as noise, as something that obscures the objects of interest. Two others saw variation as something that can be controlled, and that steps should be taken to minimize it in the design of data collection. The remaining four teachers did not exhibit sufficient evidence to characterize their conception of variation. Five of the sixteen teachers exhibited a robust understanding of variation, including at least one teacher with each of the three conceptions, showing that the level of understanding of variation is not necessarily dependent on one’s conception of variation.

Inference. De Vetten et al. (2018) surveyed 722 first-year preservice elementary teachers to assess their content knowledge of informal statistical inference. In open-ended responses, many teachers were vague or ambiguous about the population that inferences would be made about. Many did not seem to attempt to generalize at all beyond the samples that were given. Only about 20% of respondents were estimated to have the correct population in mind when completing the tasks. Over 90% of respondents said that comparisons between groups could not be made if the size of the two groups were not similar. Nearly 40% incorrectly agreed with a statement stating that when comparing samples from two populations, generalizations could not be made about the populations because if a different sample was chosen, the result could be different. Over 70% agreed that *nothing at all* could be determined about a population based on a sample size of 15.

Liu and Thompson (2009) conducted teaching experiments with eight secondary teachers of statistics and identified several issues that these teachers seemed to have with statistical inference. Most of these teachers had compartmentalized knowledge of probability, and were unable to apply this knowledge of probability to ideas about unusualness of an event. Teachers tended to show an overzealous commitment to the null hypothesis, feeling that rejecting it was akin to *proving* it was incorrect, and that as long as an experimental result was theoretically possible, then the null hypothesis should not be rejected. They also failed to correctly reject the null hypothesis as a result of incorrect assumptions about the model used in a simulation, believing that if a result showed up infrequently, then the claim which they were testing against the null hypothesis must be incorrect. Finally, all but one of the teachers failed to identify hypothesis testing as a way to test a claim about a population in a scenario in which they were given information about a sample. Various misconceptions about inference were also identified by Lovett and Lee (2018), who administered a series of multiple-choice tasks to 217 preservice secondary mathematics teachers. Approximately 40% of the preservice teachers indicated that a large p-value would allow one to reject the null hypothesis. Nearly half indicated that the p-value would be large for a hypothesis test comparing means when they were given two dot plots with a large gap between distributions. Less than a third correctly answered a question about the meaning of statistical significance. Approximately one-fifth were able to correctly identify the meaning of a 95% confidence interval.

Sampling. Watson (2001) developed an instrument to assess teachers' needs and achievements in chance and data, and administered it to 43 teachers, the majority of which were secondary teachers. In regard to teaching the topic of sampling, many of the teachers were unable to provide high-level or sophisticated responses when asked to design lessons around sampling, and it appeared that, in general, teachers were less familiar with the topic of sampling than other areas of the curriculum. In the de Vetten et al. (2018) study mentioned above, 32% of the preservice elementary teachers expressed the incorrect understanding that random sampling was not a valid way to get unbiased samples. The majority preferred a stratified sample, with 88% of teachers believing that a stratified sample is representative of the population. Similar understandings about sampling were found in the Lovett and Lee (2018) study described above, with 30% of preservice secondary mathematics teachers indicating a preference for a stratified sample to a simple random sample. Over 58% of preservice teachers in this same study indicated that results from a volunteer sample could be generalized to a population, while 39% were not able to identify statistical questions for which a sample would be preferable to a census.

Sampling Distributions. Lovett and Lee (2018) found that, when asked to choose a distribution that could represent a sampling distribution of means, 36% of preservice mathematics teachers incorrectly chose a distribution that had the same variation as the population. Similarly, 30% of teachers did not recognize that increasing the sample size would result in a distribution with a decreased variation. Similar questions on the assessment given by Hannigan et al. (2014) produced even worse results. The authors found that 76% of preservice secondary mathematics teachers did not select a proper sampling distribution given a population and sample size, and 85% did not show understanding that statistics from large samples would vary less than statistics from small samples. In de Vetten et al. (2018), when given a dot plot of a sample means, only 35% of the preservice elementary teachers were able to create a sensible dot plot of sample means with twice as many samples.

Association. Casey and Wasserman (2015) gave a series of task-based interviews to 19 preservice and inservice teachers, asking them to place lines of best fit onto scatter plots. In general, teachers were successful at approximating the least squares regression line, particularly when the correlation was moderate to strong, but many had trouble interpreting the results. Over three-quarters of the teachers indicated that a strong correlation implies a causal relationship between the two variables. This belief was particularly prevalent for those that had yet to teach the topic. When interpreting the slope of their line of best fit, almost all of the teachers gave a deterministic explanation, discounting the variation that existed in the data. Other difficulties that have been identified in research include believing that a strong correlation implies a linear relationship and being unduly influenced by previous contextual beliefs (Engel & Sedlmeier, 2011).

Summary. These studies show that, regardless of whether we consider preservice or inservice teachers, elementary or secondary, there seems to be widespread lack of knowledge of statistics and of how to teach it. Although there is some indication that secondary teachers tend to have more sophisticated content knowledge than elementary teachers (Watson, 2001) and that pedagogical content knowledge can come with experience (Cai & Gorowara, 2002), in most cases, the level of statistical knowledge is far from ideal.

Beliefs

As with teacher knowledge of statistics, studies on inservice secondary teachers' beliefs about statistics are scarce, so we must rely in part on work with preservice, elementary, or post-secondary teachers. Several areas of teachers' beliefs can affect their instructional practice, including beliefs about teaching in general (Thompson, 1984), beliefs about their students' needs (Schoenfeld, 1998; Sztajn, 2003), and beliefs about the social make-up of the class (Thompson, 1984; van der Sandt, 2007). However, we will focus here on the two areas of beliefs that have received the most attention in statistics education research: beliefs about statistics, and beliefs about teaching and learning statistics.

Beliefs About Statistics. In the aforementioned study of Begg and Edwards (1999), the 34 elementary inservice and preservice teachers were also asked about their beliefs about statistics. Most had what the researchers describe as a "narrow view" of statistics, describing statistics as graphs, tables, numbers, information, and percentages. Most viewed the work of statisticians as number crunching, although several mentioned the task of collecting data. The teachers generally described statistics as being useful, particularly to make sense of the world or to help see patterns and make predictions. Similar sentiments about the utility of statistics were also reflected by the preservice elementary teachers in Chick and Pierce (2008), who responded positively to the prompt "To be an intelligent consumer, it is necessary to know something about statistics" (p. 3). However, these teachers did not seem to have complete faith in statistics. In response to the prompt "Because it is easy to lie with statistics, I don't trust them at all" 52% chose a neutral or affirmative response. In addition, many indicated that they would trust the opinion of a few friends over a statistical survey (Chick & Pierce, 2008). Other studies, as well, have found that teachers across all levels often have ambivalent feelings about the utility and trustworthiness of statistics (Estrada, 2002, as cited in Estrada et al., 2011; Watson, 2011).

Yang (2014) used surveys and interviews to examine the conceptions that 27 secondary mathematics teachers had about statistics and mathematics. Three different categories of conceptions emerged, only two of which were held about statistics. The most common conception of statistics, held by 70% of teachers, was the *instrumental application* conception, the view that statistics is a logical system of tools used to solve problems. The *relational application* conception was the next most common conception, held by 19% of teachers. In this conception, statistics is a system of related concepts and representations that are used to solve problems. The remaining participants exhibited a combination of the two conceptions. No participants exhibited a *relational methodology* conception of statistics, where statistics is viewed as a way of thinking, though some held this view for mathematics. In addition, no participants exhibited a hypothesized fourth conception, *contextual investigation*, where statistics is viewed as an investigative process used to explore social activity. These beliefs seemed to be influenced by the fact that assessments that the teachers' students were subject to did not test ways of thinking or ability to perform investigations, and thus, the teachers only ever taught statistics as tools and procedures to solve problems. Two-thirds of teachers expressed a different conception of statistics than they did for mathematics.

As discussed previously, there are certain aspects of statistics that makes the field different than that of mathematics. Most of the elementary preservice and inservice teachers interviewed by Begg and Edwards (1999), however, did not see a significant enough difference between the two to warrant teaching the two topics any differently. The teachers identified links that statistics had to other curricular areas, but most treated statistics as simply one of the units of the mathematics curriculum, as exemplified by this comment from a teacher: “it's part of maths, we know it’s a maths thing” (p. 6).

Like knowledge, beliefs about statistics are not fixed. Lee et al. (2017) examined changes in beliefs and perspectives about statistics of 489 classroom teachers who participated in an online professional development. The teachers participated in experiences including reading and watching material on statistics, engaging in data investigations, watching videos of expert discussions and of classroom teaching, and forum discussions. As the teachers participated in the professional development, the researchers identified four major shifts in the way that participants viewed statistics. These emerging beliefs of participants included that statistics is more than just computations and procedures, that the use of dynamic technology enhances statistics, that statistics involves real and messy data, and that statistical thinking develops along a continuum.

Beliefs About Teaching and Learning Statistics. Zieffler et al. (2012) developed the Statistics Teaching Inventory to assess the practices and beliefs of instructors of post-secondary introductory statistics courses. Using a Likert-like scale from 0 (strongly disagree) to 3 (strongly agree), one part of the survey had participants respond to statements about teaching statistics that were based on recommendations from the college-level GAISE Report (Garfield et al., 2005), which is often used as a guideline for reform-oriented (as opposed to traditionally-oriented) teaching of statistics. For example, one statement on the Statistics Teaching Inventory reads “Technology tools should be used to illustrate most abstract statistical concepts.” Other statements involve beliefs about the topics that should be included in an introductory statistics course and the efficacy of lectures versus activities. If a participant generally agrees with these statements (or disagrees with ones that are reverse-coded), then their beliefs are determined to be aligned with recommendations from the GAISE report. Another section of the survey concerned assessment beliefs based on recommendations from the GAISE report, and was structured similarly, with statements such as “It is important to assess students on their ability to successfully complete a statistical investigation (e.g., an open-ended student project).” Other statements involved beliefs about the use of traditional versus alternative assessments, and the need to regularly review assessments. For beliefs about teaching, 86% of respondents had an average response level of at least 1.5 (the midpoint of the scale), and 71% had an average response level of at least 2.0, indicating general alignment of beliefs with the GAISE report, and thus, with reform-oriented teaching of statistics. Beliefs about assessment were less extreme, with 90% having an average response level of at least 1.5, and 62.5% having an average response level of at least 2.0. These results show that at the post-secondary level, although there is variation in responses, most teachers’ beliefs in general are at least somewhat aligned with some of the more reform-oriented recommendations.

The preservice and inservice elementary teachers in Begg and Edwards (1999) were also asked about their beliefs. Regarding the importance of statistics in the curriculum, all of the teachers considered it important, though only a quarter believed it was “really” important or “one of the most important” topics. When compared to other topics taught in mathematics, none of the teachers felt it was more important than the other topics, and only about 15% ranked it as equally important to all the other topics. When asked if statistics could be omitted from the curriculum without doing any real harm, most of the teachers indicated it could not.

Umugiraneza et al. (2016) surveyed 75 mathematics teachers across grades 4-12, asking them about their beliefs regarding teaching mathematics and statistics. Responses were given on a Likert-like scale ranging from 1 (strongly disagree) to 5 (strongly agree). Some of the more strongly agreed-with beliefs include that statistics teachers should foster positive attitudes toward statistics (mean = 4.52) and an attitude of inquiry (mean = 4.47) in students. Other beliefs that participants were generally in agreement with include that using technology helps students' learning of statistics (mean = 4.05) and that effective statistics teachers enjoy learning and doing statistics themselves (mean = 3.93). However, some beliefs that may be potentially concerning were also generally agreed with, including that the participant would feel uncomfortable if a student proposed a solution the participant had not considered (mean = 4.07), that statistics is best taught in an expository style (mean = 4.04), and that the statistics work done in participants' classrooms is not relevant to the students (mean = 3.91). Beliefs for which there was mixed results (i.e., large standard deviations and means close to neutral) include that statistics is computations (mean = 2.51), and that it is difficult to teach statistics both conceptually and procedurally (mean = 3.42) or without a textbook (mean = 3.27).

Teachers' beliefs about statistics seem to be closely related to their beliefs about teaching and learning statistics. At both the elementary and secondary level, teachers who seem to have an instrumental view of statistics emphasize mostly procedural skills, such as graph construction and computing measures of center and spread, when identifying important topics of statistics to teach (Begg & Edwards, 1999; Watson, 2001). Similar preferences were seen for the preservice elementary teachers in Chick and Pierce (2008); most of the teachers designed lessons that did not expect students to seriously grapple with and engage with the data in meaningful and contextual ways.

Attitudes

In general, the elementary teachers in Begg and Edwards' (1999) study held negative views of statistics. Examples of words and phrases the teachers used to describe their attitudes toward statistics include “fear, fear,” “horrors,” “uninteresting,” “I didn't understand,” “baffling,” “boring,” “horrible graphs,” “statistics is not my forte” (p. 2). This is consistent with the findings of Estrada (2002, as cited in Estrada et al., 2011), in which preservice and inservice elementary teachers gave generally low-scoring responses to affective questions such as “I enjoy taking statistics courses.” Chick and Pierce (2008) also found that most elementary teachers gave neutral or negative responses to questions about attitudes toward statistics.

These generally negative attitudes toward statistics may be linked with teachers’ levels of confidence in learning statistics and teaching statistics. Begg and Edwards (1999) found that although most teachers were less than confident about their own ability with statistics, most were confident in their ability to *teach* statistics, at least at the elementary level. Even though nearly all the teachers felt unfamiliar with at least one statistical concept or term in the curriculum they were teaching, this did not seem to cause feelings of concern toward their teaching. This is also reflected in their preference for professional development that focuses on pedagogy and tasks, rather than on statistical content. Other studies, however, have found that many teachers are concerned about their ability to teach statistics.

For example, Lovett and Lee (2017) administered an instrument designed to measure the self-efficacy to teach statistics to 236 preservice secondary mathematics teachers. In this survey, participants were asked to report their confidence to teach 44 different items, using a Likert-like scale ranging from 1 (*not at all confident*) to 6 (*completely confident*). Overall, the mean of teachers’ responses was slightly higher than 4, which represents *somewhat confident*. When asked to rank their confidence to teach various high school mathematics subjects (e.g., algebra, statistics, calculus), 63% of respondents ranked statistics last (i.e., least confident). Chick & Pierce (2008) found that 44% of preservice elementary teachers disagreed with or chose the neutral option in response to the prompt “I feel I have sufficient knowledge [of] statistics/‘data’ for teaching in primary school” (p. 3). Watson (2001) found that secondary teachers were, in general, more confident in their ability to teach statistics than elementary teachers. However, even secondary teachers lack confidence in their ability to teach certain topics such as odds (Watson, 2001), comparing distributions, association, and developing a research question (Harrell-Williams et al., 2014).

Not all attitudes toward statistics are negative. Hannigan et al. (2013) administered the Survey of Attitudes Towards Statistics (SATS) instrument to 116 preservice mathematics teachers, both undergraduate and graduate. All of the undergraduate preservice teachers were in the top 10% of mathematics achievement in secondary school, and all of the graduate preservice teachers had a degree with a strong mathematics component. The SATS instrument measures respondents’ attitudes toward statistics with six different components--affect, cognitive competence, value, difficulty, interest, and effort--using statements with Likert-like responses ranging from 0 (strongly disagree) to 7 (strongly agree). The participants had somewhat positive feelings toward statistics, with mean scores of 4.8 and 5.0 for affect and interest, respectively, where 4 represents responses of *neither agree or disagree*. In addition, they placed a relatively high value on statistics (mean=5.5). In general, they rated their ability to do statistics fairly positively (mean=5.1), although they rated statistics as somewhat difficult (mean=3.7, with lower scores indicating higher difficulty) and as requiring some effort for them to succeed at (mean=5.8). Graduate students tended to have more positive feelings toward statistics than undergraduates, and rated statistics as slightly less difficult.

It is important to keep in mind that many participants in the studies in this section are preservice and/or elementary teachers. Elementary teachers tend to differ from secondary teachers in the number, depth, and nature of statistics courses experienced as a student (Banilower et al., 2018). It is reasonable to believe that this difference in experiences is likely to result in differences in knowledge, beliefs, and attitudes (Estrada, 2002, as cited in Estrada, 2011; Pierce & Chick, 2011). Preservice teachers also tend to have different attitudes toward statistics than inservice teachers, with preservice teachers viewing statistics as more useful, easier to understand, and more interesting (Estrada, 2002, as cited in Estrada, 2011). As the teachers in the current study are inservice secondary teachers, it is important to keep these differences in mind when considering the types of knowledge, beliefs, and attitudes these teachers possess.

Teacher Decision-Making

In the preceding two sections, we first examined several studies that suggested that knowledge and beliefs have a significant impact on teachers' instruction. Then, we examined common knowledge, beliefs, and attitudes that teachers tend to possess regarding statistics and teaching statistics. Next, we turn next to an examination of the process by which teachers make decisions in the classroom. We will focus on two theoretical models that attempt to explain the mechanisms by which knowledge and beliefs impact teachers' classroom instruction, and in particular, the instructional decisions that they make.

Teaching-in-Context

Schoenfeld (1998) presented the theoretical model of *teaching-in-context* (Figure 1) to explain "precisely why teachers make particular choices at each point of instruction and precisely which beliefs, goals, and knowledge those decisions depend upon" (p. 2). The model builds on several prior research studies that show that beliefs and knowledge are important determinants of what happens in the classroom, how beliefs and knowledge are formed in teachers, and how teachers use their beliefs and knowledge in planning a lesson. By focusing on the interactions between teachers' beliefs, knowledge, goals, and what teachers see and perceive in the classroom, the model attempts to describe the mechanisms by which teacher decision-making takes place. The model has several components, each of which will be briefly discussed, as they each inform a key part of the conceptual framework for the current study.

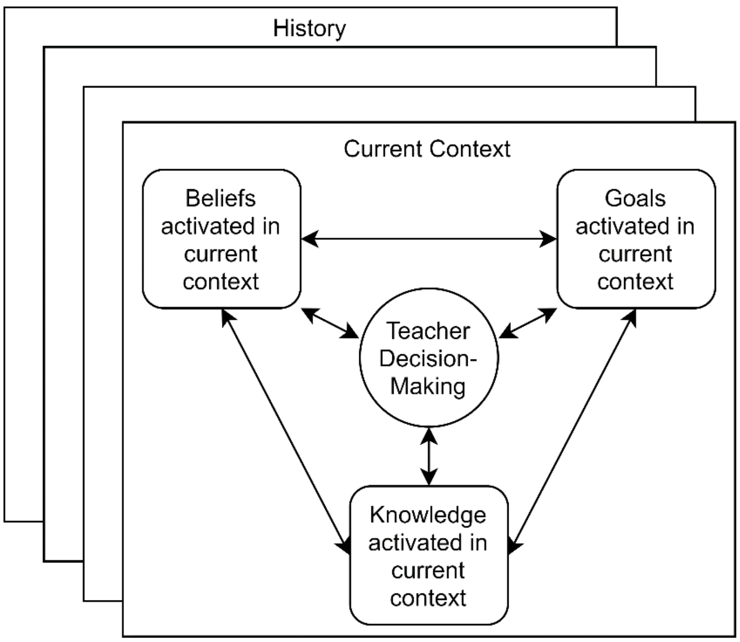


Figure 1. Teaching-in-context. Adapted from Schoenfeld (1998).

History refers both to the teacher's personal history and to the teacher's history with the students in the classroom. Relevant personal history can include past teaching experiences, or it can include such things as professional development and experiences as a student. Along with history with the students in the class, this serves as the source of teachers' beliefs and knowledge.

Beliefs include a number of different beliefs that inform teacher decision-making. This includes beliefs about how students learn, about what they need to know, about what constitutes good teaching, and various others. Similarly, *knowledge* represents the various pieces of knowledge that teachers possess relevant to teaching. This includes such types of knowledge as knowledge about the content being taught, knowledge about the students in the classroom, and knowledge of different learning activities that can be used in instruction about a particular topic. *Goals* can include long-term goals for an entire unit or course, or they can include short-term goals concerning what the teacher wants to happen immediately. Together, beliefs, knowledge, and goals serve to inform teachers' decision-making. Importantly, these three components all impact each other. If a teacher believes that knowledge of statistics is useful in everyday situations, then that teacher may have a goal of getting students to see the utility of statistics. If a teacher knows that a student has struggled in past mathematics courses, then the teacher might believe that different teaching methods may be more effective with that student than with other students.

Current context refers to the *immediate* context *as perceived by the teacher*. That is, although there may be several contextual factors that impact a teacher's instruction at one point in time or another (e.g., expectations from administration, access to technology), it is only the context that the teacher is cognizant of *in the moment* that will have an impact on the decisions that he or she makes. A critical note is that current context includes the current state of the classroom (as perceived by the teacher), including any event that may have just occurred.

The mechanism by which *teacher decision-making* happens is as follows: a teacher's knowledge, beliefs, and goals are informed by past experiences. Based on these, a teacher creates an *action plan* for the lesson. Given a specific current context (i.e., a current state of the classroom as perceived by the teacher), a subset of these knowledge, beliefs, and goals are *activated*. That is, they are given priority over other knowledge, beliefs, and goals the teacher possesses. This current context can be anticipated by the teacher (e.g., an activity has just been completed), or it can be unexpected (e.g., a student inquires about a tangential topic). Drawing on these activated knowledge, beliefs, and goals, the teacher's action plan may be modified. Based on this new (or existing) action plan, the teacher then decides how to respond to the current context of the classroom.

Hypothetical Learning Trajectories and the Mathematics Teaching Cycle

The concept of a learning trajectory and its use in mathematics education has evolved rapidly over the last 25 years. Simon (1995) developed the Mathematics Teaching Cycle (Figure 2) to describe, from a constructivist perspective, how a teacher makes decisions when planning, implementing, and revising a lesson. At the core of the Mathematics Teaching Cycle is a *hypothetical learning trajectory* (HLT). The HLT consists of three components: the teacher's learning goal, the teacher's prediction of the learning process or conceptual path students can take to reach that goal, and the teacher's plan for activities to guide students along that hypothesized path to reach the goal. These three components, according to Simon (1995) are interdependent. If the learning goal changes, then so too might the planned activities. If the teacher's hypothesized learning path changes–for example, by observing students unexpectedly struggling to progress–then that too might change the planned activities.

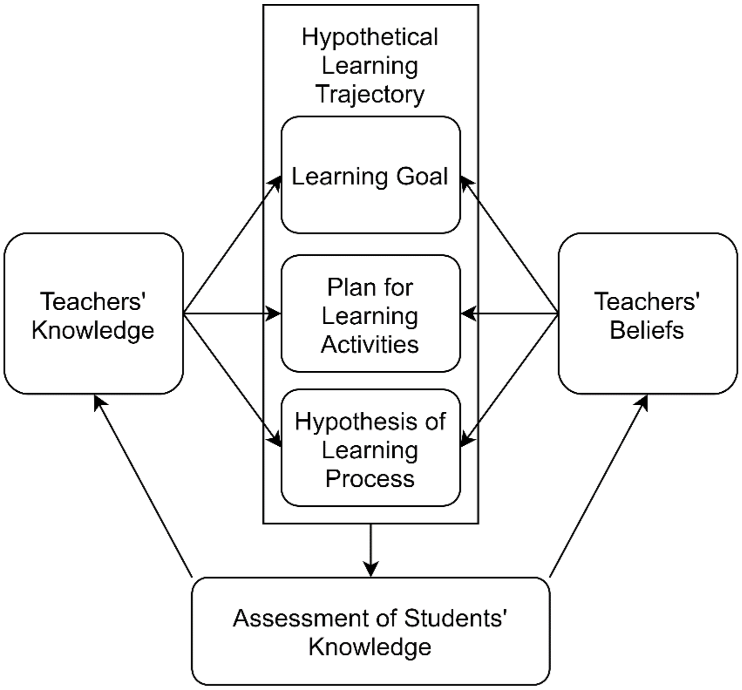


Figure 2. The mathematics teaching cycle. Adapted from Simon (1995).

The components of the HLT are informed by teachers' knowledge and beliefs. For example, a teacher's planned activities will be influenced by knowledge of the content and of potential tasks that can be used for that content. The learning goal for the lesson will depend on the hypothesized learning path that students can take from the content that has previously been taught. As teachers implement a lesson, they will assess the effectiveness of the lesson. In particular, they will assess the knowledge that students gained. This assessment can happen in real-time as the teacher observes students working, or it can happen retrospectively, after a lesson has completed or as the teacher is examining student work produced. As a result of this assessment, teachers' knowledge and beliefs are modified. Based on these new knowledge and beliefs, any of the three components of the HLT may also then be modified. This cycle of implementation and revisions can take as long as an entire course, or it can happen nearly instantaneously.

Since Simon's (1995) work, research on HLTs has expanded to encompass the more general concept of *learning trajectories* (LTs) and includes much more than the decision-making of teachers. Although learning trajectories were originally a "tool for individual teachers to make sense of their own students' day-to-day progress and to frame their moment-to-moment and day-to-day instructional planning," the phrase *learning trajectory* "increasingly signals research that aims for a systematic, detailed description of the likely progression of children's reasoning about big ideas of mathematics over long periods of time" (Maloney et al., 2014, p. xiii).

As the focus of this study is day-to-day instructional decision making of teachers, the ways in which HLTs will be incorporated into the framework of this study will be more closely aligned with the earlier description. However, using a lens of learning trajectories allows researchers to connect instructional decision-making with other aspects of instruction such as curriculum, standards, and assessment (Lobato & Walters, 2017). Because Simon's (1995) construction of an HLT plays a key part in this study's conceptual framework, we will examine how LTs have been used to examine instructional practice, paying particular attention to: 1) the role of the LT in teacher practice and decision-making, and 2) LT's applications in the statistics classroom.

Learning Trajectories and Decision-Making

Oftentimes, learning trajectories are discussed only in the sense of a conceptual path along which students' thinking may progress – i.e., only one of three components of Simon's (1995) HLT. Although studying these learning paths alone may be useful, teaching and learning are inextricably connected, and the power of LTs comes when we link these conceptual paths to instructional tasks that can help students advance down those paths (Clements & Sarama, 2004;

Daro et al., 2011). Clements and Sarama (2004) emphasize the ever-changing nature of LTs and the role the teacher has in constructing and refining LTs:

> The teacher must construct new models of children's mathematics as they interact with children around the instructional tasks and thus alter their own knowledge of children and future instructional strategies and paths. Thus, the realized learning trajectory, the taken-as-shared practices and understandings, are emergent (p. 85).

Steffe (2004) shares a similar sentiment: "learning trajectories of children must be constructed by teachers/researchers who participate first-hand in children's constructive activity" (p. 155).

Daro et al. (2011) describe how teachers use LTs (knowingly or unknowingly) in practice. Teachers start by selecting a set of ordered instructional experiences and tasks. These experiences and tasks are chosen either because the teacher believes they will help students progress down a LT toward a desired goal, or because they are part of a curriculum or instructional materials designed based on the similar beliefs of others. What separates this selection and ordering of experiences from other instructional guidelines is that the hypothesized order is based on how the teachers (or curriculum designers) conceive of students' thinking and how that thinking might develop (Clements & Sarama, 2004). This is only the beginning of the role of LTs, however, according to Daro et al. (2011). Next, the teacher must pay careful attention to whether students' knowledge is progressing in a way that matches the hypothesized learning path. This requires that teachers assess students' knowledge and thinking. This assessment is aided by the teacher's own knowledge--of mathematics, of LTs, of the students, etc. Although this assessment can happen at the conclusion of a lesson or unit (e.g., by examining test responses), the use of LTs is more successful if a teacher can continually assess students' thinking and adapt their instruction on the fly (Arnold et al., 2018; Daro et al., 2011;

Petit, 2011). What is important to note is that successful use of a LT also requires that teachers assess not only the *correctness* of student responses, but the thinking processes that caused the students to arrive at those responses (Daro et al., 2011). The teachers' use of a LT will likely be more successful if the LT focuses on a single topic of instruction, even if the reality of students' learning is a complex web of interrelated learning paths (Clements & Sarama, 2004).

Amador and Lamberg (2013) used the lens of Simon's (1995) mathematical teaching cycle and HLT to examine how four fourth-grade teachers planned and enacted mathematics lessons. Each participant took part in 10 teaching sets--a pre-lesson interview, an observed lesson, and a post-observation interview. One additional interview was conducted with each participant which asked about their planning process, influences on this process, and their views on education, the curriculum, students, and assessment. In addition, lesson plans and photographs of teaching artifacts were collected for analysis. For three of the four teachers, the pressures of high-stakes testing were a driving force behind the planning of their lessons--something that may not be adequately captured in Simon's model. This suggests that a component of *context* may serve as a useful addition to the model (Amador and Lamberg alternatively proposed that a *hypothetical testing trajectory* may be a better model than a hypothetical learning trajectory for some teachers). For the fourth teacher, beliefs and knowledge about how students learn the content being taught was the driving force behind her lesson planning. Other components of knowledge that appeared to be significant factors for the teachers include content knowledge and knowledge of mathematical representations.

Learning Trajectories in Statistics

Statistics educators and researchers are increasingly looking toward learning trajectories to support students' development of statistical concepts (Arnold et al., 2018). Lehrer et al. (2014)

describe the process of developing a LT for middle school students' understanding of data modeling. This LT was designed by researchers and given to teachers who were not involved in the development process, along with assessment items to use in their classroom. They found that the LTs needed heavy revisions before teachers were comfortable incorporating them into their practice. Even after extensive professional development, many teachers struggled to use the LTs when planning their instruction. Many teachers indicated that they would rather be told which tasks to use with their students and in what order.

Arnold et al. (2018) present three examples of studies in statistics education that investigate the use of learning trajectories. In the first case, presented by Confrey and Jones, the researchers designed a LT around the topic of variability, and used a series of learning activities with sixth-graders to confirm or modify the LT. This case shows that even when a LT is based on extensive research and expertise, it can still require several iterations based on assessments of students' knowledge and analysis of whether they are progressing as expected. In the second case, presented by Arnold, the research team started with a learning goal that would become the driving force for the creation of a LT for a topic (making inferences by comparing samples) that was new to the curriculum. Arnold describes the process by which the LT was repeatedly modified based on observations of students' difficulties, their reactions in the classroom, and on analysis of student data after the lessons. This case shows that this design-based research approach can be used to construct a LT and a series of instructional activities, even for new approaches to statistics. The final case, presented by Lee, shows how a LT was used to develop a sequence of tasks designed to assist adult learners in conceptualizing a repeated sampling approach to inference. When students worked through the sequence of tasks, several of them struggled with the last two tasks. Noticing this, the instructors (which included the designers of the LT) made the decision to add two additional tasks to the trajectory intended to assist students in reaching the intended learning goal.

From this collection of studies using LT in statistics instruction, we see that even in a field such as statistics where LTs are a relatively new concept, LTs can be used to construct and refine new approaches to teaching statistics. The use of LTs in a classroom can be successful regardless of whether the developers of the LT are the teacher themselves. However, these studies show that success seems much easier when the teachers are involved in the construction of the LTs, or at the very least, have sufficient knowledge of the hypothesized learning paths that the LT entails. Taken together with the other research on LTs in mathematics education, we can see the crucial role that the teacher plays in the successful implementation of a LT. All of these LTs necessitated refinement based on the teacher's assessment of student knowledge. They all showcase the iterative process of choosing or starting with a learning goal, hypothesizing a learning path for students to take to get there, selecting tasks and activities to guide students along that path, assessing students' knowledge and correspondence with the hypothesized path, and modifying the hypothesized learning path and/or the instructional activities based on that assessment.

Other Approaches to Teacher Decision-Making

Stockero and Van Zoest (2013) examined the instructional practice of six beginning mathematics teachers in an attempt to characterize *pivotal teaching moments*--events that occur in the classroom that provide an opportunity for a teacher to modify instruction to support student understanding--as well as the ways in which the teachers decided to respond to those events. Five types of pivotal teaching moments were identified: (a) *extending*, in which a student inquires about a topic that is related to the mathematical content being discussed, but was not a part of the teacher's original plan of instruction; (b) *incorrect mathematics*, in which a student shares an incorrect solution or incomplete mathematical thinking; (c) *sense-making*, in which a student verbalizes his or her attempts to understand mathematical concepts; (d) *mathematical contradiction*, in which two competing solutions or methods are presented; and (e) *mathematical confusion*, in which a student expresses a lack of understanding about a particular topic. Some of these pivotal teaching moments had the potential to have significant impact on student learning, while others had a moderate potential to do so.

Stockero and Van Zoest (2013) also examined the decisions teachers made in responding to the above pivotal teaching moments. Five types of responses were identified: (a) *extends mathematics and/or makes connections*, in which the teacher goes beyond the material that was originally planned to be discussed in order to examine a related topic; (b) *pursues student thinking*, in which the teacher attempts to find out more about what a student is thinking; (c) *emphasizes meaning of the mathematics*, in which the teacher highlights an underlying definition or underlying mathematics of the issue being discussed; (d) *acknowledges, but continues as planned*; and (e) *ignores or dismisses*. Examining the instruction of a secondary geometry teacher, Cayton et al. (2017) identify two additional types of pivotal teaching moments unique to technology-rich classrooms: *technology confusion* and *incorrect technology use*, as well as one additional type of teacher response: *repeat technology directions*.

Jacobs et al. (2010) examine in-the-moment teacher decision-making through the lens of *professional noticing*. Focusing on responding to mathematical thinking, the authors describe the process of making decisions as dependent on three skills: (a) attending to children's strategies, (b) interpreting their understandings, and (c) using these understandings in deciding how to respond. Importantly, these are *skills*, in that they can be learned and developed over time and are impacted by past teaching experiences and professional developments.

Conclusion

Teaching-in-context and the mathematics teaching cycle provide two models that attempt to explain teacher decision-making. Along with other research that describes the relationship between instructional practices and knowledge and beliefs, these will serve as the foundation of the conceptual framework for this study. Examining other approaches to examining teacher decision-making will allow us to more easily identify and describe events that have the potential to result in alterations to a teacher's initial instructional plans, as well as teachers' responses to those events. A review of common beliefs and knowledge possessed by teachers concerning statistics teaching and learning was presented, which will serve as a baseline to compare teachers to and will help us identify which beliefs and knowledge are influencing teacher decision-making. Key differences between mathematics and statistics were also reviewed, which will help us identify whether teachers of statistics are attending to these differences in the statistics classroom, which will likely affect their instructional decision-making.

CHAPTER 3. METHODS

This chapter provides a description of the methods of the study and justifications for the study's design. The conceptual framework that informed this study is presented first. This is followed by an explanation of the study's design, including information about the participants in the study. A description is then presented of several trial observations and interviews that were conducted, and the resulting modifications to the study that occurred. Possible biases and limitations resulting from the design of the study are then discussed, followed by steps that were taken to reduce the impact that these biases and limitations may have had.

Conceptual Framework

The conceptual framework of this study (Figure 3) is founded on Schoenfeld's (1998) model of teaching-in-context and on Simon's (1995) concept of the mathematics teaching cycle, incorporating other research on teacher decision-making and on the impact of teachers' knowledge and beliefs on practice. In this framework, teachers have a repository of *knowledge* and *beliefs* as a result of their history (both in and out of the classroom). The knowledge and beliefs appearing in this framework are those that have been found to influence teachers' practice. Influential knowledge includes knowledge of content, of technology, of teaching and learning, and of the students in the classroom. Influential beliefs concern many of the same areas, as well as beliefs about students' needs, desires, and actions.

Teachers use these knowledge and beliefs to construct an *initial instructional plan*. This instructional plan consists of three components: (a) a *hypothesis of learning*, a conception of how knowledge of various concepts are related and how students' knowledge progresses from one concept to the next; (b) a *learning goal* for the lesson; and (c) a set of *planned activities* to help students progress their knowledge and arrive at the learning goal. The formation of this initial instructional plan is impacted by the context in which the teacher teaches. This context could include the technology that the teacher has access to, the expectation for students to take and pass the AP Exam, the pre-requisites for the course, and so on.

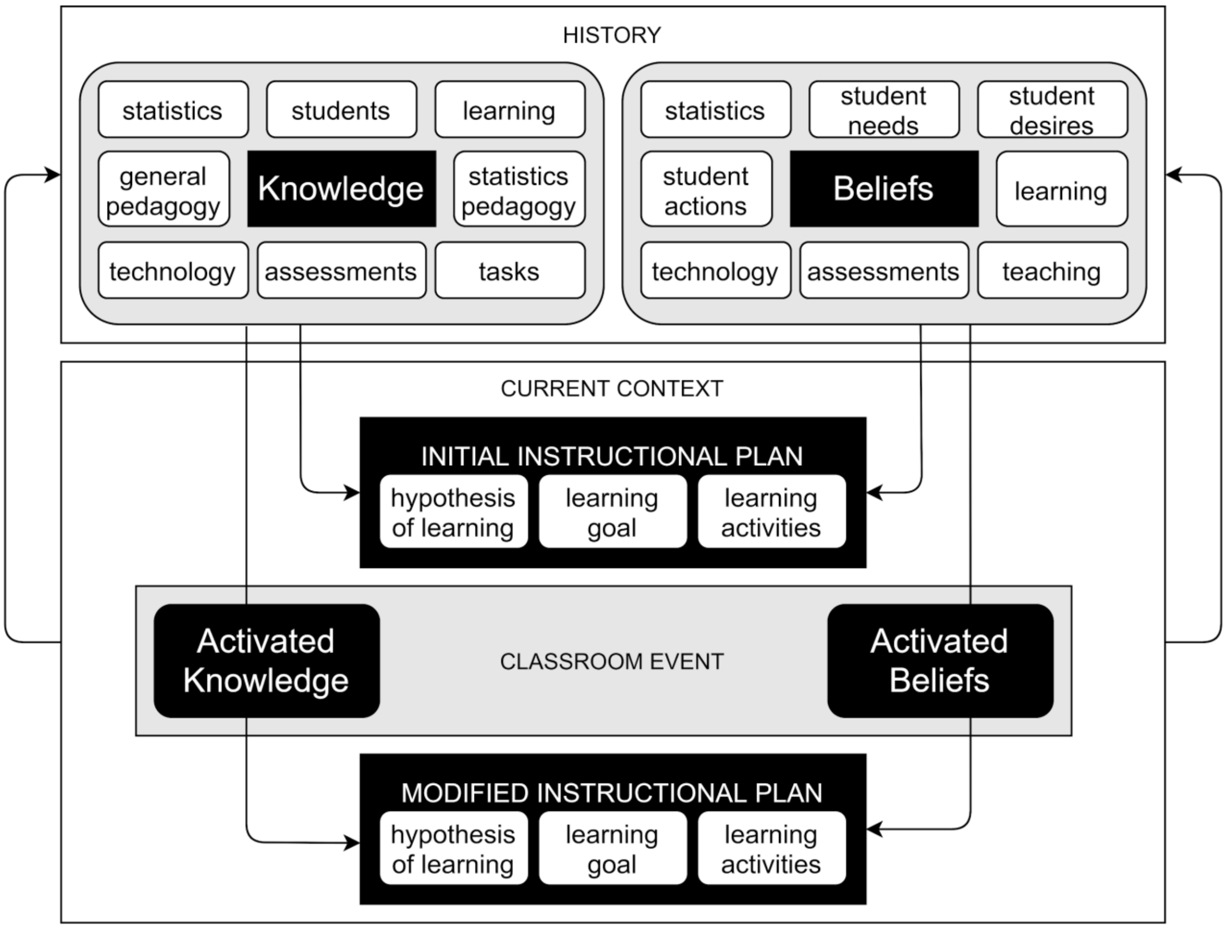


Figure 3. Conceptual framework for this study.

As the teacher implements the initial instructional plan, *events occur in the classroom* that require a response from the teacher. Some of these events may be pivotal teaching moments (Stockero & Van Zoest, 2013), the response to which may have a significant impact on students' learning. These events could be novel solutions presented by students, a disagreement about a solution among students, or a student struggling to progress in a task. Other events may have less of a direct impact on student learning (e.g., a student asking a question not related to the content, an interruption to the classroom), but nonetheless require a response from the teacher.

As a result of events that occur, a certain subset of the teacher's knowledge and beliefs become *activated*, taking a higher priority over other knowledge and beliefs. Based on these activated knowledge and beliefs, the teacher may (or may not) choose to alter their initial instruction plans, creating a new *modified instructional plan*. Any or all of the components of the instructional plan can be modified. The modification of this instructional plan is manifested in the way the teacher responds to the event. Which knowledge and beliefs get activated, and how a teacher's instructional plan is modified, can also be impacted by the context in which the teacher is teaching.

After these alterations are enacted, this experience--the event, the alteration of the instructional plan, and the teacher's response--and the aftermath become part of the teacher's *history*, perhaps altering the teacher's knowledge and beliefs. These new knowledge and beliefs may then be activated by the next classroom event during the current lesson, or they may impact the teacher's initial instructional plan for a future lesson.

This process can happen multiple times in a lesson, resulting in several different alterations to a teacher's instructional plan at various points in the lesson. During a lesson which primarily consists of teacher-led lecture with few questions asked, there may be few events that require a response from the teacher, and thus the teacher's instructional plan may be unlikely to change during the lesson. On the other hand, when there are many chances for students to express their understanding, or in lessons that are not unfolding as the teacher expects, there may be several instances where teachers modify their instructional plan as a result of different knowledge or beliefs being activated or considerations of their current context.

Study Design

The following questions guide this study:

1. For teachers of secondary statistics, how are instructional plans for teaching statistics lessons created?

   1a. What knowledge and beliefs do teachers draw upon when creating instructional plans?

   1b. What current contexts seem to influence which of these knowledge or beliefs take priority?

2. How are secondary statistics teachers' instructional plans for statistics lessons modified after creation?

   2a. What events cause teachers to modify their instructional plans for statistics lessons, either during or after implementation?

   2b. Which knowledge and beliefs take priority due to these events?

A qualitative approach was taken to answer these research questions. According to Creswell (2013), qualitative research has several key features that make it an appropriate approach to answering these research questions. Qualitative research allows one to explore an issue, getting a more complex, deeper understanding. It allows us to examine "the processes that people experience, why they responded as they did, the context in which they responded, and their deeper thoughts and behaviors that governed their responses" (p. 48), all of which is needed to help answer these research questions. Using qualitative research, we can focus on the process by which teachers' instructional plans are created, rather than only on the end result. Qualitative research is also appropriate in this study since it allows us to examine and interact with teachers in their natural setting of their classroom, as they plan, implement, and revise actual lessons.

For this research, a collective instrumental case study (Stake, 2005; Yin, 2009) was used. A collective instrumental case study is one in which we explore an issue through the examination of multiple participants. It allows us to develop themes and identify patterns between teachers, but also acknowledges the impact that specific contexts and classrooms have on each individual teacher.

Defining the Case

In this study, the case being examined was a collection of seven high school statistics teachers during the fall semester of 2019. Each teacher was teaching a course in which the primary content focus of the course was statistics and probability. Some teachers taught multiple sections (i.e., groups of students) of a statistics course during the semester, whereas others taught only a single section. The teachers taught in seven different schools across multiple school districts in a single state in the southeastern United States.

Ethical Considerations

The study received North Carolina State University Institutional Review Board (IRB) approval (Protocol No. 16688) for research involving human subjects. All participants provided their consent to participate by signing an informed consent form (Appendix A) that informed them of the purpose of the study, risks and benefits of participation, and other relevant information about the study. Written consent via email was also obtained from administration (e.g., principal, director of research, department head) at each participating school or district. For the one district that required it, a formal application for permission to conduct research was also submitted, which was subsequently approved. All students in participants' observed courses were provided with a child assent/guardian consent form (Appendix B) that informed them of the purpose of the study, the student's role in the study, and other relevant information. If a student did not return the form signed by themselves and a guardian, the student still participated in all observed lessons, but were seated outside the view of the camera, and care was taken to avoid recording the student's face.

Participants and Sampling

Participants for this study were chosen primarily via a convenience volunteer sample. To be eligible for participation in this study, a teacher had to be teaching at least one section of a course to high school (grades 9-12) students in which the content primarily consisted of statistics and/or probability. Potential participants were identified via personal contacts known to the researcher and his colleagues, as well as via websites of nearby schools and school districts that were determined to be amenable to research requiring observations and possible video recordings. From this pool of candidates, an initial set of 22 teachers was chosen, primarily based on proximity to the researcher. Each of these teachers were then invited via email to participate in the study. Additionally, for three districts or schools in which statistics teachers could not be identified, an administrator or research liaison was identified. Each of these administrators or research liaisons were then sent an email requesting permission to conduct the study and for assistance in identifying or recruiting teachers in the school or district. Once a teacher expressed interest in the study, permission was sought from the school or district administration if it had not already been obtained.

Of the 25 initial contacts (22 teachers and 3 administrators), 16 responses (15 teachers and one administrator) were received. Of these, four teachers indicated that they were not teaching statistics during the semester for which the study was planned. Three other teachers expressed interest, but permission from administration was not promptly received. One administrator granted permission but was not able to readily identify teachers who were willing to participate. One other teacher was eliminated from consideration due to district policies that were not amenable for the research. This resulted in seven remaining teachers who make up the group of participants for this study. Considering the variety of contexts that these seven teachers taught in, it was determined that no further recruitment was necessary. Information about the seven participants and their school contexts is shown in Table 1.

Table 1. Participants' teaching experience and current statistics teaching assignment.

Teacher	**Years of teaching experience[a]**	**Number of prior times teaching statistics[a]**	**Statistics course taught during study**	**School**
Ms. Andrews	8	~8	Statistical Methods I (a credit-bearing course at a local college)	Selective public early college high school
Ms. Baker	14	~12	AP Statistics	K-12 charter school
Ms. Carey	6	~9	AP Statistics	Private religious secondary school
Mr. Dennis	16	~32	Advanced Analytics and Statistics	Private college-prep secondary school
Mr. Enloe	9	~18	AP Statistics	Private college-prep high school
Mr. Fahey	7	4	AP Statistics	Selective public residential high school
Ms. Greene	4	1	AP Statistics	Traditional public high school

[a]: does not include current year/semester. Some teachers teach multiple sections at a time, each of which is included in the number of times teaching statistics.

Data Instruments and Collection

For each participant, data consisted of: (a) one pre-observation interview prior to the first observation; (b) a series of lesson observations (typically three consecutive lessons); (c) a post-observation interview after each observed lesson; and (d) artifacts of teaching. Each part of the data collection was designed to capture information relating to specific parts of the conceptual framework and to answer the research questions that guide this study, as detailed below.

Pre-Observation Interview. The first part of the three-part pre-observation interview (Appendix C) was designed to assess (a) the teacher's background and teaching experience; (b) contextual factors related to the school, course, and students; (c) norms established in the classroom; (d) overall goals for the course. The second part of the pre-observation interview was designed to assess influences on teachers' instructional planning including: (a) teachers' preparation for teaching statistics; (b) influential knowledge and beliefs; (c) available resources for teaching statistics and how they are used; and (d) influential contextual factors. The final part of the pre-observation interview was designed to assess the three components of the teacher's instructional plan—the learning goals for the students, the planned activities, and the hypothesized learning processes of the students—first for the series of upcoming lessons as a whole, and then for the next observed lesson. The interview was semi-structured, allowing for the researcher to ask follow-up questions to elicit further information. It was administered to the teacher prior to the first classroom observation, either in person or virtual, at a time convenient for the teacher and researcher. The interview typically lasted 30-45 minutes. All pre-observation interviews were audio and video recorded.

Lesson Observations. Six of the seven teachers were observed for three lessons, and the seventh was observed for two lessons. For each teacher, the lessons observed were consecutive lessons for a particular section (i.e., group of students) of the statistics course that the teacher taught, which may or may not have been on consecutive days. For example, one teacher's statistics course was observed on Wednesday, Friday, and the following Monday, since he did not teach statistics on Thursdays. The lessons were chosen by the teachers and researcher to avoid formal assessments or other activities in which limited teacher practice could be observed. Otherwise, no constraints were provided for the topic or type of lessons to be chosen. Prior to the start of the lesson, copies of any materials that the teacher planned for students to use were requested so that the researcher was aware of what materials the students were viewing. This included such materials as student handouts or web links located on the class's learning management systems. There were three main purposes of the observations: 1) to record instances that illustrated a teacher's instructional plan, including any knowledge or beliefs that the teacher seemed to be drawing upon, as well any contextual factors that seemed to be influencing teachers' instruction; 2) to identify and record events that occurred that could potentially result in a modification to a teacher's instructional plan; 3) to identify how the teacher responded to these events. Teachers were given the option to consent to video and audio recording, audio recording only, or neither. Six of seven teachers consented to video and audio-recording. However, student and guardian permissions were not able to be obtained for one of these six teachers, so only five teachers' lessons were video- and audio-recorded. One teacher's lessons were audio-recorded only, and the seventh teacher's lessons were neither audio- nor video-recorded. For all participants whose lessons were video-recorded, student assent and parental/guardian consent for recordings were obtained (or student consent if the student was over 18 years of age). If a student did not assent or their parent/guardian did not consent, the student participated in the lessons as usual, but was seated outside of the view of the video camera, and all attempts were made to ensure they were not video-recorded. Audio recordings used a wireless microphone attached to the teacher. Video recordings were focused primarily on the teachers and their interactions with students in the classroom.

Regardless of whether the observations were video- or audio-recorded, the researcher used an observation form to record field notes during each observation. In addition to logistical information (e.g., date, teacher, class period), two categories of information were recorded in the field notes. First, the researcher recorded any teacher decisions that he believed were illustrative of the teacher's initial instructional plan that was not discussed in the previous interview. These often included events such as the grouping of students for an activity, or the decision to start class with a homework review--things that the teacher may have not mentioned when describing their instructional plan, but could nonetheless be indicative of particular knowledge or beliefs held by the teacher. These decisions were recorded in the field notes so that the teacher could later be asked about them in the post-observation interview. Second, the researcher recorded any noticed event that had the potential to modify the teacher's instructional plan. Example events that were recorded include questions that students asked, novel solutions proposed by students, students struggling to complete an activity, or having limited class time available. For each of these events, the researcher recorded the time of the event (using the current timestamp of the video or audio recording if available) and a brief description of the event. In addition, he recorded a brief description of how the teacher responded to this event (see observation protocol in Appendix D). See Appendix F for an example of how this looked in practice.

Post-Observation Interviews. Either immediately after each observed lesson or later that day, the teacher participated in a post-observation interview (Appendix E). The interview consisted of three parts. The first part was a semi-structured interview focused on (a) the teacher's perception of how the lesson went; (b) deviations from the initial instructional plan; and (c) changes to the instructional plan that the teacher foresees for the next time the content is taught. The second part of the interview was a stimulated recall, using the audio or video recordings and field notes from the observations. First, teachers were asked about decisions recorded in the field notes that seemed to be part of the teacher's initial instructional plans. Teachers were asked whether these decisions were indeed part of the teacher's initial instructional plan for the lesson, or whether they were in-the-moment decisions. If teachers indicated that these decisions were pre-planned, then as in the pre-observation interview, they were asked to discuss the reasons for including that decisions in the planning of the lesson. Next, teachers were asked to discuss the events that were recorded in the field notes that the researcher deemed had the potential to alter a teacher's instructional plans. The researcher identified each event in turn for the participant. If recordings were available, the interviewer located that event in the recording, using the timestamps recorded in the field notes. After watching the recording of the event or listening to a description of the event, the teacher was asked about their thought process at the time, their response to the event, and how the event may have altered their plans for the rest of the lesson. Some recorded events that the researcher deemed redundant or less relevant to the research questions were skipped for the sake of time. The third part of the interview was similar to the final five questions of the pre-observation interview, intended to assess the teacher's instructional plan for the next observed lesson. For the post-observation interview conducted after the final day of observations, part three was omitted. Post-observation interviews generally lasted approximately 45 minutes.

Artifacts of Teaching. Prior to each observed lesson, participants were asked to provide any resources that were planned to be used during the lesson. This included handouts, links to websites, documents on course websites, and so on. These artifacts served three main purposes. First, they were used to assist the researcher during the observations, so that he was aware of material being worked on and discussed by the students and teacher. Second, these artifacts were analyzed for indications of the teacher's instructional plan, which were then sometimes discussed during the post-observation interviews. Third, some of these artifacts were also used during the stimulated recall portion of the post-observation interview, to help identify events that may have caused a modification in the participants' instructional plan.

Alignment of Interview Protocols and Conceptual Framework

The conceptual framework for this study (Figure 3) informed the design of the interview protocols used in this study. While different components of the framework emerged throughout the interviews, specific questions were designed to capture the various components, as follows. *Knowledge* and *beliefs* that influence participants' initial instructional plans are captured in questions 7-10 (knowledge and beliefs that influenced participants' overall approach to teaching statistics) and questions 15-16 (knowledge and beliefs that influenced the planning of the first observed lesson) of the pre-observation interview, as well as in question 5 (knowledge and beliefs that influenced the planning of the just-observed lesson) and question 10 (knowledge and beliefs that influenced the planning of the next observed lesson) of the post-observation interview. These questions also address any *history* that a participant had that may have given rise to these knowledge or beliefs.

The different components of the *initial instructional plans* are captured in both the pre-observation interviews and post-observation interviews. *Learning goals* are captured in questions 4 and 13 of the pre-observation interviews and question 7 of the post-observation interview. *Hypotheses of learning* are captured in questions 12, 14, and 17 of the pre-observation interview and questions 8, 9, and 11 of the post-observation interview. *Learning activities* are captured in questions 7 and 14 of the pre-observation interview and questions 5 and 8 of the post-observation interview.

Contexts that influenced initial instructional plans are captured in question 3 (curricular contexts), question 5 and 6 (student contexts), questions 8 and 15 (classroom resources), and questions 7 and 11 (various contexts) of the pre-observation interview, as well as in analogous questions from the final part of the post-observation interview. *Contexts* that influenced modifications to instructional plans are captured in questions 2, 4, and 6 in the post-observation interview.

Classroom events that had the potential to result in modifications to participants' instructional plans are captured in the post-interviews. Specifically, questions 1-3 and question 6 (the stimulated recall) prompt participants to discuss events that occurred in the classroom. *Knowledge and beliefs that are activated* as a result of the above events and contexts are captured in questions 2-4 and question 6 (the stimulated recall), which prompt participants to recall reasons for responding to events in the manner that they did. These questions also address participants' *modified instructional plans* as they describe their responses to events.

Trial Observations and Interviews

Several trial interviews and observations were performed to gain experience with data collection and to refine data collection methods and interview questions. What follows is a description of each trial and the resulting changes to the study methods that resulted.

Trial One

The first trial consisted of a pre-observation interview and a single observation followed by a post-observation interview. The pre-observation interview occurred the day before the observation and was not recorded, but was otherwise administered under similar conditions to pre-observation interviews that occurred during the study proper. The purposes of the trial pre-observation interview were to obtain feedback on the content and clarity of the questions, to obtain an estimate for the expected length of the interview, to experience how a participant might respond to questions, and to obtain experience creating and delivering follow-up questions to responses that might be given.

The teacher was then observed during a single class period in a college introductory statistics course in a classroom setting (i.e., class size, seating configuration) likely similar to many high school statistics classrooms. The lesson was not recorded, but a video camera was turned on so that the researcher could see in real-time what would have been recorded if the recording was enabled. One purpose of the observation was to see if the single researcher would encounter difficulties taking field notes while simultaneously directing the video camera.

Approximately 30 minutes after the conclusion of the lesson, the teacher participated in a post-observation interview. This interview was not recorded. Since the classroom observation was not recorded, recordings could not be used in the stimulated recall portion of the interview. Otherwise the interview was similar in nature to those that were conducted during the study proper. The purposes of the trial post-observation interview included all of those from the trial pre-observation interview, in addition to assessing the level of comfort of the participant in answering questions about decisions that she had made, and determining if the level of detail of the recorded events was sufficient for the participant to recall and respond to.

No substantial difficulties were encountered during the observation in regard to simultaneously manipulating the camera while taking field notes that would seem to impact the quality of field notes taken or the quality of the recording. Feedback given from this participant on all three parts of the trial was all positive in nature. At the conclusion of the post-observation interview, the participant expressed, without any prompting from the researcher, that none of the questions asked caused any discomfort.

Several changes to the methods and interview questions emerged due to this trial. First, the researcher was unable to hear many of the conversations the teacher had with small groups during the observation, particularly those that were not near the researcher. Thus, for future observations, a wireless microphone was placed on the teacher that fed into a listening device on the researcher.

Second, the teacher directed the students to a web link on the course website for which the researcher did not have access to. This website had resources that the students used during their group work. Having access to these resources would have enabled more detailed recording of field notes and led to better comprehension of the conversations being observed. Thus, prior to any future observed lessons, the researcher obtained from the teacher any resources (e.g., handouts, URLs, documents on course websites) that were expected to be used during the lesson.

Another resulting change regarded the observation protocol. Many of the decisions the teacher made seemed to be made prior to the start of the lesson. For example, the choice of task, group size, etc. Given that these are decisions that should likely be asked about in the post-observation interview, a section was added to the observation protocol to record any observable planning decisions that the researcher was not aware of prior to the observation. Questions asking participants about these decisions were added to the post-observation interview.

Several modifications to interview questions were made after this trial to help ensure that participants' responses could be used to answer the research questions and to address components of the conceptual framework. For example, from the data gathered during this trial, the researcher's ability to ascertain the participant's hypotheses about student learning was limited. As this is potentially an important factor in the decision-making of teachers, questions were added to the pre-observation interview asking about the planned structure and ordering of the next few lessons as a whole (rather than just a single lesson) and the expected progression of students' knowledge over that series of lessons. Other questions added to the pre-interview to help data better align with the research questions and the conceptual framework include questions about student characteristics, norms of the classroom, curricular issues specific to the school, and the participant's goals for the course. Some questions were reworded or had potential follow-up questions added, due to the participant seemingly not interpreting the question as it was intended, giving a response that did not provide information relevant to the research questions.

In the post-observation interview protocol used in this trial, participants were asked about planning decisions for the lesson that was just completed. This trial raised concerns about possible biases in participants' views on these decisions given that they had seen the impact these decisions had. Therefore, the post-observation interview protocol was adjusted so that participants were asked about their planning decision for the *next* lesson that was going to be observed (if any), rather than the one that was just completed. These new questions were also added to the end of the pre-observation interview to capture planning decisions for the first observed lesson.

The duration of pre-observation and post-observation interviews were 18 and 21 minutes, respectively. With the planned additions to the interview protocols, it was determined that the expected duration of each would be close to the 30-45 minutes that was predicted.

Trial Two

A second trial observation was performed at a later date. This observation did not involve the researcher entering a classroom. Rather, a previously video-recorded classroom lesson from an AP Statistics classroom was used. This observation was video recorded (i.e., a video camera was pointed at the screen while the video played). The purposes of this observation were to continue gaining experience taking field notes and recording events, and to assess the difficulty of quickly locating events in the recording for playback. The lesson was played in real-time, without pausing, while the researcher used the observation protocol as he would in a normal classroom observation. Like the first trial observation, the protocol used did not contain a section for instances that are illustrative of the teacher's instructional plan, so this information was not noted. For illustrative purposes, video excerpts from the lesson can be seen at https://www.youtube.com/watch?v=RRvChxY6uoQ&t=14s. This is not the same video that was used during the trial observation, but it is excerpts from the same lesson. The excerpts shown correspond to various times in the observed video. However, the only response to a captured event that appears in this excerpt is at 3:00 in the above, which occurs immediately after the event recorded at 34:20 in the video viewed during the trial observation. This event and response were noted because the discussion that followed did not seem to be part of the teacher's original plan for the lesson. In the post-observation interview, the teacher would be shown the event and the response that followed, and asked about his decisions regarding how he responded, and whether he felt the event changed his plan of activity for the lesson.

This second video-taped observation resulted in a much lower density of events recorded than the first observation, and many of the events that were recorded seemed less impactful on the teacher's instructional practice (though no post-observation interview was conducted for the second observation to confirm this). It was determined, however, that this was not problematic, as it is expected that some teachers will more often assess their students' knowledge while a lesson is unfolding and use this assessment to modify their instructional plan more frequently (Simon & Tzur, 2004). For example, many of this teacher's questions directed at the class were of the initiate-response-evaluate type (Mehan, 1979), and most of the answers given by students were correct. Thus, these responses were likely confirming the teachers' hypotheses about student learning and unlikely to impact the teacher's instructional plan for the lesson. In regard to the video recording, there was minimal difficulty in locating relevant video clips in a timely manner, and thus, video-stimulated recall was considered to be a viable method to use in the post-observation interviews.

Trial Three

The third trial consisted of only a pre-observation interview. The participant was a community college statistics instructor, a former AP Statistics teacher, a regular AP Statistics exam reader, and a statistics education researcher. The primary purposes of the third trial were to test the newly-created questions in the pre-observation interview, assess the duration of the interview, and to obtain feedback from a statistics education researcher and former AP Statistics teacher regarding possible additions or modifications to interview questions.

Some changes were made to the pre-observation interview due to answers provided by the participant. The order of some questions was changed and several questions were reworded, to help ensure that answers to the questions would be helpful in answering the research questions. Several interview questions regarding participants' decisions were reworded to increase the focus on the reasons behind these decisions, rather than on simply descriptions of the decisions. One question regarding teacher self-confidence was deemed to be of minimal relevance to the research questions and was removed.

Trial Four

Trial four consisted of a pre-observation interview, and then the following day, an observation and post-observation interview. The participant was a community college statistics teacher. The primary purposes of this trial were to get experience conducting a classroom observation using the new observation protocol and to ascertain whether the modified interview protocols would result in data that could be used to answer the research questions.

Following the post-observation interview, the participant was asked for her thoughts on the interview questions. She expressed concern that some teachers may get defensive and feel like the quality of their teaching was being questioned. Therefore additional framing was added to some questions in the interviews reassuring participants that the events and decisions were being discussed so that the researcher could better understand the experience of a statistics teacher, rather than because the events and decisions were unusual or perceived to be suboptimal. As trial five was scheduled for a few days later, any further changes were postponed until after trial five.

Trial Five

Trial five was similar to trial four, except that two consecutive lessons were observed, with a post-observation interview immediately after each observed lesson. The primary purposes of this trial were to identify any issues that might arise when observing two consecutive lessons (e.g., redundant questions) and to continue to determine whether the interview questions would result in answers that would be helpful in answering the research questions.

Two changes to the interview protocols were made as a result of this trial. In the pre-observation interview protocol, questions were added asking about participants' prior teaching experience, both in general and for statistics in particular, because it was hypothesized that teachers with more teaching experience may make instructional decisions differently than those with less experience. In the post-observation interview protocol, a question was added to ask participants if and how they felt that the material in the observed lesson connected to the next day's material, in order to gather more information about participants' hypotheses of students' learning.

At the conclusion of the five trials, it was determined that the study methods would be efficacious in answering the research questions.

Analysis of Data

Pre-Observation Interviews

Data from the pre-observation interviews were primarily used as one of the sources to help answer research question 1 and to gather demographic information about participants. Specifically, Part 1 of the pre-observation interviews was used to gather information about participants' teaching experience, the context in which they teach, and their overall goals and expectations for students in their classrooms. Though the questions in Part 1 do not directly address participants' decision-making process, these are potentially important factors that may impact that process. Part 2 of the interview examines participants' decision-making when planning for statistics lessons and the factors (i.e., knowledge or beliefs) that impact those decisions (research question 1a). It also asks about past experiences and current contextual factors that influence those decisions (research question 1b). Part 3 of the interviews were used to assess the teacher's hypotheses of students' learning and how those hypotheses were formed (research question 1a). Part 3 also addresses specific planning decisions for the first observed lesson (research questions 1a and 1b).

The recordings of the pre-observation interview questions were transcribed verbatim. Five of the seven interviews were transcribed by the researcher, while two were transcribed by a third-party transcriptionist (the transcriptionist also transcribed four of the post-observation interviews). The transcriptionist signed a confidentiality agreement, and secure password-protected means were used to share video recordings. The first transcript provided by the transcriptionist was verified by the researcher by viewing the recording while checking the provided transcript for accuracy. The quality and accuracy of the transcript was deemed to be at a sufficient level to allow the transcriptionist to continue with the other assigned transcripts.

The responses from the interviews were then coded. Passages from the interview transcripts were assigned one or more of five top-level codes representing: *knowledge*, *beliefs*, the teacher's initial *instructional plan*, *current contexts,* and *past contexts*. Each of these five top-level codes contained multiple subcodes, which were also assigned to the passages. For example, the top-level code *Beliefs* has as one of its subcodes *StudentNeeds*, indicating when a participant is expressing their beliefs about students' affective needs or desires. Note that the purpose of the study is not to describe all of a participant's knowledge and beliefs, but only those that impact a participant's instructional plan. Subcodes for the teacher's initial instructional plan consisted of the three components of their instructional plan—learning goal, planned activity, hypothesized learning process—with additional subcodes representing different types of planned activities (e.g., assessment, learning activity). The five top-level codes and the initial list of subcodes were based on the components that appear in the conceptual framework. Some additional subcodes, such as one representing the context of the class schedule were added as necessary when participants expressed that beliefs, knowledge, or contexts not appearing in the conceptual framework had an impact on their initial instructional plans.

To identify the beliefs and knowledge that teachers drew upon when planning statistics instruction, as well as the source of these beliefs and knowledge, passages were then identified that included both of the following: one sub-code from *Knowledge*, *Beliefs*, or *Past Contexts*; and one sub-code from *Instructional Plan.* For example, Table 2 shows portions of selected passages that were coded with both *Beliefs about statistics* and *Planned student activities*. In total, 770 passages were identified that contained a relevant pair of codes (passages containing multiple pairs of relevant codes were counted once for each pair). For each of these passages, if the participant discussed the impact that knowledge or belief had on his or her instruction, a short summary was written that captured the essence of this relationship, as well as the source of that knowledge or belief, if stated. This resulted in 393 summaries. Each of these summaries were then analyzed along with summaries from other passages containing the same pair of codes to identify common themes, first for each individual teacher and then across all teachers. These themes were then used to answer research question 1a.

Table 2. Selected passages coded for Beliefs about statistics and Planned student activities.

Participant	Quotation coded for *Beliefs about statistics* and *Planned student activities*
Ms. Greene	I wanted students to not be given super specific instructions about how to collect data, because that's like real life. People that collect real data have to decide what is the best way to do this. And I wanted them to try things, and say "Oh, that doesn't work very well. Let's try this differently." So, I tried not to give a lot of instructions.
Ms. Greene	I think algebraically working through problems is very different than manipulating data, and interpreting data, and having conversations about [data]. There's not but so much stats you can do, there's a lot of stats that you have to talk about. And so [I have] my students work in pairs to talk about different things.
Mr. Enloe	In algebra classes, they just kind of they get an answer, they circle it and they move on. And the thing about statistics is that we're taking those numbers and we're interpreting what it means in the context of the problem. And so early on in the year, I give them these template sentences.
Mr. Enloe	I'm gonna give them a couple of examples of this computer output that I showed them today, and make them, just real quick, interpret the slope, interpret the intercept, interpret s, interpret r^2, write the equation of the least squares regression line....If you were doing statistics out in the real world, this is how it's gonna be anyway. You're not gonna actually sit and calculate stuff. You're gonna put it into a computer program and have to read the output. So, I think that's a huge skill and I want to make sure they can do that.

Then, to identify contextual factors that may have impacted whether teachers' knowledge and beliefs were put into practice, passages were identified that included both of the following: one sub-code from *Context*, and one sub-code from *Planned activities* or *Intended learning goal.* In total, 332 passages were identified that contained one such pair. For example, Table 3 shows selections from passages that were coded with both *External assessment* (context) and *Planned student activities.* If the passage was also coded with a sub-code from *Knowledge*, *Beliefs*, or *History*, this was noted along with the passage to assist with identification of influential factors that were mediated by the context. Each of these passages were summarized to indicate the way in which the context in which the participant taught impacted his or her instruction, along with any ways that the context mediated the impact that the participant's knowledge or belief had on his or her instruction, if these relationships were discussed. This resulted in 206 summaries. As before, the summaries were then analyzed along with other passages containing the same pair of codes to identify common themes for each teacher and across all teachers. These themes were used to answer research question 1b.

Table 3. Selected passages coded for External assessment and Planned student activities.

Participant	Quotation coded for *External assessment* and *Planned student activities*
Ms. Baker	All of my students will sit for the exam…Huge influence on how I teach the class. I often tell people that we do teach to the AP statistics test. Fortunately, I think it's a really good test…. So, I feel like that's something--it's worth living up to in terms of my material.
Mr. Dennis	I used to [teach probability] just with…some generic problems I got out of textbooks. But over the years, I've found that's never how those questions are asked on the AP exam, so I'm not gonna present them that way [anymore].
Mr. Enloe	We learn the material, then we dive deeper into some multiple-choice [questions]…some free response questions. And we try to get as deep as we can, so that when they go to take the test and then eventually the AP Exam, they've seen a lot of examples of how things are supposed to be asked, and how things are supposed to be answered.

Table 3 (continued).

Ms. Greene	This is AP statistics, and so I'm preparing them to take the AP exam….And so every unit, I try to give a handful of practice AP problems, because it's the best way to assess how they're going to do on the actual AP exam.

Classroom Observations Recordings and Field Notes

The primary purpose of the recordings of classroom observations and field notes were their use during the video- or audio-stimulated recall portions of the post-observation interview. Any events that were identified to be illustrative of a teacher's instructional plan that had not been previously discussed in the pre-observation interviews were recorded in the field notes and discussed in the subsequent post-observation interview. Similarly, any events that occurred that had the potential to alter a teacher's instructional plan were also noted in the field notes and discussed during the subsequent post-observation interview. Thus, the interview transcripts were generally the preferred data source used to analyze events that occurred in the classroom. However, the classroom recordings did serve other purposes during the data analysis.

First, each recording was watched or listened to to identify any instances that were illustrative of a participant's instructional plan that were neither captured in the field notes nor discussed in any of the interviews. The specific component of the instructional plan (learning goal, planned activity, hypothesized learning process) that was illustrated was noted. Because it was no longer practical to ask participants about these events, limited conclusions could be made about the knowledge, beliefs, or past contexts that influenced these instructional plans. However, these events were used as a means of triangulating other data--to build a more complete description of a teacher's instructional plan, and to support inferential claims about the relationship between a teacher's beliefs and knowledge and their instructional plan (research question 1a). Similarly, any events in the video or audio recordings that may have potentially impacted instructional plans--and a teacher’s responses to those events--that were neither recorded in field notes nor discussed during interviews were also noted. Again, though these events were not discussed with participants, they were used to triangulate other data. Specifically, the type of event and type of response was noted to help understand how a teacher generally responds to events in the classroom and if their instructional plans were altered (research question 2a). The video and audio recordings were also used to ensure that the events and responses discussed in the interviews were being faithfully categorized (i.e., assigned a type during the first round of coding) whenever there was uncertainty based on the interview transcripts alone. Finally, the video and audio recordings were used when writing passage summaries or identifying themes from the summaries, when the passages from the interview were ambiguous or lacked detail.

Post-Observation Interviews

Data from the post-observation interviews were used to answer all research questions. The first part of the interview asks participants to identify unexpected events (research question 2a), real-time alterations to instructional plans or planned alterations for the future (research question 2), and their reasons for those alterations (research question 2b). Part 2a of the interview asks participants to discuss decisions that seemed to be illustrative of their initial instructional plans and their reasons for those decisions (research question 1). Part 2b comprises the stimulated recall portion of the interview, and asks participants to discuss events that resulted in alterations to their instructional plans (research question 2a), their responses to these events (research question 2), and the reasons that they responded in the way that they did (research question 2b). Part 3 of the interview is essentially identical to the final five questions of the pre-observation interview, addressing planning decisions for the next observed lesson, if any (research question 1).

The post-observation interviews were coded twice. The first coding was identical to the pre-observation interviews, coding for *knowledge*, *beliefs*, the teacher's initial *instructional plan*, *current contexts*, and *past contexts*. These codes were primarily applied to part 2a of the interviews in which the teachers discussed events that appeared to be illustrative of their initial instructional plans and part 3, in which teachers discussed their instructional plans for the upcoming lesson, but were also applied to other parts of the interview when participants discussed influences on their initial instructional plans. These interviews were then analyzed along with the pre-observation interviews to help answer research question one, using the process described in the section on pre-observation interviews.

The post-observation interviews were then coded a second time, focusing on portions of the interview in which participants discussed events that occurred during the observed lessons and the participants' responses to those events. This was generally parts 1 and 2b of the interviews, but also sometimes included part 2a if participants expressed that the instances discussed were deviations from their original instructional plans or part 3 if teachers discussed how their plans for the next observed lesson were influenced by the lesson that had just occurred. The codes *knowledge, belief, context,* and *hypothesized_learning_process* (and their sub-codes) were applied similarly to how they were applied for the pre-interviews, but only when participants expressed that these had an influence on their in-the-moment decision-making in the classroom. In addition, new codes were added to capture information relating to research question two. The code *event* was applied to the discussion of any event that had the potential to result in alterations to participants' instructional plans, and sub-codes were applied to categorize the type of event (e.g., being short on time, a student presenting an answer to the class). The code *response* was applied to the discussion of teacher's responses to these events to record how participants' instructional plans were altered. Sub-codes were used to categorize the type of alteration (e.g., altering a whole-class discussion, postponing an activity until a future lesson, altering one's hypothesized learning process). The initial list of sub-codes for events and responses were based on research regarding teacher decision-making (e.g., Cayton et al., 2017; Stockero & Van Zoest, 2013), with other sub-codes emerging as common types of events and responses were observed. Events were then paired with their corresponding responses. These pairs were additionally labeled with any beliefs or knowledge identified in the previous round of coding that impacted the teacher's response. The event-response pairs were then categorized by the type of response. Each of these response categories were analyzed to determine which types of events tended to result in each type of response. These groups were then further analyzed to determine if certain beliefs or knowledge tended to impact how teachers responded to these events.

Artifacts of Teaching

The artifacts of teaching (e.g., lesson plans, handouts) served several purposes. First, they were used to assist the researcher during the observation. Having a copy of the problems that students were working on, for example, allowed the researcher to better understand the conversations that students or the teacher were having, resulting in better identification and description of key events to record in the field notes. Second, they were used during the pre-observation interview when discussing teachers' instructional plans or during the stimulated recall portion of the post-observation interview when discussing events in the classroom, teachers' responses to those events, and teachers' alterations to instructional plans. For example, in the pre-observation or post-observation interview, a teacher may have been asked why he or she included a particular problem or set of problems in the task. During the post-observation interview, a student handout may have been used to assist in the discussion concerning a student's approach to solving a problem. A lesson plan may have been used in a discussion about why a part of the lesson that was planned did not get implemented during the observed lesson. Finally, like the classroom recordings, all classroom artifacts were reviewed during data analysis as a means to triangulate other data sources. Any instances in the classroom artifacts that were indicative of a teacher's instructional plan that had not been previously discussed were noted, along with the part of the teacher's instructional plan (learning goal, planned activity, or hypothesized process of learning) that the artifact seemed to illustrate. Though it was no longer feasible to ask teachers about these artifacts at this phase of the study, the artifacts were used to strengthen results from the interviews regarding the creation of teachers' instructional plans and the knowledge and beliefs that they drew upon to create those plans (research question 1a).

Possible Biases and Limitations

Though teachers' initial instructional plans were discussed during the pre-observation interviews (or during the previous lesson's post-observation interview), it was not feasible to discuss every aspect of a teacher's instructional plan. Thus, during the classroom observations, it was not always possible to determine whether a decision made by a teacher was an alteration to their initial plans or whether the teacher simply did not mention that part of the plan during the pre-observation interview. While deciding which events to record in the field notes, it is possible that some events that resulted in alterations to a teacher's instructional plan were not recorded due to my assumptions about what was or was not included in the teacher's initial instructional plan. For example, several teachers were observed asking students in the class if they understood a concept or solution that was just explained. The students often responded in chorus that they did, and the teacher moved on to another part of the lesson. This event was typically not recorded in the field notes since it did not seem like this was an alteration to the teacher's instructional plan. However, it is possible that the teacher was expecting a less positive response and had planned to spend more time explaining the concept.

Even when an event and response were recorded, for the sake of time, I often had to be selective about which events and responses to ask the teacher to discuss in the post-observation interview. This is another area where my personal biases may have come into play. Particularly for lessons that were lengthy or in which there were a high number of recorded events, several events had to be omitted from the discussion. In selecting the events to discuss, I had to decide, often in real-time, which events to ask the teacher about. This decision was based on several factors, such as which events and responses I felt were different enough from other events that were being discussed, which events were the most likely to be unexpected, the responses in which the teacher seemed to be making deliberate decisions, or responses in which the teacher seemed likely to be drawing on their beliefs or knowledge. All of these factors required that I make assumptions about the teacher's decision-making process, and thus, biases could have been introduced.

As with most studies involving classroom observations, the observed lessons run the risk of not being representative of a teacher's entire instruction. Indeed, one teacher noted that his instruction tended to look quite different at the beginning of a unit--with more teacher-centered lecture--than in the latter half of a unit, with more student-centered explorations. Though teachers were all observed for multiple (typically three) lessons, these lessons for each teacher all occurred in a single unit and in the same part of the school year. It is feasible that instruction on probability distributions, for example, may look different than instruction on hypothesis testing. Later in a school year, as teachers and students become more acquainted with each other, interactions in the classroom may look quite different. Teachers were asked to discuss their instructional planning in general (used to answer research question 1), but their discussions on real-time decision-making (used to answer research question 2) was generally limited to the observed lessons.

Participants were asked to retrospectively discuss their actions and decisions in the classroom. As discussed in the literature review, interviewees' retrospective retellings of their actions are not always reflective of their actual actions. It is possible that teachers in this study did not accurately recount events that happened in the classroom, their internal decision-making, or their observable responses to events. This may have been the result of teachers' own biases when retelling, of failing to accurately remember events, or of trying to present a curated or doctored version of events to the researcher.

Lastly, my presence in the classroom may have had an impact on the actions of both the teacher and the students. One teacher, for example, remarked that he likely put more effort into planning an observed lesson than he did for a typical lesson. Another teacher remarked after her first observed lesson that her students seemed more eager to share their thoughts and engage in whole-class discussions than they typically were.

Validity

With these limitations in mind, several steps were taken in the design of the study to help ensure the validity of the results. First, all participants were observed over multiple class periods. Though this still only amounted to a small portion of a teacher's entire instruction for a course, observing multiple class periods increases the chances that a teacher's "typical" instruction would be observed. It also allows for both the teacher and the students in the class to become accustomed to my presence, reducing any impact that my presence may have had. For example, the teacher above who remarked that her students seemed to be acting differently during the first observed lesson later remarked after the third observed lesson that the students were back to acting like their typical selves and that they seemed to be ignoring my presence. Teachers were also reassured that the quality of their instruction was not being assessed, and that nothing observed would be reported to any colleagues or administration at their school. Likewise, students were reassured before observations began that the purpose of the research was not to assess their actions, but rather to observe the teacher's actions. These reassurances hopefully resulted in the observed lessons being more representative of typical lessons for that course, and in teachers being more willing to openly and truthfully share their decision-making processes.

In addition to the above reassurances, other steps were taken to increase congruence between teachers' retelling of events or decision-making and the actual events or decision-making that occurred. All post-observation interviews occurred the same day as the observations, and in most cases, immediately afterward. For five of the seven teachers, post-observation interviews employed the use of video-stimulated recall. The use of video-stimulated recall and a short amount of time between the observation and the teachers' retelling have both been shown to reduce discrepancies between the retelling and the actual occurrences (Lyle, 2003).

The data that was primarily used to answer the research questions were recorded responses to interview questions. In other words, the data were the participants' own words. Although my subjectivity may still have affected coding and identification of themes, starting with participants' own words should increase the validity of the results. Also, although the interviews served as the primary data source, recordings of classroom observations and collected artifacts of teaching served as triangulation sources, to help confirm findings from the interviews. This helped ensure that information obtained during the interviews, such as decisions the teacher made or events that happened, were reflected in the instructional practice that occurred.

Steps were also taken to reduce the bias inherent in the researcher choosing which events to record and to discuss during the stimulated recall portion of the post-observation interview. Prior to the stimulated recall portion of the interview, the teachers are asked to identify any instances of unexpected events, any part of the lessons where they deviated from their original plans, or any revisions that they plan to make to the lesson for the next time it was taught. This allows the teachers to first discuss events that they themselves feel are most relevant or influential, perhaps even noting events that the researcher failed to record. Also, the researcher was liberal in choosing which events to record. That is, even events that seemed to have a relatively low probability of resulting in alterations to instructional plans were recorded. These events could then be mentioned during the stimulated recall portion of the post-observation interview, allowing the teacher, not the researcher, to decide if these events did actually result in alterations to the teacher’s instructional plans. However, the discussion of some events still had to be omitted for the sake of time.

Finally, codes, summaries of coded passages, and the themes that emerged were shared with an experienced researcher in statistics education. Discussions were held concerning the derivation of the themes from the summaries of coded passages. When it was unclear how the themes were derived from the summaries or when the connection seemed tenuous, the themes were revised to more accurately reflect the accounts of the participants. When needed, the data were revisited to identify additional passages or recorded classroom instruction that could provide additional corroborating or disconfirming evidence for the themes.

CHAPTER 4: INFLUENCE OF BELIEFS AND KNOWLEDGE ON INSTRUCTIONAL PLANNING OF SECONDARY STATISTICS TEACHERS

Journal

This chapter presents a journal-ready manuscript to be submitted to the *Journal of Mathematics Teacher Education*. The *Journal of Mathematics Teacher Education* (JMTE) publishes articles focused on the improvement of the education of mathematics teachers and on the development of teaching methods that promote students' learning of mathematics. The journal covers all stages of the professional development of mathematics teachers and mathematics teacher educators, as well as factors that influence the learning of mathematics teachers and their students. Research papers published in JMTE should have an interest beyond the local or national level, and should be no longer than 10,000 words, not including references. The following manuscript adheres to the purpose of the journal, as its findings can help mathematics teacher educators provide instruction and professional development that will have an impact on the practice of teachers of statistics.

Introduction

Recent technological advances and world events make it imperative that every student should develop an understanding of basic statistics, just as they should have an understanding of algebra and geometry (Wild et al., 2018). Policymakers around the world have recognized this, as statistics now has a prominent place in many secondary curricular standards including Australia (Australian Curriculum, Assessment & Reporting Authority, 2011), New Zealand (New Zealand Ministry of Education, 2015), and Germany (Kultusministerkonferenz, 2012). Over half of high schools in the United States now offer a standalone course in probability or statistics, a number which has more than doubled since 2000 (Banilower et al., 2018; Weiss et al., 2001). Many of these courses are Advanced Placement (AP) Statistics (College Board, 2020), a course with similar content to a college-level introductory statistics course. Nearly 220,000 students took the AP Statistics Exam in 2019, twice as many as in 2008 (College Board, 2008, 2019). Many high schools offer non-AP statistics courses, as well, often allowing students to obtain college credit from a local community college or university. With the rapid rise in secondary schools offering courses in statistics, greater numbers of teachers are being tasked with planning and implementing instruction for these courses, often with no other teachers in the school teaching or having taught the course before. Inservice and preservice secondary mathematics teachers often feel unprepared and lack confidence to teach statistics (Banilower et al., 2018; Lovett & Lee, 2018), adding to the difficulty of this task.

When planning for and implementing statistics courses, there are a number of decisions teachers have to make about what to teach and how to teach it, many of which are unique to a statistics course. These decisions include which tasks to give students, which datasets students will work with, which data analysis or visualization tools the teacher and students will use, and how to assess students' learning, among others (Wild et al., 2018). These decisions ultimately have a significant impact on students' learning (Eichler, 2011; Hiebert & Grouws, 2007). When making these decisions, teachers draw upon knowledge and beliefs, but are also influenced by the context in which they teach (Hoy et al., 2006). In order to prepare secondary teachers to teach statistics, a better understanding is needed of these different influences. To that end, this study attempts to answer the following research questions:

1. What knowledge and beliefs do secondary teachers draw on when planning statistics instruction?

2. What contextual influences impact whether secondary teachers' beliefs are put into practice or whether they are able to effectively utilize their knowledge?

Literature Review

Teaching Statistics in Secondary Schools

There are various issues secondary teachers face when teaching a course in statistics. Content that was once reserved for post-secondary statistics courses is now being taught to students pre-college (Zieffler et al., 2018). Many secondary teachers did not themselves take a statistics course until university, if at all (Banilower et al., 2018). Statistics teachers are often less likely to receive regular support for teaching statistics and have fewer collegial conversations about statistics teaching, compared to other mathematics subjects (Whitaker, 2016). The amount of statistical technology is expanding, not only for doing statistics, but also for teaching and learning statistics. Because of these issues that statistics teachers face, the various instructional choices they have to make, and the varied contexts in which these teachers teach, there is a large variation in the content and setting of statistics courses being offered to secondary students in the United States (Zieffler et al., 2018).

Many teachers rely on what they know about teaching mathematics, either from experience or from professional development, to inform their teaching of statistics (Zieffler et al., 2018). This has its advantages and disadvantages. Many recommended pedagogical strategies for teaching mathematics are also effective for teaching statistics. Student-centered instruction, collaborative learning, encouraging and facilitating student discourse, building on students' prior knowledge, and scaffolding students' learning are all strategies that can be effective for teaching statistics (Ben-Zvi et al., 2018; Watson et al., 2018). On the other hand, taking a mathematical approach to teaching statistics can be detrimental. There are key differences between statistics and mathematics that necessitate different approaches to teaching the topic, including the omnipresence of variability, the non-deterministic nature of topics like sampling and inference, and the importance of data and context (Franklin et al., 2007, Rossman et al., 2006).

While research in statistics education has become much more prominent in the last two decades, most of this research has focused on two main areas: 1) curricular issues, including assessment, standards and guidelines, and technology, and 2) students' (and sometimes teachers') knowledge of statistics (Petocz et al., 2018; Watson, 2016). Students' and teachers' affect and beliefs toward statistics have also received some attention (e.g., Lee et al., 2017; Zieffler et al., 2012). What has received considerably less attention is what is actually happening on a day-to-day basis in classrooms and why, particularly at the secondary level (Eicher, 2011; Pearl et al., 2012).

Influences on Instructional Practice

Teachers' knowledge and beliefs have a significant impact on instruction, but this impact is often mediated by the context in which they teach. Schoenfeld (1998) proposed a model of *teaching-in-context* (Figure 4), in which teachers' instructional decision-making depends on a certain subset of their knowledge and beliefs that are activated by the context in which they are teaching. A teacher's past experiences (e.g. professional development experiences, past teaching experiences, existing relationships with students) also impact which goals, beliefs, and knowledge take priority at any given moment.

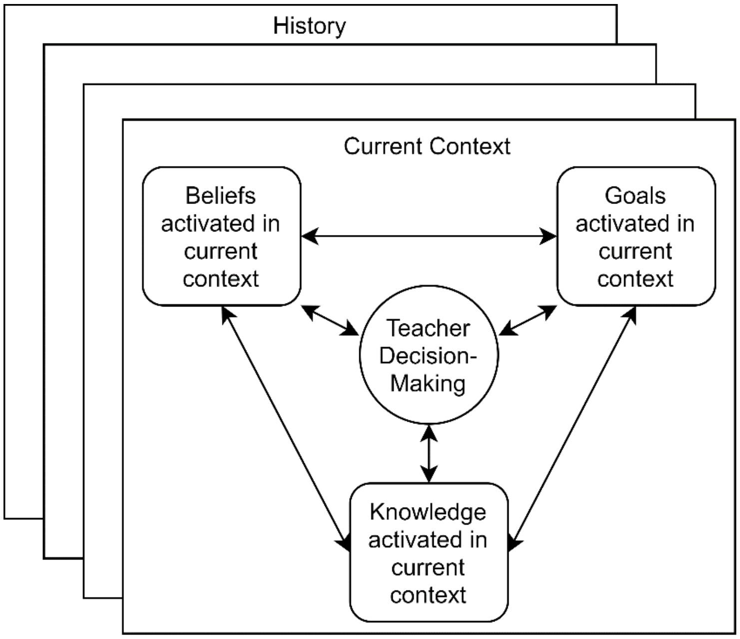


Figure 4. Teaching-in-context. Adapted from Schoenfeld (1998).

In order to effectively teach mathematics, knowledge is needed beyond the mathematical content itself. Mathematical knowledge for teaching (MKT) is required, which includes, for example, knowledge of why mathematical procedures work, common errors students make, and representations that can help students understand a topic (Hill et al., 2008). Charlambous et al. (2012) found that teachers with higher levels of MKT were more likely to show more skillful use of representations, use more appropriate explanations during instruction, were more able to capitalize on student contributions while directing the class toward the learning goal, and were better able to compensate for curriculum materials that were lacking. Groth (2007) extends this idea by describing a *statistical knowledge for teaching*, which includes knowledge specific to teaching statistics, such as being able to appraise the fruitfulness of a student-posed statistical question.

Although secondary teachers tend to have more sophisticated knowledge of statistics than elementary teachers (Watson, 2011), numerous studies have suggested that teachers' knowledge of statistics and how to teach it is less than ideal, though most have focused on preservice teachers (e.g., Biehler et al., 2018; Burgess, 2002; Hannigan et al., 2013; Lee et al., 2014). Even when considering inservice AP Statistics teachers, however, Peters (2009) found that only five of sixteen teachers showed a robust understanding of variation, a key concept in statistics. What is less known is how this lack of knowledge impacts teachers' statistics instruction.

Teachers' beliefs have also been shown to influence instructional practice. For example, Stipek et al. (2001) found that teachers who espoused more traditional beliefs about mathematics (e.g., that mathematics was a set of operations and procedures that needed to be learned) were more likely to emphasize performance and speed when assessing students. Other beliefs that have been found to influence teachers' instructional practice include beliefs about the social makeup of the classroom, beliefs about students' affect and needs, and beliefs about the role of the teacher (Beswick, 2007; Sztajn, 2003; Thompson, 1984). However, the impact that teachers' beliefs have on their instructional practice can be limited by the context in which they teach. For example, Chen (2008) found that teachers' beliefs about the use of technology in teaching were often not reflected in their actual instruction. Reasons that teachers' beliefs were not put into practice included a lack of resources or access to resources, insufficient planning time, inadequate administrative or technical support, and discomfort with technology. Other contextual factors that have been found to inhibit teachers' beliefs from being put into practice include a lack of instructional time, pressures of student assessments, social norms of a school, and student behavior (Johnson, 2006; Skott, 2009; van der Sandt, 2007).

Few studies specifically examine how teachers' beliefs impact their teaching of statistics, however. In addition to the above beliefs, there is large variation in the beliefs about statistics and beliefs about teaching statistics that teachers hold (Chick & Pierce, 2008; Lee et al., 2017; Umugiraneza et al., 2016). This study examines how these different beliefs and knowledge may be reflected in the teaching of statistics and which contextual factors may inhibit teachers' knowledge and beliefs from being put into practice.

Methods

Participants

To answer the research questions, a collective instrumental case study approach (Stake, 2005; Yin, 2009) was used. The cases in this study consisted of seven secondary mathematics teachers who were currently teaching at least one section of a statistics course. The teachers taught in a variety of contexts, including public, private, and charter schools across five different counties in a state in the southeastern United States. The students in their courses were likely ages 15-18. Information about participants' statistics courses and schools can be found in Table 4 (all names are pseudonyms).

Table 4. Participants' teaching experience and current statistics teaching assignment.

Teacher	Years of teaching experience[a]	Number of prior times teaching statistics[a]	Statistics course taught during study	School
Ms. Andrews	8	~8	Statistical Methods I (a credit-bearing course at a local college)	Selective public early college high school
Ms. Baker	14	~12	AP Statistics	K-12 charter school
Ms. Carey	6	~9	AP Statistics	Private religious secondary school
Mr. Dennis	16	~32	Advanced Analytics and Statistics	Private college-prep secondary school
Mr. Enloe	9	~18	AP Statistics	Private college-prep high school

Table 4 (continued).

Mr. Fahey	7	4	AP Statistics	Selective public residential high school
Ms. Greene	4	1	AP Statistics	Traditional public high school
[a]: does not include current year/semester. Some teachers teach multiple sections at a time, each of which is included in the number of times teaching statistics.				

Data Sources

As part of a larger study, each participant was observed for three consecutive lessons (one teacher was observed for two lessons). Prior to the first observation, each teacher participated in an interview designed to identify factors that influenced the teacher's statistics instruction and to identify their learning goals and planned activities for the upcoming lessons. After each observed lesson, an additional interview was conducted with the participant, which included questions that asked the participant to discuss reasons behind observed practices that were not discussed in the prior interview and to discuss the planning of the next observed lesson, if any. Part of the post-observation interviews consisted of a stimulated recall (Calderhead, 1981), in which teachers were shown video recordings or were read a description of an instructional practice, and asked to recall reasons behind decisions that were made. These seven pre-observation interviews and 20 post-observation interviews served as the corpus of data for this study.

Analysis of Data

The 27 interviews were transcribed verbatim, and passages in the interviews were assigned top-level codes derived from the literature: knowledge, beliefs, context, previous experiences, planned activities, and intended learning goal. Each passage was also assigned one or more sub-codes using open coding and the constant comparative method (Glasser, 1965; Kolb, 2012) to identify different categories that emerged. For example, within the top-level code of beliefs, some of the sub-codes that emerged included beliefs about statistics, beliefs about learning, and beliefs about technology.

To identify beliefs and knowledge that teachers drew upon when planning statistics instruction, as well as the source of these beliefs and knowledge, passages were identified that included both of the following: one sub-code from *Knowledge, Beliefs*, or *Previous experiences*; and one sub-code from *Planned activities* or *Intended learning goal.* For example, Table 5 shows portions of selected passages that were coded with both *Beliefs about statistics* and *Planned student activities*. In total, 770 passages were identified that contained a relevant pair of codes (passages containing multiple pairs of relevant codes were counted once for each pair). For each of these passages, if the participant discussed the impact that knowledge or belief had on instruction, a short summary was written that captured the essence of this relationship, as well as the source of that knowledge or belief, if stated. This resulted in 393 summaries. Each of these summaries were then analyzed along with summaries from other passages containing the same pair of codes to identify common themes, first for each individual teacher and then across all teachers.

Table 5. Selected passages coded for Beliefs about statistics and Planned student activities.

Participant	Quotation
Ms. Greene	I wanted students to not be given super specific instructions about how to collect data, because that’s like real life. People that collect real data have to decide what is the best way to do this. And I wanted them to try things, and say “Oh, that doesn't work very well. Let's try this differently.” So, I tried not to give a lot of instructions.

Table 5 (continued).

Ms. Greene	I think algebraically working through problems is very different than manipulating data, and interpreting data, and having conversations about [data]. There's not but so much stats you can do, there's a lot of stats that you have to talk about. And so [I have] my students work in pairs to talk about different things.
Mr. Enloe	In algebra classes, they just kind of they get an answer, they circle it and they move on. And the thing about statistics is that we're taking those numbers and we're interpreting what it means in the context of the problem. And so early on in the year, I give them these template sentences.
Mr. Enloe	I'm gonna give them a couple of examples of this computer output that I showed them today, and make them, just real quick, interpret the slope, interpret the intercept, interpret s, interpret r^2, write the equation of the least squares regression line....If you were doing statistics out in the real world, this is how it's gonna be anyway. You're not gonna actually sit and calculate stuff. You're gonna put it into a computer program and have to read the output. So, I think that's a huge skill and I want to make sure they can do that.

Then, to identify contextual factors that may have impacted whether teachers' knowledge and beliefs were put into practice, passages were identified that included both of the following: one sub-code from *Context*, and one sub-code from *Planned activities* or *Intended learning goal*. In total, 332 passages were identified that contained one such pair. If the passage was also coded with a sub-code from *Knowledge*, *Beliefs*, or *Previous experiences*, this was noted along with the passage to assist with identification of factors whose influences were mediated by context. Each of these passages were summarized to indicate how the context in which the participant taught impacted his or her instruction, along with any ways that the context mediated the impact that the participant's knowledge or belief had on instruction, if these relationships were discussed. This resulted in 206 summaries. As before, the summaries were then analyzed along with other passages containing the same pair of codes to identify common themes for each teacher and across all teachers.

Codes, summaries of coded passages, and themes that emerged were shared with an experienced researcher in statistics education. Discussions were held concerning the derivation of the themes from the summaries of coded passages. When it was unclear how the themes were derived from the summaries or when the connection seemed tenuous, the themes were revised to more accurately reflect the accounts of the participants. When needed, the data were revisited to identify additional passages to provide additional corroborating or disconfirming evidence for the themes.

Results

Five different areas of beliefs were identified that influenced the statistics instruction of participants: learning, teaching, students and their needs, statistics, and technology. Knowledge of teaching statistics was also found to impact participants' statistics instruction, and originated from a variety of past experiences. The context and environment that the teachers taught in often mediated this relationship between teacher's beliefs or knowledge and their instruction, resulting in instructional decisions that were sometimes at odds with participants' beliefs and knowledge. Such contexts included limited planning time, limited instructional time, the presence of external assessments, large class sizes, and others. What follows elaborates on these findings.

Influential Beliefs about Learning

For all participants, beliefs about how students learn impacted how participants structured their classroom, norms they established for interactions, as well as larger curricular structures and goals of a lesson. Beliefs about learning that impacted instruction included beliefs about the role of discourse in learning, about needed supports for learning, and about the types of activities that best supported students' learning.

All seven teachers **believed that group work and student-to-student discourse was beneficial for students' learning**, and this impacted the physical arrangement of their classroom. However, these beliefs sometimes differed in subtle ways that were reflected in their instruction. Five teachers had students seated in groups of four. For three of these teachers, this arrangement was driven by the belief that students benefited from hearing multiple voices and perspectives as well as from explaining concepts to others, and thus students in these classes were encouraged or expected to work collaboratively with other members of their group on most assignments outside of formal assessments. Two of these three teachers also randomly regrouped students to increase the number of perspectives these students hear, increase the sense of community in the classroom, and let students discuss novel approaches and ideas from the previous lesson that they may have otherwise not gotten the chance to see. A fourth teacher, Mr. Fahey also had his students arranged in groups of four. Like the other three teachers, he saw value in group discussions in supporting learning, so he encouraged students to work with other students in the group during activities. However, he believed it was difficult for students to demonstrate their understanding of a topic in a group, so for any graded assignments, he restricted collaborations to pairs. Ms. Andrews also had her students arranged in groups of four, but believed that a pair arrangement would result in students remaining more focused and contributing more equally to the assignments. She felt, however, that the large number of students in the class and a lack of physical space prohibited this type of arrangement.

The remaining two teachers, Mr. Enloe and Ms. Greene, had students arranged in pairs. Both of these teachers believed that pairs allowed students to bounce ideas off each other and help each other when one was stuck. Ms. Greene believed that larger groups would work better for statistics, but because she taught three classes that were not statistics and believed that those classes worked better with students in pairs, she settled for having students in pairs in the statistics class so as to avoid having to rearrange desks for one period of the day.

For some teachers, **beliefs about the types of discourse that were important for student learning** impacted how the teachers interacted with these groups or pairs. When students were working in groups or pairs, four of the teachers expressed that they consciously tried to limit the amount of student-teacher discourse in favor of student-student discourse. Some of them did so by simply avoiding interjecting into students' conversations, while others would respond to students' questions by prompting them to discuss the question with other students in the group first. Their reasons for doing so were similar to the beliefs described by Mr. Dennis:

> It's through that conflict that they have--of reading the question, not understanding it, and asking their neighbor--that they start to really solidify their understanding, much more so than me just talking and doing it. And so I found that they can internalize the concepts at a much deeper rate, at a much faster rate, if they go through that productive struggle in their groups without me interjecting the answer.

On the other hand, Ms. Carey believed that her students could better overcome obstacles for learning with her assistance. She tried to spend time with every group while they were working so that she could assist with learning that was occurring. Since this class had more students than usual, this also meant that she felt the need to schedule larger chunks of time for group work than she had in previous years, to ensure that she could interact with as many groups as possible. She would also often schedule time during class for students to work on their homework so that they had the opportunity to individually ask her questions if they got stuck.

Similar types of beliefs also impacted whole-class discussions. Most teachers **valued discussions that centered around student-generated solutions and explanations**. For several teachers, this meant that students would be in front of the class, presenting and explaining a solution to a problem and answering questions from other students. Like small-group discussions, these student-led whole-group discussions were driven by the belief that students can learn by explaining and that it is often beneficial for students to hear a concept explained in different ways. Most teachers also valued presentation of multiple solution strategies, believing that it validated students' work or perhaps helped them see an approach they understood better. Some expressed that they welcomed incorrect or incomplete solutions, since they believed the process of refining an answer was more beneficial to students' learning than simply seeing the correct answer.

Teachers' **beliefs about activities that support students' learning** impacted the general structure of a unit or lesson. Two teachers talked about the importance of starting each day with a short review of the previous day's material. Ms. Baker felt that doing so helped students retain information and draw connections between old material and new, which she felt was an important part of student learning. Ms. Carey shared similar beliefs, but often felt she had to skip this review since class periods were only 45 minutes long and she wanted to ensure there was enough time to get through the day's activities. Three other teachers began most classes with a homework review, which they believed would help students solidify knowledge in order to build upon it and construct new knowledge when they were introduced to new material.

Several of the teachers also talked about the importance of ending each lesson with a closing or summative discussion. These teachers believed that these discussions would help students make connections and would solidify learning. For Mr. Dennis, one of the biggest constraints on his instruction was limited class time, and he felt that the summative discussion at the end of a class period--and thus his students' learning--was often negatively impacted because of limited class time. In addition to closing discussions each day, which Ms. Baker described as “really important”, she also scheduled “buffer days” between each content chapter, in which the class looked back over not only the chapter just completed, but over several weeks’ worth of content. She believed that these buffer days helped students see how the content they are learning was connected, a belief that she says was inspired by the curriculum and the professional development she has received based on it.

All seven teachers in the study **valued “hands-on” experiences for their students**. Most participants believed that “doing” statistics contributed more to students’ learning than watching or listening to the teacher. They believed that when learning was situated in active experiences, it was more likely to lead to students retaining information and making connections in later units between old material and new. The specific purpose and timing of these experiences varied between teachers, however. Some believed that these active experiences were more effective at creating a solid foundation of knowledge that could be later built upon more formally. Ms. Baker, for example, eschewed teacher-centered lectures in favor of hands-on experiences to introduce new concepts to students. Mr. Enloe, on the other hand, felt that his students needed to have an understanding of basic concepts and vocabulary before “diving deep” into student-centered investigations, so hands-on experiences typically came in the latter half of a unit. Several of the teachers utilized a mixture of these activities, using hands-on experiences to initially explore concepts that were then solidified with a discussion or lecture, and then giving students problems to practice to ensure they understood the concepts. However, class time often impacted these teachers’ decisions about the amount of hands-on activities their students would do:

> There are a couple times where time is an issue. And while I would love to go deep into some activity, I have to get it done, you know? And so, sometimes a lecture is the quickest way to do that. (Mr. Fahey)

Even when Mr. Fahey did plan hands-on activities for his students, he would sometimes feel the need to modify them so that they took less time. For example, in one particular activity, he would have preferred that students were able to collect enough data to use in an analysis. But because of limited class time, he instead had students only collect a small amount of data to understand the data collection process, and then used simulated data in the analysis. Similarly, while Ms. Greene admired other teachers for having students collect their own data, she felt it took a lot of class time, so she instead used existing data sets that she felt would have a personal connection to her students. In addition to limited class time overall, some teachers felt that the class schedule impacted their use of hands-on activities. Mr. Fahey's course had recently changed from 90-minute class periods to more frequent 50-minute class periods. He found that the shorter class periods meant he either had to cut activities short or split them up over two days, neither of which he felt was optimal. Ms. Carey (45-minute class periods) and Ms. Baker (53-minute class periods) also mentioned that short class periods made it harder for students to do things like gather data.

Planning time was also a factor for some teachers including Ms. Greene, who taught three other classes besides statistics. She had been given a set of instructional presentation slides by teaching colleagues at another district. Even though she believed that hands-on experiences resulted in deeper understanding and better retention, she often resorted to using these slides in a more teacher-centered type of instruction due to lack of time outside of class to focus on planning for statistics. Having a large number of students in a class was another factor for many teachers that made it more difficult to have their students participate in hands-on activities, despite participants seeing the value in those activities. For some, this was because it was harder to interact with all students to ensure they were learning and were staying on track. For Ms. Baker, having classrooms with limited physical space meant that with larger class sizes, it was difficult to have students moving around, gathering data, working on whiteboards, and for her to circulate around.

Figure 5 shows a summary of the most influential beliefs participants held about learning, along with ways in which these beliefs impacted their instruction, or would have, according to participants, if mediating factors did not exist. Also shown are the mediating factors, if any, that inhibited these beliefs from being put into practice.

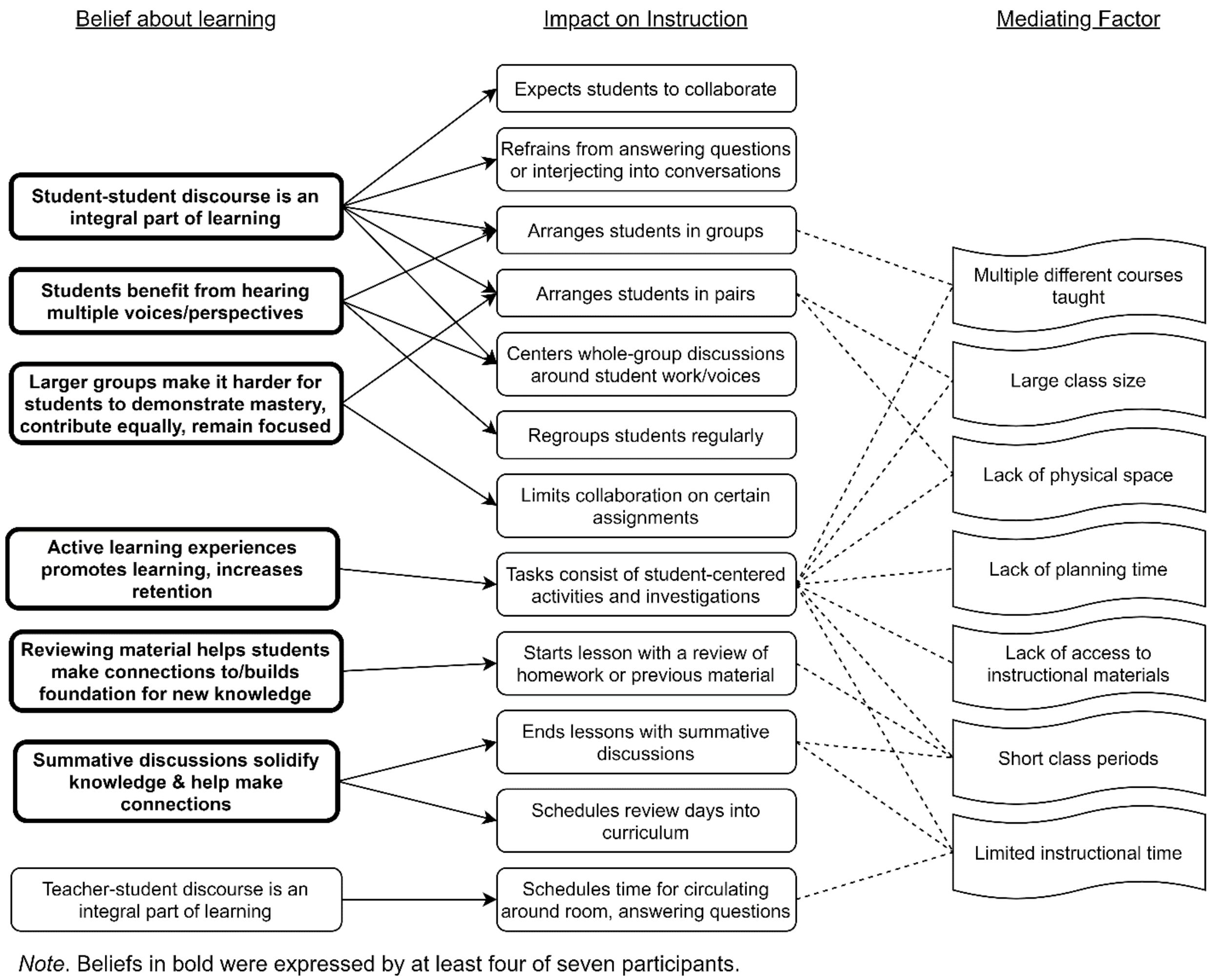


Figure 5. Influence of beliefs about learning on participants' instruction and mediating factors.

Influential Beliefs About Teaching

Several teachers indicated that beliefs about what constituted good teaching impacted their instruction. These beliefs generally concerned the role and responsibilities of the teacher and how to effectively assess students' learning.

Several participants discussed **the role and responsibilities of a good teacher**. Several participants believed that a teacher should not be the sole or primary authority of what is right and wrong. These teachers tended to encourage student-to-student conversations. For example, Mr. Enloe often refused to tell students whether their answers were correct, and instead prompted students to discuss the answer with other students around them. Similarly, several teachers held beliefs about who the primary source of knowledge in the classroom should be. Rather than explaining concepts to students, some of these teachers tried to set students up in situations where the students could construct their own knowledge. On the other hand, some teachers like Ms. Carey believed part of her role was to be a source of knowledge when students got stuck, helping them overcome any obstacles to their learning that might arise. Because of this belief (along with her aforementioned belief about the value of student-to-teacher discourse for learning), when students were working on an activity, Ms. Carey tried to interact with as many groups as possible.

All participants expressed **beliefs about assessing students' knowledge** that impacted their instruction. Three teachers--Ms. Baker, Ms. Carey, and Ms. Greene--expressed that watching, listening, and conversing with students while they work is often the best way to assess what they know. These teachers tended to treat in-class activities as learning experiences rather than formal assessments, typically not collecting or assigning grades to daily classwork. Another teacher, Ms. Andrews, felt that collecting classwork was a good way to assess what students knew, but whether or not she assigned a grade sometimes depended on the characteristics of students in the class, and whether assigning a grade would motivate them or whether it would cause them stress.

The other three teachers believed that more formal mid-unit assessments were needed to ensure everyone was where they needed to be or to determine which material might need more focus in the remainder of a unit. For example, Mr. Fahey noted after a class observation:

> Their assessment is planned for a week from now, and I need to have some mid-unit check for them to see if they're doing what they need to do….If the results on the quiz

tomorrow are disastrous, I will take time on the following class period and definitely reteach and focus on the topic.

These teachers believed that these mid-unit assessments also benefited students, letting them know if they needed to seek additional help, making it clear what information they needed to know in order to effectively engage with the material in the latter part of the unit, and helping them prepare for any end-of-unit assessment.

Figure 6 shows a summary of the most influential beliefs participants held about teaching, along with ways in which these beliefs impacted their instruction, or would have, according to participants, if mediating factors did not exist. Also shown are the mediating factors, if any, that inhibited these beliefs from being put into practice.

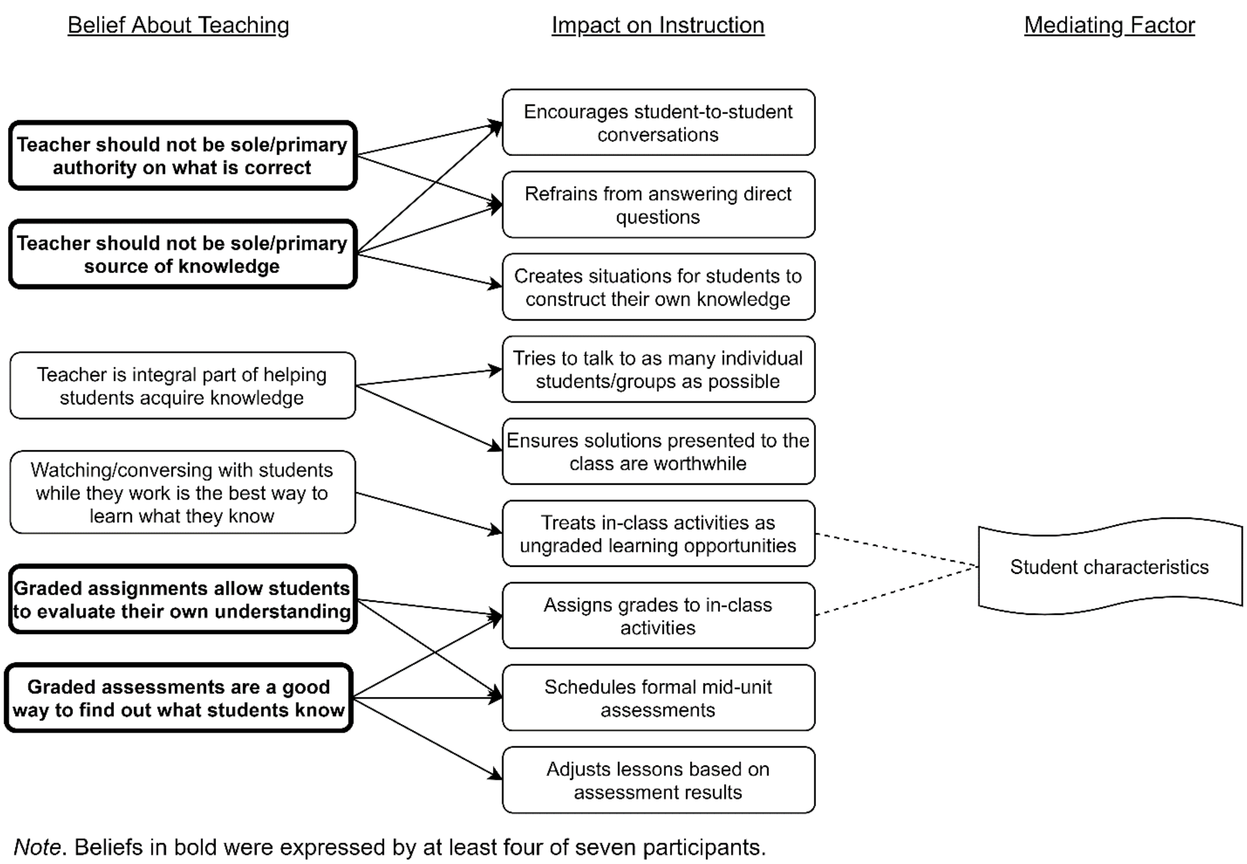


Figure 6. Influence of belief about teaching on participants' instruction and mediating factors.

Influential Beliefs about Students and Their Needs

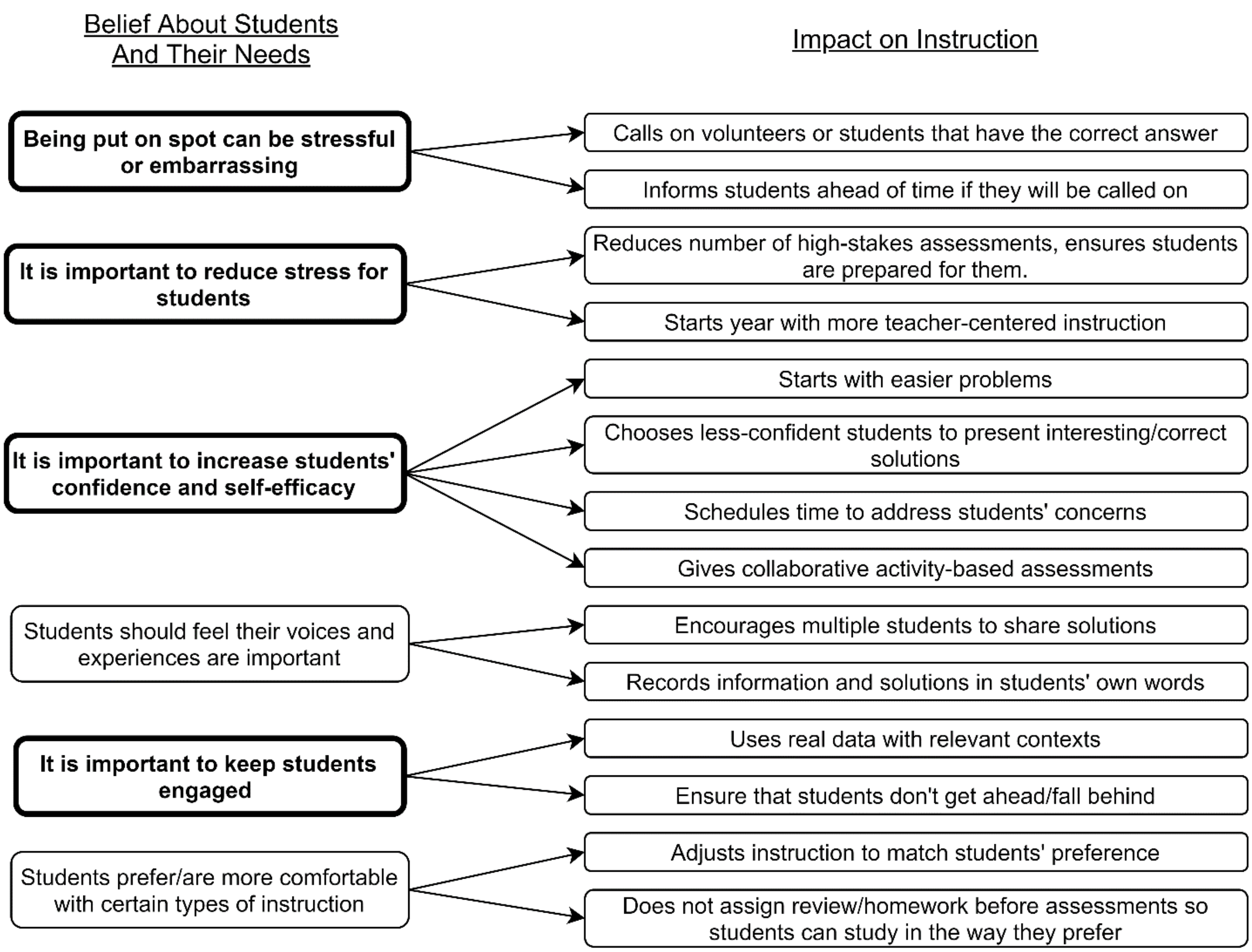
In addition to beliefs about student learning and teaching, all participants also expressed that beliefs about student affect impacted their instruction. These included beliefs about students' emotional needs, their level of engagement, and their preferences for instruction.

Five participants stated that they planned aspects of their instruction to meet the **emotional needs of students**. Some teachers believed that being put on the spot to answer a question or give an explanation in front of their peers could be potentially embarrassing or stressful for students. These teachers would try to curtail these emotions by only calling on students if they felt the student knew the answer, or by letting the student know ahead of time if they expected them to share something with the class, or by calling on volunteers for more difficult questions. Some teachers attempted to reduce stress for their students by reducing the number of graded assignments or high-stakes tests, by ensuring students felt prepared for high-stakes tests, or by slowly transitioning from teacher-centered instruction to student-centered activities as the year progressed. Similarly, several teachers felt it was important to take steps to increase students' self-efficacy and decrease anxiety. These steps included moving more slowly through material early in the year, starting a lesson with easier problems, choosing typically less-confident students to explain their answers to the rest of the class when they had correct or novel solutions, reserving class time for students to get individual assistance, drawing attention to the fact that other students had similar struggles, and giving collaborative activity-based assessments. Several teachers felt it was important for students to feel that their voices, experiences, and strategies were important and valuable, and did so in various ways. For Ms. Baker, this meant recording information and answers on the whiteboard in students' own words, even if the language used was less formal than she herself would have used. For Ms. Greene, this meant encouraging students to share alternate strategies for solving a problem, even after a correct solution was already presented and discussed. For Mr. Enloe, this meant avoiding calling on the same student repeatedly and encouraging others to speak up, even if one student was eager to share.

Most teachers expressed **beliefs about students' levels of engagement in the class** and took steps to try to increase this engagement. Four teachers expressed beliefs that students would be more engaged if they were presented with real data about contexts that were relevant to them, and would be more likely to wonder about or come up with predictions about the conclusions that would be found. Some teachers expressed concerns that students would "check out" or be a distraction to other students if either they felt they didn't understand what was going on or they felt like they already understood something that was being discussed. These teachers tried to find ways to keep these students engaged, for example, by having them explain problems to other students or share their solutions with the class.

Several participants expressed **beliefs about the type of instruction that their students preferred**. Sometimes they adjusted their instruction to meet these preferences. Mr. Enloe, for example, believed that students have their own preferred methods of studying for a test. So, in the days leading up to a test, he would not assign any mandatory homework or review problems, so that students could devote their time to preparing for the test in their preferred way. At times, students' perceived preferences did not match the teacher's own preferences or beliefs about how to teach. For example, Ms. Carey believed that her students were more comfortable with the lecture-style, teacher-centered instruction that they were used to, and that they would feel abandoned if she "let them go", replacing lectures with group activities to let them discover concepts on their own. But she also believed that student-centered instruction was more beneficial for students. To reconcile these beliefs, she decided to start out the year with more lectures, and "wean them off" lectures as the year went on.

Figure 7 shows a summary of the most influential beliefs participants held about students and their needs, along with ways in which these beliefs impacted their instruction. There were no contextual factors mentioned by participants that inhibited any of their beliefs about students and their needs being put into practice.



Note. Beliefs in bold were expressed by at least four of seven participants.

Figure 7. Influence of beliefs about students and their needs on participants' instruction.

Influential Beliefs about Statistics

Three teachers--Mr. Enloe, Ms. Greene, and Ms. Baker--regularly expressed that beliefs about statistics impacted how they taught statistics. They felt that **statistics was different from**

mathematics, and that this meant that it needed to be taught and learned in a different way. The other four teachers did not express beliefs about statistics in the interviews (this was not explicitly asked); thus, it is unclear if they brought their beliefs about statistics explicitly into their instructional decision making.

For Mr. Enloe, statistics was not about performing calculations, but about interpreting calculations and data representations, and then communicating results. He also felt that writing in statistics was different from writing in mathematics. He spent a lot of class time focused on helping students write good answers. Class discussions would often involve taking a student's response and then refining it to be clearer and more concise. Additionally, the class would often assess written responses using the AP Statistics Exam scoring rubric. He felt that this helped students see what was expected of them when talking about statistics, and in particular, on the AP Exam. The beginning of every unit in Mr. Enloe's class included instruction on relevant vocabulary, which was followed by a vocabulary quiz. According to Mr. Enloe, this ensured that he and the students had a common language with which to communicate before they dove deeper into the material.

Ms. Greene's beliefs about vocabulary and communication in statistics also impacted her instruction. She felt that statistics involved "a lot of words", more so than other mathematics courses. She felt that knowing the vocabulary was important when learning statistics, but she did not want them focused on memorizing terms and definitions. She gave students a list of definitions at the beginning of a unit, but felt that their time would be better spent on "playing with data", and that their understanding of the vocabulary would emerge through those activities and the following discussions. Like Mr. Enloe, Ms. Greene felt that statistics was not about following procedures, but about communicating:

> I think just, like, algebraically working through problems is very different than manipulating data….There's not but so much stats you can *do*, there's a lot of stats that you have to talk about….I have stressed to my students that stats is not about following procedures. Procedures will only get you so far. Whereas a lot of, I think, algebra and calculus is, like, building procedures that you can then use to solve different problems. But statistics is a little bit different from that.

Because of these beliefs, Ms. Greene used direct instruction less often than she did in other mathematics courses. However, she expressed that she still used direct instruction more often than she preferred. One reason for this is that she feels there is a lot of material that needs to be covered before the AP Exam, and that direct instruction is often a quicker way to get through material. Ms. Greene has tried to replace some of the direct instruction with student-centered activities, but she also teaches four different courses and feels that the time available for her to plan and revise the statistics course is limited, so she often ends up repeating the direct instruction she did the semester before.

Like Mr. Enloe and Ms. Greene, Ms. Baker believed that statistics is not about performing calculations. She had pushed for the removal of mathematical prerequisites for the statistics course, because she feels that students can be successful in the course even if they have not previously experienced much success in mathematics:

> One of the things I love about statistics is I get so many students in that class who've struggled with kind of the algebra sequence and never had a—really, a chance to see themselves as capable mathematicians….They're able to kind of have a new experience and see themselves as a quantitative thinker in a way they maybe weren't before.

These beliefs were also reflected in Ms. Greene's instruction, as she actively tried to create situations where "thinking like a statistician" was prioritized over mathematical skills. Similar to Mr. Enloe and Ms. Baker, she believed that statistics is a social discipline, so she encouraged students to talk amongst themselves while they worked.

Figure 8 shows a summary of the most influential beliefs participants held about statistics, along with ways in which these beliefs impacted their instruction, or would have, according to participants, if mediating factors did not exist. Also shown are the mediating factors, if any, that inhibited these beliefs from being put into practice.

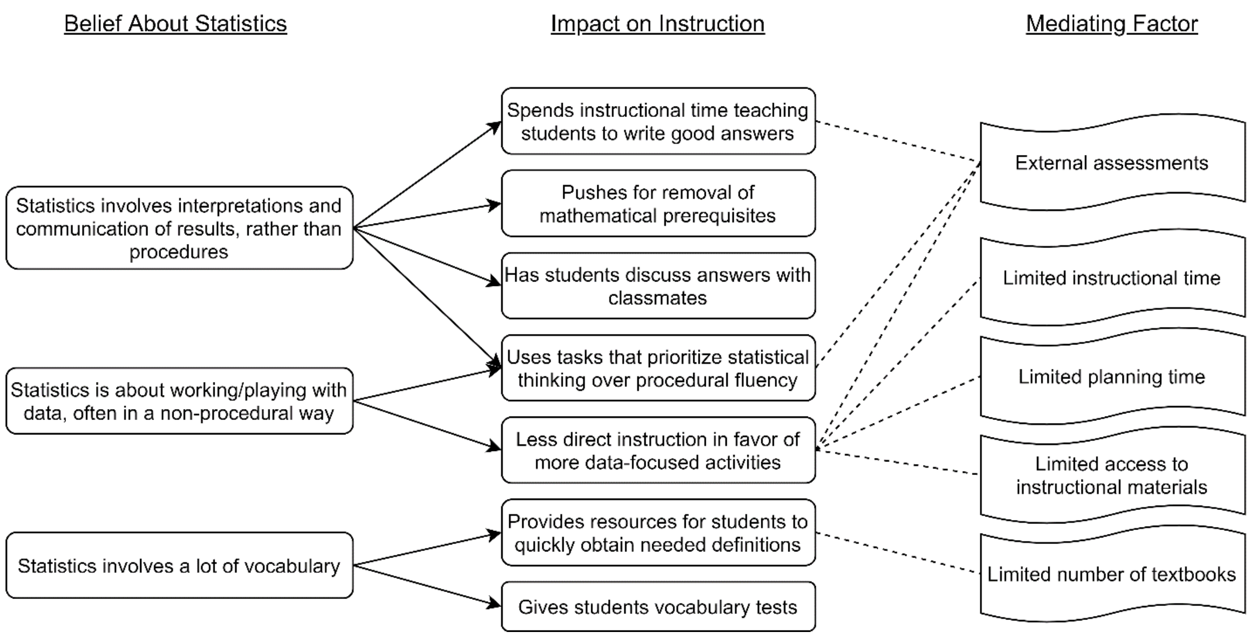


Figure 8. Influence of beliefs about statistics on participants' instruction and mediating factors.

Influential Beliefs about Technology

All participants expressed that beliefs about technology impacted their instruction. Four participants expressed concerns that certain types of **technology, if not used carefully, could hinder students' learning of statistics**. Ms. Carey, for example, felt that once students know how to solve something using a calculator, they tend to memorize steps and forget what numbers actually mean. For this reason, she first makes sure students are proficient at finding solutions with minimal help from the calculator. Ms. Carey does see the value in using technology other than the calculator, particularly in the real world, but she is concerned that students will become over-reliant on technology that they cannot use on AP-exam aligned assessments, so she ensures that students can solve problems without other technologies. Ms. Greene, Mr. Fahey, and Ms. Andrews share similar concerns that students can get into the habit of following procedures when they use technology, which can interfere with their understanding of what the calculations and results actually mean. In general, Ms. Andrews preferred using graphing calculators only if the focus of the lesson was on interpretations of results, rather than on finding results. However, she realized that using calculators can make solving problems more efficient, so because class time is often limited, she allowed students to use calculators to perform calculations, even though she felt "something is lost" when students don't solve problems by hand. Mr. Fahey was concerned that the indiscriminate use of simulation technology can result in students' losing some understanding of the data generation and collection process. He felt that technology can cause a level of abstraction from the context of the data, so he preferred that students first had hands-on experiences before technology-enhanced parts of a lesson.

Ms. Greene had concerns about students simply following procedures on the calculators but also felt that having a good understanding of what the calculator was doing could be beneficial to students. There were certain calculator skills that she had previously taught students to do, such as making a residual plot, that she felt gave students a better understanding of the statistics and of the technology. However, she no longer teaches some of these skills if they are not assessed on the AP Exam, because available instructional time is limited. Ms. Greene does see value in technology other than calculators that let students explore concepts. However, she does not have ready access to computers for students, but rather has a set of shared Chromebook computers that must be checked out in advance. Because she cannot get them on an everyday basis, and because it takes a lot of time to distribute them to students and get students all logged in, she rarely uses them. Ms. Greene said that she "would probably use more technology if [she] had it." However, she also felt that because students cannot use other technology on the AP Exam, she needs to prioritize the use of the calculator over other technologies.

The other three teachers expressed that technology allows students to focus less on the calculations, and more on interpreting the results and understanding the underlying statistical concepts, as seen in Ms. Baker's explanation of why she uses calculators and other technology frequently:

> The biggest role that that technology in statistics can play for me is, we're not having to talk about "this is how you make the dot plot by hand" and "this is how you can use this table." We're able to really think about and focus on the interesting parts of the material. And more of those mechanical and logistical issues are just sort of dealt with by the technology.

For some teachers, limited available class time impacted how frequently students would use technology other than calculators. For example, Mr. Fahey said that he would sometimes manipulate data on his own computer and project it to the students rather than let students manipulate the data themselves. He would do this in order to save time, despite the fact that he "[doesn't] think they'll get as much out of that."

Figure 9 shows a summary of the most influential beliefs participants held about technology, along with ways in which these beliefs impacted their instruction, or would have, according to participants, if mediating factors did not exist. Also shown are the mediating factors, if any, that inhibited these beliefs from being put into practice.

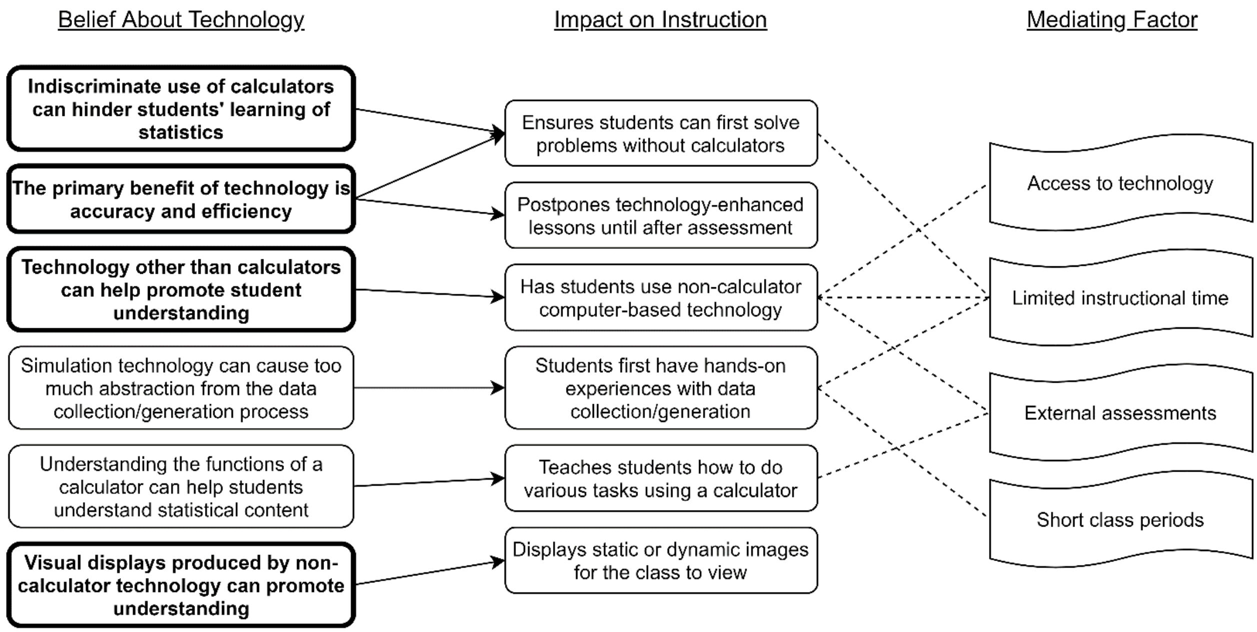


Figure 9. Influence of beliefs about technology on participants' instruction and mediating factors.

Sources of Knowledge That Influenced Teaching Statistics

Aside from their beliefs, participants discussed their knowledge of statistics and of teaching statistics that impacted their instruction. The seven participants expressed that their knowledge came from a variety of sources, including workshops, conferences, and other professional development experiences, personal and work colleagues, and university coursework.

The five teachers who were teaching AP Statistics had all attended at least one AP Statistics Summer Institute, typically a four- to five-day content-rich training provided by College Board-endorsed consultants. Most participants described the Summer Institutes as having a large impact on their teaching. Some mentioned that it helped them know the expected way for students to communicate their answers on the AP Exam, so they adjusted their instruction and assessment practices to help students' responses become aligned with those expectations. Others mentioned how it impacted the order that they taught topics and helped them make connections between different topics. Ms. Greene, however, described the Summer

Institute as “not super helpful”, although she still used some of the tasks and activities that she was introduced to, as did most of the teachers who attended. In addition to attending the Summer Institute, Mr. Enloe had thrice been an AP Statistics Exam reader, someone who evaluates and scores students’ free-response questions on the AP Exam. He described this as “by far the best professional development I’ve ever gotten.” Participating as an AP Exam reader influenced the way he taught toward the exam and the selection of problems students worked on.

Outside of the AP Summer Institute or AP Exam readings, five of the seven teachers mentioned attending workshops, conference sessions, or other professional development opportunities specifically targeted at teaching statistics. These professional developments tended to have a significant impact on these teachers’ instruction, as Mr. Baker described:

> I think they've definitely allowed me to move away from lecturing and into more hands-on collaborative workshopping experiences. I think that when I was less comfortable with what I was teaching and how I was teaching it, it was much easier for me to stand at the board or at a document camera and work examples and be like “This is what you need to know.” And now I've been able to shift from me telling them what to think to giving *them* the chance to think.

For Mr. Dennis, attending the United States Conference on Teaching Statistics caused him to restructure his entire course. Rather than teaching inference only during the last third of the course, he now injects opportunities for inferential reasoning throughout the course. Most participants returned from these professional development opportunities and tried out new tasks that they learned about, often using technologies such as GeoGebra or JMP that they had not attempted to use in their statistics classroom before. Ms. Greene had attended a workshop on using a technology to teach statistics, but because the technology was not available on the computers that were shared throughout the department, she did not implement what she had learned.

Most participants are the only statistics teacher at their school, so their interactions with other statistics teachers are infrequent. However, for those that do have other statistics teachers at their school, these other colleagues are often a source of knowledge and ideas regarding the teaching of statistics. Mr. Fahey, for example, estimated that 80% of the investigative tasks he used with students were a result of weekly meetings that he had with the statistics teachers at his school.

Five participants mentioned that coursework they took during their teacher preparation program also had an influence on their teaching of statistics, though which courses were most impactful varied from teacher to teacher. Ms. Andrews felt that her mathematics education courses that included content on teaching statistics were the most influential, resulting in an increased focus on informal inference, on data collection, and on statistical investigations. Ms. Carey, on the other hand, felt that her undergraduate courses in economics prepared her better to teach statistics than any other undergraduate or graduate-level courses. She had adapted a task, for example, from an econometrics course, and felt that these experiences allowed her to help students see the applications and connections that statistics has to things outside the course. For Mr. Enloe, the influence was more pedagogical in nature, impacting more general teaching strategies or ways to assess students' learning. Mr. Fahey cited his coursework as introducing him to various types of technology that he now uses in his statistics classes.

Occasionally, participants reported that their own understanding of a particular statistics topic influenced how they taught it. For example, Ms. Andrews explained why she chose not to teach her students the binomial formula:

> I've never really liked the formula. I'm not sure why. Like, I don't remember it. I don't use it on a regular basis. So, I think it's more just, like, I use the calculator. It makes sense to me. I can interpret it. So, I don't know...it just seems the easiest to me.

Though Ms. Greene was eager to increase her knowledge of teaching statistics, her teaching was sometimes influenced by her limited experience--at the time of this study, she was in her second semester of teaching statistics. She admitted that a lot of what she does in the statistics course is “trial and error”, taking notes each semester on what works and what doesn’t in order to make improvements for the future. Because planning time was often limited, however, she sometimes found it difficult to implement a large number of changes from her previous semester to her current one. Ms. Greene also tended to rely on past AP Exam questions for her assessments, because she felt she was not adept enough to create good assessment questions for statistics.

Figure 10 shows a summary of impact that participants’ knowledge had on their instruction, or would have, according to participants, if mediating factors did not exist. Also shown are the mediating factors, if any, that inhibited this knowledge from being put into practice.

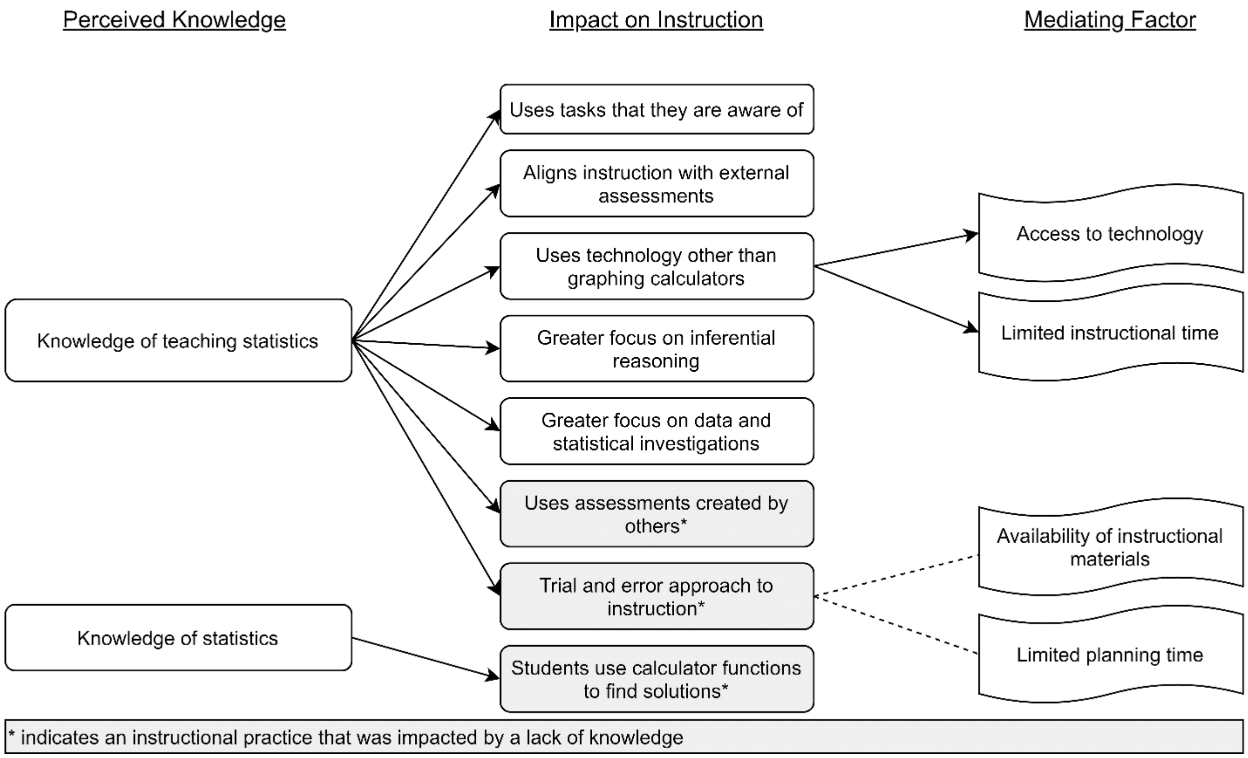


Figure 10. Influence of participants' perceived knowledge on their instruction and mediating factors.

Summary and Discussion

For all participants, beliefs that they held had a significant impact on how they taught statistics, as previous studies have found is the case with teachers of mathematics (e.g. Beswick, 2007; Chen, 2008; Stipek et al., 2001; Sztajn, 2003; Thompson, 1984). The specific beliefs held by participants were far from homogenous, however. Sometimes, these differences were subtle, but in many cases, they were decided--for example, beliefs about the necessity of the teacher in students' construction of knowledge, or whether efficiency with a calculator detracted from or enhanced student learning of statistics. These differing beliefs typically resulted in differences in instruction. For example, participants that felt scaffolding from a teacher played an integral role in students' learning would ensure that there was time built into lessons for teacher-student interactions to occur, while participants that felt it was preferable for learning to occur more naturally would consciously limit the amount of teacher-student interactions and would refrain from directly answering questions.

Even when two teachers held similar beliefs, however, the level to which these beliefs were reflected in their instruction often differed. The seven teachers in this study all taught in different environments and in different contexts. These differing contexts (e.g., class size, access to technology, length of class periods) often played a mediating role in whether and how teachers' beliefs or knowledge were put into practice. For example, Ms. Baker and Ms. Greene both believed that learning statistics was best done through active experiences where students were manipulating or "playing with" data. Ms. Baker's course used a curriculum that emphasizes these types of experiences, and many of the data-rich activities she used in her classroom came from this curriculum. However, Ms. Greene did not have easy access to these types of tasks, but instead had instructional materials from another teacher that mainly consisted of presentation slides. Because she taught four different courses, the amount of time she had to plan for her statistics course was limited. Because she was only in her second semester of teaching statistics, she had not had much time to find or design data-rich activities. Even though Ms. Greene believed that data-rich activities would serve her students better, she often ended up using the presentation slides and instructional materials that were more readily available.

Four of the five influential areas of beliefs identified align with those previously found to impact instruction of mathematics teachers: beliefs about learning (Beswick, 2007), beliefs about teaching (Beswick, 2007; Thompson, 1984), beliefs about students and their needs (Sztajn, 2003; Thompson, 1984), and beliefs about technology (Chen, 2008). The fifth area of belief--beliefs about statistics--could be considered analogous to beliefs about mathematics that have previously been found to be influential for teachers of mathematics (Beswick, 2007; Thompson, 1984).

The most common mediating factors that inhibited teachers' instruction from being aligned with their beliefs included limited instructional time, the presence of external assessments such as the AP Exam, and short class periods. Teachers often felt the need to consider or grapple with these factors when planning instruction, often resulting in instruction that they felt was less than ideal. In some cases, a participant held two beliefs, while not contradictory, would potentially lead to conflicting practices. For example, Ms. Carey believed that her students preferred more teacher-centered instruction and would be stressed out with student-centered activities, but also believed that student-centered activities were more beneficial for their learning. Ms. Andrews believed that assigning a grade to an assignment could motivate students to work harder but could also cause stress for students and negatively affect their work. In these cases, which of these beliefs were ultimately reflected in a teacher's instruction seemed to depend on contextual factors. For Ms. Carey, it depended on how far along in the school year the class was and on the relationship that she had built with students. For Ms. Andrews, it depended on the particular characteristics of the students in her class. This would seem to support Schoenfeld's (1998) model of *teaching-in-context*, in which beliefs may take higher or lower priority, depending on current context.

The belief that had the most mediating factors was that active learning experiences promote learning and increase retention. This belief was held, to some extent, by all seven participants. Yet, six participants reported that contextual factors resulted in them implementing less active learning experiences than they would have desired. These factors included limited instructional time, short class periods, limited access to tasks which emphasized active learning and a lack of planning time to design or find such tasks, large class sizes, and limited space. For many teachers, the belief in the value of active learning experiences seemed to be a core belief that they felt strongly about, and yet many struggled regularly with ensuring students were partaking in these kinds of experiences. Other beliefs that had a large number of mediating factors inhibiting them from being put into practice include the belief that statistics involves working or playing with data and the belief that technology other than calculators can help promote student learning. These beliefs are desirable ones for statistics teachers to hold, and putting these beliefs into practice can have positive impacts on student learning (Franklin et al., 2015; GAISE College Report ASA Revision Committee, 2016). Even if teachers hold these beliefs, however, students may not see these benefits if contextual factors prevent teachers from acting on these beliefs.

None of the teachers reported any mediating factors that inhibited beliefs regarding students and their needs from being put into practice. One possible explanation is that these beliefs may be so important to teachers that they refused to compromise on them. Another possible explanation is that these beliefs may be easier to put into practice and that contextual factors are less likely to inhibit teachers from doing so. Similarly, beliefs about teaching saw significantly fewer mediating factors compared to beliefs about learning, statistics, or technology. More research would need to be done to determine why teachers more readily put these areas of beliefs into practice.

It is noteworthy that only three of seven teachers--Mr. Enloe, Ms. Greene, and Ms. Baker--readily discussed how they approached the teaching of statistics different from the teaching of mathematics. Teachers were not directly asked if they approached these subjects differently. However, for these three teachers, these differences emerged in their interviews rather frequently, while for the other four teachers, these differences were almost never mentioned. There are some key differences between statistics and other mathematical topics

(Rossman et al., 2006) that are reflected in recommended guidelines for teaching statistics (e.g., Franklin et al., 2007). Considering these differences is an important part of planning effective statistics instruction. Most participants reported receiving little to no instruction on teaching statistics during their preservice education. Introducing this instruction into preservice teacher education could help teachers recognize some of these differences and positively impact their teaching of statistics. Da Ponte (2011) and Franklin et al. (2015) provide several suggestions for preparing teachers to teach statistics. These include having teachers analyze statistical tasks and real students' work, challenging teachers' beliefs about statistics, building deep conceptual understanding of statistics, and engaging teachers in the statistical problem-solving process. Franklin et al. (2015) also suggest improvements at the program level, such as increasing the importance and relevance of statistics teacher education and creating partnerships across institutions and disciplines. Promising efforts are being made in developing materials to assist with statistics teacher education, including Enhancing Statistics Teacher Education with E-Modules (ESTEEM, 2020) and Mathematics Of Doing Understanding Learning and Educating for Secondary Schools (MODULE(S^2), 2020). These projects seek to create research-based statistics teacher education materials that can be adapted into various contexts and courses. These materials focus on improving preservice teachers' knowledge of statistics and of teaching statistics, as well as shifting preservice teachers' perspectives and beliefs about statistics and teaching statistics.

Regardless of whether beliefs about statistics were influencing teachers' instruction, many of the planning decisions teachers made are particularly meaningful in statistics classrooms. Soliciting two different solutions to a problem, for example, can play a very different role in a statistics course than it can in many other mathematics courses—the non-deterministic nature of statistics allows for two different solutions to both have merit, whereas in mathematics, one of two competing solutions is likely to be incorrect. Choosing contexts for problems that are relevant to students can have a larger impact in statistics, where context plays a more vital role in the problem-solving process (Cobb & Moore, 1997; delMas, 2004). Giving students the opportunity to engage in non-procedural tasks, though important in other mathematics courses, is crucial in statistics courses to allow students' statistical thinking to develop (GAISE College Report ASA Revision Committee, 2016). As technology has a significant impact on statistical analyses that can be performed and on the learning that can result, the choice of which technology to include in the classroom is also an important one (GAISE College Report ASA Revision Committee, 2016). These decisions are among the many that teachers of statistics must consider if they are to meet the learning needs of their students.

Teachers acquired knowledge on how to teach statistics from a variety of sources. For the two participants that taught in schools with other statistics teachers, these other teachers were a valuable source of knowledge. The other five participants, like many secondary statistics teachers, did not have another statistics teacher in the school that they could regularly consult with, and thus, acquiring knowledge on teaching statistics required putting in more effort. Some reported turning to the internet, searching "far and wide" for advice on teaching statistics. Others tried to seek out colleagues at previous schools or through personal contacts that had experience teaching statistics. Professional development opportunities, particularly those directed specifically at teaching statistics, were cited by several teachers as having a positive influence on their instruction. In addition to the AP Statistics Summer Institutes, participants mentioned conference sessions and workshops on teaching statistics as being particularly beneficial. For Mr. Enloe, participating as an AP Statistics Exam Reader was especially helpful, which Jacobbe et al.

(2013) found was also the case for most AP Statistics teachers. These professional development opportunities also provided opportunities for teachers to create relationships with other statistics teachers that often extended beyond the workshop or conference, which the teachers reported had a positive impact on their teaching. These findings reinforce those of Lee, Mojica, and Lovett (2020), Peters (2014), and Whitaker (2016), who found that teaching in a district that supports professional development opportunities concerning the teaching of statistics and that having access to a community of statistics educators were key factors that allowed teachers to obtain robust statistical knowledge and to have positive statistics teaching identities. However, past research (Lee & Harrison, 2020; Watson, 2001) has shown that relatively few secondary statistics teachers are receiving professional development targeted at teaching statistics. Supporting and encouraging teachers to pursue these professional development opportunities would presumably improve teachers' confidence and ability in teaching statistics, as well as put them in touch with other teachers of statistics, ultimately benefiting the instruction that their students are receiving. Making it easier for statistics teachers to connect with each other outside of one-time professional development opportunities would also seem like a promising way to promote community and knowledge-sharing among statistics teachers. Some participants mentioned the benefit of the large online discussion forums found at the AP Online Teacher Community, but more localized efforts similar to Math Circles (American Institute of Mathematics, 2020) could put statistics teachers in regular contact with other nearby statistics teachers.

All seven participants had to continually grapple with contextual factors, ranging from limited planning time, large class sizes, short class periods, external assessments, and limited access to technology. Even though these teachers may have had the knowledge needed to effectively teach statistics under ideal circumstances, these factors often resulted in the teachers having to either compromise their beliefs or be inventive in the ways in which they dealt with these factors. However, preservice teachers are often tasked with lesson planning without regard to these contextual factors. They are often given weeks to plan a lesson, given access to technology that may not be accessible in many schools, and are given the freedom to choose their class length and characteristics of students. These utopian conditions unfortunately do not match the reality that these teachers will likely face in real classrooms. Making preservice teachers aware of these contextual factors that they will have to contend with and giving preservice teachers practice planning in more realistic situations (e.g., an hour to plan a 45-minute lesson for a class of 30 students with only graphing calculators) can better equip these teachers with the skills they need to succeed in the classroom.

CHAPTER 5: DECISION-MAKING IN SECONDARY STATISTICS CLASSROOMS

Journal

This chapter presents a journal-ready manuscript to be submitted to the *Cognition and Instruction* journal. *Cognition and Instruction* publishes articles focusing on the mental, socio-cultural, and mediational processes and conditions of learning and intellectual competence. Articles focusing on the "how" of learning and intellectual practices are preferred. There are no length limits on articles. The following manuscript adheres to the purpose of the journal, as it focuses on the mental processes of teachers' decision-making, as well as social factors that impact these processes.

Introduction

A teacher's *intended curriculum*—their plan for instruction—goes through a number of transformations in order to become *enacted curriculum*—what actually happens in the classroom (Stein et al., 2007). Many of these transformations are a result of decisions that teachers make in real time, as they are implementing a lesson. Teachers respond to events that occur in the classroom, analyze the current situation, draw on their beliefs and knowledge, and choose whether to alter their lesson, or proceed as planned (Schoenfeld, 1998). This study takes a closer examination of these decisions in the context of secondary statistics classrooms to identify how and why teachers make real-time alterations to their lessons. Specifically, the study attempts to identify the types of events that result in different types of alterations, as well as which knowledge or beliefs are being drawn upon when teachers decide to alter their plans.

Background

Statistics and Statistics Education Research at the Secondary Level

The number of statistics and probability courses offered at the secondary level is at an all-time high. In the United States, for example, over half of all high schools now offer a course in statistics or probability, twice as many as 20 years prior (Banilower et al., 2018; Weiss et al., 2001). Many of these high schools choose to offer Advanced Placement (AP) Statistics (College Board, 2020), with a similar curriculum to a college-level introductory statistics course. In 2019, nearly 220,000 students took the AP Statistics Exam, twice as many as in 2008 (College Board, 2008, 2019); the number of students enrolled in AP Statistics (not all students take the exam) is estimated to be 50% higher (Warne, 2017). Many high schools offer non-AP statistics courses as well, sometimes allowing students to get course credit at a local college or university. Along with this rapid rise in course offerings, statistics education research has also become more prominent over the last two decades.

Various areas of statistics education at the secondary level have been explored, including students' knowledge of statistics (e.g., Ben-Zvi, 2004; Noll & Shaughnessy, 2012), teachers' knowledge of statistics (e.g., Hannigan et al., 2013; Liu & Thompson, 2007; Peters, 2014), curricular issues (e.g., März & Kelchtermans, 2013; Siswono et al., 2018), and teachers' beliefs and perceptions of statistics and statistics education (e.g., Lee et al., 2017; Yang, 2014). However, relatively little research has been conducted on what is actually happening on a day-to-day basis in high school statistics classrooms, and why (Petocz et al., 2018).

Models of Teacher Practice

In the process of implementing a lesson, a teacher faces a number of decisions that they must make in real time. They must decide how to respond to student questions, how and when to assist students that seem to be struggling, how to assess an unexpected solution from a student, what to do when an activity takes longer than expected, and countless others. Schoenfeld (1998) proposes a model of *teaching-in-context* (Figure 11) to explain why and how these decisions unfold. In this model, a teacher's decisions are a function of the current instructional context—the current state of a classroom and events that have just occurred. This instructional context causes certain beliefs or knowledge to be brought to the forefront in a teacher—beliefs and knowledge that may or may not have been at the forefront when planning the lesson. Drawing on these beliefs and knowledge, the teacher responds to events that happen in the classroom, perhaps making alterations to the original plans for the lesson.

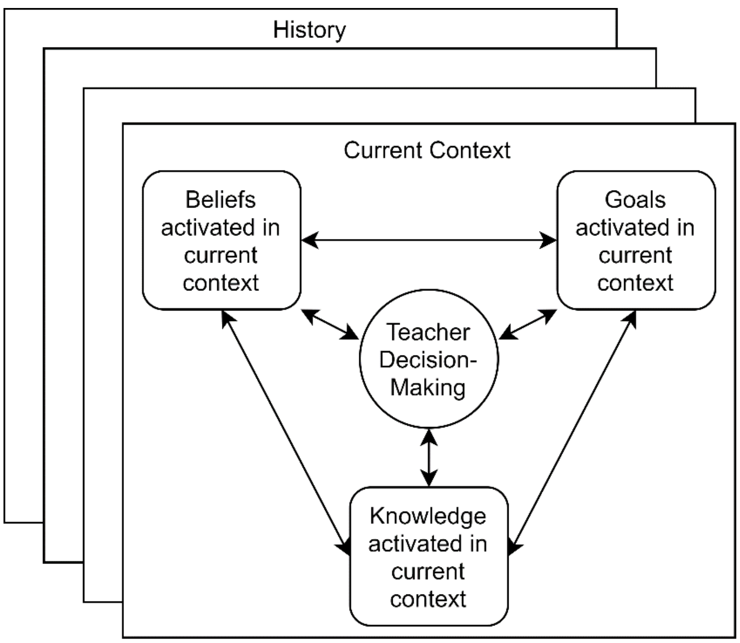


Figure 11. Teaching-in-context. Adapted from Schoenfeld (1998).

Implementing a lesson provides a teacher with an opportunity to assess the quality and effectiveness of the lesson, including how students responded to the lesson and the impact it had on their learning. Teachers can then choose to make alterations to the lesson in order to improve its effectiveness. In the mathematics classroom, Simon (1995) refers to this iterative process as the *mathematics teaching cycle* (Figure 12). These cycles can be a long as an entire course, or nearly instantaneous as a teacher makes real-time modifications to a lesson. Cycles are self-reflective and each cycle can contain multiple smaller cycles.

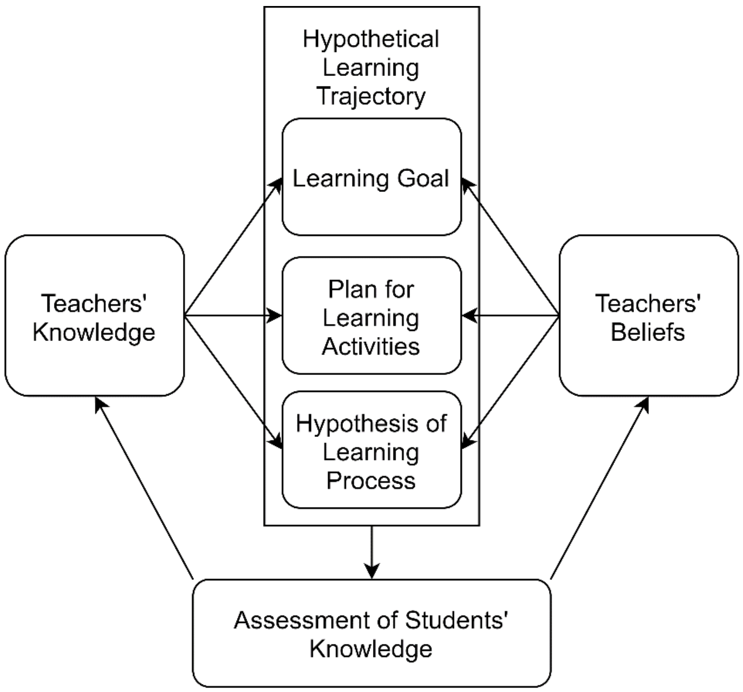


Figure 12. The mathematics teaching cycle. Adapted from Simon (1995).

Using these two frameworks—*teaching-in-context* and the *mathematics teaching cycle*—this study examines the mechanics by which secondary statistics teachers make alterations to their instructional plans. In particular, it attempts to identify which events result in a teacher making alterations to different parts of a lesson, as well as which knowledge or beliefs are brought to the forefront when teachers decide to alter their plans.

Teacher Decision-Making

By teacher decision-making, we refer to the *in-the-moment* ways in which teachers respond to events that occur in the classroom. We can examine this process in three parts: the events that occur, teachers' observable responses to these events, and the mental processes behind these responses--the knowledge and beliefs that are drawn upon to make decisions about how to respond.

Pivotal Events

Not every event that occurs in the classroom requires a response from the teacher. Even events that do require responses are often anticipated by the teacher, and thus the teacher's response does not require in-the-moment decision-making and typically requires little or no alterations to a teacher's instructional plans. Other events, however, are more critical, requiring the teacher to make quick decisions about how to proceed. Stockero and Van Zoest (2013) examined the instructional practice of six beginning mathematics teachers in an attempt to characterize *pivotal teaching moments*--events that occur in the classroom that provide an opportunity for a teacher to modify instruction to support student understanding--as well as the ways in which the teachers decided to respond to those events. Five types of pivotal teaching moments were identified: (a) *extending*, in which a student inquires about a topic that is related to the mathematical content being discussed, but was not a part of the teacher's original plan of instruction; (b) *incorrect mathematics*, in which a student shares an incorrect solution or incomplete mathematical thinking; (c) *sense-making*, in which a student verbalizes his or her attempts to understand mathematical concepts; (d) *mathematical contradiction*, in which two competing solutions or methods have been presented; and (e) *mathematical confusion,* in which a student expresses a lack of understanding about a particular topic. Some pivotal teaching moments have the potential to have significant impact on student learning, while others have a moderate potential to do so.

Teacher Responses

Stockero and Van Zoest (2013) also examined the decisions teachers made in responding to the above pivotal teaching moments. Five types of responses were identified: (a) *extends mathematics and/or makes connections*, in which the teacher goes beyond the material that was originally planned to be discussed in order to examine a related topic; (b) *pursues student thinking*, in which the teacher attempts to find out more about what a student is thinking; (c) *emphasizes meaning of the mathematics*, in which the teacher highlights an underlying definition or underlying mathematics of the issue being discussed; (d) *acknowledges, but continues as planned*; and (e) *ignores or dismisses*. Examining the instruction of a secondary geometry teacher, Cayton et al. (2017) identify two additional types of pivotal teaching moments unique to technology-rich classrooms: *technology confusion* and *incorrect technology use*, as well as one additional type of teacher response: *repeat technology directions*.

Jacobs et al. (2010) examine in-the-moment teacher decision-making through the lens of professional noticing. Focusing on responding to mathematical thinking, the authors describe the process of making decisions as dependent on three skills: (a) attending to students' strategies, (b) interpreting their understandings, and (c) using these understandings in deciding how to respond. Importantly, these are *skills*, in that they can be learned and developed over time and that are impacted by past teaching experiences and professional developments.

Influential Knowledge

A teacher's knowledge takes many forms, including content knowledge, pedagogical knowledge, knowledge of particular students, curricular knowledge, technological knowledge, and so on. Teachers draw upon this knowledge not only in planning a lesson, but throughout implementation of a lesson as they respond to in-classroom events. Hill et al. (2008) identified several impacts on teachers' practice from their *mathematical knowledge for teaching (MKT)*--the knowledge of mathematics that supports not only the *doing* of mathematics, but also the *teaching* of mathematics. Teachers with a higher level of MKT tended to avoid mathematical errors, better responded to student thinking, used richer examples, and made connections between students' work and the underlying mathematics (Hill et al., 2008). Teachers with higher levels of MKT are also more likely to make quality adaptations to a task, assess novel solution strategies from students, keep the lesson moving toward the learning goal (Hill & Charalambous, 2012), compensate for inadequate curriculum (Charalambous et al., 2012), use multiple representations more skillfully, and incorporate student thinking into instruction (Lewis & Blunk, 2012). Groth (2007) adapts the concept of MKT into the statistics classroom, describing *statistical knowledge for teaching*, which includes knowledge specific to teaching statistics, such as knowledge of the differences in how students decode box plots versus stem-and-leaf plots.

Other types of knowledge also impact the decisions that teachers make in the classroom. Teachers with higher levels of general pedagogical knowledge, for example, are more likely to notice and respond to students' comprehension problems, maintain the pace of instruction, manage the students in the classroom, give clearer explanations, and maintain positive relations with students in the class (König and Pflanzl, 2016; Voss et al., 2011). Teachers draw upon their knowledge about students in the classroom when deciding the importance of keeping the lesson well-structured (Sztajn, 2003). Knowledge about technology (Cayton et al., 2017) and knowledge of curriculum (Ball et al., 2008) have also been shown to impact the decisions teachers make.

Influential Beliefs

Like knowledge, the beliefs of teachers entail many different things, including beliefs about mathematics, beliefs about effective teaching, beliefs about how students learn, and more. Several areas of beliefs have been shown to impact the decisions that teachers make in the classroom. Beliefs about mathematics seem to have a significant impact on teacher's decision-making, particularly on how teachers incorporate student thinking into instruction. Teachers who view mathematics as a set of procedures are less likely to have students share novel solution strategies, to cognitively challenge students as they work, or to allow students' input to impact instruction (Beswick, 2007; Muijs & Reynolds, 2002; Stipek et al., 2001). Beliefs about students' social and emotional needs impact the sharing of student solutions and the amount of student-to-student conversation and debate allowed in the classroom (Sztajn, 2003; Thompson, 1984). Teachers' views about effective teaching and their epistemological beliefs impact how likely they are to attend to student engagement and understanding and to allow student thinking to drive discussions (Beswick, 2007).

Teachers also hold beliefs that are particular to statistics and to teaching statistics. Some of these beliefs concern the nature of statistics itself. These include beliefs about the importance and relevance of statistics, the role of computations and procedures in statistics, as well as the role of context, of uncertainty, of data, of technology, of measurement, and of interpretation in statistics (Chick & Pierce, 2011; Harrison et al., 2020; Lee, Mojica, & Lovett, 2020). Reid and Petocz (2002) categorize these beliefs about statistics into six distinct conceptions about what statistics is: individual numerical activities, using individual statistical techniques, a collection of statistical techniques, the analysis and interpretation of data, a way of understanding real-life using different statistical models, and an inclusive tool used to make sense of the world and develop personal meanings.

Other beliefs held by teachers concern the teaching and learning of statistics. These include beliefs about what constitutes a good statistical learning task--for example, whether tasks should be hands-on, include real data or data collected by students, use contexts relevant and engaging to students, or require students to communicate their findings to others (Harrison et al., 2020). Beliefs also concern statistics' place in the school curriculum and which material is important for students to learn (Chick & Pierce, 2011), as well as how students' statistical thinking develops (Lee, Mojica, & Lovett, 2020).

These beliefs about statistics and teaching statistics impact both the planned instruction of teachers and the decisions they make in the classroom (Eichler, 2011). For example, teachers who view statistics as a set of mathematical procedures tend to place less emphasis in their instruction on the context of the situations being examined (Eichler, 2008). Teachers who view statistics as a way to understand the world tend to use real data sets and tend to introduce new statistical methods as a way to explain a real-world phenomenon or solve real-world problems (Eichler, 2008).

Impact on Students' Statistics Learning and Beliefs

The decisions that teachers make during implementation of statistics lessons ultimately have an impact on students' learning and beliefs (Eichler, 2011). Pfannkuch and Horring (2005) examined the development of students' statistical thinking in regard to comparing data distributions during a series of lessons implemented by a secondary mathematics teacher. They found that students' perspectives and attitudes toward the investigative cycle were positively influenced by the emphasis that the teacher placed on the cycle during instruction. However, many students had trouble drawing evidence-based conclusions, a weakness that the researchers attributed to the imprecise language used by the teacher during discussions concerning drawing conclusions and the infrequent modeling of the process of drawing appropriate conclusions.

Cobb et al. (2003) examined the development of eighth-grade students' understanding of statistical covariation over a series of 41 lessons. The researchers found that the development of students' understandings was heavily influenced by the social and sociomathematical norms established by the actions of the teacher and the teacher's responses to students' actions.

Teachers' actions found to influence students' understanding included pressing students to explain and justify their reasoning, to ask clarifying questions, and to express agreement or disagreement with others. How the teacher responded to competing student solutions or to students' statistical arguments were also found to be influential for students' learning.

Eichler (2008) found that the instruction of four teachers with different beliefs regarding statistics resulted in differing knowledge and beliefs in their students. For example, one teacher tended to introduce statistical concepts by having students explore real-world situations--his students tended to believe that statistics was useful for society, though they saw little relevance for their everyday lives. Additionally, his students typically did not give formal ways of explaining statistical concepts, but rather emphasized meaning of the concepts. On the other hand, a different teacher emphasized the theory behind statistical concepts and the logical nature of statistics during his instruction, often focusing on dice, cards, and the like rather than real data sets or realistic situations. His students tended to give formal explanations of concepts but were often unable to make connections between different statistical concepts. His students typically assigned little to no relevance to statistics, either for society or for their everyday lives. Thus, the decisions that statistics teachers make in the classroom can have a significant impact on student outcomes.

Conceptual Framework

The conceptual framework of this study (Figure 13) draws on Schoenfeld's (1998) model of teaching-in-context and on Simon's (1995) concept of the mathematics teaching cycle, incorporating other research on teacher decision-making and on the impact of teachers' knowledge and beliefs on practice. In this framework, teachers have a repository of *knowledge* and *beliefs* as a result of their history (both in and out of the classroom). The knowledge and beliefs appearing in this framework are those that have been found to influence teachers' practice. Influential knowledge includes knowledge of content, of technology, of teaching and learning, and of the students in the classroom. Influential beliefs concern many of the same areas, as well as beliefs about students' needs, desires, and actions.

Teachers use these knowledge and beliefs to construct an *initial instructional plan*. This instructional plan consists of three components: (a) a *hypothesis of learning*, a conception of how knowledge of various concepts are related and how students' knowledge progresses from one concept to the next; (b) a *learning goal* for the lesson; and (c) a set of *planned activities* to help students progress their knowledge and arrive at the learning goal. The formation of this initial instructional plan is impacted by the context in which the teacher teaches. This context could include the technology that the teacher has access to, the expectation for students to take and pass the AP Exam, the pre-requisites for the course, and so on.

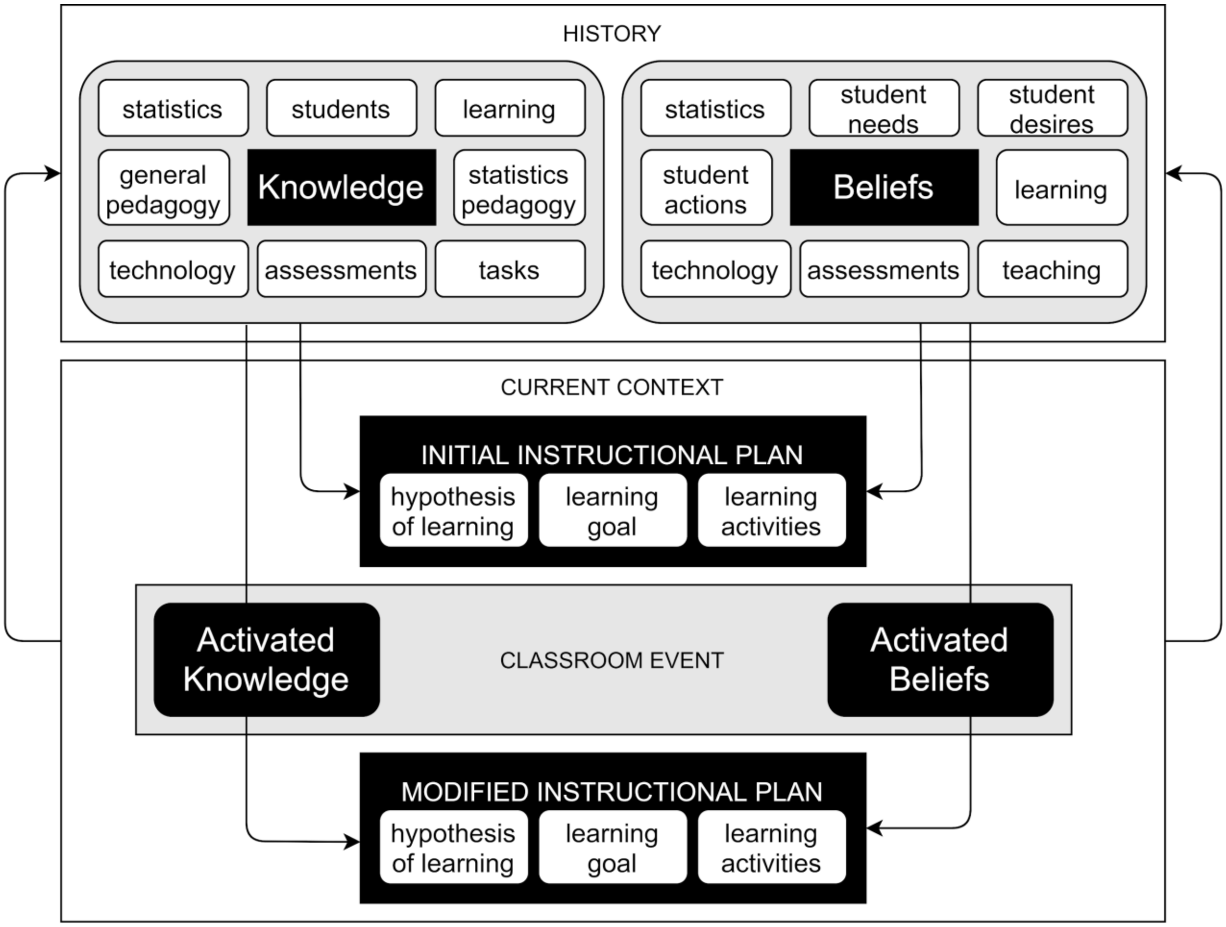


Figure 13. Conceptual framework for this study.

As the teacher implements the initial instructional plan, *events occur in the classroom* that require a response from the teacher. Some of these events may be pivotal teaching moments (Stockero & Van Zoest, 2013), the response to which may have a significant impact on students' learning. These events could be novel solutions presented by students, a disagreement about a solution among students, or a student struggling to progress in a task. Other events may have less of a direct impact on student learning (e.g., a student asking a question not related to the content, an interruption to the classroom), but nonetheless require a response from the teacher.

As a result of events that occur, a certain subset of the teacher's knowledge and beliefs become *activated*, taking a higher priority over other knowledge and beliefs. Based on these activated knowledge and beliefs, the teacher may (or may not) choose to alter their initial instruction plans, creating a new *modified instructional plan*. Any or all of the components of the instructional plan can be modified. The modification of this instructional plan is manifested in the way the teacher responds to the event. Which knowledge and beliefs get activated, and how a teacher's instructional plan is modified, can also be impacted by the context in which the teacher is teaching.

After these alterations are enacted, this experience--the event, the alteration of the instructional plan, and the teacher's response--and the aftermath become part of the teacher's *history*, perhaps altering the teacher's knowledge and beliefs. These new knowledge and beliefs may then be activated by the next classroom event during the current lesson, or they may impact the teacher's initial instructional plan for a future lesson.

This process can happen multiple times in a lesson, resulting in several different alterations to a teacher's instructional plan at various points in the lesson. During a lesson which primarily consists of teacher-led lecture with few questions asked, there may be few events that require a response from the teacher, and thus the teacher's instructional plan may be unlikely to change during the lesson. On the other hand, when there are many chances for students to express their understanding, or in lessons that are not unfolding as the teacher expects, there may be several instances where teachers modify their instructional plan as a result of different knowledge or beliefs being activated or considerations of their current context.

Methods

Participants and Data Sources

Participants in this study consisted of seven secondary teachers. Each teacher was teaching at least one section of a statistics course at the time of the study. The teachers taught in a variety of school contexts, including public, private, and charter schools (Table 6). Students in these teachers' statistics classes typically ranged from 15-18 years of age.

Table 6. Participants' teaching experience and current statistics teaching assignment.

Teacher	**Years of teaching experience[a]**	**Number of prior times teaching statistics[a]**	**Statistics course taught during study**	**School**
Ms. Andrews	8	~8	Statistical Methods I (a credit-bearing course at a local college)	Selective public early college high school
Ms. Baker	14	~12	AP Statistics	K-12 charter school
Ms. Carey	6	~9	AP Statistics	Private religious secondary school
Mr. Dennis	16	~32	Advanced Analytics and Statistics	Private college-prep secondary school
Mr. Enloe	9	~18	AP Statistics	Private college-prep high school
Mr. Fahey	7	4	AP Statistics	Selective public residential high school
Ms. Greene	4	1	AP Statistics	Traditional public high school

[a]: does not include current year/semester. Some teachers teach multiple sections at a time, each of which is included in the number of times teaching statistics.

Data for this study were part of a larger study examining statistics teachers' decision-making during their planning and implementation of statistics lessons. Each teacher first participated in a 30-40-minute video-recorded interview in which the teacher discussed influences on his or her statistics teaching, including beliefs, knowledge, and past experiences such as professional development or previous statistics teaching experiences. Teachers also discussed their goals for the lessons that would be observed and specifically for the lesson that was planned for the first observed class period.

Each teacher was then typically observed for three class periods (one teacher was observed for two class periods). For each teacher, the observed periods were consecutive lessons for a particular section of a statistics course. Observations for five of the seven teachers were video recorded, and the other two teachers' observed lessons were audio recorded. Field notes were used during the observations to capture several types of information. First, the researcher recorded any general implementation decisions that the teacher had not previously discussed being part of their plans for the lesson. This often included decisions such as grouping students or demonstrating how to solve an example problem—decisions that may have been part of a teacher's initial instructional plan, but that the teacher did not mention when they discussed their plan during the prior interview. Second, the researcher recorded any events that had the potential to alter the teacher's instructional plan. These were events that the researcher determined may have been unanticipated by the teacher and that required a response from the teacher. These events could originate from a student (e.g., by asking a question) or they could originate from other sources (e.g., the result of a random simulation). Lastly, the researcher recorded a description of the teacher's response to each of these events.

Following each observation, the observed teacher participated in a post-observation video-recorded interview. Each interview consisted of multiple parts. Part 1 asked teachers to discuss their initial thoughts about the lesson, and to identify any unforeseen events that occurred during the lesson, any deviations from their initial plan for the lesson, and any changes the teacher would consider making the next time the lesson was taught. Part 2 and 3 of the post-observation interview concerned the information recorded in the field notes.

In Part 2 of the interview, the teacher was asked to discuss each of the general implementation decisions recorded in field notes that had not been previously discussed in interviews. For example, one teacher was asked to discuss the decision to collect homework, when she had not done so in previous classes. In particular, the teacher was asked to indicate whether the decision regarding how to implement the lesson was made before the class or whether it was a decision made in real-time (i.e. an alteration to the teacher's initial instructional plans). If the decision was made in real-time, the teacher was asked to discuss the reasons behind that decision.

Part 3 of the interview consisted of a stimulated recall (Calderhead, 1981). The researcher presented the teacher with a video or audio recording (if available) or read a description of an event that was recorded in field notes that the researcher determined had the potential to alter the teacher's instructional plans, along with the teacher's response to that event. The teacher was then asked to recall their thoughts of the event when it happened and then describe the reasons that they decided to respond in the way that they did. This was repeated for several different events and responses that had been recorded in the field notes. Because of limited time during the interviews, not all events recorded in the field notes were discussed. When selecting events to discuss, the researcher prioritized those events that seemed to require a conscious, thoughtful response from the teacher rather than a reflexive response. The researcher also attempted to select a diverse set of events to discuss, sometimes omitting those events which were similar to ones already discussed. It is acknowledged that this process of selection by the researcher may privilege certain events over others, and that results should be considered with these potential biases in mind. Part 4 of the interview asked the teacher to discuss their plans for the next observed lesson, if applicable. Post-observation interviews typically lasted 30-40 minutes.

In total, the data for this study consisted of seven video-recorded pre-observation interviews, 20 audio- or video-recorded classroom observations, field notes taken during those observations, and 20 video-recorded post-observation interviews.

Data Analysis

Transcriptions of the seven pre-observation and 20 post-observation interviews were first analyzed to identify beliefs that were influential on teachers' instructional practice. Responses to interview questions were assigned codes identifying the area the belief concerned (e.g. beliefs about learning, beliefs about technology). Codes were based on previous literature identifying beliefs that influence teachers' practice. A similar process was used to identify interview responses in which a teacher discussed knowledge that they held that influenced instructional practice (e.g., knowledge about statistics, knowledge about particular students).

The 20 post-observation interviews were then analyzed to determine which events in the classroom caused teachers to alter their instructional plans, and how. Each observed event that was discussed during the post-observation interviews was classified according to the type of event (e.g. being short on time, a student asking a question). The teacher's response to this event was classified according to how the teacher's instructional plan was altered (e.g., postponing a learning activity, altering the class discussion). The initial list of classifications for events and responses were based on research regarding teacher decision-making (e.g., Cayton et al., 2013; Stockero & Van Zoest, 2013), with other classifications emerging as common types of events and responses were observed. Events were then paired with their corresponding responses. These pairs were additionally labeled with any beliefs or knowledge identified in the previous round of coding that impacted the teacher's response. The event-response pairs were then categorized by the type of response. Each of these response categories were analyzed to determine which types of events tended to result in each type of response. These groups were then further analyzed to determine if certain beliefs or knowledge tended to impact how teachers responded to these events.

The video and audio recordings of the classroom lessons and field notes taken during the lessons were primarily used during the stimulated recall portion of the post-observation interviews, but were also used to supplement and triangulate data from the interviews. First, because not all instances captured in the field notes were discussed during the interviews due to time constraints, any such instances in the field notes were coded similarly to the interview data to identify teachers' instructional plans, events that had the potential to modify instructional plans, teachers' responses to events and alterations to instructional plans. Similarly, each classroom recording was watched or listened to and any instances not captured in the field notes nor discussed in the interviews were coded in the same way. Because it was no longer practical to ask participants about these events, limited conclusions could be made about the knowledge and beliefs that influenced these decisions. However, these data were used as a means of triangulating interview data--to build a more complete description of a teacher's instructional plans, to identify additional evidence of teachers' beliefs and knowledge, and to understand how teachers generally responded to events in the classroom and determine if their instructional plans were altered. The video and audio recordings were also used to ensure that the events and responses discussed in the interviews were being faithfully categorized whenever there was uncertainty based on the interview transcripts alone. Finally, the video and audio recordings were used to corroborate conclusions made from the interview data whenever the interview data were ambiguous or lacked detail.

Results

During the 20 post-observation interviews, 232 events recorded in the field notes that had the potential to change teachers' instructional plans and 227 responses to those events were discussed with participants (sometimes two simultaneous or successive events resulted in a single response, and some events resulted in multiple responses). By analyzing these event-response pairs and the accompanying beliefs or knowledge of the participants that influenced responses, events were identified that caused teachers to alter various aspects of their instructional plans, including the student learning activity, the lecture, and the whole-class discussion. Not all the 227 responses to events consisted of alterations to participants' instructional plans, but at least one alteration occurred in every observed lesson, with some participants altering their instructional plans more frequently than others. Some of the difference in the frequency in which participants altered their instructional plans can be explained by differences in instructional style--student-centered instruction tended to result in a higher rate of events that had the potential to result in alterations to instructional plans. However, even teachers who encountered similar events often responded to those events differently. These different ways of responding can be attributed to differences in the beliefs and knowledge that participants were drawing upon, as will be elaborated below.

What follows are the results concerning the types of events that resulted in alterations to participants' instructional plans--specifically, alterations to student learning activities, to lectures, and to whole-class discussions. Then, to illustrate how differences in beliefs and knowledge impacted participants' alterations to instructional plans, a closer examination of the decision-making of two participants--Ms. Baker and Mr. Dennis (pseudonyms)--is presented.

What Caused Teachers to Alter, Add, Cut, or Postpone a Student Learning Activity?

As part of their instructional plans, all teachers in this study had a series of learning activities for students to engage in to help them reach the intended learning goal. At times, events that occurred in the classroom resulted in teachers modifying these activities, or in removing, adding, or postponing a learning activity. The most common types of events that resulted in these alterations to learning activities were observing evidence of student thinking, students' work pace, and having minimal class time remaining.

Observing Evidence of Student Thinking

Several teachers in the study made alterations to learning activities after observing evidence of student thinking. In one example, a teacher had planned an activity in which students would complete a practice free-response question from a past AP Statistics Exam without any teacher assistance. At multiple points during the activity, students asked the teacher to assess their work up to that point. Based on the work that she saw, the teacher ended up providing assistance to these students, helping them arrive at a more complete solution. When asked about her reasons for altering her plan, the teacher cited her knowledge of these particular students and her beliefs about the quality of work and level of effort these students would produce if she had refused to provide assistance. In another example, conversations that a teacher had with students just prior to class about their lack of understanding of the homework resulted in the teacher modifying the homework review from a whole-class activity to one where she circulated around the room, having one-on-one conversations with students. This teacher believed that these one-on-one conversations would increase self-efficacy and reduce feelings of panic in students, which she felt was one of her duties as a teacher.

Students' Work Pace

Students progressing through activities faster or slower than expected also resulted in teachers making changes to student activities. Often, this simply meant adjusting the amount of time given to students to complete an activity. Other times, the nature of an activity was modified. For example, seeing that a group of students were nearing the end of an activity earlier than other groups resulted in one teacher presenting the group of students with an additional problem and asking them to brainstorm possible approaches to solving it with other students in the group. The teacher had originally intended to present the problem as part of a whole-class discussion. Another teacher, seeing that certain students were progressing through a set of problems more quickly than other students, paused all students and instructed students that had completed certain problems to discuss them with other students at their table before proceeding. During the post-observation interview, the teacher justified this decision based on his beliefs that discussing statistics with others benefited learning and that students need practice verbally describing statistical concepts such as distributions.

Minimal Remaining Class Time

Teachers also altered activities when class time was running short. Sometimes, this simply meant turning a classwork activity—often meant to be a paired or group activity—into a homework activity. In another example, a teacher switched the order of tasks in an activity that students were working on. In this instance, conversations that the teacher had been hearing among students made him believe that students could be successful on problems that had originally meant to be completed after a hands-on simulation activity. As time was short, he decided to postpone the hands-on simulation until the next day. When asked about this decision, the teacher remarked that he felt that performing the hands-on simulation without time immediately afterwards for a whole-class discussion would result in a less satisfactory learning experience for students. For some teachers, being short on time resulted in simply getting rid of a planned activity. Teachers cited several reasons for choosing the activities to get rid of, including the activity not being required for students to reach the learning objective, the purpose of the activity being to practice skills rather than build knowledge, and the teacher being able to assess students' knowledge more easily with other activities. Only one instance was observed in which a teacher added a learning activity that was not previously planned to a lesson. This came about as a result of having some time left at the end of class after having completed the planned learning activities. This teacher believed that low-stakes, low-stress assessment activities provided students with a way to engage with material, so she gave students a flashcard activity to complete with their neighbors.

Reflecting on Core Learning Beliefs

In some cases, teachers seemed to modify an activity solely based on their beliefs about how students learn in general rather than reacting to an observed event. In one instance, a teacher changed an activity from students working with their neighbor to students being randomly assigned someone to work with. In another, a teacher changed a whole-class discussion into a small-group discussion, asking each group to discuss a scenario among themselves. In both cases, the teachers described their in-the-moment realizations based on their beliefs about the types of discourse that best supported student learning that resulted in changing how students were grouped for a discussion to better support their learning.

What Caused Teachers to Alter a Lecture?

As part of their lesson, several teachers included a teacher-centered presentation of new material, demonstration of how to solve problems, or review of previously learned material. For those that did, events sometimes occurred in the classroom that resulted in these teachers making modifications to these lectures. Observing evidence of student thinking and questions that students asked were the most common events that prompted these changes, though other types of events also did so.

Observing Evidence of Student Thinking

Three teachers used information gathered from the previous lesson's homework to make modifications to their lectures. These teachers either reviewed submitted homework just prior to class or during the first few minutes of class, or they had conversations with students prior to class about the homework. Most often, these alterations to instructional plans involved a teacher identifying where students unexpectedly struggled and then adding to or modifying the planned lecture to better address those content areas. Sometimes, however, students performed better on their homework than a teacher had expected. This resulted in teachers deciding not to spend time during the lecture reviewing that material, and instead either spending more time reviewing other material or giving students more time to work on the day's task. One teacher had planned a lecture that would build off what students had done for homework, but having seen that most students made little progress on the homework—either from lack of comprehension or lack of effort—she chose to modify the lesson to focus more on a review of the fundamental concepts.

Questions Asked by Students

Questions that students asked, either during or before a lecture, also often resulted in modifications to those lectures. Like students' responses to the homework, these questions often helped a teacher identify material students may not be comprehending as well as a teacher expected. This would result in the teacher spending more time during the lecture going over that material. Other times, however, students asked questions about topics that the teacher had not planned to cover. In one example, students had approached the teacher before class and asked why the standard normal distribution uses a mean of zero and a standard deviation of one. Because the teacher felt students in her class were "unusually curious" and because she believed that students having unanswered questions could make them feel "panicky", she decided to modify her lecture to cover that topic. This also caused her to modify a different part of her lecture to use a data set that students had previously seen instead of the planned data set that used a new context. She explained that she did this so as to not have to spend time explaining the new context, making up for time spent on the previous additions to her lecture. In another example, as a teacher was handing out a practice free-response AP Statistics Exam question, a student asked about the format of the AP Exam and whether the problem they were about to work on was the type of problems they would see on the AP Exam. This resulted in the teacher spending the next few minutes explaining the format of the AP Exam, something she had not planned to do. When asked about her reasons for this change, the teacher cited both knowledge and belief: the knowledge that unlike most mathematics courses, the students in the statistics course have had a wide range of previous coursework and varying experience with other AP Exams; and the belief that one of her responsibilities is to prepare students for the AP Statistics Exam.

Other Events

Other events also caused teachers to modify their lectures. Being short on time resulted in one teacher deciding to cut a part of his lecture, instructing students to instead review the notes at home in order to have sufficient time to complete the rest of the day's tasks. This teacher believed that students practicing problems and then discussing them as a class was more beneficial to students' learning than a teacher-centered lecture. Another teacher's lecture included using technology to take a random sample of data. Once the teacher saw the plot of the distribution of the sample, she decided to address characteristics of the distribution such as the center and shape, topics that were not originally part of the lecture.

What Caused Teachers to Alter a Whole-Class Discussion?

Most teachers in the study regularly included some form of whole-class discussion in their instructional plans. These discussions took a variety of forms, including students presenting solutions to the class, conversations about what was observed during an activity, and conversations to address any lingering questions or misconceptions that students may have had. For some teachers, these discussions tended to be fairly structured with specific planned topics to discuss, while for others, discussion topics tended to organically emerge based on what occurred during a student-centered activity. Even structured discussions, however, often ended up being modified due to events in the classroom. The most common types of events that resulted in teachers modifying whole-class discussions include observing evidence of student thinking, a student presenting a solution, a student asking a question, and students' pace of work.

Observing Evidence of Student Thinking

As students worked, teachers would often circulate around the room, either listening to student conversations that were occurring or stopping and engaging in conversations with individual students or groups. The information about students' thinking that teachers gathered from these conversations often informed and modified the subsequent whole-class discussion. For example, during an activity on probability and informal inferential reasoning, students discussed among themselves the implications of an event occurring when it had a three percent probability of occurring under a certain probability model. The teacher overheard several students commenting that they were not willing to discount the viability of the probability model, given that three percent is "still a chance." As a result, the teacher decided to add a section to the whole-class discussion in which he proposed increasingly smaller probabilities and asked the class for their opinions about the viability of the model. According to the teacher, this instructional move was guided by his belief that having students consider these types of issues would promote the development of inferential reasoning.

As the teachers circulated around their room while students were working, several of them attempted to identify common student mistakes or misconceptions. They would then make a point to emphasize these concepts during the whole-class discussion, in order to ensure students had more opportunities to develop stronger conceptions. Other teachers attempted to identify novel or interesting student solution methods or small group conversation topics and would bring these to the fore during a whole-class discussion, often asking specific students to share their strategy or small-group discussions with the class.

Teachers also attended to student thinking during whole-class discussions, often resulting in real-time modifications of those discussions by the teacher. Often, this happened during a student’s response to a question asked by the teacher. That is, teachers often asked questions during the whole-class discussions as a way to assess students’ knowledge and then, based on that assessment, steered the discussions in a way they felt would benefit students’ learning. At other times, students would make a comment about something being discussed that would then result in the teacher modifying the discussion. In one example, the class was discussing a problem in which students were asked about the relative impact that adding two different points to a bivariate data set would have on the linear correlation coefficient. One student commented that when one considered the formula for the correlation coefficient, it seemed like adding a data point near the mean of each variable would decrease the correlation rather than increase it, contradicting the correct answer to the problem. The numeric calculation of the correlation coefficient was not a planned part of the discussion for this teacher. However, following the student's comment, a somewhat lengthy conversation occurred concerning the correlation coefficient formula and how it related to the graphical representation of the data set and the least-squares regression line. Eventually, the teacher decided to cut the conversation short in order to proceed with the rest of the planned discussion. When asked about this series of events, the teacher referenced multiple factors that influenced his decisions. He felt that the student that made the comment was "a very strong math student" and might benefit from a conversation about the computations involved, but also knew that other students in the course had a wide variety of mathematical backgrounds. He was also knowledgeable enough about the statistical content to know how to relate the computation of the correlation coefficient to the graphical representations, but believed that computations were less important in statistics than interpretation and communication were, so ultimately decided to table the discussion and proceed with other planned parts of the lesson.

Student Presentation of a Solution

Several observed lessons included a portion where students presented solutions to problems to the rest of the class. Oftentimes, the content of these solutions resulted in the teacher adding something to the whole-class discussion that was not originally planned. For example, a student in one teacher's class presented a solution in which she computed the mean and standard deviation of a binomial random variable but did not justify the use of the formulas utilized. Because the teacher believed that students often blindly use formulas without knowing when they are applicable, the teacher decided to add a segment to the subsequent discussion about how to differentiate binomial random variables from other types of random variables. In another example, a student provided a verbal answer which contained some slight errors. Rather than immediately correcting the student, the teacher recorded the student's answer on the board, and proceeded with a whole-class discussion about the student's answer. In the post-observation interview, the teacher referenced two factors that impacted this decision. First, he believed that this particular student would be able to handle having her answer analyzed in front of the class. Second, he believed that it is beneficial for students' learning to not always see a correct answer, but instead to build on or refine an answer to make it more complete.

Questions Asked by Students

Questions that students asked, even if they were not directly related to students' understanding of the statistical content, also resulted in modifications to whole-class discussions. For example, after a class discussion on an activity that students had completed, a student asked what a potential assessment question would look like that covered that content. Although this was not part of the teacher's instructional plan, he decided to add a new problem as part of the class discussion. This teacher believed that students' curiosity and engagement was important, and that responding by giving students an example of a problem that would be on the assessment was a way to promote those characteristics in students. In multiple observed instances, teachers responded to questions that students asked during a discussion by requesting that another student in the class answer the question. Sometimes this request was directed at a particular student that the teacher felt could provide a sufficient answer to the question, and at other times, it was an open request to everyone else in the class. In the interviews, these teachers expressed their beliefs about the value of student-to-student conversations, and their beliefs that having students explain concepts to each other benefited both the explainer and the listeners.

Students' Work Pace

Discussions were also modified in various ways as a result of students progressing through activities at a different rate than expected. When only some groups were able to complete an activity during one teacher's lesson, that teacher modified the post-activity discussion so that those groups shared their results with the rest of the class. In another instance, students were progressing slower than expected, so the teacher decided to take some of the planned post-activity discussion and have a discussion after students had completed only the first part of the activity, in order to help them be focused and successful on the rest of the activity. Another teacher, after seeing an activity was going to take longer than expected, decided to incorporate the planned post-activity discussion into the discussion that was planned for the start of the next class in order to give students more time to complete the activity.

Summary

Across the 20 observed lessons, 232 events that required a teacher response and had the potential to result in an alteration to participants' instruction plans were recorded in the field notes and discussed during the interviews, with each lesson having as few as five such events or as many as 18. The actual number of events that had the potential to result in alterations to instructional plans was likely higher, as some recorded events that were deemed to be redundant or less relevant were not included in the stimulated recall portion of the interviews, and other events may have not been recorded at all. Thus, the data from the post-observation interviews includes primarily events privileged by the researcher in the discussion. For these 232 events, 227 participant responses were discussed, though not all resulted in alterations to student activities, lectures, or whole-class discussions. At least one alteration was discussed by participants for each observed lesson, however, with one lesson having over a dozen different alterations that were discussed.

Regardless of what portion of the instructional plans were altered (student activity, lecture, or whole-class discussion), seeing evidence of student thinking was one of the most common events that resulted in alterations. This was often a result of a conversation that a teacher had with a student or listening to a conversation that students had with each other. Other common events resulting in alterations include questions that students asked, often resulting in alterations to lectures and discussions to address these questions. Students' work pace often resulted in alterations to student activities and discussions, either to give students more time to finish the activity or build understanding, or to modify the activities or discussion so that they took less time. Solutions presented by students often shaped the discussions that followed in ways that participants had not originally anticipated. Being short on time resulted in various types of alterations to student activities, including cutting or postponing activities or modifying them so that they could fit into the minimal remaining time.

The Impact of Teachers' Beliefs and Knowledge on Alterations to Instructional Plans

It should be noted that different teachers responded differently to similar events. If an event resulted in a teacher modifying a lesson, a similar event may have resulted in a different modification—or no modification at all—for a different teacher. During post-observation interviews, teachers were asked about the reasons for modifications that they made. Teachers often cited beliefs they held (e.g. beliefs about learning, beliefs about statistics) or knowledge that they had (e.g. knowledge about particular students, knowledge of statistics) when explaining why they made the changes they did. Teachers also expressed their beliefs and knowledge throughout other parts of the interviews. Even if teachers did not cite their beliefs or knowledge when discussing specific decisions, they each seemed to have an underlying system of beliefs and set of knowledge that informed the decisions they were making. To illustrate how a teacher's system of beliefs and set of knowledge consistently impacted how that teacher modified their instructional plans, a closer examination of two teachers' decisions will be presented.

On the surface, Ms. Baker and Mr. Dennis planned lessons that seemed to be structured very similarly to each other. Both teachers typically planned to start class with a short whole-class discussion to review previously learned material. Both then planned to have students, typically arranged in groups of four, complete a student-centered activity to build new knowledge. Finally, both planned to end each class with a whole-class discussion to solidify the knowledge that was meant to be gained from the activity. Occasionally, Mr. Dennis would also include a short (less than 10 minutes) teacher-centered lecture prior to the student activity, but both teachers expressed that they preferred activities over lectures as a way for students to build understanding. Despite these similarities in their instructional plans, their implemented lessons often turned out very differently, as the two teachers tended to differ in the quantity and type of modifications made to their planned lessons. In the next sections, a closer analysis of these two teachers will illustrate how differences in instructional modifications could often be traced back to differences in their systems of beliefs and sets of knowledge.

Ms. Baker

According to Ms. Baker, learning statistics happens through having experiences and making observations on those experiences. She juxtaposed this with being told new information, which she believed resulted in comparatively lower comprehension and retention. Of the classifications in Reid and Petocz (2002), Ms. Baker's conception of statistics might be categorized as *an inclusive tool used to make sense of the world and develop personal meaning.*

Ms. Baker believed that finding a solution to a problem was not the primary means by which learning occurred, but rather that learning resulted from explorations and processing the result of those explorations. She believed that having these learning experiences in a collaborative setting allowed for authentic discussions to occur and built a sense of community—"statistics is social", according to Ms. Baker. As a teacher, Ms. Baker believed her role was to set up these experiences for students to have using relatable scenarios, and then to orchestrate discussions that help students turn their observations into concrete conclusions that they could carry forward. She also believed that a teacher's role was to help students notice connections between the new understandings they were developing and their previous knowledge and experiences. Ms. Baker highly valued students' voices and tried to center all aspects of a lesson around student thinking. She wanted students to feel comfortable sharing their thinking even if it was incomplete and to feel that whatever point their understanding was at was a valuable foundation upon which learning could occur.

Ms. Baker taught AP Statistics in a relatively small K-12 charter school. At the secondary grade levels, there were approximately 40-46 students per grade level. Class sizes were generally small, with only 12 students in Ms. Baker's statistics class. Ms. Baker had 15 years of prior teaching experience and had taught approximately a dozen statistics courses in that time. Ms. Baker taught a variety of other courses at her current school, as well, including Computer Science and Math 1 (a course based on an integrated approach to have concepts of algebra, geometry, number, and statistics introduced), the latter of which all students in her statistics course had taken in the past. As a result, Ms. Baker had established relationships with most, if not all, of the students in her statistics course, and had previously taught many of them. Despite the observations for this study taking place less than a month into the school year, Ms. Baker had considerable knowledge of her students, both on an academic level and a personal level. Ms. Baker also possessed an advanced knowledge of statistics, having an undergraduate degree in pure mathematics and having studied operations research and actuarial sciences.

From the very beginning of each class, Ms. Baker attempted to incorporate students' voices into the lesson. One noticeable example of this was early in the second observed lesson when the class was reviewing what was learned the day before. Ms. Baker had asked students to list characteristics of a quantitative univariate data set that were not impacted by adding a constant to each data value. One student suggested an answer ("layout of the graph") that was less precise or formal than the answers Ms. Baker had in mind. Nevertheless, Ms. Baker decided to include the student's answer in the list being compiled on the board. In the interview after the lesson, she indicated she did this because she believed that students needed to feel their thinking was worthwhile and that teachers should try to honor students' thinking when possible. She wanted to encourage "rough draft thinking" because she believed that it served as a valuable foundation upon which to learn. She also cited knowledge that this particular student had not always been successful at mathematics despite working hard and being eager to engage in her classes. Thus, she wanted to have that student feel that her thinking was driving the class discussion.

As most of the activities in Ms. Baker's class were student-centered and collaborative, she had ample opportunities to observe student thinking. During classwork, she often responded to a student's thinking by pressing the student to verbalize their reasoning or to think more deeply about a situation. This occurred irrespective of the level of sophistication of a student's thinking. For example, during an activity on linear transformations of data, a group of students were able to successfully explain to Ms. Baker how various statistics changed when data were transformed. Ms. Baker responded by pressing them for an explanation of why certain statistics acted similarly or differently to other statistics. When the students were unable to readily provide an explanation, she suggested that they explore the data beyond the confines set by the task, by using technology to explore visual representations of the data. Ms. Baker often suggested such alterations or additions to the task to students as they worked. She later stated that students being able to produce an answer to the problem *was* part of her instructional goal, but that she believed that being able to produce a correct answer was not always a reflection of deeper understanding.

Ms. Baker also believed that she should not be the sole or ultimate authority of what was correct in the classroom. Because of this, when students asked her questions while they were working, she often responded with questions of her own or asked the students to first discuss the question with other students at their table. During an interview, Ms. Baker was asked about a particular incident in which she responded to a student’s question with “Have you asked your table first?” Ms. Baker knew that this student was the only tenth-grade student in the class and was likely to have had fewer interactions with other classmates in the past, and thus was more likely to turn to the teacher as a resource instead of to other students. Because of that, Ms. Baker said she was more deliberate about redirecting that student’s questions to other students at the table in order to engender more student-to-student conversations than might naturally happen.

Whole-class discussions in Ms. Baker’s class were heavily shaped by the student thinking that she observed during preceding activities. She often asked particular students or groups to share their methods or reasoning with the rest of the class. The resulting conversations typically constituted the majority of the whole-class discussions that occurred. These conversations were often ones that Ms. Baker did not anticipate having prior to the lesson, as the students she chose to share often had novel solutions or explanations. Indeed, when asked about how she typically plans for a whole-class discussion, she remarked that she will sometimes have topics in mind to address but that, in general, her plans for the whole-class discussion take form while the class activity is going on. She believed that effective teaching meant being able to take students' thinking and use it to direct the class toward the learning goal.

Ms. Baker's decisions in the classroom were driven by her belief that centering the lesson on student thinking was the most effective way to help students learn, as well as by her knowledge of the students in her class. She believed that statistics, and learning in general, was social, and that having student-to-student conversations would benefit students' learning. Although Ms. Baker had clearly defined learning goals for each lesson, she did not believe that the path to that learning goal was clearly prescribed. Thus, she very readily made changes to a lesson in real time, based on what she saw happening in the classroom.

Mr. Dennis

Like Ms. Baker, Mr. Dennis believed that learning statistics was best accomplished through experiencing statistics. For Mr. Dennis, this typically meant presenting students with a problem for which statistics is needed or useful, and having students apply their own knowledge and reasoning to work toward a solution. Of the classifications in Reid and Petocz (2002), Mr. Dennis's conception of statistics might be categorized as *a way of understanding real-life using different statistical models.* Mr. Dennis viewed students' struggles as conducive to their learning, as long as they eventually reached the intended learning goal. One way that students' struggles could remain productive was by having other people to bounce ideas off of and hear other perspectives from. Mr. Dennis believed that theoretical concepts were better understood if students first saw them in action. He believed that students needed to "experience" statistics and talk about it in order to learn it. This could mean seeing the results of a probability simulation or manipulating data, and then having conversations with others about what was observed. As a teacher, Mr. Dennis felt his role was to ensure students did not get too lost as they were working and to ensure that students ultimately reached the desired learning goal. The value of whole-class discussions was in making it clear to students what they should have learned and clearing up any lingering misconceptions that students may have.

Mr. Dennis taught in a private, college-preparatory secondary school (grades 9-12) that enrolled approximately 460 students across the four grade levels. The course observed was Advanced Statistics--though not an AP course, it largely followed the AP curriculum, and most enrolled students took the AP Statistics Exam. Mr. Dennis was the most experienced of the teachers in the study, both overall and in teaching statistics, with 16 years of teaching experience and having taught over 30 statistics courses. Like Ms. Baker, Mr. Dennis likely had more knowledge of statistics than most teachers, having double-majored in statistics during his undergraduate education. In the years leading up to the study, statistics courses had made up the bulk of Mr. Dennis's teaching load, teaching up to four sections of statistics a year (the school offered two different versions of a statistics course). Thus, this statistics course was the first time that Mr. Dennis had taught most students in his class. As the observed lessons happened less than a month into the school year, Mr. Dennis was still learning about his students—when asked to describe any interesting characteristics of the students in the observed class, he noted that "It's hard for me to answer that because all the classes blend together."

Mr. Dennis's students submitted their homework digitally. Prior to class, Mr. Dennis would often review submitted homework to identify common errors or misconceptions and use those to inform the start-of-class review. While students worked on the day's activities at their tables, Mr. Dennis often circulated around the room, listening to conversations that were happening. In general, if he was satisfied with the conversations he was hearing, he would refrain from entering into student conversations unless he was asked a question. When asked about that decision, Mr. Dennis cited his belief that students working with other students to make progress toward solving a problem resulted in a deeper understanding than being told how to solve it. However, if he heard an incorrect explanation, he would often step in to point out something the student may be missing or to ask the student to clarify their explanation. Even though he refrained from entering into conversations that he saw as productive, Mr. Dennis did spend a significant amount of time during classwork having conversations with students or groups, particularly with students that he noticed were struggling or that requested assistance. These conversations typically consisted of students explaining their reasoning or their struggles and Mr. Dennis attempting to help students build on their knowledge and further their thinking. Mr. Dennis often referred to these conversations as helping students get past their struggles or helping them make progress toward the intended learning goal. When he saw that multiple students were struggling more than he expected, he would extend the amount of time students had to work to ensure they could reach that learning goal.

Along with advancing students' thinking, Mr. Dennis stated that one of his main goals during classwork was to learn about students' reasoning, particularly their misconceptions. When he saw misconceptions that were common, he made a point to address or emphasize them during the post-activity discussion. Beyond that, however, what he saw during classwork generally had little impact on the resulting discussion. No instances were observed where he purposely selected a student or group to share their work—any sharing that occurred was by volunteers. Thus, if students had novel solutions strategies or astute insights about the content, they may not have been shared with the rest of the class. When students did share their solutions to a problem, Mr.

Dennis often followed up the students' explanation by clarifying anything he felt may have been unclear or by emphasizing the key points of the solution. In general, class discussions were relatively structured and planned out before the lesson began. When discussions started to drift toward subjects that Mr. Dennis felt were not relevant to the lesson's learning goal, he was quick to steer the conversations back on topic. For example, during one class discussion on a problem about joint probability, one student proposed a modification to the problem and asked how it would impact the solution. Mr. Dennis responded by suggesting that they refrain from exploring the hypothetical problem to avoid causing confusion for the rest of the students in the class.

Mr. Dennis's actions in the classroom were driven by his desire for all students to reach the intended learning goal, and by his beliefs about how to ensure that happened. Each part of the lesson was fairly structured, and real-time alterations were relatively minor. However, this does not mean that student thinking was ignored. Despite being the most experienced statistics teacher in the study, Mr. Dennis was continually revising his conceptions of how students learn statistics, perhaps more so than any other participant. In the interviews, Mr. Dennis frequently discussed what he learned about student thinking, and was already considering multiple ways to improve the lesson for the next time that he taught it. Several of the activities that were used in the observed lessons were activities that Mr. Dennis was using for the first time or activities that had been altered from the previous year, an indication that he was continually taking information learned from what happens in the classroom and using it to make improvements that would benefit student learning.

Summary

In many ways, Ms. Baker and Mr. Dennis were similar as teachers. They were two of the more experienced teachers in the study, and unlike some other participants, both expressed high levels of comfort with their understanding of statistics and their ability to teach statistics. Ms. Baker and Mr. Dennis structured their lessons similarly, with similar types of activities intended to support students' learning. Both teachers were devoted to learning about and assessing students' thinking, as they felt this was an important component of effective teaching. However, Ms. Baker was much more likely to incorporate students' thinking into the lesson, making alterations to the lesson in the moment in response to what she heard and saw. Mr. Dennis also used his assessments of students' thinking to improve lessons, but generally did so from one lesson to the next, rather than altering a lesson in real time. In the framework of the mathematics teaching cycle (Simon, 1995), Ms. Baker could be said to have much more frequent and smaller sub-cycles.

These differences could often be traced back to their differing beliefs and knowledge. Mr. Dennis more strongly believed in the importance of reaching the intended learning goal for the class, and his in-class decisions were often made to help students reach that goal. Although he believed in the value of productive struggle, he was quicker to correct or address misconceptions to get students back on the intended track. Like Ms. Baker, he believed that working collaboratively and hearing multiple perspectives benefited students, but he more often served as an authority on what was ultimately correct. Ms. Baker was more comfortable with what she called "rough draft thinking" and more often let students' thinking drive the class discussions. She also more often leveraged her knowledge of students to adapt the lesson to support the learning and emotional needs of each individual student. It was more acceptable to Ms. Baker if there were certain aspects of a topic that students did not fully understand by the end of a lesson. Part of the reason for this is the curriculum she used, which she described including "spiraled review" and "looping back", and she believed that through the repeated exposures to concepts that this curriculum provided, students would ultimately build a deeper understanding of those concepts.

Discussion

All the teachers in this study had put in time and effort into planning their lessons—some even admitted having put more time than usual knowing that they would be observed. However, after nearly all of the 20 lessons observed, teachers remarked that their lesson did not go exactly as planned. While some teachers like Ms. Baker and Mr. Dennis structured their lesson similarly, overall, there was a wide variety in the pedagogy and in the types of learning activities used. Despite this variety, all observed lessons contained multiple events which required the teacher to assess the situation and respond accordingly. Not all of these events resulted in a teacher altering their instructional plan, and some teachers were more apt than others to make significant changes to their instructional plans. However, every observed lesson contained at least one instance in which the teacher made a significant alteration to the original plan, and most contained several such instances.

Observing evidence of student thinking was the most common classroom event that resulted in alterations to a lesson. This is perhaps not surprising—though teachers often try to predict how students will approach a problem, this is one aspect that the teacher has little control over. Depending on the structure of a lesson, however, there may be more or less opportunities for teachers to observe student thinking. Ms. Baker's and Mr. Dennis's lessons, for example, included large periods of time where students were working in groups on open-ended problems and having conversations with each other and with the teacher. Thus, the teachers had ample opportunity to observe student thinking. Some other teachers' lessons tended to include more individual or teacher-centered activities or included tasks which allowed for less variety in student approaches. These teachers tended to have less opportunity to observe student thinking and the thinking that they did observe was more likely to be anticipated by the teacher. Teachers who had more opportunities to observe student thinking tended to encounter more unexpected events that could potentially cause them to alter their lesson.

Several of the most common types of events that resulted in alterations to instructional plans are similar to the pivotal teaching moments found in Stockero and Van Zoest (2013). Events classified as *observing evidence of student thinking* in this study often consisted of teachers observing students making statistical errors, students expressing confusion or struggling on a problem, or students verbalizing their attempts to make sense of a statistical concept. These instances could correspond to the classifications of *incorrect mathematics, mathematical confusion,* or *sense-making*, respectively, by Stockero and Van Zoest (2013). *Questions asked by students* that impacted lesson alterations in these statistics classrooms often consisted of students asking about topics that were related to the statistical content intended to be covered in a lesson, but outside the scope of the intended learning goal. These questions would correspond to the *extending* classification of Stockero and Van Zoest (2013). Teachers also made alterations in this study, however, to questions asked by students that would not appear to be categorized as *extending* the statistical content, such as in the case of Ms. Baker, when a student asked about the format of the AP Exam, which is highly related to the context of her classroom and the realities of pressure students feel in an advanced placement class with an exam that impacts ability to earn college credits.

Events classified as *students presenting a solution* in this study could correspond to Stockero and Van Zoest’s (2013) classifications of *incorrect mathematics* or *mathematical contradiction,* depending on the content of the solution presented. However, teachers in this study often altered their instructional plans based on unanticipated correct or novel solutions, as well. Teachers in this study also had to respond to other types of events that appear to not be included in the classifications of pivotal teaching moments by Stockero and Van Zoest (2013). In particular, *students' work pace* and *being short on time*, though not directly related to the content being taught, were common events that still required in-the-moment decisions often resulting in alterations to instructional plans that could have significantly impacted the learning of students.

Teachers in this study responded to events in a variety of ways, not all of which were alterations to their instructional plans. Some of the response types in Stockero and Van Zoest (2013), such as *acknowledges but continues as planned*, *ignores or dismisses*, or *pursues student thinking* each occurred numerous times in the observed statistics lessons, but typically did not require alterations to instructional plans, so are outside the score of this study. The other two response types in Stockero and Van Zoest (2013), on the other hand--*extends mathematics and/or makes connections* and *emphasizes meaning of the mathematics*--could often be used to describe the alterations participants made to discussions and lectures. For example, the teacher who modified a whole-class discussion to include a conversation about the calculation of a correlation coefficient would be extending and making connections. In the instance where a student blindly applied formulas to calculate the mean and standard deviation of a random variable, the teacher altered the discussion to emphasize the meaning or definition of a binomial variable and how that differed from other types of random variables.

Participants also responded in additional ways, however, that do not appear to be captured by the response types in Stockero and Van Zoest (2013). For example, portions of lectures or discussions were sometimes removed or shortened, either due to time constraints or because students expressed higher levels of understanding than a teacher had anticipated. In addition, participants often made alterations to the activities students were working on, in ways often not captured in the response types of Stockero and Van Zoest (2013). Students were regrouped or instructed to share with classmates, the time given to students to complete activities were shortened or lengthened, the order of activities was altered, problems were added or removed from activities, data collection activities were turned into data simulation activities, and so on.

It was not just the type and quantity of unexpected events that determined if and how a teacher altered a lesson. Each of the seven teachers expressed certain beliefs and knowledge that impacted the decisions they made in the classroom. The specific beliefs and knowledge that they used to respond to a given event, however, varied from one event to the next. As reflected in Schoenfeld's (1998) model of *teaching-in-context*, each event seemed to trigger certain beliefs and knowledge, causing them to come to the forefront, impacting the resulting decision. Similar events caused different beliefs and knowledge to come to the forefront for different teachers. As is seen in the case of Ms. Baker and Mr. Dennis, this meant that similar types of events were likely to lead to different types of alterations for different teachers. It is not known whether a teacher's response was wholly a function of their set of beliefs and knowledge—that is, whether or not two teachers with identical beliefs and knowledge would have the same subset of beliefs and knowledge triggered by an event and then make the same decision as to how to respond. This study did not attempt to identify the entire set of beliefs or knowledge held by a teacher, but rather only those that impacted the decisions they made in response to classroom events. Of course, a teacher's knowledge and beliefs are almost certainly unique, so this question will remain difficult to answer. At the very least, this study suggests that a teacher's beliefs and knowledge have a significant impact on the decisions they make to alter a lesson.

When discussing reasons for making alterations to a lesson, the one type of knowledge that was mentioned most often by teachers was the knowledge of particular students. When an event occurred in the classroom, this was the knowledge that was most often brought to the forefront. Sometimes, this was knowledge of a student's past successes in mathematics courses. Other times, it was knowledge of a student's affect or personality. As seen in the cases of Ms. Baker and Mr. Dennis, a more intimate knowledge of students may impact how and when a teacher makes an alteration to a lesson. This seems to reflect the findings of Sztajn (2003), who found that knowledge and beliefs about students in the classroom often played a larger role in some teachers' decision-making than any other types of knowledge or beliefs.

It is noteworthy that only three teachers in this study explicitly suggested that their knowledge of statistics (or lack thereof) had an impact on any alterations they made to a lesson, and did so rather infrequently. Previous research (e.g., Hill & Charalambous, 2008) has found that a teacher's mathematical knowledge has a significant impact on instruction. While these results may appear to be in conflict, some things should be noted. First, in the present study, no attempt was made to measure teachers' statistical knowledge—results were based on teachers' verbal explanations of their decision-making process. Thus, differing levels of knowledge *may* have had an impact on teachers' decision-making, but the teachers may have been unaware that they were drawing upon their knowledge of statistics, or did not think it was worth explicitly mentioning. Indeed, several instances were observed in the classroom in which participants seemed to be drawing on their statistical knowledge, despite not indicating in the interviews that this factored into their decision to alter a lesson. For example, during one observed lesson, a student used a normal approximation to the binomial distribution to solve a problem. This was not part of the curriculum for the course, and not something that the teacher had planned to teach students to do. However, this teacher was able to modify the discussion in real time to discuss this method and why and when it works. This is something the teacher would likely have not been able to do were it not for his deep knowledge of statistics. Another possible explanation for the apparent discrepancy concerning the impact of teachers' content knowledge on instruction is that studies that investigate the impact of knowledge on teacher practice rarely differentiate between *intended curriculum* and *enacted curriculum*. It may be the case that the teachers in this study drew heavily on their statistics knowledge when planning a lesson (their intended curriculum), but when events occurred that could potentially result in alterations to the lesson, other knowledge and beliefs were more likely to come to the forefront.

Whether or not participants were actively or consciously drawing on their knowledge of statistics, some observed events and responses seemed to be particular to statistics classrooms. For example, teachers had to make real-time decisions in response to questions about how to collect data, to aberrations in simulated data, and to students creating different data representations than anticipated. Occasionally, these events resulted in complications for a teacher, but more often, these events led to rich discussions with students concerning statistical concepts such as measurement, variability, and data representations. These quick decisions by teachers would seem to require statistical knowledge for teaching (Groth, 2007), despite most teachers rarely explicitly discussing this knowledge when describing their decision-making. Specifically assessing teachers' statistical knowledge for teaching (which this study did not attempt to do) may shine more light on the relationship between a teacher's statistical knowledge for teaching and their decision-making in the statistics classroom.

Several different types of beliefs were cited by teachers when explaining reasons for making alterations to a lesson. These included beliefs about students' affective needs, beliefs about statistics, and beliefs about the role of a teacher. The most common type of beliefs that teachers drew upon when making alterations was beliefs about learning. Each teacher had their own beliefs about what types of activities best promoted student learning, about what types of supports students needed in order to learn, and about how students acquire or build new knowledge. Regardless of the type of event that necessitated a response, beliefs about learning were often in the forefront of a teacher's mind, impacting the real-time decisions they made and any alterations to a lesson. These findings reinforce those of Eichler (2008) and Eichler (2011), who found that statistics teachers' beliefs have a significant impact on their instruction.

Oftentimes, it was difficult to distinguish between knowledge and belief. For example, one teacher made a claim about the relative confidence level of two students. Another made a claim about how a student would respond if she did not receive assistance from the teacher. Yet another made a claim about what statistics looked like "in the real world." Each of these claims are based on a teacher's perception and interpretation, and yet teachers spoke of them as though they knew the claims to be true. However, in the end, teachers drew on their knowledge as they did on their beliefs. Perhaps any effort to differentiate the two is unnecessary. Both knowledge and beliefs are formed by teachers' experiences, and whether something is a knowable fact did not seem to affect how it impacted teachers' decision-making. It may be useful to consider both together as part of a teacher's *conceptions*, as suggested by Knuth (2002) and Sinclair et al. (2017), as it seems in the present study that "the separation [between beliefs and knowledge] is less distinct in reality than it is in theory" (Knuth, 2002, p. 85).

On occasion, an alteration to a lesson plan occurred simply by a teacher reflecting on their beliefs about learning without any identified event occurring that caused it. Teachers described these decisions as "seeing an opportunity" or having a "realization" that an alteration might benefit students. Whether there was some unidentified event that prompted these sudden realizations or whether the realization was truly spontaneous is unknown. It is also possible that other teachers had similar realizations, but for whatever reason, did not act upon them. More research is needed to investigate the cause of these types of alterations.

Teacher preparation programs often have a large focus on designing and planning classroom lessons. However, a teacher's planned curriculum is often different from their enacted curriculum (Stein et al., 2007), as illustrated in this study. Teachers are prepared to choose or construct activities, to anticipate student solutions, and to craft discussions around these hypothetical solutions. Participants in this study, however, often saw students produce unanticipated solutions, struggle in unexpected places, inquire about topics outside the learning goal, and progress through tasks slower or faster than anticipated. All of these events required in-the-moment decisions from a teacher, many of which were likely to have important impacts on students' understanding.

Real-time decision-making is a skill that is difficult to teach and even more difficult to master--it is however a *skill*, one that can be learned and developed over time (Jacobs et al., 2010). The most straightforward way for preservice teachers to practice real-time decision-making is during student teaching, when they are placed into real classrooms with real students. However, the demands on students during student teaching are already large, and developing skills prior to being in a classroom would be preferable. Textual or video cases of teachers in the classroom have long been used in preservice teacher education, and allow preservice teachers to analyze the decision-making of other teachers or to consider how they themselves might make decisions in the same situations (Maher, 2008). A more active approach is that of *rehearsals* (Lampert et al., 2013), in which preservice teachers teach activities to teacher educators and other preservice teachers who simulate the range of students that might be found in a classroom. This allows the preservice teacher to develop decision-making skills in a more controlled environment while also receiving guidance from a teacher educator along the way. Recent innovations in technology even allow for computer-simulated virtual learning environments, in which preservice teachers can interact with virtual representations of students in real-time (e.g., Peterson-Ahmad et al., 2018).

As some types of events occurred more frequently than others in this study, it may be productive for teacher educators to focus on these events during the above activities. Asking preservice teachers to respond to student thinking that they had *not* anticipated would appear to be a vital component of activities such as rehearsals. A range of solution strategies (both valid and erroneous) could be generated by other preservice teachers, or even K-12 school students that were asked to work through a problem ahead of time. Preservice teachers could then be asked to respond to the observed student thinking, both to support the student that had generated the thinking, as well as by deciding if and how to incorporate the student thinking into a subsequent whole-class discussion. Rather than being flexible about the amount of time given to preservice teachers during activities such as rehearsals, they could be asked to carefully manage a prescribed amount of time and have to adjust their instruction when time got short or when their planned activities took a different amount of time than anticipated. For statistics teacher education specifically, preservice teachers could be asked to respond quickly to real results of simulations, to issues concerning measurement and data collection, and to novel or unexpected data representations produced by students. By being given opportunities to practice and develop their decision-making skills, teachers should be more prepared for the realities that they may face in the classroom.

CHAPTER 6: DISCUSSION

This chapter presents a summary of the findings of this study concerning the two research questions. A discussion of these findings follows, including implications for teacher education and development and suggestions for future research.

Summary of Research Question One

For Teachers of Secondary Statistics, how are Instructional Plans for Teaching Statistics Lessons Created?

The manuscript presented in Chapter 4 examined the decision-making process behind the creation of secondary statistics teachers' initial instructional plans. Findings from the manuscript in Chapter 5 related to Research Question One are broken up into two sub-questions, as follows.

Research Question 1a: What Knowledge and Beliefs do Teachers Draw Upon When Creating Instructional Plans?

Five different areas of beliefs were identified that impacted decisions teachers made when planning statistics lessons: beliefs about learning, about teaching, about students and their needs, about statistics, and about technology.

Influential **beliefs about learning** (Figure 5) included beliefs about the role of discourse in learning, about support that students need as they learn, and about the types of activities that best support students' learning. These beliefs impacted how participants structured their classroom, norms they established for interactions, as well as larger curricular structures and goals of a lesson.

Influential **beliefs about teaching** (Figure 6) included beliefs about assessing students' knowledge and about the role of the teacher in the classroom. These beliefs impacted the types of assessments given to students, how teachers interacted with students, and the structure of assignments and discussions.

Influential **beliefs about students and their needs** (Figure 7) included beliefs about student affect and emotional needs, about their levels of engagement, and about students' preferences for instruction. These beliefs impacted the style of instruction, the nature of whole-class discussions, the types of activities and assessments assigned, and the ways in which participants interacted with students.

Influential **beliefs about statistics** (Figure 8) were generally expressed by only three of the seven participants. These beliefs included that statistics involves interpretation and communication rather than procedures, that statistics involves playing with data, and that statistics involves a lot of vocabulary. These beliefs impacted the type of activities and assessments assigned and the nature of small-group and whole-class discussions.

Influential **beliefs about technology** (Figure 9) included beliefs that the indiscriminate use of technology could negatively impact student understanding, that technology other than graphing calculators can help promote student understanding, and that understanding what technology was doing was important for student understanding. These beliefs impacted which activities included technology use and when technology-enhanced activities appeared in a unit.

Participants' knowledge also impacted the instructional plans that they created for statistics lessons. Knowledge of statistics and of teaching statistics were the most common areas of knowledge that impacted participants' instructional plans. This knowledge came from a variety of sources and past experiences.

Participants' **knowledge of teaching statistics** (Figure 10) included knowledge of tasks, of technology for teaching statistics, of pedagogical strategies for teaching statistics, and of standards and assessment expectations. This knowledge (or lack thereof) impacted the tasks and assessments used, the emphasis placed on certain topics, the technology used, and the style of instruction. This knowledge tended to come from professional development experiences such as workshops and conferences, from personal and work colleagues, from preservice teacher education, and from web searches. Participants' **knowledge of statistics** (or lack thereof) impacted if and how certain skills or topics were taught (e.g. whether students were expected to know and use formulas or use calculator functions). This knowledge tended to come from professional development opportunities and from university coursework (both within and outside of education courses).

Research Question 1b: What Current Contexts Seem to Influence Which of These Knowledge or Beliefs Take Priority?

A number of different current contexts shaped how participants' beliefs and knowledge were put into practice when planning statistics instruction. Contextual factors often made it difficult for participants to put their beliefs and knowledge into practice. This resulted in instruction that participants felt was less than ideal, or resulted in participants having to struggle to find ways to overcome obstacles resulting from these contexts.

Most participants believed in the value of student-centered activities, investigative tasks, and data-rich activities, and wanted to implement more of them. This was often inhibited by **large class sizes**, a **lack of physical space**, **limited planning time**, **limited instructional time**, **short class periods**, **limited access to instructional materials**, and the **presence of external assessments** such as the AP Statistics Exam. Similarly, some teachers believed that non-calculator technologies could be beneficial to students' learning and wished to use more technology, but were also inhibited by many the same contexts, as well as having **limited access**

to technology. On the other hand, because of limited instructional time and short class periods, some other teachers resorted to having their students use calculators, despite believing that having students solve problems without calculators may be more beneficial for their learning.

Short class periods and limited instructional time also impacted whether participants scheduled a review of previous material at the start of class, a summative discussion at the end of class, or time for teachers to address individual students' concerns, which many teachers believed were valuable. A lack of physical space, large class sizes, and **teaching multiple classes in the same classroom** caused participants to arrange classrooms in ways they believed were less than ideal. The presence of external assessments, in addition to the impacts mentioned above, also resulted in teachers placing a higher priority on procedural fluency than they would have otherwise done. External assessments, along with limited instructional time, limited planning time, and a **lack of access to instructional materials**, resulted in some teachers using more direct instruction than they believed would be ideal for students' learning.

Summary of Research Question Two

How are Secondary Statistics Teachers' Instructional Plans for Statistics Lessons Modified After Creation?

Chapter 5 focused on the real-time decision-making processes of participants as they implemented statistics lessons. In particular, it examined when, how, and why teachers made alterations to their instructional plans. Findings related to Research Question Two are broken up into two sub-questions, as follows.

Research Question 2a: What Events Cause Teachers to Modify Their Instructional Plans for Statistics Lessons, Either During or After Implementation?

Alterations to participants' instructional plans generally took one of three forms: alterations to student learning activities, alterations to lectures, and alterations to whole-class discussions. Various types of events resulted in alterations to different parts of lessons.

Observing evidence of student thinking, **students' work pace**, **minimal remaining class time**, and **reflecting on core learning beliefs** caused participants to alter student learning activities. Sometimes, these alterations consisted of cutting an activity or postponing it until a future lesson. Other times, participants made alterations to activities themselves by adding, removing, or changing the order of questions, or by changing parameters of the activity such as the grouping of students, the use of technology, or the instructions given to students.

Observing evidence of student thinking and **questions asked by students** were the most common events that resulted in alterations to lectures. Alterations most often consisted of placing increased emphasis on a topic or adding additional topics in order to address students' concerns or misconceptions. Alterations also sometimes consisted of shortening lectures if the student thinking that participants observed was more satisfactory than expected.

Observing evidence of student thinking, **students' presentation of a solution**, questions asked by students, and students' work pace resulted in alterations to the whole-class discussion that participants had planned. The topic of discussion was often altered to center around a student-generated idea or solution, or to address part of an activity that some students were not able to complete.

Research Question 2b: Which Knowledge and Beliefs Take Priority due to These Events?

Different teachers responded differently to similar events. When asked about the reasons behind decisions to respond to events in certain ways, participants often referenced beliefs and knowledge that were at the forefront, influencing their decisions. These beliefs and knowledge seem to underlie many of the decisions they made.

General **beliefs about learning** were the most common types of beliefs that participants explicitly mentioned impacted their real-time decision-making. These often impelled participants to alter their instructional plans to generate discourse and collaboration, to ensure students heard a variety of perspectives or solutions, or to ensure students had experiences beneficial to their learning. Specific **hypotheses about how students learn statistics** also often drove decisions participants made. These impelled participants to modify the instructions for a task, to choose particular student solutions to share, or to remove or add a topic to a whole-class discussion.

Beliefs about teaching that participants indicated had an influence on their real-time decision-making include beliefs about assessing students and about the role of a teacher. These beliefs impelled participants to modify an activity or shorten a lecture in favor of a student-centered activity so that the teacher could better assess student thinking, or to modify whole-class discussions to increase transparency or meet a perceived obligation to answer students' questions.

Participants indicated that **beliefs about students and their needs** had an impact on their real-time decision-making. In particular, these beliefs concerned students' self-efficacy and confidence, their levels of stress or comfort, their engagement, and their need to have their voices and perspectives validated. These beliefs impelled teachers to modify a lecture or discussion so that students' concerns could be addressed, to skip or shorten less engaging parts of a lesson, or to ask students to share their solutions and help drive discussions.

Participants' **knowledge of students** in their classroom was the knowledge that was most commonly indicated by participants as being influential on their real-time decision-making. This knowledge impacted how participants chose students' work to drive a discussion, how participants responded to questions during a discussion, and the type and amount of assistance given during a student activity.

Two participants expressly indicated that their **knowledge of statistics** impacted their real-time decision-making. This knowledge impacted the ways in which these two participants responded to students' questions and in how they orchestrated discussions around students' work.

Implications

In planning their statistics instruction, all participants drew upon both their knowledge and their beliefs. Thus, if the goal is to impact the statistics instruction that students experience in classrooms, it is important that teachers engage in preservice teacher education and inservice professional development that focuses on *both* knowledge and beliefs. Most of the teachers in this study seemed to have strong content knowledge of statistics (though their content knowledge was not directly assessed), and some had backgrounds (e.g. studies in actuarial sciences, double-majoring in statistics) that would also indicate strong content knowledge. However, teachers much more often expressed that they drew upon their *pedagogical* content knowledge of statistics rather than their knowledge of statistics itself when they were planning statistics instruction. Teachers used their pedagogical content knowledge when choosing tasks and assessments, when choosing when and how to use technology, when deciding which topics to place the greatest emphasis on, and when revising instruction from one implementation to the next.

Pedagogical content knowledge also seemed to be an important factor for responding to in-classroom events. There were several events that occurred that were specific to statistics classrooms that required real-time decisions from the teacher. Students had questions about data collection methods, simulated data did not clearly illustrate statistical concepts, students created data representations other than ones that the teacher intended, and students struggled with the lack of mathematical explanations given for statistical concepts. Content knowledge alone would likely be insufficient for responding to these events. For example, content knowledge may allow a teacher to identify the differences between a dot plot and a histogram. Pedagogical content knowledge, on the other hand, may allow the teacher to know when and why students might choose one representation over another. It may allow the teacher to know what students typically attend to in each of those representations, or how the choice of representation may impact students' learning, or how to help develop students' knowledge of the differences in these representations. Thus, it is imperative that statistics teachers receive preparation on how to *teach* statistics, not just instruction on statistical content.

However, several of the teachers in this study reported receiving little preservice preparation for teaching statistics, and some (particularly the two non-AP teachers) also received little inservice preparation. Of those that did, most indicated that the preparation that they received was very beneficial as they planned and implemented statistics lessons. Although some participants had backgrounds in statistics-related areas and had long had a desire to teach statistics, others started teaching statistics unexpectedly. With the rise in the number of students taking high school statistics courses (or receiving statistics instruction in integrated math courses), it is more important than ever that *all* teachers, especially at the secondary level, are receiving preparation for teaching statistics. Materials have been developed with preservice statistics teacher preparation in mind that could serve this cause, including Enhancing Statistics Teacher Education with E-Modules (ESTEEM, 2020) and Mathematics Of Doing Understanding Learning and Educating for Secondary Schools (MODULE(S^2), 2020).

During implementation, knowledge about particular students was often impactful, sometimes even more so than content knowledge or pedagogical content knowledge. Teachers often leveraged what they knew about students' personalities, preferences, and tendencies when deciding how to respond to in-classroom events. Teachers used this knowledge, for example, when choosing how and when to select students to share their ideas publicly, when choosing the types and amount of assistance to provide to students, and when choosing how to supplement or modify an activity for particular students or groups. Teachers need to be aware that there is no "one size fits all" instruction and be prepared to adapt their instruction based on the students in their classrooms.

Teacher beliefs were found to have a large impact on statistics teachers' instruction, both during planning and during implementation. Many of these beliefs were not particular to statistics classrooms. Teachers drew upon beliefs about learning, teaching, students' needs, and technology in various ways when planning and implementing statistics instruction. These are beliefs that can and should be supported and developed across various phases of teachers' preservice and inservice preparation. However, teachers also drew upon their conceptions about how students learn specific statistical content, as well as their beliefs about statistics in general. Beliefs about statistics are often shaped by pre-tertiary educational experiences or by experiences outside the classroom (Beswick, 2007), and are often of a nature that may not support effective teaching of statistics (Begg & Edwards, 1999; Chich & Pierce, 2008). Careful and targeted instruction on teaching statistics can help shift these beliefs toward those which are more likely to support effective statistics instruction (Lee et al., 2017), further evidence that more teachers need to be receiving this type of instruction.

One aspect that several participants found very beneficial to their planning of statistics lessons was conversing with other teachers of statistics. The two participants that had other teachers of statistics in their school reported that they conversed with those teachers regularly, and that those conversations helped the teachers design activities and keep on pace with the curriculum. Many of the other participants, being the only statistics teacher in their school, relied on online communities, personal acquaintances, relationships formed at conferences, and other statistics teachers at nearby schools or universities in order to have someone to discuss the teaching of statistics with. Making it easier for statistics teachers to connect with each other outside of one-time professional development opportunities would provide more statistics teachers with opportunities to discuss the teaching of statistics with colleagues. Localized efforts similar to Math Circles (American Institute of Mathematics, 2020) could put statistics teachers in regular contact with other nearby statistics teachers.

All participants had to continually grapple with contextual factors when planning their statistics lessons, ranging from limited planning time, large class sizes, short class periods, external assessments, and limited access to technology. Even though these teachers may have had the knowledge needed to effectively teach statistics under ideal circumstances, these factors often resulted in the teachers having to either compromise their beliefs or be inventive in the ways in which they dealt with these factors. However, preservice teachers are often tasked with lesson planning without regard to these contextual factors. They are often given weeks to plan a lesson, given access to technology that may not be accessible in many schools, and are given the freedom to choose their class length and characteristics of students. These utopian conditions unfortunately do not match the reality that these teachers will likely face in real classrooms. Making preservice teachers aware of these contextual factors that they will have to contend with and giving preservice teachers practice planning in more realistic situations can better equip these teachers with the skills they need to succeed in the classroom.

In every observed lesson, participants in this study encountered unanticipated situations in the classroom that they had to respond to. Students invented novel solutions, struggled in unexpected places, completed activities more slowly than anticipated, and asked unexpected questions. Responding to these events required teachers to quickly process not only what they were observing, but also their knowledge and beliefs about teaching and learning, about statistics, and about students, as well as contextual factors such as what technology was available or their perceived obligation to prepare students for external assessments. As this was something that participants were required to do in every lesson, this type of quick decision-making should be a focus of teacher education and professional development.

Some suggestions for activities that can develop preservice (or inservice) teachers' decision-making skills include using textual or video case studies (Maher, 2008), rehearsals (Lampert et al., 2013) or virtual classrooms (Peterson-Ahmad et al., 2018). Though it is important that preservice teachers try to anticipate the range of student work, ideas, and questions that they might encounter in the classroom, it is equally important that teachers are comfortable and skilled at dealing with *unanticipated* events. Observing evidence of student thinking was the event that most commonly caused teachers in this study to alter their plans, so a wide range of solution strategies or ideas can be presented to preservice teachers during the above activities, perhaps solutions generated by real students, if possible. During rehearsals, other preservice teachers can be encouraged to ask questions or propose solutions or ideas that may be unusual. Preservice teachers should practice responding in real-time to these events, as well as making decisions about whether to incorporate these ideas into whole-class discussions.

Findings of this study suggest additional steps that would be useful specifically for preparing preservice teachers to teach statistics. During planning, teachers in this study had to grapple with whether to allow students to collect their own data or whether to use simulated or preconstructed data sets. They had to decide whether students would use technology to analyze data on their own, or whether they would watch the teacher perform the analysis. They had to choose between open-ended explorations or more procedural approaches. In the absence of any contextual factors, these decisions would have been easy ones for most teachers in this study. However, realities of the classroom such as limited instructional time, short class periods, external exams, and limited planning time often made these decisions much harder to make. Some teacher education and professional development approaches have been shown to have an impact on teachers' beliefs about teaching statistics (e.g., Harrison et al., 2020; Lee, Mojica, & Lovett, 2020), but as the present study shows, mediating factors often make it difficult for those beliefs to be put into practice. Teachers should be made aware of these factors so that they have the best chance of implementing lessons that align with their beliefs.

Discussion and Future Research Directions

In the framework that guided this study, knowledge and beliefs played a similar role in the creation and modification of teachers' instructional plans. However, when describing reasons behind decisions they made, participants often discussed these two factors differently. Particularly for real-time decisions, beliefs were much more readily mentioned than knowledge was. The difference was also illustrated by the fact that participants often referenced their knowledge in indirect ways, by discussing sources of knowledge--workshops, colleagues, university courses, etc.--rather than the knowledge itself. At times, however, it was difficult to discern whether participants were expressing their beliefs or knowledge (e.g. about teaching statistics). When knowledge *was* directly and expressly mentioned as being influential, it was often knowledge about students, rather than content or pedagogical knowledge. Clearly, many of participants' decisions *were* impacted by their content or pedagogical knowledge, as many of the instructional moves performed by participants illustrated the knowledge these teachers possessed. However, participants rarely mentioned that they were relying on this knowledge when asked to describe reasons behind these instructional moves, particularly for in-the-moment decisions. For participants, the relationship between their knowledge and their practice seemed to be less noticeable or noteworthy than the relationship between their beliefs and practice. Previous studies identified in Chapter Two have identified correlations between knowledge and practice (e.g. Askew et al., 1997; Charalambous et al., 2012), but few have examined how exactly this knowledge gets turned into practice. The present study begins to identify how teachers draw on their knowledge to make decisions, but as participants were not explicitly asked about their knowledge, opportunities to examine this relationship were limited, and inferences would have to be made in order to draw any conclusions. Future research can focus more explicitly on knowledge to learn when and how teachers draw on their knowledge to make decisions.

The framework used in this study was informed, in part, by two previous frameworks--*teaching-in-context* (Schoenfeld, 1998) and the *mathematics teaching cycle* (Simon, 1995). From this study, several observations can be made about these frameworks and their utility in examining teaching decision-making. Though teachers expressed various beliefs and knowledge throughout their interviews, each event in the classroom seemed to trigger a small subset of these knowledge and beliefs that teachers drew on when they decided how to respond. These were often beliefs or knowledge that did not have a high priority when planning the lesson, resulting in teachers making alterations to the lesson in real-time. For example, in the case of Ms. Baker, her belief that one of her roles was to prepare students for the AP Statistics Exam was low priority for her planning of a lesson. When a student asked about the format of the exam, it triggered this belief resulting in her altering the lesson to better prepare students for the exam. This is in accordance with Schoenfeld’s (1998) mechanism of beliefs and knowledge being activated by the context of the classroom. However, a choice was made in the current study to separate contextual factors (e.g., short class periods, the presence of external assessments) from immediate classroom events, both of which Schoenfeld (1998) considers *current context*. As this study examined both planning decisions and real-time decisions during implementation, it made sense to consider the impact of classroom events only for the latter type of decisions. However, this separation also turned out to be pertinent when considering only the real-time decisions of the teacher during implementation. Certain contextual factors such as the presence of external assessments or large class sizes were constant and always present for some teachers’ classes. However, only sometimes did these contextual factors have an impact on the decision-making process of teachers. In other words, though contextual factors often remained the same, whether or not they factored into teachers’ decision-making depended on the event that had just occurred.

Aspects of Simon’s (1995) *mathematics teaching cycle* were often illustrated by teachers in this study. Teachers created instructional plans--analogous to Simon’s *hypothetical learning trajectories*--based on their knowledge and beliefs. As a result of implementing these plans, these teachers made alterations to these plans, either actual changes in real-time or planned changes for the next implementation. As Simon (1995) posited, teachers in this study had varying lengths of cycles and frequencies of alterations, as can be seen when comparing the instruction of Ms. Baker and Mr. Dennis. However, some parts of this process for teachers in this study unfolded slightly differently as might be suggested by Simon's (1995) mathematics teaching cycle. Though observing and assessing student thinking was the most common event that resulted in alterations to instructional plans, teachers in this study also made alterations due to their assessments of other things in the classroom. Teachers also assessed student affect, the amount of time remaining in the class, questions from students that did not illustrate their thinking, and contextual factors in which they taught. Another aspect that seems not to be captured by Simon's (1995) model is that these alterations to instructional plans did seem to require alterations to teachers' beliefs or knowledge. Certainly, in some cases, events in a classroom caused teachers to reconsider their beliefs about how students learn or to gain new knowledge about students. More often, however, teachers' responses to events in this study were based on beliefs or knowledge that the teachers already possessed. Integrating aspects of both of these frameworks--teaching-in-context and the mathematics teaching cycle--resulted in a framework for this study (Figure 13) that appears to better capture the decision-making of teachers during both planning and implementation of a lesson.

The situated nature of this research (i.e. real instruction in real classrooms) introduced some constraints on the data that could be collected and the conclusions that could be drawn. Events that happened in classrooms, responses by participants, beliefs, knowledge, and instructional contexts were all categorized based on common elements, but each of these phenomena are unique—either to a particular lesson or to a particular teacher. For this reason, it becomes almost impossible to isolate the impact that particular events, knowledge, beliefs, and contexts had on participants' responses, and can make generalizing beyond the situations being investigated difficult. The situated nature of the research also makes real-time examination of teachers' decision-making processes nearly impossible, as teachers in a classroom are typically not able to explain in real-time why they are taking the actions that they are. Both of these issues could be somewhat addressed by a simulated classroom or in a laboratory setting. Specifically, events and contexts can be tightly controlled, and teachers are more readily able to verbalize their decision-making process. There are certain benefits, however, to situating research in real classrooms (e.g., teachers' responses to events are more likely to be authentic). Therefore, a combination of these two types of research (situated and laboratory) is needed to fully explore teachers' decision-making.

For situated studies such as this one, the use of stimulated recall appears to be an effective method for examining teacher decision-making. Teachers had little trouble recalling events that occurred in the classroom, regardless of whether video recordings were shown of the events during post-observation interviews. Teachers seemed to have no reservations about discussing their decision-making, and in most cases, could readily offer explanations for the decisions that they made. The fact that all the post-observation interviews took place the same day as the observations, and often immediately afterward, likely helped teachers' recollection of the events and their decisions about how to respond. Of course, it is difficult to know how accurate teachers' descriptions of their decision-making process was. Outside of classroom settings, the use of stimulated recall has resulted in interviewees describing their decision-making in less detail (Kuusela & Paul, 2000) or as more logical and deliberate (Lyle, 2003) than real-time verbalizations of decision-making. There does not appear to be any similar studies assessing the accuracy of responses to stimulated recall in a classroom setting. This is another area where a combination of a simulated classroom or laboratory environment along with actual classrooms would likely be necessary, given that teachers are typically not in a position to describe their decision-making process in real-time when in a classroom setting.

Several aspects of teachers' decision-making were highlighted in this study that are worth exploring in further detail. When planning a statistics lesson, some teachers' beliefs appeared similar to those of other teachers, but were manifested in different ways. For example, many teachers in this study believed that it was important for students to hear multiple perspectives and ideas. This belief caused some teachers to plan whole-class discussions that consisted of multiple students presenting their work to the rest of the class. For another teacher, this belief resulted in regrouping students on a daily or near-daily basis. In some cases, these different manifestations could be attributed to contextual factors--one teacher would have placed her students in larger groups to hear multiple perspectives, but the fact that she taught other classes in which she wanted students in pairs made it more difficult to place students in larger groups in her statistics course. In other cases, teachers held other competing beliefs--one teacher thought it was important for his students to hear multiple perspectives but also believed that it was hard for students to demonstrate understanding of a concept in larger groups, so he restricted collaboration to pairs for graded assessments. In several cases, however, there was no clear explanation for the differences in how beliefs were manifested. Most teachers expressed the above belief about the benefit of hearing multiple perspectives, but only one teacher regularly regrouped students. Perhaps there were other contextual factors or competing beliefs that were not discussed in the interviews. Or perhaps teachers' beliefs about hearing multiple perspectives differed in some way, despite teachers' verbalizations of those beliefs being similar. A closer examination of exactly *how* teachers' beliefs and knowledge get turned into practice--rather than simply *which* knowledge and beliefs do so--would further illuminate the decision-making process of teachers, and why teachers with similar beliefs and knowledge in similar contexts often make instructional decisions that differ.

If we consider decision-making as a skill (Jacobs et al., 2010), one possible explanation for these differences is that teachers have different skill sets when it comes to decision-making. Just as two students may solve an algebra equation differently due to different learned techniques or different fluency with different procedures, two teachers may make decisions differently based on the development of their decision-making skills. Though there have been several suggestions on how to develop teachers' decision-making skills via methods such as rehearsals or virtual or simulated classrooms (Lampert et al., 2013; Peterson-Ahmad et al., 2018), there has been little research investigating the impact that these strategies have on actual classroom practice. The current study did not attempt to assess the quality of the decisions that teachers made, though other studies have done so (e.g. Stockero and Van Zoest, 2013). It may be worthwhile to adopt similar analyses to investigate whether and how teachers' decision-making skills can be improved by some of the methods mentioned above (e.g. rehearsals, simulated classrooms).

Several issues related to teacher decision-making were left unexplored by this study. One of the most important is the impact that teachers' decisions may have had on student learning or other outcomes such as student affect. It is worth exploring whether certain types of responses tended to result in certain outcomes or improvements in students' learning. Some teachers in this study tended to make more frequent alterations to their instructional plans. In most cases, teachers felt these alterations would be beneficial for students' learning, but were they correct?

For teachers who made fewer alterations, was this because their initial instructional plans needed fewer adjustments to be exemplary, or were these teachers missing opportunities to improve their instruction? Some teachers seemed to expect and even welcome having to make adjustments and starting with less-defined plans, but other teachers had more definitive and structured plans. Were the latter group of teachers better at anticipating what would happen in the classroom, or were they less adaptive to events that might occur? Knowing the answers to these questions and the impact that these alterations have on student learning are important steps toward transforming this research into actionable plans for teacher professional development.

REFERENCES

Abd-El-Khalick, F., Bell, R. L., & Lederman, N. G. (1998). The nature of science and instructional practice: Making the unnatural natural. *Science Education*, *82*(4), 417-436.

Amador, J., & Lamberg, T. (2013). Learning trajectories, lesson planning, affordances, and constraints in the design and enactment of mathematics teaching. *Mathematical Thinking and Learning*, *15*(2), 146-170.

American Institute of Mathematics (2020). *Math Teachers' Circle Network: Building Mathematical Communities.* https://www.mathteacherscircle.org/

American Statistical Association (2019). ASA newsroom. https://www.amstat.org/ASA/Newsroom.aspx

Arnold, P., Confrey, J., Jones, R. S., Lee, H. S., & Pfannkuch, M. (2018). Statistics learning trajectories. In D. Ben-Zvi, K. Makar, & J. Garfield (Eds.), *International handbook of research in statistics education* (pp.295-326). Springer.

Artzt, A. F. & Armour-Thomas, E. (1998). Mathematics teaching as problem solving: A framework for studying teacher metacognition underlying instructional practice in mathematics, *Instructional Science*, *26*, 5–25.

Askew, M., Brown, M., Rhodes, V., Wiliam, D., & Johnson, D. (1997). *Effective teachers of numeracy: Report of a study carried out for the Teacher Training Agency*. King's College, University of London.

Aubrey, C. (1997) *Mathematics teaching in the early years: An investigation into teachers' subject knowledge*. Falmer Press.

Australian Curriculum, Assessment and Reporting Authority. (2011). Australian curriculum: Mathematics. https://www.australiancurriculum.edu.au/

Baker, S., Gersten, R., Dimino, J. A., & Griffiths, R. (2004). The sustained use of research-based instructional practice: A case study of peer-assisted learning strategies in mathematics. *Remedial and Special Education*, *25*(1), 5-24.

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

Banilower, E. R., Smith, P. S., Malzahn, K. A., Plumley, C. L., Gordon, E. M., & Hayes, M. L. (2018). *Report of the 2018 NSSME+*. Horizon Research, Inc.

Begg, A., & Edwards, R. (1999). Teachers' ideas about teaching statistics. In *Proceedings of the 1999 combined conference of the Australian Association for Research in Education and the New Zealand Association for Research in Education*. Australian Association for Research in Education & New Zealand Association for Research in Education.

Ben-Zvi, D. (2004). Reasoning about variability in comparing distributions. *Statistics Education Research Journal*, *3*(2), 42-63.

Ben-Zvi, D., Gravemeijer, K., & Ainley, J. (2018). Design of statistics learning environments. In D. Ben-Zvi, K. Makar, & J. Garfield (Eds.), *International handbook of research in statistics education* (pp. 473-502). Springer.

Beswick, K. (2007). Teachers' beliefs that matter in secondary mathematics classrooms. *Educational Studies in Mathematics*, *65*, 95-120.

Biehler, R., Frischemeier, D., & Podworny, S. (2018). Elementary preservice teachers' reasoning about statistical modeling in a civic statistics context. *ZDM Mathematics Education*, *50*(7), 1237-1251.

Blanton, M. L., & Kaput, J. J. (2005). Characterizing a classroom practice that promotes algebraic reasoning. *Journal for Research in Mathematics Education*, *36*, 412-446.

Borko, H., Stecher, B. M., Alonzo, A. C., Moncure, S., & McClam, S. (2005). Artifact packages for characterizing classroom practice: A pilot study. *Educational Assessment*, *10*(2), 73-104.

Burgess, T. (2002). Investigating the "data sense" of preservice teachers. In B. Phillips (Ed.), *Proceedings of the Sixth International Conference on Teaching Statistics, Cape Town.* International Statistical Institute.

Burgess, T. (2008). Teacher knowledge of and for statistical investigations. In C. Batanero, G. Burrill, & C. Reading (Eds.), *Joint ICMI/IASE Study: Teaching Statistics in School Mathematics. Challenges for Teaching and Teacher Education. Proceedings of the ICMI Study, 18*.

Cai, J., & Gorowara, C. C. (2002). Teachers' conceptions and constructions of pedagogical representations in teaching arithmetic average. In B. Phillips (Ed.), *Proceedings of the Sixth International Conference on Teaching of Statistics.* International Statistical Institute.

Calderhead, J. (1981). Stimulated recall: A method for research on teaching. *British Journal of Educational Psychology*, *51*(2), 211-217.

Canada, D. (2006). Elementary pre-service teachers' conceptions of variation in a probability context. *Statistics Education Research Journal*, *5*(1), 36-63.

Casey, S., Hudson, R., Harrison, T., Barker, H., & Draper, J. (2020). Preservice teachers' design of technology-enhanced statistical tasks. *Contemporary Issues in Technology and Teacher Education (CITE Journal)*, *20*(2).

Casey, S. A., & Wasserman, N. H. (2015). Teachers' knowledge about informal line of best fit. *Statistics Education Research Journal*, *14*(1), 8–35.

Cayton, C., Hollebrands, K., Okumuş, S., & Boehm, E. (2017). Pivotal teaching moments in technology-intensive secondary geometry classrooms. *Journal of Mathematics Teacher Education*, *20*(1), 75-100.

Chance, B. L. (2002). Components of statistical thinking and implications for instruction and assessment. *Journal of Statistics Education*, *10*(3).

Chance, B., Garfield, J., Ben-Zvi, D., & Medina, E. (2007). The role of technology in improving student learning of statistics. *Technology Innovations in Statistics Education Journal 1*(1)

Charalambous, C. Y., Hill, H. C., & Mitchell, R. N. (2012). Two negatives don't always make a positive: Exploring how limitations in teacher knowledge and the curriculum contribute to instructional quality. *Journal of Curriculum Studies*, *44*(4), 489-513.

Chen, C. H. (2008). Why do teachers not practice what they believe regarding technology integration? *The Journal of Educational Research*, *102*(1), 65-75.

Chidziva, J. (2017). *Peer observation on the pedagogical content knowledge of grade 11 novice teachers of statistics in a circuit* (Doctoral dissertation).

Chick, H. L., & Pierce, R. (2008). Teaching statistics at the primary school level: Beliefs, affordances and pedagogical content knowledge. In *ICMI study 18 and 2008 IASE round table conference* (Vol. 1, pp. 1-6).

Clements, D. H., & Sarama, J. (2004). Learning trajectories in mathematics education. *Mathematical Thinking and Learning*, *6*(2), 81-89.

Cobb, G. W., & Moore, D. S. (1997). Mathematics, statistics, and teaching. *The American Mathematical Monthly*, *104*(9), 801-823.

Cobb, P., McClain, K., & Gravemeijer, K. (2003). Learning about statistical covariation. *Cognition and Instruction, 21*(1), 1–78.

Cohen, D. K., & Ball, D. L. (1990). Relations between policy and practice: A commentary. *Educational Evaluation and Policy Analysis*, *12*(3), 331-338.

College Board (2008). Program summary report 2008. https://research.collegeboard.org/programs/ap/data/archived/2008

College Board (2019). Program summary report 2019. https://research.collegeboard.org/programs/ap/data/participation/ap-2019

College Board (2020). *AP Statistics*. AP Central. https://apcentral.collegeboard.org/courses/ap-statistics

Copur-Gencturk, Y., Hug, B., & Lubienski, S. T. (2014). The effects of a master's program on teachers' science instruction: Results from classroom observations, teacher reports, and student surveys. *Journal of Research in Science Teaching*, *51*(2), 219-249.

Creswell, J. W. (2013). *Qualitative inquiry and research design: Choosing among five approaches* (3rd ed.). Sage.

Daro, P., Mosher, F., & Corcoran, T. (2011). *Learning trajectories in mathematics: A foundation for standards, curriculum, assessment, and instruction*. Consortium for Policy Research in Education.

da Ponte, J. P. (2011). Preparing teachers to meet the challenges of statistics education. In *Teaching statistics in school mathematics: Challenges for teaching and teacher education* (pp. 299-309). Springer.

da Silva, C. B.& Countinho, C. (2008). Reasoning about variation of a univariate distribution: A study with secondary mathematics teachers. 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.

delMas, R. (2004). A comparison of mathematical and statistical reasoning. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning, and thinking* (pp. 79–95). Kluwer Academic Publishers.

Desimone, L., Hochberg, E. D., & McMaken, J. (2016). Teacher knowledge and instructional quality of beginning teachers: Growth and linkages. *Teachers College Record*, *118*(5).

de Vetten, A., Schoonenboom, J., Keijzer, R., & van Oers, B. (2018). Pre-service primary school teachers' knowledge of informal statistical inference. J*ournal of Mathematics Teacher Education, 22(6), 639-661*.

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*.

Eichler, A. (2011). Statistics teachers and classroom practices In C. Batanero, G. Burrill, C. Reading (Eds.), *Teaching statistics in school mathematics: Challenges for teaching and teacher education* (pp. 175-186). Springer.

Eichler, A., & Erens, R. (2015). Domain-specific belief systems of secondary mathematics teachers. In B. Pepin, B. Roesken-Winter (Eds.), *From beliefs to dynamic affect systems in mathematics education,* (pp. 179-200). Springer.

Engel, J., & Sedlmeier, P. (2011). Correlation and regression in the training of teachers. In C. Batanero, G. Burrill, C. Reading (Eds.), *Teaching statistics in school mathematics: Challenges for teaching and teacher education* (pp. 247-258). Springer.

Enhancing Statistics Teacher Education with E-Modules (2020). https://hirise.fi.ncsu.edu/projects/esteem/

Estrada, A. (2002). *Análisis de las actitudes y conocimientos estadísticos elementales en la formación del profesorado (Analysing attitudes and elementary statistical knowledge in the training of teachers)*. Unpublished Ph.D. dissertation, Universidad Autónoma de Barcelona, Spain.

Estrada, A., Batanero, C., & Lancaster, S. (2011). Teachers' attitudes towards statistics. In C. Batanero, G. Burrill, C. Reading (Eds.), *Teaching statistics in school mathematics: Challenges for teaching and teacher education* (pp. 163-174). Springer.

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

Franklin, C., Kader, G., Mewborn, D., Moreno, J., Peck, R., Perry, M., & Scheaffer, R. (2007). *Guidelines for assessment and instruction in statistics education (GAISE) report.* American Statistical Association.

GAISE College Report ASA Revision Committee. (2016). *Guidelines for Assessment and Instruction in Statistics Education College Report 2016*. http://www.amstat.org/education/gaise.

Garet, M. S., Heppen, J. B., Walters, K., Parkinson, J., Smith, T. M., Song, M., ... & Wei, T. E. (2016). *Focusing on mathematical knowledge: The impact of content-intensive teacher professional development.* National Center for Education Statistics.

Garfield, J. (Chair), Aliaga, M., Cobb, G., Cuff, C., Gould, R., Lock, R., Moore, T., Rossman, A., Stephenson, R., Utts, J., Velleman, P. & Witmer, J. (2005). *Guidelines for assessment* *and instruction in statistics education: College report.* American Statistical Association, http://www.amstat.org/Education/gaise/.

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

Graham, A. (1987). *Statistical investigations in the secondary school*. Cambridge University Press.

Grolemund, G., & Wickham, H. (2014). A cognitive interpretation of data analysis. *International Statistical Review*, *82*(2), 184-204.

Groth, R. E. (2007). Toward a conceptualization of statistical knowledge for teaching. *Journal for Research in Mathematics Education 38*(5), 427-437.

Guskey, T. R. (1986). Staff development and the process of teacher change. *Educational Researcher*, *15*(5), 5-12.

Hahn, G. J., & Doganaksoy, N. (2012). *A career in statistics: Beyond the numbers*. Wiley.

Hamilton, L. S., McCaffrey, D. F., Stecher, B. M., Klein, S. P., Robyn, A., & Bugliari, D. (2003). Studying large-scale reforms of instructional practice: An example from mathematics and science. *Educational Evaluation and Policy Analysis*, *25*(1), 1-29.

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.

Harland, J., & Kinder, K. (1997). Teachers' continuing professional development: framing a model of outcomes. *British Journal of Inservice Education*, *23*(1), 71-84.

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 preservice 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*. International Statistical Institute.

Harrison, T. H., Azmy, C. N., & Lee, H. S. (2020). Impacts of online instruction on statistics education perspectives. Submitted for publication.

Heaton, R. M., & Mickelson, W. T. (2002). The learning and teaching of statistical investigation in teaching and teacher education. *Journal of Mathematics Teacher Education, 5*(1), 35-59.

Hiebert, J., & Grouws, D. A. (2007). The effect of classroom mathematics teaching on students' learning. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 371–404). Information Age.

Hill, H. C., Blunk, M. L., Charalambous, C. Y., Lewis, J. M., Phelps, G. C., Sleep, L., & Ball, D. L. (2008). Mathematical knowledge for teaching and the mathematical quality of instruction: An exploratory study. *Cognition and Instruction*, *26*(4), 430-511.

Hill, H. C., & Charalambous, C. Y. (2012). Teaching (un)connected mathematics: Two teachers' enactment of the pizza problem. *Journal of Curriculum Studies*, *44*(4), 467-487.

Hoy, A. W., Davis, H., & Pape, S. J. (2006). Teacher knowledge and beliefs. In P. A. Alexander & P. H. Winne (Eds.), *Handbook of educational psychology* (p. 715–737). Lawrence Erlbaum Associates Publishers.

Huffman, D., Thomas, K., & Lawrenz, F. (2003). Relationship between professional development, teachers' instructional practices, and the achievement of students in science and mathematics. *School Science and Mathematics, 103*(8), 378-387.

Jacobbe, T. (2012). Elementary school teachers' understanding of the mean and median. *International Journal of Science and Mathematics Education*, *10*(5), 1143-1161.

Jacobbe, T., & Carvalho, C. (2011). Teachers' understanding of averages. In C. Batanero, G. Burrill, C. Reading (Eds.), *Teaching statistics in school mathematics: Challenges for teaching and teacher education* (pp. 199-209). Springer.

Jacobbe, T., Hartlaub, B. A., & Whitaker, D. (2013). Motivation for and benefits from attending the AP statistics reading. *Journal of Education and Training Studies*, *1*(2), 32-37.

Jacobs, V. R., Lamb, L. L., & Philipp, R. A. (2010). Professional noticing of children's mathematical thinking. *Journal for Research in Mathematics Education, 41*(2), 169-202.

Johnson, C. C. (2006). Effective professional development and change in practice: Barriers science teachers encounter and implications for reform. *School Science and Mathematics*, *106*(3), 150-161.

Justice, N., Zieffler, A., Huberty, M. D., & delMas, R. (2018). Every rose has its thorn: Secondary teachers' reasoning about statistical models. *ZDM Mathematics Education*, *50*(7), 1253-1265.

Koehler, M. S., & Grouws, D. A. (1992). Mathematics teaching practices and their effects. In Grouws, D. A. (Ed.), *Handbook of research on mathematics teaching and learning*, 15-126. Macmillan Publishing Company.

Kolb, S. M. (2012). Grounded theory and the constant comparative method: Valid research strategies for educators. *Journal of Emerging Trends in Educational Research and Policy Studies*, *3*(1), 83-86.

König, J., & Pflanzl, B. (2016). Is teacher knowledge associated with performance? On the relationship between teachers' general pedagogical knowledge and instructional quality. *European Journal of Teacher Education*, *39*(4), 419-436.

Knuth, E. J. (2002). Teachers' conceptions of proof in the context of secondary school mathematics. *Journal of Mathematics Teacher Education*, *5*(1), 61-88.

Kultusministerkonferenz. (2012). Bildungsstandards im Fach Mathematik für die allgemeine Hochschulreife. http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2012/2012_10_18-Bildungsstandards-Mathe-Abi.pdf.

Kuusela, H., & Paul, P. (2000). A comparison of concurrent verbal protocol analysis retrospective. *American Journal of Psychology*, *113*(3), 387-404.

Lampert, M. (2004). Response to Teaching Practice/ Teacher Learning Practice Group. In J. Spillane, P. Cobb, & A. Sfard (organizers). *Investigating the practice of school improvement: Theory, methodology, and relevance*.

Lampert, M., Franke, M. L., Kazemi, E., Ghousseini, H., Turrou, A. C., Beasley, H., ... & Crowe, K. (2013). Keeping it complex: Using rehearsals to support novice teacher learning of ambitious teaching. *Journal of Teacher Education*, *64*(3), 226-243.

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

Lee, H. S., & Harrison, T. R. (2020). Trends in teaching advanced placement statistics: Results from a national survey. Manuscript in preparation.

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).

Lee, H. S., Lovett, J. N., & Mojica, G. M. (2017). Characterizing impacts of online professional development on teachers' beliefs and perspectives about teaching statistics. In Galindo, E., & Newton, J., (Eds.), *Proceedings of the 39th annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education*. Hoosier Association of Mathematics Teacher Educators.

Lee, H. S., & Mojica, G. F. (2008). Examining how teachers' practices support statistical investigations. In C. Batanero, G. Burrill, C. Reading & A. Rossman (Eds.), *Joint ICMI/IASE Study: Teaching Statistics in School Mathematics. Challenges for Teaching and Teacher Education. Proceedings of the ICMI Study 18 and 2008 IASE Round Table Conference*.

Lee, H. S., Mojica, G. F., & Lovett, J. N. (2020). Examining how online professional development impacts teachers' beliefs about teaching statistics. *Online Learning*, *24*(1), 5-27.

Lehrer, R., Kim, M. J., Ayers, E., & Wilson, M. (2014). Toward establishing a learning progression to support the development of statistical reasoning. In A. P. Maloney, J. Confrey, & K. H. Nguyen (Eds.), *Learning over time: Learning trajectories in mathematics education*, (pp. 31-59).

Lewis, J. M., & Blunk, M. L. (2012). Reading between the lines: Teaching linear algebra. *Journal of Curriculum Studies*, *44*(4), 515-536.

Liu, Y., & Thompson, P. W. (2007). Teachers' understandings of probability. *Cognition and Instruction, 25*(2), 113–160.

Liu, Y., & Thompson, P. W. (2009). Mathematics teachers' understandings of proto-hypothesis testing. *Pedagogies: An International Journal*, *4*(2), 126-138.

Lobato, J., & Walters, C. D. (2017). A taxonomy of approaches to learning trajectories and progressions. In J. Cai (Ed.), *Compendium for research in mathematics education* (pp. 74-101).

Lovett, J. N., & Lee, H. S. (2017). New standards require teaching more statistics: Are preservice secondary mathematics teachers ready? *Journal of Teacher Education*, *68*(3), 299-311.

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.

Lyle, J. (2003). Stimulated Recall: A report on its use in naturalistic research. *British Educational Research Journal*, *29*(6), 861–878. 569

MacKay, R. J., & Oldford, R. W. (2000). Scientific method, statistical method and the speed of light. *Statistical Science*, *15*(3), 254-278.

Maher, C. A. (2008). Video recordings as pedagogical tools in mathematics teacher education. In D. Tirosh, & T. Wood (Eds.), *Tools and Processes in Mathematics Teacher Education* (pp. 65-83). Brill Sense.

Makar, K., & Confrey, J. (2005). Variation talk: Articulating meaning in statistics. *Statistics Education Research Journal*, *4*(1), 27-54.

Maloney, A. P., Confrey, J., & Nguyen, K. H. (2014). *Learning over time: Learning trajectories in mathematics education*. IAP.

Martínez, J. F., Borko, H., & Stecher, B. M. (2012). Measuring instructional practice in science using classroom artifacts: Lessons learned from two validation studies. *Journal of Research in Science Teaching*, *49*(1), 38-67.

März, V., & Kelchtermans, G. (2013). Sense-making and structure in teachers' reception of educational reform. A case study on statistics in the mathematics curriculum. *Teaching and Teacher Education*, *29*, 13-24.

Mathematics Of Doing Understanding Learning and Educating for Secondary Schools (2020). https://www.mtsu.edu/jstrayer/modules/modules2.php

Mayer, D. P. (1999). Measuring instructional practice: Can policymakers trust survey data? *Educational Evaluation and Policy Analysis, 21*(1), 29-45.

McCaffrey, D. F., Hamilton, L. S., Stecher, B. M., Klein, S. P., Bugliari, D., & Robyn, A. (2001). Interactions among instructional practices, curriculum, and student achievement: The case of standards-based high school mathematics. *Journal for Research in Mathematics Education*, 493-517.

Mehan, H. (1979). Learning lessons: Social organization in the classroom. Harvard University Press.

Muijs, D., & Reynolds, D. (2002). Teachers' Beliefs and Behaviors: What Really Matters? *The Journal of Classroom Interaction, 37*(2), 3-15.

National Governors Association Center for Best Practices and Council of Chief State School Officers. (2010). Common Core State Standards for Mathematics. Author.

New Zealand Ministry of Education (2015). The New Zealand Curriculum for English-medium teaching and learning in years 1–13. https://nzcurriculum.tki.org.nz/The-New-Zealand-Curriculum

Noll, J., & Shaughnessy, M. (2012). Aspects of students' reasoning about variation in empirical sampling distributions. *Journal for Research in Mathematics Education, 43*(5), 509–556.

Ormrod, J. E., & Cole, D. B. (1996). Teaching content knowledge and pedagogical content knowledge: A model from geography education. *Journal of Teacher Education, 47*(1), 37.

Paparistodemou, E., Potari, D., & Pitta, D. (2006). Prospective teachers' awareness of young children's stochastic activities. In A. Rossman & B. Chance (Eds.), *Proceedings of the Seventh International Conference on Teaching Statistics*. International Statistical Institute and International Association for Statistical Education.

Pearl, D.K., Garfield, J.B., delMas, R., Groth, R.E., Kaplan, J.J., McGowan, H., and Lee, H.S. (2012). Connecting Research to Practice in a Culture of Assessment for Introductory College-level Statistics. www.causeweb.org/research/guidelines/ResearchReport_Dec_2012.pdf

Peters, S. A. (2009). Developing an understanding of variation: AP statistics teachers' perceptions and recollections of critical moments (Doctoral dissertation). The Pennsylvania State University, University Park, PA.

Peters, S. A. (2014). Developing understanding of statistical variation: Secondary statistics teachers' perceptions and recollections of learning factors. *Journal of Mathematics Teacher Education, 17*(6), 539-582.

Peterson-Ahmad, M. B., Pemberton, J., & Hovey, K. A. (2018). Virtual learning environments for teacher preparation. *Kappa Delta Pi Record, 54*(4), 165-169.

Petit, M. (2011). Learning trajectories and adaptive instruction meet the realities of practice. In P. Daro, F. Mosher, & T. Corcoran. (Eds.), *Learning trajectories in mathematics: A* *foundation for standards, curriculum, assessment, and instruction*. Consortium for Policy Research in Education.

Petocz P., Reid, A., & Gal, I. (2018). Statistics Education Research. In D. Ben-Zvi, K. Makar, & J. Garfield (Eds.), *International handbook of research in statistics education* (pp. 71-99). Springer.

Pfannkuch, M. (2018). Reimagining curriculum approaches. In D. Ben-Zvi, K. Makar, & J. Garfield (Eds.), *International handbook of research in statistics education* (pp. 387-413). Springer.

Pfannkuch, M., & Horring, M. (2005). Developing statistical thinking in a secondary school: A collaborative curriculum development. In G. Burrill & M. Camden (Eds.), Curricular Development in Statistics Education International Association for Statistical Education 2004 Roundtable (pp. 163–173). International Statistical Institute.

Pierce, R., & Chick, H. (2011). Teachers' beliefs about statistics education. In C. Batanero, G. Burrill, C. Reading (Eds.), *Teaching statistics in school mathematics: Challenges for teaching and teacher education* (pp. 151-162). Springer.

Porkess, R. (2012). *The Future of Statistics in our Schools and Colleges*. Royal Statistical Society and The Actuarial Profession.

Reid, A., & Petocz, P. (2002). Students' conceptions of statistics: A phenomenographic study. *Journal of Statistics Education, 10*(2).

Rossman, A., Chance, B., & Medina, E. (2006). Some important comparisons between statistics and mathematics, and why teachers should care. In G. F. Burrill & P. C. Elliott (Eds.), Thinking and reasoning with data and chance: Sixty-eighth NCTM yearbook (pp. 139-150). National Council of Teachers of Mathematics.

Sánchez, E., & García, J. (2008). Acquisition of notions of statistical variation by inservice teachers. 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.*

Schoenfeld, A. H. (1998). Toward a theory of teaching-in-context. *Issues in Education*, *4*(1), 1–94.

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

Simon, M. A. (1995). Reconstructing mathematics pedagogy from a constructivist perspective. *Journal for Research in Mathematics Education*, *26*(2), 114. doi:10.2307/749205.

Sinclair, N., Cirillo, M., & De Villiers, M. (2017). The learning and teaching of geometry. In J. Cai (Ed.), *Compendium for research in mathematics education*, (pp. 457-488).

Siswono, T. Y. E., Hartono, S., & Kohar, A. W. (2018). Effectiveness of project based learning in statistics for lower secondary schools. *Eurasian Journal of Educational Research*, *75*, 197-210.

Skott, J. (2009). Contextualising the notion of 'belief enactment'. *Journal of Mathematics Teacher Education*, *12*(1), 27-46.

Stake, R. E. (2005). Qualitative case studies. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (3rd ed., pp. 443-466). Sage.

Steffe, L. P. (2004). On the construction of learning trajectories of children: The case of commensurate fractions. *Mathematical Thinking and Learning* *6*(2), 129-162.

Stein, M. K., Remillard, J., & Smith, M. S. (2007). How curriculum influences student learning. In Lester (Ed.), *Second handbook of research on mathematics teaching and learning*, (pp. 319-370).

Stipek, D. J., Givvin, K. B., Salmon, J. M., & MacGyvers, V. L. (2001). Teachers' beliefs and practices related to mathematics instruction. *Teaching and Teacher Education*, *17*(2), 213-226.

Stockero, S. L., & Van Zoest, L. R. (2013). Characterizing pivotal teaching moments in beginning mathematics teachers' practice. *Journal of Mathematics Teacher Education*, *16*(2), 125-147.

Sztajn, P. (2003). Adapting reform ideas in different mathematics classrooms: Beliefs beyond mathematics. *Journal of Mathematics Teacher Education*, *6*(1), 53-75.

Thompson, A. G. (1984). The relationship of teachers' conceptions of mathematics and mathematics teaching to instructional practice. *Educational Studies in Mathematics*, *15*(2), 105-127.

Umugiraneza, O., Bansilal, S., & North, D. (2016). Teachers' confidence and beliefs in teaching mathematics and statistics concepts. *International Scientific Researches Journal*, *72*(9), 1-16.

van der Sandt, S. (2007). Research framework on mathematics teacher behaviour: Koehler and Grouws' framework revisited. *Eurasia Journal of Mathematics, Science & Technology Education*, *3*(4).

Voss, T., Kunter, M., & Baumert, J. (2011). Assessing teacher candidates' general pedagogical/psychological knowledge: Test construction and validation. *Journal of Educational Psychology, 103*(4), 952.

Walkowiak, T. A., Berry, R. Q., Meyer, J. P., Rimm-Kaufman, S. E., & Ottmar, E. R. (2014). Introducing an observational measure of standards-based mathematics teaching practices: Evidence of validity and score reliability. *Educational Studies in Mathematics, 85*(1), 109-128.

Warne, R. T. (2017). Research on the academic benefits of the Advanced Placement program: Taking stock and looking forward. *SAGE Open*, *7*(1).

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

Watson, J. (2016). Whither statistics education research? In White, B., Chinnappan, M. & Trenholm, S. (Eds.), *Opening up mathematics education research (Proceedings of the 39th annual conference of the Mathematics Education Research Group of Australasia)* (pp. 33–58). MERGA.

Watson, J., Fitzallen, N., Fielding-Wells, J., & Madden, S. (2018). The practice of statistics. In D. Ben-Zvi, K. Makar, & J. Garfield (Eds.), *International handbook of research in statistics education* (pp. 105-138). Springer.

Weiss, I. R., Banilower, E. R., McMahon, K. C., & Smith, P. S. (2001). *Report of the 2000 national survey of science and mathematics education.* Horizon Research, Inc

Whitaker, D. (2016). The development of a professional statistics teaching identity. In Wood, M. B., Turner, E. E., Civil, M., & Eli, J. A. (Eds.), *Proceedings of the 38th annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education*. The University of Arizona.

Wild, C. J., & Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. *International Statistical Review*, *67*(3), 223-248.

Wild, C. J., Utts, J. M., & Horton, N. J. (2018). What is statistics? In D. Ben-Zvi, K. Makar, & J. Garfield (Eds.), *International handbook of research in statistics education* (pp. 5-36). Springer.

Yang, K. L. (2014). An exploratory study of Taiwanese mathematics teachers' conceptions of school mathematics, school statistics, and their differences. *International Journal of Science and Mathematics Education*, *12*(6), 1497-1518.

Yin, R. K. (2009). *Case study research: Design and methods* (4th ed.). Sage.

Zazkis, R., & Leikin, R. (2010). Advanced mathematical knowledge in teaching practice: Perceptions of secondary mathematics teachers. *Mathematical Thinking and Learning*, *12*(4), 263-281.

Zieffler, A., Garfield, J., & Fry, E. (2018). What is statistics education? In D. Ben-Zvi, K. Makar, & J. Garfield (Eds.), *International handbook of research in statistics education* (pp. 37-70). Springer.

Zieffler, A., Park, J., Garfield, J., delMas, R., & Bjornsdottir, A. (2012). The statistics teaching inventory: A survey on statistics teachers' classroom practices and beliefs. *Journal of Statistics Education*, *20*(1).

APPENDICES

Appendix A. Participant Informed Consent Form

North Carolina State University
INFORMED CONSENT FORM for RESEARCH

Title of Study: Instructional Decision-Making of Secondary Statistics Teachers
Principal Investigator: Taylor Harrison

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 this research study is to gain a better understanding of the decisions secondary statistics teachers make.

You are not guaranteed any personal benefits from being in a study. Research studies also may pose risks to those who 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 or the NC State IRB office as noted below.

What is the purpose of this study?
The purpose of this study is to examine the instructional decision-making of secondary statistics teachers. This research hopes to help educators better understand the current practice mathematics teachers, to inform curriculum development, to assist in teacher preparation, and to better understand the professional development needs of secondary statistics teachers.

Am I eligible to be a participant in this study?
In order to be a participant in this study you must be currently teaching a course with a primary content focus of statistics and/or probability.

What will happen if you take part in the study?
If you agree to participate in this study, you will be asked to participate in an interview regarding your preparation and approach to teaching statistics. You will then be observed by a researcher while teaching a lesson or series of lessons during your regular class time. After each observation, you will take part in an interview asking you about the decisions you made before the lesson, during the lesson, and after the lesson. Each interview is expected to take about 40 minutes.

Photos and video
As a part of this research, we would like your consent to video record and/or audio record your classroom teaching. You may still participate in the study if you do not consent to audio or video recording.
____ I consent to have my classroom teaching video-recorded and audio-recorded
____ I consent to have my classroom teaching audio-recorded *only.*
____ I do not consent to have my classroom teaching video-recorded or audio-recorded

As a part of this research, we would like your consent to video record and/or audio record your interviews with a researcher. You may still participate in the study if you do not consent to audio or video recording.
____ I consent to have my interviews video-recorded and audio-recorded
____ I consent to have my interviews audio-recorded *only.*
____ I do not consent to have my interviews video-recorded or audio-recorded

It may be beneficial to share recordings made during this study during professional presentations or for educational purposes. You may still participate in the study if you do not consent to having your recordings shared with others:

____ I do not want my video or audio recordings to be viewed by others during professional presentations or for educational purposes (derivative works -- e.g., digital or text-based case studies -- based on these recordings may be used instead).
____ I consent to having my video or audio recordings viewed by others during professional presentations or for educational purposes. Your real names, city, or schools will not be used.

Risks and Benefits

There are minimal risks associated with participation in this research. There are limited direct benefits to your participation in the research. Indirect benefits may include reflection on your teaching practices and beliefs, leading to a better self-understanding and awareness. At the conclusion of the study, you will be provided with information on resources that are available that may assist in your teaching of statistics.

Confidentiality

The information in the study records will be kept confidential to the full extent allowed by law. Data will be stored securely on an NC State managed computer and/or password-protected Google accounts via Google Drive. No reference will be made in oral or written reports which could link you to the study. Any identifiable information collected as part of this research will not be used or distributed for future research purposes without your consent.

To help maximize the benefits of your participation in this project, by further contributing to science and our community, your de-identified information, in addition to video and audio recordings if consented to above, will be stored for future research and may be shared with other people without additional consent from you

Compensation

For participating in this study you will receive a gift card. To be eligible for compensation, you must participate in the pre-observation interview, be observed for at least one lesson, and participate in at least one post-observation interview. Compensation will be in the amount of [amount]. The gift card will be sent via email after the completion of data collection. If you decide to withdraw from the study before the interviews and observations are complete, the compensation will not be received.

What if you have questions about this study?

If you have questions at any time about the study itself or the procedures implemented in this study, you may contact the researcher, Taylor Harrison, [email], or at [address], or at [phone number].

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 the NC State IRB Office via email at irb-director@ncsu.edu or via phone at 1.919.515.8754. You can also find out more information about research, why you would or would not want to be in research, questions to ask as a research participant, and more information about your rights by going to this website:
http://go.ncsu.edu/research-participant

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."

Participant's printed name ____________________________________

Participant's signature ______________________________________ **Date** _________________

Investigator's signature _____________________________________ **Date** _________________

Appendix B. Student Assent/Guardian Consent Form

Informed Consent Form: Parents of Students in Mathematics Classroom Sessions

Instructional Decision-Making of Secondary Statistics Teachers

Taylor Harrison, Principal Investigator

What are some general things you should know about research studies?

Your child is being asked to take part in a research study. Your child's participation is voluntary. Your child has the right to be a part of this study, to choose not to participate, or to stop participating at any time without penalty.

You or your child are not guaranteed any personal benefits from being in this study. Research studies also may pose risks to those who participate.

In this consent form you will find specific details about the research in which your child is being asked to participate, including risks and benefits. 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 child's participation, do not hesitate to contact the researcher named above or the NC State IRB office (contact information is noted below).

What is the purpose of this study?

The purpose of this study is to investigate the decisions that teachers make when planning and implementing a statistics lesson.

What will happen if you take part in the study?

Your child will contribute to this research by participating in lessons with their regular classroom teacher during regular class periods. The lessons may be video recorded with the best efforts of avoiding your child's appearance. Up to 8 lessons during the spring semester of 2019 may be observed and recorded. The recordings will attempt to capture the students from their back preventing their identifications.
If you do not want your child to participate in this study, they will still participate in the class sessions and engage in the same tasks as other students, but they can be seated outside the view of the video camera.

Video recordings may be retained for use in further research studies.

Risks & Benefits

There will be minimal risk associated with their participation in the study. Participation in this study is not a course requirement and your child's participation or lack thereof, will not affect their class standing or grades in the course. You are free to withdraw your child from the study at any time; however, they will still participate in all of the activities that are class requirements. To withdraw your child from the study, contact the researcher, Taylor Harrison, at [email], or at [phone number], or at [address].

The knowledge we gain from their experiences will add to the knowledge base in statistics education, especially with regard to preparing teachers to teach statistics, and in how teachers make decisions in planning and implementing statistics education.

Confidentiality

Your child's name and other personal identifying information will not be collected during this study. The information in the study records will be kept confidential to the full extent allowed by law. No reference will be made in the videos, oral or written reports, or transcripts which could link them to the study. Though their face may be visible on videos, their real names, city, or schools will not be used. Data will be stored securely on an NC State managed computer. Individual data with identifiable details removed may be made available to the public as required by a professional association or journal. Selected video clips or derivative works (e.g., digital or text-based case studies) in which your child's face may be visible may be used for professional presentations or educational purposes.

What if you have questions about this study?

If you have questions at any time about the study itself or the procedures implemented in this study, you may contact the researcher, Taylor Harrison at [email], or at [phone number], [address].

What if you have questions about your rights as a research participant?

If you feel your child has not been treated according to the descriptions in this form, or your child's rights as a participant in research have been violated during the course of this project, you may contact the NC State IRB (institutional Review Board) Office via email at irb-director@ncsu.edu or via phone at 1.919.515.8754. You can also find out more information about research, why you would or would not want to be a research participant, questions to ask as a research participant, and more information about your rights by going to this website: http://go.ncsu.edu/research-participant

---

CONSENT for parent: "I have read and understand the above information. I have received a copy of this form. I agree to allow my child to participate in this study with the understanding that I may choose not to have him/her participate or to stop participating at any time without penalty or loss of benefits to which I am otherwise entitled."

ASSENT for child: "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."

Child's name (print): ____

Child's signature:____ date: _______________

Parent's name (print):____

Parent's signature:____date:________________

Appendix C. Pre-observation Interview Protocol

Part 1:

1) Can you start by telling me how long you've been teaching?
2) And how many times have you taught statistics?
3) Can you tell me a little bit about the [AP] Statistics course at this school--who tends to take it and when, where does the course fall in the rest of the math curriculum?
4) For the [AP] Stats course that you teach--what are your overall goals for students in this course? What do you want students to walk away with at the end of the course?
5) This particular section that I'll be observing--can you describe this group of students and any interesting characteristics that I might want to know? (Prompts, if needed: quiet, talkative? seem to be engaged? like working in groups? eager to offer their thoughts?)
6) Are there any particular norms that you've established with the students in this class? (Prompts if needed: do students raise their hands to answer? do you call on students randomly? any stated expectations for group work? students pick their own groups or are they assigned?)

Part 2: The next part of the interview will have you talk about some of the general ways you approach the teaching of statistics in this course.

7) Let's talk about some different aspects of your planning and instruction for this course. What I'm really interested in in this section is not just a description of what you do, but the reasons why you do what you do.
    a) So, first, the source of tasks. Where do you get tasks that you use in the classroom, and why?
    b) Next, technology. What do you and your students use, and why?

c) When you're deciding how much lecture vs. individual or group work to include in a lesson, what goes into that decision?

d) Finally, assessment. How do you assess your students' learning, and why?

8) What resources are available to help you teach statistical topics? How often do you use them?

   a. Follow-up, ask about: websites, textbooks, district resources, colleagues, instructional support

9) How do you feel your pre-service teacher preparation has had an impact on how you teach statistics?

10) Have you had any professional development or professional learnings experiences that impact how you teach statistics?

11) Finally, let's talk about some of the constraints that you might face in your particular classroom. If you could envision the perfect statistics lesson, how closely do you think the lessons that you plan to teach will actually come to that? If you envision there might be some differences, what might be some reasons?

   a) Examples, if needed: planning time, class size, class time, student characteristics, adherence to testing

Part 3: The next part of the interview will ask you about some more specific thoughts and plans for the upcoming statistics topics that you plan to teach.

12) Let's consider the sequence of lessons that you plan on teaching over the next several days. If we consider a student in your classroom. What statistical or mathematical knowledge do they need to have today in order to be successful over the next few days? And what knowledge or skills do you want them to have by the end of this sequence of

lessons [unit, etc.]? Walk me through how you expect their knowledge to develop during that time between here and there.

13) Let's specifically talk about the lesson that you plan to teach tomorrow [or during the first observation]. What is the learning goal that you have for your students for this lesson?

14) What will I see you and your students doing tomorrow? Tell me a little bit about how you think the lesson that you have planned will help students reach that learning goal.

    a. Follow-ups: Use of technology, group size, amount of time devoted to activity vs lecture

15) Were there any particular resources that you used while planning your lesson, such as textbooks, websites, other tasks, etc.? If so, how did you use them, and how did they influence your plans for the lesson?

16) Are there any previous experiences that you've had that helped you plan this lesson or influenced how you planned it? For example, maybe a past teaching experience, or some professional development that you had, or conversations with others, and so on. If so, how did those experiences help or influence your plan for this lesson?

17) You mentioned that your learning goal for tomorrow is ___________. What are some of the different ways you expect students might reason about this topic?

    a. Follow-up: What do you anticipate students might have trouble with?

Appendix D. Observation Protocol

Date:	
Teacher/Period:	
Start time:	

Planning decisions to address:

Type	Time	Event/Response	Notes

Appendix E. Post-observation Interview Protocol

Part 1:

1) Any initial thoughts about today's lesson? How do you think it went?
    a. Was there anything that didn't go as well as you would have liked?
2) Was there anything unexpected that happened during the lesson? If so, how did you respond, and why?
3) Was there any point where you decided to deviate a bit from how you initially envisioned the lesson would go?
4) If you were to teach this lesson again, is there anything that you would change?

Part 2a: The next part of the interview will ask you about some of the general ways the lesson was structured, and then after that, we'll discuss some of the specific things that happened during today's lesson. Note that I'm not bringing things up because I think they're unusual or substandard, but I'm simply curious and want to better understand the reasons behind what you do.

5) I noticed that you decided to ____________. Can you tell me a little about that decision, and how you think it might impact students' learning?

Part 2b: Alright, let's talk about some particular things that happened during today's lesson. I'm going to [show you some short video clips/let you listen to some short audio clips/describe some things that I saw] from today's lesson. As you're [watching/listening] each [clip/description], I'd like for you to try to remember what you were thinking when this happened, and how you decided how to respond. I'd like you to try to remember what you were thinking *at the time,* and not necessarily what you think about the event looking back.

(After watching/listening to each clip/description):

6) At this point in the lesson, can you describe to me what you were thinking? How did you decide how to respond?
    a. Do you feel this event changed the trajectory of the lesson, either in the short-term or long-term? If so, how and why do you think it changed?

Part 3:

7) Let's specifically talk about the lesson that you plan to teach tomorrow [or during the next observation]. What is the learning goal that you have for your students for this lesson?
8) What is the structure of the lesson that you have planned for tomorrow? Tell me a little bit about how you think the lesson that you have planned will help students reach that learning goal.
    a. Follow-ups: Use of technology, group size, amount of time devoted to activity vs lecture
9) Do you think students need to have a good understanding of today's material to be successful with tomorrow's material? If so, why?
10) Are there any previous experiences that you've had that helped you plan this lesson or influenced how you planned it? For example, maybe a past teaching experience, or some professional development that you had, or conversations with others, and so on. If so, how did those experiences help or influence your planning?
    a. Follow-up: Were there any particular resources that you used while planning your lesson, such as textbooks, websites, other tasks, etc.? If so, how did you use them, and how did they influence your plans for the lesson?

11) You mentioned that your learning goal for tomorrow is ___________. What are some of the different ways you expect students might reason about this topic?

    a. Follow-up: What do you anticipate students might have trouble with?

Appendix F. Sample Field Notes

Date:	[Day 1]
Teacher/Period:	TeacherX
Start time:	2:10

Planning decisions to address:

random grouping

Type of event	Time of event	Description of event/response	Notes
I*	0:00-6:20	Reviewing old material	What's the purpose of this part of the class? What are you looking for?
I	6:50-8:15	Talks about grade conference, and the three main things that contribute to the discussion: GoFormatives, Learning Log, Personal Progress Checks	
E*	9:10-11:00	Shoes on or shoes off?	Why did you have this as a class discussion, rather than choosing one for them?
E*+	15:45-15:55	What does Sx stand for?	Why did you direct them toward their tablemates?
E*+	17:20-21:11	Discussion between sample statistics and population statistics	
E*+	22:30-23:20	Asks students why did they use a histogram?	
E*+	24:00-24:10	Gives students the option of doing a histogram even though instructions say dot plot. Introduces binned dot plot.	Why did the teacher discourage the use of dot plots?
E*	27:15-29:00	Ask students what the IQR is and how to calculate it	Teacher carries around a marker to write on tables

E*	29:10-31:30	Overhears a student making a prediction	Off-camera
E*+	31:40	"Why are you asking me?"	
E*+	32:15-32:30	*Is* that what Sx is?	
E	33:00-	Quartile discussion – teacher responds that she likes that a group used box plot to find/estimate IQR.	
E*+	33:35-34:00	So, what is Sx?	
E	35:25	I have a question. "Have you already asked them?"	
E	36:30	Calculation	
E	38:00	Discussion about subtracting 2 from a list	
E/I*+	39:20-40:50	Discussion about outliers	Why go through this discussion instead of just telling them?
E/I	42:15	Discussion about bi-modal shape	
E*		Was a longer close planned?	The necessity of a close was mentioned in the pre-interview (mentioned ~10 mins desired)

Legend:

- E = event that has the potential to impact teacher's instructional plan
- I = instance that illustrates a teacher's instructional plan

- * = events asked about during post-observation interview
- + = events for which video-recording was shown during post-observation interview