SUBJECT MATTER KNOWLEDGE FOR TEACHING

STATISTICAL ASSOCIATION

Stephanie A. Casey

243 Pages August 2008

A practice-based approach was used to study the knowledge secondary teachers need to teach statistical association. Through this study a description of the knowledge needed by teachers was created.

SUBJECT MATTER KNOWLEDGE FOR TEACHING

STATISTICAL ASSOCIATION

Stephanie A. Casey

241 Pages August 2008

This study sought to describe the subject matter knowledge needed for teaching statistical association at the secondary level. Taking a practice-based approach, three teachers were observed as they taught statistical association with an analytical focus on the demands upon teachers' subject matter knowledge involved in the practice of teaching. Following each observation, an interview was conducted with the teacher to discuss the subject matter knowledge used to teach the observed class session. Records of practice, including observation transcripts, interview transcripts, handouts, textbook pages, teacher's guide pages, and copies of student work, were assembled to create a compilation document for each of the observed class sessions. Through analysis of the compilation documents by the researcher and additional analysis by a statistician and a statistics education expert, a detailed description of the subject matter knowledge needed for teaching statistical association was created.

APPROVED:

Date	Cynthia Langrall, Chair
Date	Edward Mooney
Date	Tami Martin
Date	Amee Adkins
Date	Jinadasa Gamage

SUBJECT MATTER KNOWLEDGE FOR TEACHING STATISTICAL ASSOCIATION

STEPHANIE A. CASEY

A Dissertation Submitted in Partial
Fulfillment of the Requirements
for the Degree of

DOCTOR OF PHILOSOPHY

Department of Mathematics

ILLINOIS STATE UNIVERSITY

2008

DISSERTATION APPROVED:

Date Cynthia Langrall, Chair

Date Edward Mooney

Date Tami Martin

Date Amee Adkins

Date Jinadasa Gamage

ACKNOWLEDGMENTS

This dissertation would not have been possible without the support of a number of individuals. All of them have been lifelines along this journey and I would like to formally thank them here.

I would first like to thank my chair, Cynthia Langrall, who helped me form the ideas for this study and has seen me through to its end. Thanks also to the members of my committee, Edward Mooney, Tami Martin, Amee Adkins, and Jinadasa Gamage, for your time spent on my behalf.

Next I want to thank the three participating teachers. To protect your anonymity I can not name you here, but I am forever indebted to you for your willingness to open your classrooms and minds to me for this study. It was a pleasure working with such high quality teachers.

A number of experts in the fields of statistics and statistics education helped during the analysis process and thanks to them this study is much stronger. Thanks to statistician Julie Clark and statistics education expert Tim Burgess for serving as analysts for the study. Also I extend my gratitude to Allan Rossman for answering all of the content questions generated during my analysis of the data. I learned so much about the topic of statistical association from my work with all of you and for that I am very grateful.

I am also indebted to my family and friends for all of their support. This was truly a communal effort, from my friends and family who babysat my kids to my personnel director who approved my leave of absence from teaching so I could collect the data for this study, and I thank all of you. Lastly I must thank my husband Jim and children Elizabeth and Connor, who have provided unconditional support and too much time away from them to do this study. You're the best and I love you very much. I look forward to reading this dissertation to you as your new bedtime story.

S.A.C.

CONTENTS

FIGURES

CHAPTER I

THE PROBLEM AND ITS BACKGROUND

Background and Rationale

The students of the United States did poorly on a recent international comparison of mathematics achievement, the Third International Mathematics and Science Study (TIMSS) (Schmidt et al., 1999). The TIMSS' inclusion of a video study of teaching (Stigler et al., 1999) brought to the forefront the possibility that teaching and teachers were to blame for United States students' achievement problems (Ferrini-Mundy & Graham, 2003). This supposition, along with evidence suggesting that United States teachers do not have essential knowledge for teaching mathematics (e.g., Ball, 1990; Ma, 1999); and the widespread belief that teachers' subject matter knowledge affects student learning (Hill, Schilling, & Ball, 2004); and the realization that effective implementation of reform standards and curricula depends upon the work of knowledgeable teachers (Ball, Hill, & Bass, 2005), has resulted in significant attention on teachers' knowledge for teaching.

This attention can be seen in the 2001 No Child Left Behind Act (NCLB, 2002) which requires all teachers to be highly qualified, with highly qualified defined as having subject-matter competency. Policy documents (e.g., Conference Board of the Mathematical Sciences [CBMS], 2001; National Board for Professional Teaching Standards, 2002; National Commission on Teaching and America's Future, 1996; National Council of Teachers of Mathematics [NCTM], 2005; National Mathematics Advisory Panel, 2008) have also emphasized the importance of knowledgeable teachers, with particular emphasis upon subject matter knowledge. However, despite this attention, mathematics educators still lack a full understanding of the mathematical knowledge needed for teaching mathematics effectively (Ball, Lubienski, & Mewborn, 2001; RAND Mathematics Study Panel, 2003).

The lack of clarity and precision regarding what knowledge teachers must have has resulted in differing claims regarding the mathematics that teachers need to know for teaching. For instance, the United States Department of Education (2002) asserted that a teacher's capability in general mathematics is the most important qualification for teaching. NCTM (2005) stated that all mathematics teachers need extensive knowledge of mathematics, the mathematics curriculum, and how students learn. Little empirical research has been done in this area to help judge the validity of these claims (Hill et al., 2004); both are based primarily on beliefs.

The nature of the knowledge needed for teaching is largely under-specified and unresearched (Ball & Bass, 2003; Ball et al., 2001; National Research Council, 2001; RAND Mathematics Study Panel, 2003). While it is obvious that mathematics teachers need to know mathematics, what is unknown is exactly what aspects of mathematics teachers need to know, how they need to know it, and how and where this mathematics knowledge is used in practice (Ball & Bass). Research is needed to systematically study these aspects of teachers' subject matter knowledge (National Research Council). There is also a call for distinguishing the mathematics needed by teachers from the mathematics needed by others who use it in their work which will result in a clear definition of the specialized knowledge of mathematics needed for teaching (RAND Mathematics Study Panel).

The need for such research is particularly acute for the teaching of statistics. As a result of the NCTM Curriculum and Evaluation Standards for School Mathematics (1989) and Principles and Standards for School Mathematics (2000) and its growing importance in today's world (Ben-Zvi & Garfield, 2004), statistics has become an accepted strand of mainstream school mathematics curricula in the United States (Moore, 2004). NCTM (2000) and statistics educators (Ben-Zvi & Garfield) have called for student learning to go beyond procedural knowledge to emphasize statistical thinking, reasoning, and literacy. Therefore more teachers with more varied backgrounds are currently called to teach students to think and reason statistically (Moore). However, many teachers have never studied statistics, and those who have typically experienced a course that emphasized procedural knowledge (Franklin, 2000) rather than the type of understanding described by the NCTM (2000) Principles and Standards for School Mathematics. This has led to concern regarding the knowledge involved in teaching statistics and whether teachers in the United States have this knowledge (CBMS, 2001; Franklin; Kettenring, Lindsay, & Siegmund, 2003; Makar & Confrey, 2004). Shaughnessy (2007) identified the knowledge necessary for teaching statistics as a needed research emphasis based on his review of current research on statistics learning and reasoning. Some have suggested (e.g., Franklin; Kettenring et al.) that the number one priority in statistics education is to address the shortage of K-12 teachers prepared to teach statistical concepts.

Pre-service secondary teachers are assumed to have learned some statistics content knowledge during their college education, predominantly through the standard elementary statistics course offered by the mathematics or statistics department (Schaeffer, 2005). It is doubtful that this type of course addresses the content knowledge needs of pre-service teachers (CBMS, 2001; Schaeffer; Usiskin, 2000). The vague description of the statistics curricula for pre-service secondary mathematics teachers contained in The Mathematical Education of Teachers (CBMS) does not provide much help to teacher educators either, as it lacks a clear specification of what teachers need to know in order to be prepared to enter the workforce. Following the suggestion of the CBMS, Usiskin, Peressini, Marchisotto, and Stanley (2003) created a textbook for a capstone course in mathematics for pre-service secondary mathematics teachers, but it does not include any data analysis or statistics. Therefore, it is questionable whether pre-service secondary teachers are gaining the subject matter knowledge needed for teaching statistics at the secondary level.

In-service teachers' content knowledge of statistics, particularly that used for teaching, is perhaps even more questionable due to the relatively recent emergence of the statistics strand in secondary curricula and the lack of professional development opportunities available to teachers through their school districts to master the statistics they are expected to teach (Makar & Confrey, 2004). Some mathematics teachers have reported that they are not comfortable teaching statistics due to their inadequate background in the area (Franklin, 2000).

What do teachers need to know about the subject in order to teach statistics? David Moore (2004), a leading figure in the field of statistics education, explained: "Teachers at all levels must understand their subject matter [referring to statistics], and at a depth at least somewhat greater than that of the content they actually teach" (p.ix). Wilson (2004) listed "extremely good statistical knowledge base, knowledge of connections between statistical concepts, knowledge of applications of statistics" (p. 7) when describing the knowledge of statistics necessary for teaching. What is this depth? What is extremely good? These types of indistinct descriptions need specification in order to be useful and meaningful.

Statement of the Problem

The prevalent hypothesis regarding teachers' subject matter knowledge is that they need to know the content for the level of schooling they will be teaching plus N levels more knowledge (Ball & Hill, 2005). However, the common practice in recent years of increasing and making more rigorous the mathematics requirements for teaching (e.g., CBMS, 2001) has not been particularly effective at helping teachers gain the knowledge needed for teaching mathematics (Ball, Bass, Sleep, & Hoover, 2004). One explanation is that the gap between the mathematics a teacher studies and the mathematics a teacher teaches usually becomes wider the more mathematics courses a teacher takes (Usiskin, 2000). As a result, teachers are commonly no more knowledgeable regarding the content they will teach than when they were students learning that content for the first time. Another explanation is that there is a unique, specialized knowledge of mathematics needed by teachers in their work that the common hypothesis does not recognize. At the elementary school level, Hill et al. (2004) found evidence for the existence of this specialized knowledge of content for mathematics teachers and established it as distinct from common content knowledge.

The purpose of the current study was to develop a practice-based theory of the subject matter knowledge needed for teaching statistical association at the secondary level. Through a practice-based approach, I observed teachers as they taught statistical association, with an analytical focus on the demands upon teachers' subject matter knowledge involved in the practice of teaching. A practice-based approach was the most appropriate way to research the topic in order to study firsthand the knowledge of content used in teaching (Ball & Bass, 2000a; Fennema & Franke, 1992). This approach connected the theoretical constructs to practice (Ball & Bass, 2003) and closely connected mathematics and teaching in ways that contribute to improving teacher practice and policy regarding teacher knowledge (Ball et al., 2001).

Following the work of Ball and Bass (2000a), and extending it to the secondary level, I looked through the lens of the discipline of mathematics in its study of teachers involved in their practice. Since I was interested in unraveling, identifying, and understanding the mathematical work that teachers are engaged in as they teach, this lens was most appropriate. It should be noted, however, that I expanded the meaning of mathematics utilized by Ball and Bass to include the triad of mathematical knowledge, statistical knowledge, and context knowledge, the subject matter knowledge components considered necessary for statistical literacy by Gal's (2004) statistical literacy model. These three knowledge components are described below.

Gal (2004) described mathematical knowledge as knowledge of the mathematics underlying statistical procedures as well as numeracy skills that enable correct interpretation of numbers used in statistical reports. The second type of knowledge, statistical knowledge, has five essential parts: knowing why data are needed and how data can be produced; familiarity with basic terms and ideas related to descriptive statistics; familiarity with basic terms and ideas related to graphical and tabular displays; understanding basic notions of probability; and knowing how statistical conclusions or inferences are reached. The third type of knowledge, knowledge of the context, is necessary for proper interpretation and analysis of statistical statements or graphs. Statistics is a unique subject in that it has a complex relationship with mathematics and context (Gal, 2004) as opposed to topics like number and operations that Ball and Bass have studied which could be considered strictly mathematical. Therefore, in my study I used this triad as a starting point for analyzing the subject matter knowledge needed by teachers in their instruction dealing with statistical association.

Statistical association refers to the correspondence of variation of two statistical variables that are both numerical, both categorical, or involving one categorical and one interval variable (Moritz, 2004). Because statistical variables refer to data that are “not merely numbers, but numbers with a context” (Moore, 1990, p.96), statistical association involves a relationship of measured quantities of distinct characteristics from a context (Moritz). The topic of statistical association was identified as one of eight big ideas in statistics (Garfield & Ben-Zvi, 2004) and is a main component in the secondary (grades 9-12) data analysis and probability standard for school mathematics (NCTM, 2000).

Research Question

Using a practice-based approach, I focused on the following question:

What is the subject matter knowledge for teaching statistical association at the secondary school level?

Theoretical Perspective

This study used the theoretical construct of teacher knowledge as developed by the researchers at the University of Michigan’s Mathematics Teaching and Learning to Teach Project following their empirical research of elementary school instruction in mathematics. Ball and Hill (2005) divided teacher knowledge into two main parts: pedagogical content knowledge and subject matter knowledge. Pedagogical content knowledge consists of two subcategories: (1) the knowledge of students in interaction with particular content, and (2) knowledge of teaching and the curriculum. Subject matter knowledge also consists of two components: common content knowledge and specialized content knowledge. In this study, I limited my attention to teachers' subject matter knowledge due to recent calls for its focus in mathematics education research (e.g., Ball et al., 2001; National Research Council, 2001; RAND Mathematics Study Panel, 2003) and the lack of work in this area in statistics education. The rest of this section will therefore elaborate more upon the construct of teachers' subject matter knowledge and related research.

As one component of subject matter knowledge, *common mathematical knowledge* refers to the knowledge of mathematics content one would expect the average person to possess in order to say that he or she knows that content (Ball & Hill, 2005). It is expected that teachers would possess this content knowledge as well. However, that knowledge alone is insufficient for teachers. Teachers also need another component of subject matter knowledge: a *specialized mathematical knowledge*. This is a unique knowledge of mathematics needed for teaching which is different from the type of knowledge needed in other occupations where mathematics is used (Ball & Bass, 2000a; Ball et al., 2004; Ball & Hill, 2005). Through their analyses of elementary teachers' work in teaching mathematics, Ball and Bass (2003) determined that teaching mathematics is a form of mathematical work requiring substantial mathematical knowledge and reasoning. It demands a depth and detail of knowledge well beyond what is needed to carry out procedures reliably. Teachers' mathematical reasoning needs to be strong to address the consistent stream of on-the-spot mathematical problems that teachers must solve. The types of mathematical problems teachers must solve are also specialized, such as determining if a student's solution method is correct and generalizable, and so require of teachers a specialized knowledge of mathematical problem solving (Ball et al.). Therefore Ball and Hill (2005) have defined mathematical knowledge for teaching as the "mathematical knowledge, skill, [and] habits of mind that are entailed by the work of teaching" (p.9).

Ball and her colleagues (Ball & Bass, 2000b, 2003; Ball et al., 2004; Ball et al., 2005) have identified a number of teaching activities as occasions when mathematics teachers need to utilize their mathematical knowledge for teaching. Mathematics teachers need to apply their mathematical knowledge for teaching when they do the following:

- Understand and judge student created methods, explanations, claims, or solutions;
- Understand reasons behind methods used in mathematics;
- Use, choose, and judge representations;
- Model operations and concepts;
- Choose examples;
- Choose and develop usable, appropriate definitions;
- Judge and modify treatments of topics in instructional materials;

- Choose and develop usable, appropriate definitions for mathematical terms;
- Determine what counts as mathematical explanation;
- Create mathematical explanations that are accurate, comprehensible, and useful for students;
- Reconcile multiplicities of interpretations, definitions, representations, solution methods, or solutions;
- Explain learning goals and mathematical purposes to others;
- Design assessments;
- Assess students' mathematical learning and take appropriate next steps;
- Determine mathematical questions and problems that are productive for student learning.

I used these activities in this study to inform the focus of the classroom observations and interviews.

An important property of mathematics teachers' specialized knowledge is that it must be usable for the mathematical work that teachers do (Ball & Bass, 2000a). Teachers are expected to facilitate discussion in the classroom, which means they need a deep knowledge of their subject matter in order to make sense of student solutions generated during such discussions (Franke, Kazemi, & Battey, 2007). This specialized knowledge differs from pedagogical content knowledge which bundles knowledge of mathematics with knowledge of learners, learning, and pedagogy. Knowledge in this packaged, intertwined form may not equip a teacher with the flexibility needed to manage the real-time complexity of teaching mathematics. Teachers need a type of mathematical knowledge that is useful and ready for the real-time problem solving that teaching requires, not one that is bundled in advance as an anticipatory resource. For example, pedagogical content knowledge in the area of statistical association could include knowing the research finding that students often believe that the only possible type of association is a direct or positive one (Batanero, Estepa, Godino, & Green, 1996) and having ready explanations as to why inverse or negative association exist as well. However, this type of knowledge may not be helpful to the teacher as he or she thinks about the mathematics in a student's original idea about how to measure the association of two quantitative variables. This situation calls for real-time problem solving on the part of the teacher, calling upon his or her subject matter knowledge when considering and evaluating the student's idea. The teacher must reason and cannot just access his or her repertoire of strategies and answers. Teachers' specialized mathematical knowledge must be usable in the inherently uncertain environment of the classroom.

In order to be useful, qualities of the specialized mathematical knowledge for teaching have been identified (Ball & Bass, 2000a, 2003; Ball et al., 2004). Particularly distinctive to teaching, teachers' mathematical knowledge needs to be unpacked (Ball & Bass, 2003). While mathematics in general aims for packing or compressing information into abstract forms, teaching requires a kind of decompression or unpacking of mathematical ideas since it deals with students in the process of learning the subject. Teachers' subject matter knowledge needs to be connected, both within and across mathematical domains and across time as mathematical concepts develop for students (Ball & Bass, 2003). Ball and Bass (2003) have consistently found that the significant mathematical issues at work in teaching a lesson are not just related to the curricular topic of the lesson. Teachers' subject matter knowledge also needs to be held flexibly, so that teachers are able to utilize it with a wide variety of students across a range of environments (Ball & Bass, 2000a). Despite the need for teachers to teach all students, a lack of flexibility in their knowledge often prevents them from truly hearing students or thinking about mathematical topics in ways other than their own.

This theory was empirically tested with elementary school teachers (Hill, Rowan, & Ball, 2005; Hill et al., 2004). Measures of the mathematical knowledge needed for teaching were developed, tested, and then used in a large-scale assessment of elementary teachers (Hill et al.). Empirical evidence established the existence of specialized content knowledge for teaching, separate from common content knowledge. This specialized knowledge included the ability to analyze alternative algorithms or procedures; show or represent numbers or operations using manipulatives; and provide explanations for common mathematical rules. The research also indicated that mathematical knowledge for teaching is at least somewhat domain-specific rather than indicative of a general factor like overall intelligence or mathematical ability. In my research study, I assumed that Ball and her colleagues' theory regarding the mathematical knowledge for teaching (Ball & Bass, 2000a; Ball & Hill, 2005) also applies to the secondary level, where this study took place.

CHAPTER II

REVIEW OF RELATED LITERATURE

Introduction

This chapter presents literature relevant to the topic of this study: subject matter knowledge used in high quality teaching of statistical association at the secondary level. Teacher knowledge is the focus of the first section. The first portion of this section presents the historical background on the topic of teacher knowledge over the past thirty years to show where the field was thirty years ago, how it has developed in its frameworks and research regarding teacher knowledge, and to help set the stage for current frameworks and research regarding teacher knowledge.

In the next portion, I describe current literature specific to mathematics teachers' knowledge. Three recently developed frameworks regarding mathematics teachers' knowledge are detailed along with relevant research findings regarding mathematics teachers' knowledge. It is important to know the current research emphases and frameworks regarding mathematics teacher knowledge in order to situate this study in the environment in which it occurred. In the final portion of the first section, I concentrate on frameworks and research regarding mathematics teachers' knowledge of statistics. This portion narrows the topic of teacher knowledge to the content of statistics, which this study focuses on, and illustrates the lack of research done in this area.

Statistical association is the topic of the second section of the review of related literature. Within the subject of statistics, this study focused upon the idea of statistical association and thus it is relevant to detail the pertinent literature on this topic. The first part of this section describes the student learning objectives regarding statistical association for the secondary level to help articulate the content students are expected to learn at this level of schooling and, therefore, help define what is meant by the study of statistical association at the secondary level. The objectives also provide a preliminary idea of what might comprise the subject matter knowledge base for teaching statistical association.

Research on students' understanding of statistical association, including their conceptions and misconceptions, is summarized in the next part of the section. This helps to anticipate the understanding secondary students will have, either prior to instruction or following instruction, to which the teacher will be called upon to analyze and respond. The final part of this section describes research on teachers' knowledge regarding statistical association on order to speak to the current understandings that teachers have of the topic of focus for this study.

Teacher Knowledge

Historical Background

Research regarding teachers' knowledge for teaching mathematics has been an active area of research for the past thirty years (Ball et al., 2001). In the 1960's and 1970's, the prevalent view of the knowledge needed for teaching mathematics emphasized the importance of a strong basis in the subject matter as evidenced by related reports (e.g., Mathematical Association of America Committee on the Undergraduate Program in Mathematics, 1961, 1971), requirements for teacher education programs, and the content of the National Science Foundation institutes for professional development (Ferrini-Mundy & Graham, 2003). With this framework of teacher knowledge, research studies known as educational production function studies measured teachers' content knowledge and determined whether it was related to student achievement (Hill et al., 2005).

In educational production function studies, the measurement of mathematics teachers' knowledge was done indirectly through proxy variables, such as the number of post-secondary mathematics courses taken or scores on certification tests regarding content knowledge (Hill et al., 2005). In Begle's (1979) well publicized meta-analysis of such research done between 1960 and 1976, he came to the surprising conclusion that these measures of mathematics teacher knowledge had little relationship with student achievement. Begle concluded that the prevalent belief that the more a teacher knows about the subject matter, the more effective he or she will be as a teacher, demanded “drastic modification” (p. 51). He recommended that further research of this type be suspended until the ways in which teacher knowledge were related to student achievement were better understood.

Possible reasons for the lack of association found between mathematics content knowledge and student achievement in these studies include the imprecise definition and indirect measurement of teachers’ knowledge (Hill et al., 2005). It is debatable whether variables such as the number of post-secondary mathematics courses taken or performance on tests of basic mathematical ability accurately measure teachers’ knowledge of mathematics, especially the specialized knowledge that is used for teaching the subject (Ball & Bass, 2000a; Ball & Hill, 2005).

Recommendations regarding the mathematical preparation of teachers (e.g., Mathematical Association of America, 1983; NCTM, 1981) began to reflect the idea that knowledge of mathematics was not enough for teachers. These organizations began pushing for a blending of both mathematical and pedagogical knowledge in teacher preparation. A swing to an emphasis upon pedagogical knowledge in the 1980’s could be seen in the policies regarding the evaluation and testing of teachers for certification and the prevalent educational research paradigms (Shulman, 1986). In the 1980's, most states had teacher evaluations that emphasized the assessment of a teacher's capacity to teach, largely ignoring his or her knowledge of the subject matter (Shulman). Process-product and teacher effectiveness studies, which were dominant research paradigms of the time, focused solely on teachers' pedagogical behavior to account for students' academic achievement. This resulted in what Shulman called the "missing paradigm problem" (p.6). A missing component of the research paradigms of the time was subject matter, which led to state policies regarding teacher evaluations for certification that emphasized pedagogy and contained little assessment of subject matter knowledge. The research was also generic, rather than content specific, in its description of effective teaching, teacher knowledge, and teacher behaviors. Shulman and his colleagues attempted to rectify this situation through their research program concerning knowledge growth in teaching, which focused on teachers' subject matter knowledge.

As a result of these events, in the mid-1980's new frameworks regarding teacher knowledge began to emerge. The most influential framework, in terms of its impact upon future frameworks, research, and teacher education (Hill et al., 2005), was Shulman's (1986) framework of teachers' content knowledge. This framework described three components of teachers' content knowledge: subject matter, pedagogical, and curricular. Subject matter content knowledge referred to the amount and organization of knowledge in the mind, including knowledge that something is true and why it is true. Pedagogical content knowledge, a new construct to the field, went “beyond knowledge of subject matter per se to the dimension of subject matter knowledge for teaching” (p.9). Curricular knowledge involved knowing both the horizontal and vertical curriculum for a topic and its associated curriculum materials. With this framework, Shulman called for attention to be shifted back to teachers’ content knowledge.

Shulman’s (1986) idea of pedagogical content knowledge has been significant to the work of understanding teacher knowledge. Its introduction gave a new conceptualization to the content and nature of teachers’ subject matter knowledge in ways that the previous focus on measuring teacher subject matter knowledge did not (Ball & Bass, 2003). It proposed that there was a distinction between the ways content needed to be understood in order to teach and the ways an average adult understood content; it suggested that expert knowledge of mathematics may be inadequate for teaching the subject.

Building upon Shulman’s (1986) work, other frameworks regarding teacher knowledge were developed. Peterson (1988) kept Shulman’s three components and added a fourth: teachers’ meta-cognitional knowledge. Grossman (1990) reorganized teacher knowledge into four categories: subject matter knowledge, pedagogical content knowledge, general pedagogical knowledge, and context knowledge. Ball (1991) described teachers' mathematics understanding as having two components: substantive knowledge or knowledge of mathematics, and knowledge about the nature and discourse of mathematics. Ball described properties of this understanding, including that it must be connected, rather than compartmentalized; and explicit, rather than tacit. Fennema and Franke (1992) developed a model of teachers' knowledge that centered on teachers' knowledge as it happens in the context of the classroom. It included the components of knowledge of mathematics, pedagogical knowledge, and knowledge of learners' cognitions in mathematics, all in an interactive, dynamic relationship with teachers' context-specific knowledge. All of these frameworks refocused attention on the centrality of subject matter in teachers' knowledge.

The situated cognition perspective and the growing respectability for qualitative research during this time period led to the emergence of field-based, qualitative studies regarding teachers' mathematical knowledge that utilized the new teacher knowledge frameworks of the time (e.g., Ball, 1990; Simon & Blume, 1994; Thompson & Thompson, 1994). A dominant methodology was case study, including interviews of teachers. These studies were focused on teacher knowledge, rather than its characteristics as seen in the previous time period, and were narrowed to specific mathematical topics. In these respects, these studies serve as precursors to the current study. A weakness of most of these studies was their lack of a practice-based approach, which meant they were not truly studying teacher knowledge in the context in which it was used and needed.

A new type of research study, a self-study of professional practice, emerged in the work of Lampert (1990) and Ball (1993). These mathematics educators each documented their teaching of a full school year with the intention of showing the specific relationships between subject matter knowledge and teaching mathematics. Their work brought out the complexity of mathematics knowledge needed by mathematics teachers at the elementary level and helped inform current frameworks regarding teacher knowledge for teaching mathematics.

Current Frameworks and Related Research

Recent frameworks regarding teacher knowledge are content specific (i.e., developed with specific reference to mathematics teachers) and emphasize the existence of teachers' mathematics as a unique area of content knowledge (Ball & Bass, 2000a, 2003; Ball & Hill, 2005; Ma, 1999; Usiskin, 2000). Usiskin argues that the specialized content teachers need to know should be considered a branch of applied mathematics deserving of its own field called "teachers' mathematics." It includes mathematics and pedagogical content knowledge, with components such as mathematical generalizations of the content of K-12 school mathematics and analysis of concepts.

Liping Ma (1999) developed the notion of profound understanding of fundamental mathematics as a result of her study comparing the mathematical knowledge of Chinese and United States elementary teachers. This type of content knowledge is considered profound because of its depth, breadth, and thoroughness. The term “fundamental mathematics” is defined as mathematics that is foundational for advanced mathematics, primary in its inclusion of important concepts, and elementary because it is introduced at the beginning of students’ formal learning of mathematics. This type of understanding is supported by the structure of mathematics as a subject and is composed of well-developed, interconnected knowledge packages. A teacher with profound understanding of fundamental mathematics teaches in a manner that makes connections among mathematical concepts and procedures, promotes multiple perspectives of ideas or solutions, emphasizes basic ideas, and has longitudinal coherence.

The perspectives of Ma (1999) and Usiskin (2000) are in accord with the selected framework for this study, which is a newer framework of teacher knowledge proposed by Ball and her colleagues (Ball & Bass, 2003; Ball & Hill, 2005) based upon analyses of elementary mathematics teaching and its mathematical entailments (Ball & Bass, 2000a, 2000b, 2003). In this framework, teachers’ knowledge is divided into two parts: subject matter knowledge and pedagogical content knowledge. Within subject matter knowledge, two types are specified. One is common content knowledge, the understanding that we would expect the average person who knows this content to have. A second type is specialized content knowledge for teaching. Knowing mathematics for teaching requires a kind of depth and detail that goes well beyond what is needed to carry out procedures correctly, and includes the ability to evaluate student claims and solutions, analyze student errors, and choose mathematical representations (Ball et al., 2005). Pedagogical content knowledge is also composed of two components: knowledge of students and content, and knowledge of teaching and curricula. This framework was described in greater depth in the theoretical perspective section.

Research related to Ball and her colleagues' framework for mathematics teachers' knowledge (Ball & Bass, 2003; Ball & Hill, 2005) has been recently conducted. Ball et al. (2005) examined the work of teaching elementary mathematics through a practice-based approach to create and test a hypothesis about the mathematical knowledge for teaching the elementary school mathematics topics of number and operations; and patterns, functions, and algebra. They began by defining a domain map for each of these topics based upon their previous study of elementary school teaching (Ball & Bass, 2000b). Next, they assembled a team of experts, including mathematicians, mathematics educators, and classroom teachers, to write assessment items to test teachers' common content knowledge and specialized content knowledge of the aforementioned topics. The items posed questions related to situations that teachers deal with in their work in the classroom. Over 250 multiple-choice items designed to measure teachers' mathematical knowledge for teaching were developed.

Next a research study was done to test the items created to measure teachers' knowledge of mathematics for the elementary school level (Hill et al., 2004). Using factor analysis, evidence for the existence of a specialized content knowledge of mathematics for teaching that is not equivalent to common mathematical knowledge was found. This type of specialized knowledge was found to be partially domain specific and consists of both general knowledge of content and knowledge in more specific domains such as number and operations.

Hill et al. (2005) empirically established that mathematics teachers' subject matter knowledge, including common content knowledge and specialized mathematical knowledge for teaching, was significantly related to student achievement gains in the first and third grades after controlling for important student- and teacher-level covariates. They measured teachers' subject matter knowledge with a multiple-choice test composed of items developed and analyzed during the previous phases of the research program. They measured students' mathematics achievement gains by their Terra Nova scores. In addition, they collected information about the students (e.g., socioeconomic status) and teachers (e.g., years of experience) for use in predicting the size of student gain scores.

Their analysis found that teachers' performance on the test, which measured their knowledge for teaching, significantly predicted the size of student gain scores after accounting for differences in other variables such as student socioeconomic status, student absence rate, teacher experience, and average length of mathematics lessons. In particular, students of teachers in the top quartile showed gains in their scores that were equivalent to that of an additional two to three weeks of instruction. This research brings us back full circle to where the educational production function studies began, relating the knowledge of teachers to their students' achievement. But whereas many of those studies did not find a significant relationship between teacher knowledge and student achievement (Begle, 1979), this one did perhaps because of its redefinition and measurement of teacher content knowledge.

Teacher Knowledge of Statistics

Frameworks regarding the knowledge secondary teachers specifically need in order to teach statistics have not been created. Gal's (2004) model of statistical literacy partially speaks to this knowledge in that it points out the importance of teachers being statistically literate in order to teach statistics in a way that aims for students to be statistically literate. In Gal's model, there is a knowledge component composed of five cognitive elements: literacy skills, statistical knowledge, mathematical knowledge, context knowledge, and critical questions.

There is also a dispositional component, including beliefs, attitudes, and critical stance, which is not relevant to the current discussion. This study focused on the elements of statistical knowledge, mathematical knowledge, and context knowledge because they are most closely related to the type of content knowledge referred to in Ball and colleagues' (Ball & Bass, 2003; Ball & Hill, 2005) framework of mathematics teachers' content knowledge.

Gal (2004) identified five components of the statistical knowledge base: knowing why data are needed and how they can be produced; familiarity with descriptive statistics' elementary terms and ideas; awareness with basic terms and ideas about graphs and tables which display data; understanding of basic probability; and knowledge of how statistical conclusions, including inferential ones, are reached. An adult who has all five parts of the statistical knowledge base possesses the knowledge needed to function in today's world.

Gal (2004) is less specific about the content of the mathematical knowledge base and cautions that expectations regarding the mathematical background needed to understand statistical ideas have been changing in recent years. Number sense, how an arithmetic mean is computed, and percentages are some examples Gal provides in his description of mathematical knowledge. However, he does not create a definitive list of this area.

Context or world knowledge base is the last of the three knowledge components of Gal's (2004) framework which is relevant to this study. He argues that a correct interpretation of a statistical message requires that the reader possess an ability to place the message in a context and to access his or her knowledge of the world. With this knowledge, the reader can critically reflect on the statistical message and be able to understand the implications of the findings or the reported numbers in the message.

Burgess (2007) proposed a framework to describe the components of teacher knowledge regarding statistical thinking and investigating needed by elementary teachers to teach statistics through investigations. The framework was presented as a matrix. The columns of the matrix referred to the four types of knowledge needed by teachers as identified by Ball and Hill (2005): common knowledge of content, specialized knowledge of content, knowledge of content and students, and knowledge of content and teaching. The elements of statistical thinking and empirical inquiry described by Wild and Pfannkuch (1999) defined the rows of the framework. These were the need for data, transnumeration, variation, reasoning with models, integration of statistical and contextual thinking, investigative cycle, interrogative cycle, and dispositions. Evidence of the need for all the types of knowledge identified in the matrix framework was found with the exception of dispositions and the need for data. Burgess' research represents an initial step toward the development of a general framework of teacher knowledge for teaching statistics.

Limited research has been conducted that describes teachers' knowledge of specific statistical topics (e.g., Casey, 2005; Makar & Confrey, 2004; Mickelson & Heaton, 2004; Thomas, 2002). Thomas (2002) described the knowledge that preservice elementary teachers have of data analysis topics. Makar and Confrey (2004) studied four in-service secondary teachers' statistical reasoning as they compared two distributions. A case study of a third-grade teachers' knowledge of data and distribution as used in her teaching was done by Mickelson and Heaton (2004). Casey (2005) studied secondary in-service teachers' conceptions of statistical association through task-based interviews. All of these studies found that teachers' knowledge was less than complete, missing some important components, and/or functioning at a surface level of understanding. A weakness of these studies is that they do not help investigate the type of knowledge needed by teachers and how they use it in their work (Ball, 2002). Simply describing teachers' statistical knowledge does not bring the field any closer to understanding what content knowledge teachers need to be high quality teachers.

Sorto (2004) conducted a study regarding the content knowledge needed for teaching data analysis at the middle school level. As with the current study, Sorto used the framework of teacher knowledge of Ball and her colleagues (Ball & Bass, 2003; Ball & Hill, 2005) for her theoretical perspective. Written documents, including policies, state and national standards regarding teacher knowledge, and a teachers' guide to a middle school mathematics curriculum, were analyzed to identify the important aspects of statistical knowledge for teaching data analysis to middle school students. The central aspects were data representation, particularly graphical representations like histograms and line plots, and measures of center and spread. Prospective middle school teachers were then assessed through a written test regarding their content knowledge of data analysis. Results of this assessment showed that their common content knowledge was better than their specialized content knowledge for teaching data analysis, including judging students' comments and identifying mistakes made by students.

A weakness of Sorto's (2004) study was its use of written documents to determine teachers' knowledge needed to teach data analysis at the middle school level. These resources are indirectly connected with what actually happens in the teaching of data analysis at the middle school level: while a state standard may say that teachers need to know, for example, measures of spread, the standard does not provide evidence of what knowledge of measures of spread teachers need in their practice or how they have applied it in practice. Ball et al. (2001) pointed out that subject matter knowledge is better manifested in the actual practice of teaching and studied with a practice-based approach, which was the approach taken by the current study.

Burgess (2007) took a practice-based approach to study the knowledge for teaching statistical investigations at the elementary level. Through videotapes and interviews he documented four elementary teachers as they taught four lessons involving statistical investigations. The data were analyzed using his aforementioned proposed framework for teacher knowledge. His analysis process included classifying the knowledge called for by each teaching incident as one of the types described by the framework. He concluded that the presence of evidence for all cells of the framework (with the exception of those associated with two statistical thinking elements absent due to the nature of the chosen investigations) showed that the framework accurately represented the necessary knowledge for teaching statistical investigations at the elementary level. I also used a practice-based approach in the current study but focused on the secondary level where studies regarding teacher knowledge of statistics are minimal and research regarding the content knowledge teachers use for teaching statistical association at the secondary level is non-existent. Burgess' methodology was also different from my study. He started with a proposed framework for teacher knowledge, then sought to validate it using data from practice. I started from scratch, using the data documenting practice to generate a description of the subject matter knowledge for teaching statistical association.

Statistical Association

Student Learning Objectives

Student learning objectives for statistics at the secondary level have been developed by the National Council of Teachers of Mathematics (2000) and Franklin et al. (2005). Together these objectives could be considered a description of the minimal common content knowledge a teacher should have to teach statistical association at the secondary level. Those objectives pertaining to statistical association are summarized below.

In the data analysis and probability standard for grades 9-12, the National Council of Teachers of Mathematics (2000) articulated their learning expectations for all secondary students. Relevant to statistical association, these include the following:

- Understand the meaning of measurement data and categorical data, of univariate and bivariate data, and of the term variable;
- Understand histograms, parallel box plots, and scatterplots and use them to display data;

- For bivariate measurement data, be able to display a scatterplot, describe its shape, and determine regression coefficients, regression equations, and correlation coefficients using technological tools;
- Display and discuss bivariate data where at least one variable is categorical;
- Identify trends in bivariate data and find functions that model the data or transform the data so that they can be modeled. (p.325)

A Curriculum Framework for PreK-12 Statistics Education (Franklin et al., 2005) was written as a result of a strategic initiative of the American Statistical Association in response to the CBMS' (2001) The Mathematical Education of Teachers (Franklin, 2004). Their learning objectives were described for three levels, A, B, and C, which are based upon development rather than age (Franklin et al.). Since these levels roughly parallel NCTM's (2000) grade bands (Franklin), objectives for levels B and C are considered relevant to the secondary level.

Students at level B should be able to do the following:

> quantify the strength of association between two variables, develop simple models for association between two numerical variables, and use expanded tools for exploring association including contingency tables for two categorical variables, time series plots, the quadrant count ratio as a measure of strength of association, and simple lines for modeling association between two numerical variables. (p.36).

When interpreting results, students are expected to know basic interpretations of measures of association and be able to start distinguishing between association and cause and effect.

At level C, students analyzing data should be able to “summarize numerical and categorical data using tables, graphical displays, and numerical summary statistics; recognize association between two categorical variables; and describe relationships between two numerical variables using linear regression and the correlation coefficient” (p.60). For interpreting results, “students should understand the meaning of statistical significance, the difference between statistical significance and practical significance, the role of p-values in determining statistical significance, and interpret the margin of error associated with an estimate of a population characteristic (parameter)” (p.60). Taking the objectives for both levels B and C together, the student learning objectives developed by Franklin et al. (2005) for the secondary level can be determined.

Student Conceptions of Statistical Association

In this section, I review the research that has been done regarding students’ preconceptions (including knowledge and beliefs) and learning of statistical association, both with two quantitative variables and two categorical variables. The review is mostly limited to studies involving secondary level students, but pertinent studies with students at the middle school and university levels are included if they describe students' understanding of concepts considered part of the secondary curriculum according to the National Council of Teachers of Mathematics (2000) or Franklin et al. (2005).

Batanero et al. (1996) identified 17 and 18-year-old students' preconceptions about association of two categorical variables prior to instruction. They found that three factors influenced how difficult a task was for the student: the size of the two-way table, the students' previous beliefs about the relationship between the variables, and the lack of understanding about inverse association. In general they found most students had a good intuitive ability to judge association in 2 x 2 tables; however, most students could not generalize their strategy for a 2 x 2 table to make a correct judgment about a 3 x 3 table. Another influence on students' judgments was their previous expectation concerning the association between the variables. If the data provided did not agree with their beliefs about the relationship between the variables, then the students often disregarded the data. Many students confused an inverse association between variables with independence, believing that the only type of association that could exist was a direct or positive one. Students' lack of proportional reasoning was also identified as a general reason they had trouble judging association.

Student strategies for judging association in the Batanero et al. (1996) study were classified as correct, partially correct, or incorrect. The three correct strategies were comparing all conditional relative frequency distributions, comparing one conditional relative frequency with the marginal frequency, and comparing the odds. Partially correct strategies were comparing absolute conditional frequency distributions rather than relative frequency distributions and comparing the sum of the frequencies in the diagonals. The use of numbers rather than proportions was a prevalent strategy of students prior to instruction.

Three incorrect conceptions of association were identified (Batanero et al., 1996). In the determinist conception, students believed that an association meant all cases must show an association with no exceptions. These students believed that the cells in the two-way table that did not agree with the association should have zero frequency. The second incorrect conception was a unidirectional conception of association where students believed dependence occurred only when it was direct. This could be explained by the tendency of students to give more relevance to positive cases than negative cases that confirm a given hypothesis. In the localist conception of association, students looked at part of the data to determine if an association existed, often only looking at the cell with the highest frequency or at only one conditional distribution. In follow-up research by Batanero, Estepa, and Godino (1997) as reported in Konold and Higgins (2003), they gave secondary students approximately thirty hours of instruction on group differences and association. Before instruction, the students solved about 20% of such problems correctly. After instruction, the percent of problems correct increased slightly to 30%.

Research has been done with both secondary and university students regarding their understanding of association of two quantitative variables (Estepa & Batanero, 1996; Estepa & Sanchez-Cobo, 2003). Before instruction, secondary students' strategies for determining association included looking to see if the pattern held true for each data point (consistently increasing or decreasing), basing their decisions on one part of the data, or relating their decisions to their personal beliefs about whether an association existed between the variables (Estepa & Batanero). There also needed to be a strong correlation between the variables before the students detected it.

In a teaching experiment, Estepa and Sanchez-Cobo (2003) worked with first-year university students to build their meaning of statistical association. Initially these students wanted to solve problems by comparing isolated values in the two data sets. With intervention, these students dropped this behavior and used the complete data sets to determine association. They also came to understand that judging association should be done in terms of intensity rather than existence, overcoming the determinist misconception. As with the categorical case, students had a difficult time understanding inverse association as compared to direct association. They did not understand how to interpret a negative correlation coefficient and failed to understand the concept by the end of the teaching experiment.

Estepa and Sanchez-Cobo (2003) described their earlier work (Estepa and Sanchez-Cobo, 1998) in which they examined secondary school textbooks' presentation of correlation and regression. They found that nearly all the examples and exercises in the textbooks had strong positive correlations. Very few examples or exercises displayed negative correlations, moderate correlations, or correlations near zero, which the authors claimed could lead students to believe only strong positive correlation exists.

Teacher Conceptions of Statistical Association

My own work (Casey, 2005) is the only study I am aware of regarding teachers' conceptions of statistical association. I interviewed eight secondary mathematics teachers, asking questions regarding their conceptions of statistical association, their ideas about the nature and role of statistical association in the secondary mathematics curriculum, and their perceptions about the instruction of the topic at their schools. A portion of the interview was task-based, where each of the teachers was given five problems and asked to discuss whether an association existed between the variables for each of the different problems. Three of the problems involved two quantitative variables, such as students' weights and the weights of their backpacks, and the other problems involved two categorical variables, such as whether or not a person smoked or had a bronchial disease. They were asked to talk through each problem as if they were teaching a secondary student with no prior background of the topic.

Through analysis of their responses to the task-based portion of the interview, I determined that the participants had moderate knowledge of how to solve problems involving statistical association. There was a surprising prevalence of judgments based on the context of the variables in the problem, even with all of the question prompts stating to use the information presented in that problem to make a decision. Many of the teachers either struggled to answer the problem analytically, therefore, using a context-based judgment to make a decision, or they were unable to temporarily suspend their prior beliefs to look at the data and examine what association might be present.

Although analytic-based judgments were more often correct than context-based ones, some of the judgments were incorrect. Mistakes included incorrectly visually analyzing a scatterplot with no association between the variables, using numbers rather than percents in analyzing categorical data, believing that an association needs to be quite strong before it can be deemed to exist, and believing that if there was an association between two variables it must be from a cause-effect relationship. Analytic-based judgments used multiple methods for both quantitative and categorical variables. For quantitative data, the most common method was visually looking for the presence of a general trend. The analytic methods for analyzing categorical data focused on the numeric method of finding conditional percents.

The teachers' comments during the task-based portion of the interview revealed two interesting points. First, multiple teachers remarked that they did not understand why they would show students an example of data that were not associated. Second, seven of the participants made a comment that they were not comfortable solving the problem(s) or confident with their answer(s). These two points speak to current teachers' understanding of statistical association.

Relationship to This Study

This study focused on the subject matter knowledge for teaching statistical association. Thus, to guide the development of this study it was important to examine the literature in mathematics education on two topics: teacher knowledge and statistical association.

My review of the literature related to teacher knowledge led me to the decision to use the framework of teacher knowledge put forth by Ball and her colleagues (Ball & Bass, 2003; Ball & Hill, 2005) as the perspective adopted for my study. However, there is a difference between mathematics and statistics (Moore, 1990) and thus I needed to adapt the framework so that it would work for my study. Gal's (2004) model of statistical literacy, which includes a knowledge component, provided me with a way to make this possible. Within the knowledge component of statistical literacy, Gal included statistical knowledge, mathematical knowledge, and context knowledge. These three components of knowledge were then adopted for this study to make the teacher knowledge framework of Ball and her colleagues (Ball & Bass, 2003; Ball & Hill, 2005) applicable to the study of teacher knowledge for teaching statistical association.

A review of the research pertaining to statistical association, including its components and the learning of the topic, was conducted to aid the description of the study and its conduct. Detailed descriptions of learning objectives concerning statistical association for students at the secondary level provided a list of topics that should be documented as they were taught by the participating teachers. This was taken into consideration as the participating teachers and their classes were chosen.

Reading research accounts regarding students' conceptions of statistical association helped me prepare for my observations of classrooms. Not only did it help me understand the type of thinking students exhibited, but it also spoke to the need for knowledge on the part of the teacher to be able to respond appropriately to students as they learned the topic. It also documented that students come to the topic with flawed conceptions and that it is difficult to correct those. My study (Casey, 2005) of eight teachers' conceptions of statistical association provided me with insight regarding current teachers' understandings of the topic and provided me with the impetus for this study.

CHAPTER III

RESEARCH DESIGN AND METHODOLOGY

Introduction

Chapter III provides the particulars as to how the study was designed and carried out in order to investigate the subject matter knowledge for teaching statistical association at the secondary school level. Information is included regarding the pilot study; selection of the practice-based approach as a methodology; choice of participating teachers; procedures, including use of observations and interviews as methods; and analysis of the data.

Pilot Study

A pilot study was done to determine how to effectually use a practice-based approach to study the subject matter knowledge needed for teaching a statistical topic (in this case sampling and design of experiments) at the secondary level. This included determining sample selection, data collection, and analysis procedures; identifying ways to make the data collection process most productive while being unobtrusive; and honing my skills as a qualitative researcher of teachers' subject matter knowledge. I focused on understanding and refining the process of conducting practice-based research.

This research was done over a period of three weeks with a high school mathematics teacher and her Advanced Placement Statistics class during a unit on statistical sampling and design of experiments. I observed twelve class sessions over a three-week period and interviewed the teacher following each class session regarding her planning, teaching, and assessment of the lesson. All of the class sessions and teacher interviews were audio-taped, and portions of the audiotape relevant to my research were later transcribed for further analysis. I also took fieldnotes during the class sessions and interviews and wrote down my reflections upon the research process each day.

The results of the pilot study had implications for the design of the present study. The inclusion of an introductory session with the participating teachers, the use of a voice recorder with a microphone, and the use of an interview protocol of questions were all done due to the findings of the pilot study. These and other implications from the pilot study will be highlighted in the following sections.

Methodology

Since the purpose of this study was to describe the subject matter knowledge for teaching statistical association at the secondary level, a practice-based approach was used to study the work of teaching statistical association at the secondary level directly. Using a practice-based approach meant that secondary teachers were studied on the job as they taught the topic of statistical association through observations of class sessions and interviews with each teacher following their observation. Records of practice, including observation transcripts, interview transcripts, copies of student work, and textbook pages, were used to document the work of the participating teachers. These records of practice then became the focus of the empirical analysis done for the study. This type of approach was most appropriate and thorough for understanding the subject matter entailments of teaching and to analyze where subject matter knowledge is used in that work (Ball & Bass, 2000a, 2003; Ball et al., 2001).

Other common approaches to studying the subject matter knowledge for teaching have strong limitations (Ball et al., 2001; Stylianides & Ball, 2004). The examination of policy documents that set standards for teacher education or school mathematics provides information regarding what policy-makers consider important content knowledge for teaching, but the content of these documents is largely based on the beliefs of the authors rather than disciplined inquiry (Ball et al.). Also, there is often a lack of connection between the intended content described by curriculum policy documents (e.g., NCTM, 2000) and the implemented content in the classroom (Stylianides & Ball). Policy documents that set standards for teacher education (e.g., CBMS, 2001) contain lists of mathematical topics important to teaching, but they are often vague, and understanding how the specific knowledge of these topics becomes integrated into the practice of teaching is more difficult to understand (National Research Council, 2001). The same weaknesses can be identified regarding the approach of performing a textbook analysis of college mathematics textbooks developed for mathematics courses taken by pre-service teachers (Stylianides & Ball).

Another approach to studying the subject matter knowledge needed for teaching mathematics is to analyze the research regarding teachers' content knowledge. This approach uses two objects of analysis: the topics selected by the researchers as important for study, and the teachers' knowledge of specific mathematical ideas (Stylianides & Ball, 2004). Ball et al. (2001), in reviewing the research literature on teachers' mathematical knowledge, pointed out that researchers have studied the knowledge they thought would matter in teaching, but the choices of topics have not always had practice-based justifications. Studies of what teachers know and do not know have been criticized for their inability to advance the field's understanding of what teachers need to know and how to use that knowledge in their work (Ball, 2002; Stylianides & Ball).

Because mathematics instruction in the United States is largely driven by textbooks (Romberg & Webb, 1993), an analysis of school mathematics textbooks could be used to propose the content that teachers are expected to teach and therefore need to know (Stylianides & Ball, 2004). At best, this approach can indicate the minimum knowledge base that teachers need to have, but it does not help answer the question of how much more or what depth of content knowledge teachers use in effective instruction of that content (Ball & Bass, 2000a; Stylianides & Ball). There is also the issue of the mismatch of the intended and implemented curriculum (Flanders, 1994) which suggests another weakness of this approach.

A major limitation of all of these approaches is that they use as data sources items that are indirectly connected with the practice of teaching mathematics (Stylianides & Ball, 2004). In these approaches, the data sources are used to infer the content knowledge teachers use in their practice, but none of them actually view the work of teaching to determine if their inference is accurate. That is the major advantage of the practice-based approach over all other approaches: it views the subject of interest in its natural context: the teaching that utilizes it (Ball et al., 2001). This approach is necessary to view the less exposed subject matter requirements for teaching, requirements that could not be seen by studying the content students are working on outside of the classroom context (Ball & Bass, 2000a). This approach also connects the theoretical constructs that result from it to the practice of teaching, an outcome desired in mathematics education research (Romberg & Carpenter, 1986).

The practice-based approach is not without its limitations. The social complexity of the classroom and the interactive relationship between knowledge and beliefs makes it difficult to detach and isolate teachers' content knowledge for investigation (Stylianides & Ball, 2004). As outsiders looking in, all of the analysts had to envision what subject matter knowledge the teacher needed to do his or her job rather than experiencing the teaching themselves. Studying this from the outside, the analysts may have noticed things that the teacher would not have because these things had become routine to the teacher (Merriam, 1998) or the teacher was not cognizant of them with everything else he or she had to attend to in the teaching process. On the other hand, it also meant that an analyst may not have been able to identify all of the subject matter knowledge components called for by a teaching incident because the analyst did not live the experience.

This study involved case studies of three secondary mathematics teachers as they taught the topic of statistical association. A case study is characterized by its intensive description and analysis of a single unit or bounded system (Smith, 1978, as cited in Merriam, 1998) and is an approach particularly suited to studies in which it is impossible to separate the phenomenon's variables from the context, and in which the interest is in a process (Merriam). Since this study took a practice-based approach to study a process, the work of teaching, in order to focus upon the subject matter knowledge needed by teachers in that process, the case study was an appropriate methodology. Because it involved multiple cases, this study was a collective case study (Stake, 1995). "The more cases included in a study...the more compelling an interpretation is likely to be" (Merriam, p.40). The validity of the conclusions of the study may be limited due to the dependence of the study upon the particular lessons observed. It is possible that in a different classroom, a class session on statistical association could occur that calls for different aspects of content knowledge than those identified by this study. However, through the inclusion of three teachers using different curricula in different settings, it is reasonable to assume that the conclusions represent the subject matter knowledge needed by most secondary teachers as they teach statistical association.

Restricting the study to a particular school level was essential due to the complexity of the topic of subject matter knowledge for teaching and its changing components at different levels of schools (Stylianides & Ball, 2004). This study was limited to the secondary school level. For the viability of this study, it was also necessary to restrict the inquiry to a particular topic in statistics: statistical association.

Participants

I selected a purposeful sample of three secondary mathematics teachers in the Chicago area for this study. Since generalization from a statistical perspective is not a goal of qualitative research, random sampling was not necessary or justifiable for this study (Merriam, 1998). Instead, a purposeful sample was chosen to select information-rich cases for in-depth study (Patton, 1990). There were two criteria used in the selection of the participants: high quality teaching and the commitment to teach statistical association during the 2006-2007 school year.

Different approaches to teaching bring with them different knowledge demands for the teacher. Study of teaching practice where students follow memorized rules leads researchers to speculate about basic requirements regarding teacher knowledge, but study of teaching practice which involves students in situations of justification can result in conclusions describing standards of excellence (Stylianides & Ball, 2004). It was not my intent in this study to describe the subject matter knowledge needed for teaching statistical association in general; rather, I hoped to describe the subject matter knowledge needed for high quality teaching of this topic at the secondary level.

For this study, in agreement with the perspective of NCTM (2000), high quality teaching was defined as that which fosters a view of mathematics as sense-making activity (Yackel & Hanna, 2003) and aims for students to learn to think statistically, reason statistically, and become statistically literate citizens. Teaching students to think statistically means helping students gain statistical knowledge and learn when and how to apply it (delMas, 2004). Instruction geared to teach statistical reasoning asks students to explain why results were obtained or why a conclusion is justified. Teaching students to be statistically literate means helping them learn how to interpret and critically evaluate statistical information, data-related arguments, or stochastic phenomena, and communicate their reactions to such items to others (Gal, 2004). There has been an oft-repeated and intensifying call for educators to emphasize these three components--statistical reasoning, statistical thinking, and statistical literacy--over the past ten years (Ben-Zvi & Garfield, 2004), and thus in this study concerning statistics education, I will use these components to help define the goals of high quality teaching. Further support for these criteria came from the pilot study, where it became apparent that the most informative data came from situations where the students were involved in sense-making activities involving statistical topics as opposed to solely memorizing definitions or procedures. In these situations in particular, the demands upon the teacher's subject matter knowledge came to the forefront.

The second criterion was that the topic of statistical association was taught in at least one of each participating teacher's courses during the 2006-2007 school year, the year this research took place. The second criterion was necessary to fulfill the purpose of the study to describe the subject matter knowledge for teaching statistical association at the secondary level by studying teachers as they taught the subject.

I began the search for participating teachers by contacting mathematics department chairs at Chicago area secondary schools to gather information regarding their mathematics curricula and teachers. The schools contacted were chosen because they offered at least one course that thoroughly covered the topic of statistical association. Comparisons to lists of student objectives regarding statistical association from NCTM (2000) and Franklin et al. (2005) were made to assess the coverage of the topic within each curriculum. All of the participating teachers were recommended by their department chairs because they met the criteria I had established for selection of participants.

Next I contacted the recommended teachers to determine their willingness to participate in the study, confirm they would be teaching statistical association during the 2006-2007 school year, and to arrange an initial visit to observe their teaching. During the observations of their teaching, I looked for evidence of high quality teaching as defined by this study. This included providing opportunities for students to explain why results were obtained, providing activities where students developed their own methods for solving problems, or requesting that students critically evaluate information. These observations along with discussions with the teachers and department chair enabled me to establish that all three participating teachers met the first criterion of high quality teaching for the sample.

Participation in a study is a two-way street, as I not only needed to select the teachers to participate but the teachers needed to decide if they wanted to participate. I gave the participants a full description of the study and an opportunity to ask questions of me. The study was described as an opportunity for both the teachers and me to learn and grow professionally by focusing our attention on the content knowledge involved in teaching the topic of statistical association. I conveyed to the teachers that the intent was not to document their content knowledge or compare what it is to what it should be; instead, the study was meant to use the practice of teaching statistical association as a means to describe what content knowledge is needed by teachers during their work. I also told the teachers about the use of member checks later in the research process, where they would be given the opportunity to provide me with feedback regarding the transcripts and findings of the study. I provided participants full disclosure regarding the intent of the study and its procedures, and by agreeing to participate in the study, the teachers agreed to participate in this way. Through mutual agreement the three teachers were selected for the study.

The first participant was Mr. Glass (all names used are pseudonyms), a veteran white male teacher of secondary mathematics with approximately thirty years of teaching experience. He taught the topic of statistical association in his Pre-Calculus/Statistics course composed of juniors attending a high school in a suburb of Chicago. The textbook used for the course was *Functions, Statistics, and Trigonometry* (Senk et al., 1988), which was written by members of the University of Chicago School Mathematics Project. This textbook series is known for its extensive use of the graphing calculator, its emphasis upon students reading the textbook, and its problem sets consisting of three sections: covering the reading, applying the mathematics, and review. The relevant material for this study was taught in the unit titled “Functions and Models” which lasted sixteen days.

Mr. Tablet, another participating teacher, is a white male with approximately twenty years of teaching experience. He taught the Advanced Placement Statistics course to juniors and seniors at a high school in the greater Chicago area. The College Board-sponsored Advanced Placement Program enables high school students to take college-level courses and exams to earn college credit (College Board, 2006). The Advanced Placement Statistics course is intended to be equivalent to a one-semester non-calculus-based introductory statistics course. Its function is to teach students the methods and concepts for collecting, analyzing, and drawing conclusions from data. Mr. Tablet used a textbook designed for Advanced Placement Statistics, *The Practice of Statistics: TI-83 Graphing Calculator Enhanced* (Yates, Moore, & McCabe, 1999), for this course.

He was involved in this study during his teaching of two units. The first unit involved the study of relationships between two variables, including both bivariate quantitative data and bivariate categorical data, and lasted sixteen days. This covered chapters three and four of *The Practice of Statistics: TI-83 Graphing*

Calculator Enhanced (Yates et al., 1999). The next relevant unit began approximately five months after the first unit ended and covered inferential procedures concerning association of statistical variables. This included textbook chapter thirteen on chi-square procedures and chapter fourteen on inference for regression procedures and lasted ten days.

The final participating teacher, Ms. Tuck, is a white female who has taught for approximately five years at a City of Chicago magnet school for college-preparatory students. At this school the mathematics curriculum is the Interactive Mathematics Program, whose development was sponsored by the National Science Foundation in an effort to create curricula that embodied the NCTM (1989) *Curriculum and Evaluation Standards for School Mathematics.* It is an integrated four-year curriculum organized into problem-based units which students work through together in groups (Key Curriculum Press, 2008).

I observed Ms. Tuck's Algebra class for freshmen and her Geometry class for sophomores as they studied the topic of statistical association. The concepts of association of quantitative variables and linear regression were included in a unit titled "The Overland Trail" (Fendel, Resek, Alper, & Fraser, 2004) in the Algebra course and taught over five days. "Is There Really a Difference?" (Fendel, Resek, Alper, & Fraser, 1998), part of the Geometry curriculum, delved into the topic of association of categorical variables and the chi-square test statistic. Three days were observed during the teaching of this unit.

Procedures

Introductory Session

I held an introductory session with each participating teacher in August or September of 2006 in advance of the 2006-2007 school year during which the data were collected. The purposes of this session included helping the participating teacher understand the function of the study and the type of data that would be needed in order to answer the study's research questions; developing in the teacher an understanding of the teacher knowledge framework (Ball & Hill, 2005); developing rapport with the teacher; and determining the logistics of the study.

I realized the importance of these preliminary activities from the pilot study. The participating teachers needed to be in tune with the study's purpose as well as my focus as an observer and interviewer in order for the methods to be most informative. For example, during the initial teacher interviews of the pilot study, the teacher continued to talk and reflect on pedagogical knowledge rather than subject matter knowledge. While there is not always a clear distinction between the two types of knowledge, by working with the teachers to help them understand the study's theoretical framework, I hoped that they would be more aware of their use of subject matter knowledge in their work. It was also important for me, as the researcher, to establish rapport with participants (Creswell, 1998), and these sessions helped me work towards the development of a productive relationship between the participating teachers and myself.

I began each introductory session with an overview of the study, including its research question, definitions of terms used in the research question, and methods. Throughout the overview I emphasized the difference between subject matter knowledge for teaching and pedagogical content knowledge for teaching, using the graphical representation of Ball and Hill's (2005) teacher knowledge framework and examples from the classroom to emphasize the distinction. Next I asked the teacher to read and sign the consent form to participate in the study. Finally, I discussed the logistics of the study with the teacher and made sure each teacher knew my expectations of him or her, including completion of consent forms by all participating students and their parents. A copy of the handout provided to the participating teachers at the introductory session is in Appendix A.

Methods

Observation

In accordance with a practice-based approach, I observed the participating teachers during their class sessions involving the teaching of statistical association. The focus of this study, the subject matter knowledge for teaching statistical association, is best studied in the context of the classroom in which it functions

(Stylianides & Ball, 2004). Each teacher notified me ahead of time when they would be teaching about statistical association, and I planned my observations so that I could be there for all relevant days. By the end of the study I observed sixteen days with Mr. Glass, twenty-six days with Mr. Tablet, and eight days with Ms. Tuck. Figure 1 presents a summary of the days observed and the topic(s) taught that day.

Mr. Glass **Day 1**	Association between quantitative variables: scatterplot, Least-squares regression line (LSRL), correlation
Day 2	Association between quantitative variables: explanatory and response variables, correlation
Day 3	Association between quantitative variables: best fit line, LSRL, correlation, coefficient of determination, residual
Day 4	Association between quantitative variables: LSRL, center of gravity
Day 5	Association between quantitative variables: review
Day 6	Association between quantitative variables: exponential regression models, residual
Day 7	Association between quantitative variables: exponential regression models
Day 8	Association between quantitative variables: quadratic regression models
Day 9	Association between quantitative variables: review, correlation
Day 10	Association between quantitative variables: comparing models, residual plots
Day 11	Association between quantitative variables: comparing models, residual, residual plots
Day 12	Association between quantitative variables: comparing models, extrapolation
Day 13	Association between quantitative variables: review, interpolation, extrapolation

Day 14	Association between quantitative variables: review, impressionistic and theoretical models, comparing models
Day 15	Association between quantitative variables: review
Day 16	Association between quantitative variables: comparing models, unit test
Mr. Tablet Day 1	Association between quantitative variables: linear relationships
Day 2	Association between quantitative variables: scatterplots, median trace
Day 3	Association between quantitative variables: correlation
Day 4	Association between quantitative variables: correlation
Day 5	Association between quantitative variables: correlation, LSRL
Day 6	Association between quantitative variables: LSRL, influential points, coefficient of determination
Day 7	Association between quantitative variables: residual, residual plot, coefficient of determination
Day 8	Association between quantitative variables: read computer output, vocabulary
Day 9	Association between quantitative variables: quiz, exponential regression
Day 10	Association between quantitative variables: exponential regression through linear transformation
Day 11	Association between quantitative variables: power regression through linear transformation
Day 12	Association between quantitative variables: comparing models
Day 13	Association between quantitative variables: comparing models, extrapolation, average data Association between categorical variables: lurking variables
Day 14	Association between categorical variables: side-by-side boxplots, two-way tables, segmented bar graph
Day 15	Association between categorical variables: two-way tables, Simpson's paradox Association between quantitative variables: regression using linear transformation
Day 16	Association between quantitative variables: review Association between categorical variables: review
Day 17	Association between categorical variables: chi-square statistic and distribution

Day 18	Association between categorical variables: chi-square statistic, distribution, and probability chart
Day 19	Association between categorical variables: chi-square statistic
Day 20	Association between categorical variables: chi-square test for independence
Day 21	Association between categorical variables: chi-square test for independence
Day 22	Association between categorical variables: chi-square test for independence
Day 23	Association between quantitative variables: t-test for slope of regression line
Day 24	Association between quantitative variables: confidence interval for slope of regression line
Day 25	Association between quantitative variables: confidence intervals and prediction intervals for regression response
Day 26	Association between quantitative variables: review
Ms. Tuck Day 1	Association between quantitative variables: data and types of variables
Day 2	Association between quantitative variables: best fit line, predictions
Day 3	Association between quantitative variables: best fit line
Day 4	Association between quantitative variables: best fit line
Day 5	Association between quantitative variables: calculator line fitting
Day 6	Association between categorical variables: chi-square test statistic, calculating expected values for two-way tables
Day 7	Association between categorical variables: chi-square test of independence
Day 8	Association between categorical variables: chi-square test of independence

NOTE: Days with bold day numbers were studied by additional analysts.

Figure 1. Observed days and the topics of instruction

During the observations my attention was primarily upon the teacher, but I also noted student comments, questions, solutions, and claims in that they may have generated a response from the teacher that called for the teacher to apply his or her subject matter knowledge. My fieldnotes were also useful for me to refer to during the interview following each observed class session. During an interview I usually talked through the class with the teacher, asking his or her subject matter thoughts concerning specific teaching incidents. The observation fieldnotes were invaluable for this.

During the observations, I looked through the lens of the discipline of the subject, including statistical knowledge, mathematical knowledge, and context knowledge, to focus on the subject matter content involved in the practice of teaching. In my initial conceptualization of my pilot study, I described my work as a study of teachers' mathematical knowledge, following the work of Ball et al. (2005). However, this conceptualization was inadequate in characterizing the work of the teacher in the pilot study. Thus, I expanded the meaning of mathematical knowledge used by Ball and her colleagues to include the triad of statistical knowledge, mathematical knowledge, and context knowledge. This resonates with the subject matter knowledge I observed the pilot study teacher use in her teaching and with a current model of statistical literacy (Gal, 2004).

It is impossible to observe knowledge directly as it is not visible. However, it is possible to observe directly the actions, verbalizations, and objects teachers and students use in the classroom. Observation of these things in the natural environment in which they occur provokes discussion of the subject matter knowledge used by the teacher or what might have been useful for the teacher to know in each situation. In conducting the pilot study, I perfected my skills as a researcher to focus upon subject matter knowledge in such observations.

Interview

Following each observed class session, I interviewed the participating teacher regarding his or her subject matter thoughts involved in the lesson. This included the planning of the lesson, implementation of the lesson, classroom occurrences during the teaching of the lesson, and assessment of student learning. The teachers were encouraged to have any personal written documentation of the lesson available during the interview to help cue their recall.

I used the interviews to work towards understanding the subject matter knowledge used in teaching statistical association from the participating teacher's perspective rather than mine, the researcher's (Merriam, 1998). To find out what was occurring in the minds of the participating teachers, since that is where their subject matter knowledge resides and is used, interviewing teachers regarding their thoughts was an appropriate and necessary method (Merriam). Through my experience in the pilot study, I learned the importance of interviews in the data collection process, as they provided significant additional information regarding the knowledge used by the teacher that was not apparent from the observations. It should be noted that while the intent was to give an authentic account of the participating teachers' perspectives, my account of the interview resulted from my construction of their perspectives.

Each interview was semistructured in that it was a mix of predetermined questions and exploratory questions created at the time of the interview (Merriam, 1998). An interview protocol was developed during the pilot study. It was designed to provide starting points for discussion, but as the interviewer I had the task of choosing the questions which were relevant to the observed class period. Some of the questions dealt with the planning of the lesson, such as "How did you choose the definitions of terms and their symbols used in the lesson?" Others dealt with the class session itself; for example "As students posed ideas and questions, what were you thinking about from a subject matter perspective as you listened and responded to them?" Another portion focused on assessment, asking questions related to the subject matter knowledge used by teachers in the design and evaluation of assessments. A copy of the interview protocol is located in Appendix B.

During each interview, the predetermined questions were supplemented with questions regarding specific occurrences in the classroom session observed. This format allowed for me to respond to the situation observed and to new ideas on the topic (Merriam, 1998). Usually I read through my observation fieldnotes and formulated questions for the teacher regarding what they were thinking about from a subject matter perspective related to a specific incident or document they had used in class. To illustrate, questions regarding a teacher's choice of a data set, what a teacher was thinking about in responding to a student question, or how he or she understood a topic discussed in class that day, were asked during the interviews for this study.

Records of Practice

The practice of teaching statistical association was documented through records of practice (Ball & Bass, 2000b). These included audiotapes of the observed class sessions; transcripts based on the audiotapes; fieldnotes taken during observations of the class sessions; handouts given to students by the teacher involving the topic; presentations used by the teacher, such as Geometers Sketchpad files or overhead slides; assessments given to the students; and copies of student work. In addition copies of the relevant textbook pages were made for each class. Since each data source uncovers different aspects of reality, multiple sources must be employed (Denzin, 1978). Additionally, the multiple sources help to provide cross-data validity checks (Patton, 1990). The audiotapes, transcripts, and fieldnotes taken together represent a data triangulation (Denzin) that documented what occurred during the class session.

I made a conscious decision to audiotape rather than videotape the observed class sessions and interviews. While videotapes can preserve more aspects of the classroom's events, such as gestures and computer displays, the observational fieldnotes and documentation from the teacher helped with the reconstruction of events involving such aspects so that potential ambiguities were avoided. Another advantage to videotaping is its preservation of aspects of interaction between persons (Roschelle, 2000), such as body language, but the non-verbal components of interpersonal interactions were not needed for this study.

Audiotaping has many of the same advantages of videotaping, including repeated observation of the same event, documentation that grounds the researcher's analysis, observation of changes over time, and the ability to bring together disjointed segments for analysis (Hall, 2000; Lesh & Lehrer, 2000; Roschelle, 2000). An advantage of audiotaping over videotaping is its reduced interruption to the classroom environment. Every participating teacher wore a microphone connected to a voice recorder during each observed class session and its subsequent interview. The pilot study emphasized the importance of this equipment for capturing all of the interactions between the teacher and students during the class sessions.

The audiotapes from all of the observed class sessions and interviews were transcribed for documentation and later analysis, a common practice in qualitative research (Kvale, 1996). There were two advantages to this process. First, it allowed me to concentrate on the topic and dynamics of each observed class session and interview. Second, it created a written document which could be line numbered and enhanced with inserted pictures and additional information such as copies of handouts. Such a document was needed for the analysis of the data, particularly by other analysts who did not participate in the observations or interviews. The fieldnotes from the observations were a helpful resource during the creation of the transcripts, providing information regarding a setting or observed activity that the audiotape did not include.

Three additional data sources were used in this study: student handouts, copies of student work, and copies of textbook pages. A handout given to students during an observed class session aided the reconstruction of the session, but it also provided a document to be analyzed regarding the subject matter knowledge of the teacher involved in its creation, implementation, and assessment. Student work was also studied to determine the subject matter knowledge needed by the teacher to assess it and respond appropriately to it. The textbook provided the basis for the organization of the curriculum and the topics covered by all three teachers. Therefore copies of relevant textbook pages helped the analysts to understand the curriculum which provided the basis for the observed class sessions, as well as the homework assignments for students.

The transcripts of the class sessions and interview sessions were given to each participating teacher for a process called member checking, where the participants are requested to examine writing where the actions or words of the person are featured (Stake, 1995). The process tests the accuracy of the interpretations of the researcher by indicating to what extent the participants agree with them (Tobin, 2000). Following this process, modifications to the transcripts and notes were made based on the participants' feedback with the intent of creating an accurate account of the classroom and interview sessions. In addition I revised the observation and interview transcripts by removing any portions irrelevant to the subject matter content for this study. For example, Mr. Tablet first taught about the chi-square statistic as the test statistic for the goodness of fit test. While the meanings of the chi-square statistic and its distribution were relevant to this study, the goodness of fit test was not, and so the portions of the transcript only concerning the goodness of fit test were edited out. In the interview transcripts, teachers comments not related to their subject matter knowledge were also removed.

Analysis

The analysis focused on the activities of teaching statistical association for their subject matter knowledge components and demands upon each teacher. Seeking to locate and analyze content knowledge as it was used in practice, the process was similar to a job analysis (Ball & Bass, 2000a).

Compilation Documents

A compilation document, containing the portions of the data identified for further study, was created for each teacher for each day of the study. Each compilation document began with a written re-creation of the class period observed. The observation transcript was enhanced to include other information, such as handouts, calculator screen-shots, or written material on the board, to make it possible for someone reading a compilation document to follow the dialogue that occurred that day in the class session. At the beginning of the observation transcript, the day's topics were listed. Throughout the observation transcript, settings were identified to help the reader understand transitions made throughout the class period.

Next in a compilation document came the interview transcript. No changes were made to the transcript other than those previously made to edit out irrelevant material. Scanned images of the relevant textbook pages followed the interview transcript. Ms. Tuck also provided me with a copy of the teacher's guide for the textbook, so the compilation documents for her lessons included copies of the relevant pages from the teacher's guide.

Relevant assessment materials comprised the last component of a compilation document. Most often this was a homework assignment, and in this case the textbook pages or handouts containing the assignment were included. Other types of assessments included quizzes, tests, or projects, and those documents were included as well. Occasionally I asked a participating teacher for copies of student work for assignments where the subject matter knowledge needed to assess the assignment was different than that used in the observed class period. Copies of student work were added as the final component to those compilation documents. Finally the lines were numbered in each compilation document to provide a point of reference during the analysis process. Together, a total of fifty compilation documents were created, corresponding to the fifty class sessions observed as a part of this study. An example of a compilation document can be found in Appendix C.

Analysis Process

Following the creation of the compilation documents, I took on the role of an analyst, studying the data sources in the compilation documents for the subject matter knowledge involved in teaching statistical association. I followed the analytic process described by Ball and Bass (2000b) in their similar work regarding teacher knowledge for teaching elementary school mathematics with the exception that I did not differentiate between common subject matter knowledge and specialized subject matter knowledge in my work.

As I studied the records of practice contained in a compilation document, I considered what subject matter knowledge on the part of the teacher was involved in the teaching activities referred to in the data source. As I read a compilation document, I broke apart the sequence of events or items into teaching incidents. A teaching incident refers to a teaching activity or objects that can be extracted from the rest of the compilation document and still make sense standing alone. This could be any incident involved in teaching, including a student question, an explanation provided by the teacher, student work on a homework problem, or a test question. Then for each teaching incident, I analyzed the subject matter knowledge needed by the teacher. The teachers' knowledge could not be documented directly through the data sources, nor was the purpose to analyze any particular teacher's knowledge. Instead, I used the data sources as a catalyst for developing conjectures regarding the subject matter knowledge a particular teaching incident might entail (Ball & Bass, 2000b). The data sources grounded the analysis in the practice of teaching, helping me to generate claims regarding the subject matter knowledge needed by teachers involved in the tasks of their profession. My resulting descriptions of the subject matter knowledge went beyond merely creating lists of topics; instead I aimed to explain the particular understanding of the topic that the teacher needed, with an end goal of creating a description that would be useable and useful to teachers, educators, and researchers.

A piece of knowledge does not stand alone; there exists a learning hierarchy for each topic which describes the progression of prerequisite skills needed to master the topic (Gagné, 1985). A limit had to be set regarding how far back or forward to go in describing the knowledge needed for understanding a topic. It was decided that the description of knowledge needed should generally go no further than two classes or grade levels, either previous or future, from the current class. This allowed me to unpack the knowledge needed for a topic so that a rich description would be provided without getting bogged down in the process by having to go all the way back to the most elemental prerequisite knowledge components.

I considered two perspectives with respect to time during the analysis process. One perspective took into account the knowledge used by the teacher to respond immediately to the current real-time demands of teaching (Ball & Bass, 2000a). This included teaching incidents such as student questions, creating examples, and evaluating student work. Instruction over time (Ball & Bass, 2003), formed the second perspective. This included knowledge of the larger picture of the statistical topic and its associated practices and how these connect with other topics. For example, teachers of linear association should understand that the coefficient of determination, r^2, is only a relevant statistic for linearly associated data but that it has a similar counterpart, R^2, which can be used regardless of the type of association between the variables.

As I analyzed the compilation documents, questions arose for me regarding the content of statistical association and meanings of terms used by the statistics community. I compiled a list of these questions for each of the three observation sites and sent them to two expert statisticians: Dr. Jinadasa Gamage of Illinois State University and Dr. Allan Rossman of California Polytechnic State University. I learned much about statistics from their answers, and their insightful responses contributed greatly to my investigation and the analysis reports I created.

To illustrate, in the curriculum unit "The Overland Trail" (Fendel, Resek, Alper, & Fraser, 2004) taught by Ms. Tuck, students were instructed to draw the line of best fit for different sets of data. This best fit line was meant to be determined by criteria that the students created and remain this way for the entire unit (i.e., no formal mathematical models were taught). This led me to wonder what statisticians consider the best fit line. The responses from the statisticians informed me that by convention statisticians consider the term "best fit line" to be synonymous with least-squares regression line. Therefore when teachers are teaching about best fit lines or evaluating student-created best fit lines and criteria, it is important that they are knowledgeable about the least-squares regression line and its criterion.

I created a written report, called an analysis report, for each compilation document. It contained knowledge descriptors for each teaching incident and the line numbers of the compilation document referring to the teaching incident. This provided an explicit correspondence between the practice of teaching and the subject matter knowledge that is needed in that practice. Each knowledge descriptor was intended to provide enough detail so that an unambiguous explanation of the concept that the teacher needed to know was created. One knowledge descriptor was sometimes used for multiple teaching incidents if it described the knowledge needed by the teacher for all of the referenced teaching incidents. A compilation document and its accompanying analysis report can be found in Appendices C and D.

Additional Analysts

Twelve of the compilation documents were analyzed by an additional two persons: a statistician and a statistics education expert. This constituted a type of triangulation known as triangulating analysts (Patton, 1990). Due to the complex and dynamic phenomena of teaching and learning mathematics where much remains hidden and needs interpretation and analysis, it was important to use an interdisciplinary group of experts to conduct the analysis of the data (Ball & Bass, 2003). Using multiple perspectives to analyze the data enhanced the quality and credibility of the analysis and its findings (Patton).

I could not reasonably expect these experts to analyze all 50 compilation documents. Thus, I needed to identify a subset of documents that would reflect the breadth of topics addressed in the lessons and the richness of the data collected. Working from my analysis reports, I constructed a list (see Figure 1) of the primary topics covered in each of the compilation documents. I cross-referenced each topic to identify the compilation documents in which it occurred. Whenever a topic was identified in more than one compilation document, I reviewed the compilation documents and their analysis reports to identify the one, or sometimes two, documents that contained the most fruitful data. This resulted in the selection of twelve compilation documents that included nearly all of the primary topics. The topics not included were often very closely related to another topic that was included (e.g., exponential regression was included while quadratic regression was not) or what I considered to be a minor topic (e.g., Simpson's Paradox). Of the twelve compilation documents I selected for the experts to analyze, three of them were from Mr. Glass' observations, four were from Mr. Tablet, and the remaining five documents were from Ms. Tuck's classes. The days corresponding to these documents are listed in bold print in Figure 1.

The two analysts were sent the materials they needed to perform the analysis process themselves. This mailing included a letter and a CD. The letter contained specific directions for the analysts and definitions of some terms used, such as "compilation document" and "statistical association." The intent was for both of these experts to go through the same analysis process I had gone through with these same compilation documents, and so I tried to describe that process in as much detail as possible in the letter. The analysts were also encouraged to ask questions as they arose through the analysis process. A copy of the directions sent to the additional analysts can be found in Appendix E.

The CD contained a number of documents the analysts needed to perform their analysis. The twelve compilation documents they were asked to analyze were included on the CD. I also provided lists describing the topics included in a compilation document. Finally, I provided a copy of the analysis report I had created for Day 3 of Ms. Tuck's compilation documents. This was meant to provide an example for the analysts in order to help them further understand what they were being asked to do, particularly with respect to the level of detail and format of the report.

Both analysts completed an analysis of the twelve compilation documents they received and constructed an analysis report for each one. I merged their reports with ones I constructed, thus creating a meta-analysis report for these twelve documents. Using the line numbers as a guide, I put together the knowledge descriptions that were similar. Appendix F contains one of the meta-analysis reports with the analyses of all three experts. It is the report for the compilation document in Appendix C.

Following the creation of these twelve meta-analysis reports, I examined them to determine the level of agreement among the analysts. I began this process by creating a chart for each of the meta-analysis reports. It had four columns: knowledge descriptor, statistics education expert, researcher, and statistician. Each row was labeled with a different knowledge descriptor and underneath the column of an analyst, I listed the line numbers that analyst had given which referred to that knowledge descriptor. For example, for the meta-analysis report in Appendix F, one row of the chart named "Modeling: signal in noisy process" as the knowledge descriptor, then lines 283-303 were listed under the statistics education expert column, and lines 154-158, 283-303 were listed under the researcher column.

Next I analyzed the level of agreement between the analysts with respect to the knowledge descriptors. A total of 189 knowledge descriptors were given in the meta-analysis reports. By report, I determined how many analysts listed the same knowledge descriptor. Of the 189 knowledge descriptors, 80 (42%) were included by all three analysts, 47 (25%) were included by two of the three analysts, and 62 (33%) were listed by one of the three analysts. Thus, 67% of the descriptors were listed by multiple analysts. This showed a relatively strong agreement between the analysts' findings regarding the knowledge needed to teach statistical association.

I further analyzed the 62 knowledge descriptors listed by a single analyst within a report. Fifteen of them came from knowledge descriptors I wrote with reference to the relevant textbook pages for that class day. The other analysts rarely included knowledge descriptors from the textbook pages, so it appeared that perhaps I had looked more closely at the textbook as a record of practice during my analysis process. I did not see this as a point of concern, however, because all of the knowledge descriptors from the textbook were either also generated from a different record of practice in another compilation document (e.g., the least-squares regression line always goes through the point [x-bar, y-bar] also was a knowledge descriptor in three other compilation documents) or minor points to be added to larger topics (e.g., it is not always possible to find a model for a phenomenon such as the stock market).

The statistics education expert wrote five knowledge descriptors regarding terminology and three descriptors regarding context knowledge that the other two analysts did not. He was adept at noting when proper use of terminology and context knowledge were relevant, and when I reviewed the portions of the compilation documents he had cited for those descriptors, I agreed with his analysis.

Fifteen of the knowledge descriptors listed by a single analyst involved topics that were later determined to be part of the prerequisite knowledge needed by teachers. I concluded that these discrepancies were minor since they dealt with prerequisite knowledge and most likely were included by only one of the analysts because the topic did not directly concern statistical association. For example, proportional reasoning was listed in four different meta-analysis reports. In one of the reports it was listed by all three analysts because the teaching incident brought the need for it to the foreground. In the other reports it was listed by a single analyst, a reflection of the fact that the teaching incidents on those classroom days involved proportional reasoning at a background level. The other 47 knowledge descriptors that referred to knowledge concerning statistical association were listed as knowledge descriptors in other analysis reports as well, and thus none of them presented unique information.

During my analysis I noted that of the 62 descriptors cited by a single analyst, 43 of them came from the first page of the compilation document or the textbook pages at the end of the compilation document. Thus, approximately 69% of them came from teaching incidents outside the meat of the lesson that occurred in the classroom that day. As a result of this analysis I concluded that there were no major points of disagreement regarding the knowledge descriptors that needed to be addressed. The strong agreement among our analyses made me confident of the validity of the analysis reports I created for the compilation documents that were not analyzed by the other analysts.

I also analyzed the charts to determine the level of agreement among the analysts regarding the number of documented teaching incidents. There were 336 teaching incidents noted by the analysts in the meta-analysis reports. Of these teaching incidents, 103 (31%) were noted by all three analysts, 125 (37%) were documented by two of the analysts, and 108 (32%) were listed by a single analyst. Overall 68% of the teaching incidents were identified by multiple analysts. I interpreted this to mean that there was general agreement among the analysts regarding which teaching incidents called for documentation of subject matter knowledge needed by the teacher. I found this to be a respectable level of agreement as I did not expect there to be total agreement among the analysts due to the nature of the analysis process and the individual determinations of teaching incidents.

Synthesis of the Analysis Reports

In the next phase of analysis, I synthesized the data from the twelve meta-analysis reports (which merged the data from my analysis report and the reports of the two additional analysts) and the 38 analysis reports that I constructed alone. This was the first time the data from all three participating teachers were brought together. Provisional categories or tags of the subject matter knowledge for teaching statistical association at the secondary level were made using the constant comparative method (Glaser & Strauss, 1967). I began by creating a tag which identified the subject matter topic for each row of the analysis reports that only I had created. Some knowledge descriptors received multiple tags if they linked together multiple ideas. I used my mathematical and statistical knowledge that pertained to the included topics and the given description to determine the area within statistical association to which a description referred. The knowledge descriptors were continually compared and sorted into groupings that had something in common using this method (Merriam, 1998). This technique was useful for seeking patterns in the data that could then be arranged in relationship to one another.

As I went through this process, I tried to minimize the number of tags that I used. Thus, when I read a knowledge descriptor, I first determined if it could be classified under one of the previously identified tags. If it could not, then I created a new tag. However, some of the topics had so many knowledge descriptors under them that I decided to give them their own separate category even though they fell under another one. This occurred with the algebra category. I originally had only the tag “algebra,” but because there were so many knowledge descriptors regarding slope and lines, I later decided to make those their own categories. The same thing occurred with the mathematical modeling category. Due to the large number of knowledge descriptors related to the best fit line, I split that off to be its own category.

There were also some knowledge descriptors related to general topics that I wanted to keep together in their own tag. For example, in reading the knowledge descriptors, I found there were overarching descriptors of knowledge concerning association, including its distinction from causation and general meaning, so I made a general category tag “association.” Also during the synthesis process the importance of knowledge concerning the language of statistics came to the forefront. Therefore I gave some of the knowledge descriptors an additional tag, “terminology,” so that I could group all of these together for further analysis and future description.

Next I moved to the twelve meta-analysis reports that contained the knowledge descriptors from all three analysts. I proceeded to give tags to all of the rows in those reports as well, again categorizing the knowledge descriptor with a tag. At the end of this process of applying the constant comparative method, I had established thirty-four tags or categories.

It was at this point in the analysis that the knowledge descriptors were re-sorted to be listed in tables with the other knowledge descriptors of the same category. For example, all of the knowledge descriptors with the category tag “data” were put together in one table because they described what teachers need to know about data. This was done for each of the thirty-four categories. During this re-sorting I read again the knowledge descriptors and made sure they were tagged appropriately. I amended those I felt needed new or additional tags and sorted them accordingly.

After this re-sorting process, five of the categories stood out as minor or peripheral to the topic of statistical association. Three of them (analysis of variance, the difference between parametric and non-parametric tests, and Venn diagram) came from one isolated comment made by a teacher regarding a topic which was not really related to statistical association topics covered at the secondary level. Another topic, chi-square goodness-of-fit, is another kind of chi-square test, but knowledge of it is not really needed to teach the relevant chi-square test for independence. The final topic was sampling, an enormous and important topic, but nonessential to the teaching of statistical association. Thus they were removed from the rest to create a final total of twenty-nine categories.

The twenty-nine categories and their related knowledge descriptors comprise the findings of this analysis and encompass the subject matter knowledge for teaching statistical association. The findings are presented in the next chapter as a concept map and narrative description and in Appendices G and H as detailed summary outlines.

Feedback from Participating Teachers and Analysts

The summary document outline was given to the other analysts and the participating teachers for their feedback. This constituted another round of member checking (Stake, 1995). I asked all of the analysts and the participating teachers to review the work, indicating specific areas of agreement and disagreement (Tobin, 2000). This process provided a control for research bias (Moschkovich & Brenner, 2000).

All of the teachers and one of the analysts provided me with feedback on the summary document outline. I made minor changes made to the outline based on their feedback. Based on one teacher's suggestion, I added real-world situation as another representation of a line. Another teacher suggested that I clarify that the mean is used as the measure of center in the calculation of standard deviation, and so I added that clarification to the description of standard deviation. An analyst noticed that I had made a mistake in describing parallel lines, so I corrected the description to read vertical translation instead of horizontal translation. The consensus was that the outline was a thorough and detailed description of the knowledge needed for teaching statistical association at the secondary level.

Chapter Summary

This chapter details the specific methods used to answer the research question. A practice-based approach in which teachers were observed as they taught statistical association was determined to be the ideal methodology for the study. I also did interviews following each observation to allow the teacher to talk out loud about his or her internal thoughts throughout the teaching process. I conducted observations and interviews with three different teachers throughout the 2006-2007 school year as each teacher taught the topic of statistical association. A compilation document was created for each day observed, including relevant portions of the observation and interview transcripts, all relevant textbook pages, and student work if applicable. In total, I made then analyzed fifty compilation documents. Twelve of these compilation documents were selected for analysis by an additional two experts in the fields of statistics and statistics education. I then merged the twelve analyst reports written by myself and the other analysts for the selected twelve compilation documents to create twelve meta-analysis reports. Following the synthesis of all fifty analysis reports, including the twelve meta-analysis reports, a concluding description was created detailing the subject matter knowledge for teaching statistical association at the secondary level.

In reflecting upon the methodology of this study, I am confident that the findings describe the knowledge needed in practice by teachers since I studied the teachers at work in their classrooms, and for that reason I am pleased with the effectiveness of the practice-based approach. I have no regrets regarding my choice of methodology because I know that this approach is the most desirable one (Ball et al., 2001). However, this approach was extremely laborious and time-consuming; therefore I wonder whether the same findings could have resulted from a more efficient approach and will discuss this further in the last chapter.

I am also convinced that the findings of this study are legitimate because of the inclusion of multiple participating teachers, analyst triangulation, and member checking in my procedures. The findings reported in the following chapter represent the synthesis of data from three teachers who used three different curricula to teach statistical association, which strengthens the validity of the study. In addition, portions of the data were analyzed by an expert in statistics education and a statistician. While not all of the data were analyzed by all three analysts, the data that were analyzed by the team produced similar findings, which strengthens the conclusions drawn by me alone. In addition, all of the analysts and participating teachers reviewed the final summary document detailing the findings of the study. This process was the final verification procedure and made the results of the study all the more trustworthy.

CHAPTER IV

RESULTS

Introduction

The findings regarding the subject matter knowledge needed to teach statistical association at the secondary level are presented here in chapter IV. First, the subject matter knowledge categories are given, along with the number of teaching incidents that referenced them. The following two sections provide explanations of the organization of the findings. They describe how two of the presentations of the findings, a summary document outline and a concept map, were created.

I present a concept map next that visually displays the major components of the subject matter knowledge for teaching statistical association. That is followed by a more detailed description of the subject matter knowledge components along with excerpts from the records of practice that led to their identification.

Appendix H contains a summary document that outlines the findings of the study. This comprehensive summary provides an in-depth, thorough presentation of the subject matter knowledge needed for teaching statistical association at the secondary level as found by this study. The outline is in the appendix solely because of its length; it is not supplementary but should be considered part of the main text. This outline will be referenced throughout the chapter by its outline identification, such as II.K. for residual plot.

Knowledge Categories

I classified the knowledge for teaching statistical association into twenty-nine categories. These categories and the number of teaching incidents referring to them are given in Figure 2, by order of the number of referenced teaching incidents.

A teaching incident refers to a teaching activity or objects that can be extracted from the rest of the compilation document and still make sense standing alone. This could be any incident involved in teaching, including a student question, an explanation provided by the teacher, student work on a homework problem, or a test question. I noted each teaching incident in the analysis reports as a segment of the data defined by its line number(s). In the meta-analysis reports not all teaching incidents were identified by the exact same line numbers. In those cases, I grouped the common line numbers that referred to the same teaching activity or object and counted that as one teaching incident.

Category	**Number of referenced teaching incidents**
Best Fit Line (including LSRL)	116
Correlation Coefficient	100
Chi-square Statistic and Test (general)	96
Mathematical Modeling (excluding best fit line)	83
Chi-square Test for Independence	64
Residual Plot	61
Calculator	60
Residual	54
Predictions	52
Coefficient of Determination	46
Algebra (excluding lines and slope)	43
T-test for Slope of Regression Line	38
Association	37
Scatterplot	36
Context	35
Hypothesis Test (general)	35
Slope	34
Lines	27
Terminology	26
Proportional Reasoning	25
Data	21
Computer	19
Probability	18
Two-Way Tables	15
Confidence Interval for Slope of Regression Line	10
Outliers & Influential Points	10
Normal Distribution & Z-scores	8
Data Display (excluding scatterplot, two-way table, and segmented bar graph)	7
Average (including mean)	6

NOTE: Prerequisite subject matter knowledge descriptors in bold.

Figure 2. Knowledge Descriptor Categories

As evidenced by the 116 teaching incidents that referenced it, the line of best fit was the predominant topic of statistical association taught in the observed secondary classes. It was a major topic of all three teachers' curricula, and each teacher spent multiple days teaching the topic. Likewise the correlation coefficient, which is the statistic used to measure the strength and direction of a linear association, was a pervasive topic and referenced 100 times. Regarding association of categorical variables, the chi-square statistic, chi-square test procedures in general, and chi-square test of independence received the greatest coverage by the participating teachers.

The knowledge descriptors under nine of the categories were determined to be descriptions of prerequisite subject matter knowledge for teaching statistical association due to their foundational nature and lack of specific application to statistical association. These nine categories were proportional reasoning, slope, lines, algebra (excluding slope and lines), statistical graphs (excluding scatterplots and segmented bar graphs), average, normal distribution and z-scores, probability, and hypothesis test in general. They are listed in bold print in Figure 2. Upon reflection the nine categories were divided into two larger groups. The first four categories were grouped together under the title "algebraic reasoning and concepts." The remaining five categories were listed under the general heading "statistics and probability." Since these categories were identified as prerequisite and therefore peripheral to the findings, a narrative description and detailed outline describing the knowledge components of these critical underpinnings are both in Appendix G rather than the following sections.

Subject Matter Knowledge Summary Document Outline

The findings of the study needed to be organized in such a way that they would be accessible to a broad audience. Thus, I created a summary document outline. I began making the outline by attempting to divide the categories according to whether they pertained to quantitative or categorical data variables since the methods of analysis differ between the two and I had also classified the content for the observed classroom days according to this scheme (see Figure 1). However, there were three overarching categories (association, terminology, and context) that applied to both quantitative and categorical data variables. Therefore I decided to have a general statistical association grouping to contain these three categories. Thus I divided the twenty categories into three broader groupings: statistical association, association between quantitative variables, and association between categorical variables.

Next I arranged the categories sequentially under each of these groupings. I sequenced the outline in the same order used for an analysis of that type of data and from general to specific. So, for example, I listed the inferential methods after the numerical and graphical analysis methods in the outline because they are typically done following those methods when analyzing a data set. I also listed general mathematical modeling first, followed by line of best fit, followed by least-squares regression line (LSRL), because I wanted the outline to start with the most general category and proceed to the more specific categories.

Finally, I added the detailed descriptions of each category. I used the knowledge descriptors contained in the analysis reports to create these descriptions, making sure to include every knowledge descriptor from the analysis reports in the outline. When arranging the descriptors in the outline, I moved from the general to the specific with the more general category listed above the more specific categories that fall under it. I organized the descriptors as I had the categories, moving from general to specific and in the order in which data analysis is carried out. To illustrate, in the section about the line of best fit, the first sub-point gives the descriptors regarding the meaning of the line of best fit for the population in general. The next sub-point gives the descriptors regarding the best fit line for a sample from a population, thereby moving from general to specific. The section continues to describe best fit lines for samples, then transitions to a more specific description of the LSRL as a particular best fit line. The final product can be found in Appendix H: Subject Matter Knowledge Summary Document Outline. Although contained in an appendix due to its length, the summary document outline should be considered part of the main text.

Concept Map

A second presentation of the findings, a concept map, was created to provide a visual representation of the relationships among the categories. It is presented in Figure 3. From the headings of the summary document outline (which mostly came from the category tabs), I created the topics for the concept circles of the map. I followed the general organization of the outline, which started with overarching themes regarding statistical association, including its meaning, terminology, and knowledge of context. Next I created two other general branches: quantitative data and categorical data. Underneath each of these I used the organization I had developed for the outline. Working from the top I created the map to read like a sentence from the top circle down to any of the circles below. For example, reading along one of the threads of the concept map from top to bottom the sentence "Statistical association between categorical data sets can be analyzed graphically by a segmented bar graph" can be created. Although these knowledge components are linearly related in the concept map, in practice they are interconnected and interact during the process of teaching.

The concept map in Figure 3 depicts the components of the subject matter knowledge for teaching statistical association and their relationships with one another. Each of the components will be explicated and linked to the records of practice in the following section.

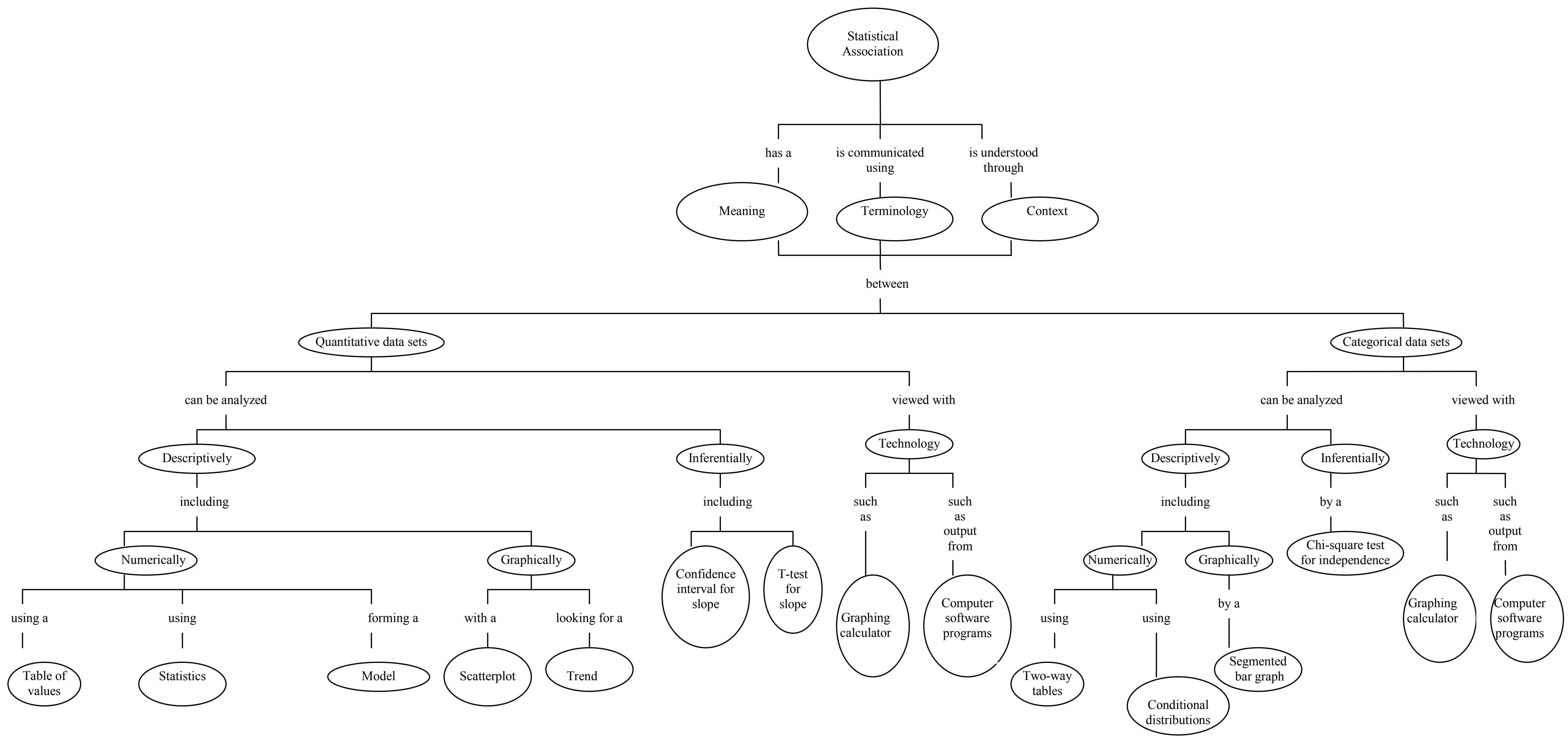


Figure 3. Concept Map

Description of Subject Matter Knowledge

This section of the chapter presents a narrative description of the subject matter knowledge needed for teaching statistical association at the secondary level. Included in this narrative description are examples from the records of practice that show where the relevant knowledge component is needed for the work of teaching. The supporting examples are referenced to the data through a code such as [GBS3 182-194]. The letters in the code correspond to one of the three participating teachers, either GBS for Mr. Glass, NCP for Ms. Tuck, or NT for Mr. Tablet. A number immediately follows that gives the day number on which the teacher was observed (the observations were numbered sequentially for each teacher). After a space, the listed range of numbers gives the line numbers in the compilation document from which the example comes. The description follows the organization of the concept map through its headings. References to the relevant portion of the summary document outline in Appendix H are also provided throughout.

Statistical Association

Meaning

An understanding of the topic of statistical association begins with knowledge of its meaning. There is an association between two statistical variables when there is a correspondence in their variations. This correspondence does not have to hold true for every piece of data, but instead describes a general trend.

The meaning of statistical association came up as Mr. Tablet was introducing the topic of statistical association to his class through a discussion of the relationship between number of wins and average attendance for Major League baseball teams. The following exchange occurred between Mr. Tablet and one of his students.

> Mr. Tablet: When we look at the Cubs, what do we see? The wins are low and the attendance is high. Would you agree? But in general that's not the case. Don't we normally, we have teams that win a lot, they have high average attendance?
>
> Student: Wouldn't you think that those aren't always linked?
>
> Mr. Tablet: Yeah, you could, but I'm just trying to show that there is an association. [NT1 68-72]

The teacher, Mr. Tablet, had to understand that association means a general trend and not an absolute rule in order to reply appropriately to the student's question.

A robust finding, in that it was documented with all three participating teachers repeatedly, was the importance of knowing the difference between association and causation. As Mr. Glass stated after analyzing a scatterplot of the cost of a mountain bike and its quality rating:

> And if you look at this thing in general, it looks like as the cost goes up what happens to the rating? The rating increases. Okay, so this mathematical model is going to somehow show the cost is related to the rating. Okay. It doesn't necessarily mean that the cost causes a rating. But it's how these things are related, Okay. [GBS1 47-51]

In addition to causation, two variables may be associated because they have a common response relationship or a confounding relationship. Teachers need to know the difference between each type of relationship and examples of each set in context. They also need to know why and how an experiment can establish causation while observational studies cannot.

Terminology

A teacher needs to be fluent in the language of statistics in order to teach statistical association with high quality. As in most fields, statistics has its own terminology that teachers must use accurately and appropriately. A list of statistical terms teachers need to have a thorough understanding of is included in section I.B. of the summary document (Appendix H). In the current section, I present illustrations from the compilation documents that show the need for this type of knowledge.

A student asked "What's the difference between correlation, association, and regression?" [NT8 224-225]. In order to respond to such a question, a teacher needs to know not only the meaning of each term but how the terms are related to one another. Another trio of similar terms is error, deviation, and residual. Mr. Glass asked his class “What does the observed y minus the predicted y compute?” [GBS3 112-113]. Any of the three terms could have been a correct answer to this question. In this case, a student responded, “error,” and Mr. Glass added that it could also be called deviation. Later in the lesson, Mr. Glass used the term residual to refer to the error in the prediction of a model [GBS3 322-324]. These are only two of numerous examples where the language of statistics came to the forefront.

<u>Context</u>

Gal (2004) previously named knowledge of context as necessary for statistical literacy. I chose to include it as a component of subject matter knowledge during the planning phase for this study, and through the analysis of the data, my decision was validated by the prevalence of the call for this type of knowledge throughout the records of practice.

One teaching incident involved the use of context knowledge to determine if it is reasonable to assume a constant rate of change between two variables. Consider the following selection from Mr. Tablet’s class:

> So we’re looking at the number of ads in a week and then the number of cars sold that week….So, well, actually anything else you might think of in

> terms of linear? Would it be always linear? Why would you say that? Because I'm actually thinking of something in business. What if I put in 2000 ads, what would eventually happen here? Level off. It's gonna level off. There's a point right where it's not going to get any better, I don't think. Again, I could be wrong on that, but again I don't know if it's going to keep going on and on and on forever, Okay? [NT6 230-237]

Here the teacher has knowledge of the context of the problem, the sales of cars, that he has applied to determine that it is unrealistic to assume that the line will continue for values well beyond the original domain of the problem.

Context knowledge is required to determine which variable should be deemed explanatory and which variable should be called response. As an activity to start class, Mr. Glass presented his students with a table of values that listed the college enrollment and faculty numbers at seventeen different institutions. He told the students "I want you to decide which one you think is the independent variable, which one you think should be the dependent variable" [GBS6 15-16]. Some students decided that enrollment was the independent variable, while others decided that faculty was the independent variable. When asked how he decided which variable was the independent variable in the interview, Mr. Glass responded "I tend not to be a Field of Dreams kind of person. You put 3000 faculty there and you'll have a large university. I don't believe that. You may be bankrupt. I think enrollment drives the number of faculty" [GBS6 223-225]. His knowledge of the context enabled him to make this informed decision. A more extensive list of examples when context knowledge was called for is in the summary document outline section I.C.Context (Appendix H).

Association between Quantitative Variables

Data

Three knowledge components related to quantitative data were identified as necessary for teachers of statistical association. First was recognition of the need for data on which to base decisions. The usefulness of data for predictions was revealed in the following activity. Students were asked to imagine that they were in a wagon train traveling the Overland Trail and needed to determine how many supplies to bring on their journey. They received a letter describing the amount of supplies brought by previous families on this stretch of the Overland Trail and the number of people in each family. They were asked to use this information to predict the amount of supplies they would need to bring for their family of twenty. The teacher, Ms. Tuck, described the usefulness of the data from previous families as follows:

> And people have gone before us, so we are like sweet; it would be helpful to know how much they took and then we can plan and we can predict from them….And you can see that maybe having more families might give us

> better information….we do know that multiple families have made this trip before. And that we have some information. If we could use that to our benefit, it will make our journey better. So, that's kind of where the act of predictions comes in. [NCP3 51-66]

The second knowledge component concerns the inherent and natural variability in data. Teachers need to be comfortable with variability, including sampling variability, and know ways to look for the trend in noisy data. Differentiating between types of data is the third knowledge component in this area; it is particularly needed for identifying bivariate quantitative data sets since that is the type of quantitative data examined for associations.

<u>Analyzing</u>

<u>Numerically.</u> In this section, I describe the knowledge necessary to analyze bivariate quantitative data using three different numerical processes. The first process is an informal assessment of the association between two variables through an examination of their values in a table. Two variables have a positive association if above-average values of the x-variable tend to occur with above-average values of the y-variable, and below-average values of the x-variable generally occur with below-average values of the y-variable. When two variables have a negative association, above-average values of the x-variable usually pair with below-average values of the y-variable, and below-average values of the x-

variable accompany above-average values of the y-variable. Teachers need to know this meaning of association and be able to apply it to evaluate numerically whether two variables have either a positive or negative association.

An activity Mr. Tablet did with his class called for this type of knowledge. He had the students in his class collect data on three variables that might be related to their height: the height of their same-sex parent, the time in minutes it takes them to get ready for school in the morning, and their lung capacity. The data for all four variables were written down in a table of values. He then asked the students “Okay now, what I want to do is take a look and guesstimate which one has the strongest association” [NT1 204-205]. To do this the teacher needed to know how to numerically analyze bivariate quantitative data.

The calculation of statistics is the second process teachers need to know to numerically assess association. The correlation coefficient r, coefficient of determination r^2, and coefficient of multiple determination R^2, are the three most important statistics for this process. A description of the knowledge needed concerning each statistic and examples calling for that knowledge follow. Further descriptions can be found in sections II.C. and II.D. of the summary document outline (Appendix H).

Teachers need to have extensive knowledge of the correlation coefficient, a statistic that measures the strength and direction of the linear relationship between quantitative variables. This includes its meaning, how to estimate the value of r from a scatterplot, the calculation of r, how to interpret its value, understanding the effect on r of adding and removing data points, and its relationships with r^2 and the slope of the LSRL.

A particularly interesting case is when r has a value approximately equal to zero. There are multiple scenarios in which this can occur, including a random scatter of points with no association and when the points follow a non-linear pattern. Another scenario is when the points follow a horizontal line. Mr. Glass knew this scenario and included the graph shown below in Figure 4 in a series of five scatterplots presented to the class. He asked the students to estimate the value of r from the scatterplot. The statements given below Figure 4 were Mr. Glass's comments on this scenario.

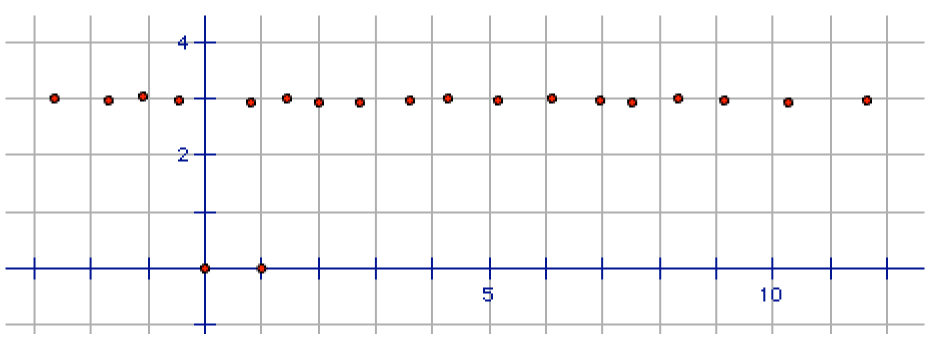


Figure 4. Scatterplot following a horizontal line

Okay. The correlation for this is 0….or pretty close to it. And that's

> because, what's happening is, this is not increasing as you go from left to right. It's not decreasing as you go from left to right. There's no way to say that there's a fixed change in the dependent variable for every change in the independent variable. Okay. The linear regression predicts an increase or a decrease. It doesn't like no change. If it can't predict a change, it doesn't like it. And actually this thing is what is weird, it is what makes most students crazy because this looks like a line and we're talking about linear functions, but it doesn't like it. It gives you a 0 correlation if you do that. Okay. Because it doesn't predict….So it just kind of messes up, it doesn't quite fit our intuition here. It's in some way an exception to that. [GBS1 131-138, 147-152]

Mr. Glass knew this scenario was counter-intuitive since the points do fall along a line and r is supposed to be close to 1 or -1 if the points closely follow a linear trend. He needed to know why statisticians decided to define r as zero when the points follow a horizontal line in order to explain this scenario to his students. Mr. Glass used the conceptualization of correlation as a measure of how much help x gives you in predicting y to justify that in the case of a horizontal line, r equals zero. In this scenario the various values of the x-variable all produce approximately the same value of the y-variable, so knowing the value of the x-variable provides no help in predicting the y-variable. Therefore, it makes sense that the value of r is zero. Other justifications for r having the value of zero in the horizontal line scenario include its direct relationship with the slope of the LSRL and its relationship as the square root of r^2. Knowledge of all these reasons is important for teachers as they teach students about a correlation coefficient value of zero.

The next statistic teachers must know is r^2, the coefficient of determination. In addition to understanding its relationship with the correlation coefficient, teachers need to understand its purpose, calculation, interpretation, and relationship with R^2.

This instance will highlight the need for teachers to understand the interpretation of r^2. Mr. Tablet gave his class a handout to review some of the material previously covered in class. It included a table of values listing the grade point average (GPA) and number of college applications submitted for nineteen of his previous students and asked the students to do a number of tasks including interpret the coefficient of determination in the context of the problem. During the class discussion, two students offered two different interpretations of r^2. The first student said “I thought that r^2 told you, like, the percent…what percent of the dependent one is explained by the other one” [NT7 254-255]. Another student asked “If we know the GPA, aren’t we saying that we have a 56.4% chance of knowing the number of college apps?” [NT7 263-264]. Mr. Tablet needed to have an intimate understanding of r^2's interpretation to assess these student interpretations and determine an appropriate response.

The third and final statistic teachers should know to teach statistical association is R^2. An advantage of this statistic is its applicability to any model type, as opposed to r^2 which can only be used when the data follow a linear trend. Teachers also need to know this statistic because it is included in standard computer and calculator output for regression.

A teacher could use this type of knowledge to respond to a student's question, as called for in Mr. Tablet's class when they were discussing the value of the correlation coefficient for a scatterplot which was curved like a parabola opening down. Mr. Tablet made the following remark:

> You might look at that and say, well, wait a minute that looks like it, there's a pattern there. Well, as far as the linear pattern, I would say that I can't tell if that's going to be a positive or a negative…it [r] must be zero. [NT3 186-191]

A student then asked "Is there a way to know…of the curve, a, ummm, like, not necessarily linear?" [NT3 193-194]. In other words, the student wanted to know if there is a statistic which measured the strength of the association for non-linear relationships. Mr. Tablet needed to know about the statistic R^2 to answer this student's question, and, in fact, used his knowledge to fashion the following response. “No, we don’t have correlations for non-linear models, but we do have a statistic called R^2 that would tell us to what extent is this data, you know, does it fit the model that you’re talking about” [NT3 199-201].

Modeling is the third process included in the knowledge component for numerically analyzing quantitative data and is located in section II.E. of the summary document outline (Appendix H). Mathematical modeling is a process used to find mathematical equation(s) that mimic the behavior of bivariate quantitative data. Teachers need to understand the purpose of mathematical models, their different types, methods for finding them, ways of assessing them, and means to compare multiple models.

In this study’s observed secondary classes, the best fit line was the central model taught by the participating teachers. In multiple classrooms, students were asked to draw the line of best fit on a scatterplot and/or determine their own criteria for drawing the line of best fit. One student response was to draw the line so that half of the points were above the line and the other half were below the line [NT5 203-204]. Another student drew the line “where most dots are at” [NCP3 180]. Trying to hit the most points was another student’s criterion for the best fit line [NCP6 286-287]. What does a teacher need to know to respond appropriately to conjectures such as these? The analysis findings revealed that teachers need to know the general meaning and reason for best fit lines; the criteria of established mathematical regression lines, such as the LSRL and median-median line and why those criteria were devised; and how to think about each student response in multiple scenarios (e.g., weak and strong association, positive and negative) to assess the strengths and weaknesses of each student's response.

The statistics community considers the LSRL the line of best fit, and, as such, teachers at the secondary level need comprehensive knowledge of this model. This includes knowledge of its criterion, calculations to determine its slope and y-intercept, interpretation of the line in context, and properties of the line such as the mean of its residuals is zero. A complete description of the knowledge needed by teachers regarding the LSRL is in section II.F. of the summary document outline (Appendix H).

Ms. Tuck was involved in a teaching incident involving the slope of the line of best fit. The class studied the relationship between days since being paid and the amount of money Doc had at the end of a day. They created a best fit line describing this relationship, and its slope was negative ten. The following exchange then occurred.

Ms. Tuck: The other thing to point out about this when we're assuming, like, a constant rate, does this mean like she just told me Doc spends ten dollars a day. Does that mean he always spends ten dollars a day?

Student: No, but you should put it like that.

Ms. Tuck: No? He doesn't always spend ten dollars a day?

Student: On average.

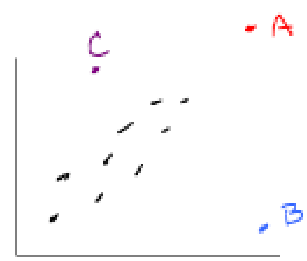
Ms. Tuck: Right. So our constant rate, we're kind of averaging it out.

[NCP5 382-385]

Applying her subject matter knowledge, Ms. Tuck chose to emphasize the point that the line's slope represents the average change in the y-variable for a one-unit change in the x-variable and that such a change in the y-variable may never actually occur in the observed data values. All teachers who teach best fit lines should have such knowledge as well.

"Outlier" and "influential" are terms that have specific meanings in statistics and are used to describe points with particular characteristics. In the context of linear regression, an outlier is a point with an unusually large error. This means it does not follow the relationship pattern shown by the other points in the data set. An influential point is one that greatly influences the location of the LSRL. Teachers need to be familiar with these terms; know how to determine if a point is an outlier or influential, and appropriate follow-up steps if one is identified; understand how to use the context to interpret such a point; and understand the relationship between an identified point and the LSRL. In the summary document outline located in Appendix H, section II.G. lists the knowledge needed by teachers related to outliers and influential points.

To create a handout for his class regarding outliers and influential points, Mr. Tablet used his knowledge of the topic to create a drawing of a scatterplot with three additional points, A, B, and C, to consider in turn as potential influential or outlying points. The drawing in Figure 5 is the one he created for his handout.



[NT4 98]

Figure 5. Scatterplot with points A, B, and C

These three points illustrate three different scenarios for consideration. Point A falls near the line fit to the rest of the data and thus is not influential. Point B has a similar x-value to Point A but because its y-value is considerably lower than that of Point A, which is where the line fit to the rest of the data would predict it to be, the line pulls down towards it and therefore Point B is influential. Point C was purposefully placed by Mr. Tablet near the average of the x-values to represent an outlier near the horizontal center of other points. When added to the data set, Point C pulls up the line slightly, but it does not exert a large influence on the location of the LSRL because there are multiple points with similar x-values. Again, Mr. Tablet needed to have a thorough understanding of these types of points and the LSRL in order to create this drawing from which his students would learn.

One of the primary uses of a model is to make predictions, and so it is important that teachers be knowledgeable about how to use models to make predictions, judge the reasonableness of those predictions, and discriminate interpolations from extrapolations. A more complete description of the relevant subject matter knowledge for predictions is in the summary document outline, section II.H. (Appendix H).

Context knowledge is often required to judge the reasonableness of an extrapolation. In the following exchange, Mr. Glass used his context knowledge to determine if a model's extrapolation to predict the temperature at the equator was realistic.

Mr. Glass: Where is zero degrees latitude?

Student: The equator.

Mr. Glass: The equator. So, this thing looks like, if we had some kind of idea what that is.

Student: It'd be really hot.

Mr. Glass: Yeah, I think I made it to, like, 100. What did I set my [graphing calculator] window at? Yeah, 100. So this one's [quadratic model] saying it's about 95,96 for prediction. But look at the linear model. It's saying, like, it's 110 and the exponential is saying 140 or something like that, way up there. So, no one has an average daily high temperature like that. So 95 maybe. Over 100, uhhuhh. [GBS12 164-170]

This illustrated another situation where context knowledge was needed to teach a topic in statistical association, which here was the topic of extrapolation.

After a model is fit to a data set, it is imperative to evaluate the appropriateness of the model. Residuals and their associated graphical display, a residual plot, can be used towards that end. A residual is a measurement of the error in a model's prediction for a given x-value. It has a sign associated with it to show if the model was under-predicting (positive residual) or over-predicting (negative residual). In addition to the definition, teachers need to know how to calculate and interpret residuals; their geometric meaning; the properties of residuals; and associated statistics calculated from residuals, including the standard deviation of residuals and the sum of squared residuals for a model. In Appendix H, sections II.I. and II.J. of the summary document outline contain more detailed descriptions of the knowledge components related to residuals and residual plots.

Following her introduction to the term “residual,” a student in Mr. Tablet’s class asked “When you measure the distance to the line in terms of, like, the residual, do you do the vertical down or do you do it, like, [student gestured perpendicular segment to the line as distance to a line is typically defined]?” [NT7 146-148]. It seems this student experienced a cognitive conflict between the definition of distance to a line she had previously learned and the distance to the line a residual was measuring. To respond to this student, the teacher needed to know that distance to a line is typically measured by the length of the perpendicular segment to the line because that is the shortest distance possible, but since a residual is a measure of the error in the prediction, it must be measured as the vertical distance, which represents the difference between the observed y-value and the predicted y-value. With this knowledge the teacher could craft an appropriate response to the student.

Graphical analysis of the residuals through a residual plot is an important technique for evaluating the appropriateness of a model and to determine whether the underlying assumptions for regression inference are met. The purposes of a residual plot, along with the knowledge of how to make and analyze it, are essential elements of this knowledge component. In a written assessment Mr.

Tablet included an item that asked the students to sketch a residual plot from their calculator and then to interpret the graph. When asked about his use of his knowledge of the subject in the grading of the quiz, Mr. Tablet had the following to say about the student responses for this item.

> Mr. Tablet: They would say it's random. Well, that's fine, but what does random mean? Or you know--
>
> Interviewer: What's the implications, they're not saying that.
>
> Mr. Tablet: Right. In other words, you're not telling me anything about what that means. [NT9 25-28]

From his knowledge of the topic, Mr. Tablet knew that it was imperative that his students understand that the randomness of the residual plot meant that the model had picked up any pattern there was in the data, and the remaining random scatter seen in the residual plot represented the expected random noise in the process.

Finally, teachers need to know how to compare mathematical models to determine which model is the best choice for a set of data. Multiple items need to be used in this decision-making process, including a scatterplot of the original data with the model(s) graphed over it; the values of statistics which measure the strength of the association such as r, r^2, sum of squared residuals, or R^2 for each model; and the residual plots. The summary document outline section II.K. found in Appendix H contains more information about the knowledge teachers need in order to teach about comparing mathematical models.

Mr. Glass included a question concerning the comparison of linear and exponential models for a data set shown in a scatterplot in a handout given to his students. Although time ran out during the class session to discuss the question, I asked him during the interview what he was thinking about as he prepared to talk about that question. He responded as follows:

> The real question is a line or an exponential curve going to give you a better r-squared value. I had ambitious plans for yesterday. I wanted to look and compare the residuals between the two. The errors between the two I cared about more than the r or r-squared value. [GBS7 124-127]

These comments showed that Mr. Glass understood the need for multiple assessments to determine the best model, including the r^2 statistic and the residuals.

<u>Graphically.</u> The scatterplot is the preferred graph for displaying quantitative bivariate data at the secondary level and hence will be the focus of this section. Teachers need to know how to construct and interpret a scatterplot. Central to the making of a scatterplot are the placement of the explanatory variable on the x-axis, the placement of the response variable on the y-axis, and the scaling

of the axes. The teaching incident transcribed below called for teacher knowledge of consistent scaling of axes for a scatterplot.

Ms. Tuck: Look where your five is.

Student: It's in the middle.

Ms. Tuck: Why is it in the middle though? And all your other ones are on the bar. Are you doing one square is one? Is a square on the graph paper one day?

Student: Uh, no, I mean I'm going to graph something different for these.

Ms. Tuck: Right, but, so, if this block is one day, there's two days, three days, four days, now all of a sudden I'm in the middle of the blocks. What's happening there?

Student: Well, what else am I supposed to do there, I mean….does it really matter?

Ms. Tuck: I think it does.

Student: Oh, Okay.

Ms. Tuck: Why wouldn't it matter in your--

Student: Well, it doesn't matter the position as long as, like, you're just trading information. [NCP4 193-201]

In this teaching incident, the teacher needed to apply her knowledge to recognize the error the student was making by having an inconsistent scale for an axis of his scatterplot and to convince the student that this was, in fact, an error.

A related knowledge piece is that changing the scale of one or both of the axes changes the pattern in the scatterplot. In fact, this is the method used for linearly transforming data. For example, if the original scatterplot of y vs. x shows exponential growth, changing the y-axis scale to the logarithm of y will result in a scatterplot that shows linear growth. Teachers need to know the method of linear transformation of data and why it works.

Teachers also need to know how to visually analyze a scatterplot to determine the form, direction (if applicable), and strength of association between the two variables. In addition they should understand how these properties can change with the addition or removal of a few data points or by looking at isolated portions of the data set rather than the data set as a whole. Mr. Tablet used this knowledge when he decided which data points in a larger data set to retain for his students to analyze. When questioned during the interview, Mr. Tablet said the following:

> But I do remember thinking I've got to pick, and actually it shows too, because I was thinking I've got to pick some of the less developed countries [x-variable] and some of the more developed countries….

> because had I picked all the ones up here [in upper left corner] who knows I might have gotten a positive slope to that set of numbers up there and then I'd really be messed up….my thought pattern too was I got to make sure that this comes out to be a negative [slope]. [NT5 431-441]

A complete description of the subject matter knowledge needed to teach scatterplots is in section II.L. of the summary document outline (Appendix H).

<u>Inference Methods.</u> Two inference methods taught at the secondary level relate to the topic of this study: a t-test for the slope of the regression line and a confidence interval for the slope of the regression line. Knowledge of hypothesis tests and confidence intervals are considered prerequisite knowledge components, and a description of the relevant prerequisite knowledge for these topics is in sections IX. and X. of the outline in Appendix G. Therefore, this section will focus on the specific knowledge needed to teach these two particular inference methods. A more complete description can be found in Appendix H: Subject Matter Knowledge Summary Document Outline sections II.M. and II.N.

The slope of the LSRL computed from sample data gives a reasonable estimate of the slope of the idealized regression line created using all the values in the population. In order to infer what the population slope is, it is important not only to have an estimate of it but also a measure of the accuracy of that estimate. To numerically assess the accuracy of the estimate, some assumptions about the model and its errors are necessary. It is important that teachers understand the need for, the particulars regarding, and how to check for the assumptions needed for these two inferential methods.

The t-test for the slope of the regression line was the final hypothesis test Mr. Tablet taught, so his students were very familiar with the format and logic of hypotheses tests by this point in the class. He started by talking about the hypotheses for the test. He asked his class “If r is equal to zero, if there’s no correlation between the variables, what’s true about the slope of the regression line?” A student responded “It’s equal to zero.” Mr. Tablet agreed, stating “It’s equal to zero too. So another thing that you could be testing here is that the slope is equal to zero or that the slope is significantly different than zero OK?” [NT23 142-145]. As seen here, a teacher needs to understand that testing whether the slope of the regression line is equal to zero is, in effect, testing whether there is a linear association between the two variables, making this a very useful and powerful test. The teacher should also understand how the assumption of the null hypothesis is used throughout the rest of the test.

Teachers need a firm foundation in the logic and computations necessary for carrying out the rest of the test, including calculating the t-test statistic, finding the p-value, and making a conclusion. Taking account of relationships between quantities, such as sample size and the t-test statistic, is also part of this knowledge component. For example, a larger sample size provides a better chance to detect an association if there really is one, and this is reflected in a formula for the t-test statistic for this test, $t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$, which numerically shows that for a fixed value of r the t-test statistic will increase as the sample size n increases. During the interview, Mr. Tablet discussed how he used this knowledge component when analyzing a scatterplot for association:

> But you know my thought is with as many points as there are, and again those are things that I'm sure it's hard for them [the students] to take into account especially after having not looked at the material, but if you have a lot of points and just the basic, you can see that there is a pattern to what's there. In my own experience I'm thinking there's a significant relationship there between these. [NT23 218-222]

The slope of the true regression line is often the most important parameter to estimate in the regression setting (Yates et al., 1999), and so knowledge concerning the confidence interval for the slope of the regression line is important for secondary teachers to have. In addition to the aforementioned knowledge concerning confidence intervals in general and the assumptions for the method, teachers need to understand the particulars of the calculation of this interval and its interpretation.

Technology

Graphing Calculator. The graphing calculator, as a technological tool, makes statistics and statistical reasoning accessible to all students (Burrill, 1996). All three teachers used graphing calculators during their teaching of statistical association. The findings point to three types of knowledge that teachers need regarding graphing calculators: commands, probable causes of errors, and limitations of the calculator.

In order to teach their students how to study statistical association of quantitative variables using the graphing calculator, teachers need to know the commands themselves. Included in these commands are those for inputting data into a list, computing a regression equation, and performing a t-test for the slope of the regression line. As teachers and students perform commands, error messages may be generated by the calculator, and it is important that the teacher know the likely causes of each message.

Ms. Tuck had an experience when she needed to know about the calculator's error message "Dimension Mismatch." She was talking the students through the commands necessary to make a scatterplot using the graphing calculator, when a student interrupted her to ask for help because her calculator gave her an error message. Ms. Tuck said the following to the student:

Dimension Mismatch, here's where you're getting if you're getting that

> error. Look at your Stat Edit, like, if you had a twelve here [List 1] and nothing here [List 2] when I go to graph it Dimension Mismatch because one column had four [entries] and another one has three. [NCP6 86-88]

Knowledge of the graphing calculator is needed by teachers to diagnose and resolve difficulties students have while using it.

Graphing calculators have memory limitations. It is important that teachers know of these limits and possible ways to work around them. A data set consisting of years 1983 through 1994 and the number of AIDS cases each of those years was used in Mr. Tablet's class. The task for the students was to calculate the exponential regression model using the graphing calculator. Knowing the graphing calculator's numerical limitations and the behavior of exponential functions, Mr. Tablet gave the following advice to his class:

> Notice, we've got the year from 1983 through 1994 Okay? Now, think about exponential on this. If we sit there and say, well, we're going to have to put 1983 in for a value as an exponent, do you understand that that could present a problem. So one of the ways whenever they use years is to say, well, you know what? Let's let 1983 be year zero, and then all of these exponents in the equation are going to be much smaller okay? I think we're going to avoid the [calculator] overflow problem when you do that. [NT10 183-189]

Such knowledge is needed by secondary teachers to be able to analyze real-world data sets, which often have years or large numbers in them, using graphing calculators. Section II.O. of the summary document outline (Appendix H) contains more information regarding what teachers need to know about graphing calculators to teach statistical association of quantitative variables.

Computer Output. The College Board (2006) included interpretation of standard computer output in its course description of Advanced Placement Statistics, which means that teachers need to have that skill themselves. Mr. Tablet made the following statement to his class regarding using reading computer output for regression: “This is the standard error right here. That’s the standard deviation [of the slope] divided by the square root of n. It’s in there already. So you don’t even have to worry about dividing by the square root of n” [NT25 60-62]. This example illustrates the importance of knowing what numbers are being reported by statistical software’s regression output so that they can be used correctly to analyze association. A list of all knowledge components related to reading computer output regarding association of quantitative variables is in section II.P. of the summary document outline (Appendix H).

Association between Categorical Variables

Data

As with data which is quantitative, teachers need to understand the usefulness of categorical data, its inherent variability, and how to differentiate between types of data. With categorical variables, these components have slightly different meanings. Data is still useful because it can help one make informed predictions, but instead of predictions of a data point's amount, the predictions will concern which particular category a data point will fall in. Part of understanding the variability of categorical variables in the area of association is knowing that an association can have exceptions. In other words, a teacher should not have the determinist misconception (Batanero et al., 1996) that all cases must show the association.

Teachers need to know the difference between quantitative and categorical data and how to classify data as one of these types. They also need to know that quantitative data can be converted to categorical data through the creation of categories based on ranges of the data's values. But there are drawbacks to this conversion. These drawbacks came out during a discussion in Ms. Tuck's class about a problem they had worked on concerning the tips left at a restaurant by adults and high school students. In the problem a tip was classified as good or bad based on an unspecified cutoff value. A student in the class brought up the point that "it doesn't describe their ability to tip well because it just says that they don't have, like, the usual amount that they should…. but a good tip could be, like, a little more than another tip" [NCP7 279-280]. This student recognized that valuable information about the quantity of the tip was lost in the categorization process. The teacher needs to be able to process this student's statement using his or her knowledge of data and their types.

Analyzing

Numerically. A numerical analysis of quantitative data typically starts by organizing the data into a two-way table to show the number of data values that fall in each cross-category. Teachers need to understand how to create a two-way table and its properties (e.g., grand total is the same for the row and column variables) in order to teach about categorical variables' association. Section III.C. of the summary document outline in Appendix H explains what teachers need to know about two-way tables.

To interpret the data values in a two-way table, teachers need to have a sense of how numbers in the table work and strong proportional reasoning skills. For example, teachers need to know that comparisons of raw numbers in the table only make sense if the totals for each category of a variable are equal; if they are not equal the numbers in the table need to be converted to proportions for comparison. Mr. Tablet needed this knowledge to respond to his student who asked “Ummm, how can you compare the “yeses” and “nos” when the total for the people that were 18 to 29 is not the same as for as over age 30?” Mr. Tablet responded “It’s…because you’re still finding a percentage,” [NT14 257-259] which showed that Mr. Tablet understood comparisons with percentages are possible since they are all ratios with the same scale. However, the student still did not understand how the comparison was possible because the student stayed after class to have the following interchange with Mr. Tablet.

> Student: Wouldn’t it make more sense, though, that they would have to have to interview the same amount of people because if they then they would have more “yeses: and more “nos” for the younger population.
>
> Teacher: Well, ideally, yes, it would be nice to do that. But if I’m ever going to compare, uh, groups that don’t have the same, you know, don’t have the same number. Because what are the chances we’re going to have the same numbers?
>
> Student: Okay, it’s not really accurate.
>
> Teacher: You have to use percentages. Well, you could still compare percentages; I mean, it’s done all the time. I mean, if you don’t have the same number you could still compare the percentages. [NT14 338-345]

From this interchange it is clear that the teacher had vital knowledge concerning proportions that the student did not.

The next step in the numerical analysis of categorical variables involves calculation of the marginal and conditional distributions. Teachers need to know each distribution's purpose, how to find it, and how to interpret it in the context of the problem. Conditional distributions are particularly important since they are used to assess whether an association exists between two variables. On the importance of conditional distributions, Mr. Tablet said "Looking at marginal distributions tells you nothing about the interaction of the two variables. Conditional distributions, you're able to see a little bit more with the association here. Okay? So, that's a key" [NT14 279-281].

Graphically. As Mr. Tablet told his class, "the graph that allows you to see anything when you're talking about two categorical variables is called the segmented bar graph" [NT15 36-37]. Teachers need to know how to make and interpret this visual display of conditional distributions. They should also understand that the conclusions are the same no matter which variable is chosen to condition on.

Inference Method. The inferential method for determining association between two categorical variables that was taught in the observed secondary classes was the chi-square test for independence. Building on their prerequisite knowledge of hypothesis tests, teachers need to know about the chi-square statistic, its distribution, and the specifics of the chi-square test for independence in order to teach students how to analyze association of categorical variables with this method. Each of these knowledge components will be described in turn in the following paragraphs. Further descriptions are available in sections III.E, III.F., and III.G. of the summary document outline in Appendix H.

There are many concepts teachers need to know to understand the chi-square statistic, including its calculation and why it is calculated as it is; the expected values that are a part of its calculation; number sense regarding the values it takes; and how to calculate it manually or with technology. In the following teaching incident involving calculating expected values, the teacher lacked the necessary knowledge for teaching this topic.

The students in Ms. Tuck's class were working on a series of problems where they were given two-way tables with only the totals filled in. They were asked to fill in the inside cells of the table with the expected values for the chi-square test for independence using their self-created methods. Below is one of the tables they were given:

	Women	Men	Total
RH			150
LH			50
Total	120	80	200

Figure 6. Two-way table with totals only

After giving the students some time to work on the problem in groups, Ms. Tuck wrote the following completed table on the board.

	Women	Men	Total
RH	90	60	150
LH	30	20	50
Total	120	80	200

Figure 7. Complete two-way table

She then said to the class "So, I've heard people tell me this should be 90, 60, 30, 20, but I need someone to explain why." A student replied "Well, I did, because earlier when you said on the first one [problem] that 150 is 75% of 200, so I did that for this one." The teacher responded "Oh, it works this way? If you do 150 out of 200?" The student said yes, after which the teacher asked, "Really?" The student finished explaining his work by saying "And it's 75% of 120, and it's 90." The teacher's comment to the student was "Oh, I didn't know it worked this way on this one" [NCP7 534-540]. When asked about this incident during the interview, Ms. Tuck said "And this is one where I was a bit confused on it myself. But I swear in …[a similar problem], if you look at it this way and take these numbers, take like the 150 over 200 and apply that percent to this [the numbers in the bottom row], it doesn't work." I then asked her "So you have to go along the bottom to get it right?" to which she responded, "It's unclear to me.…I was under the impression that you have to go along the bottom" [NCP7 825-844]. This teacher was missing knowledge pieces regarding proportions and expected values needed to understand that expected values can be calculated either direction; that is, the proportions of the row variable can be multiplied by the totals of the columns or the proportions of the column variable can be multiplied by the totals of the rows.

The p-value for the chi-square test for independence is the area under the chi-square distribution from the value of the test statistic to positive infinity. Thus it is important for teachers to have knowledge of the distribution of the chi-square statistic, the second knowledge component included in this section. This includes knowledge of degrees of freedom, the shape of the distribution, and the relationship between the p-value and the distribution of the chi-square statistic. The relationship between the chi-square statistic and its p-value came up in Ms. Tuck's class. Ms. Tuck explained the utility of the probability chart for the chi-square statistic with the following statements:

> All this is telling us is our goal is to try and find meaning in chi-squared. Like, I compute chi-squared, I get a four, what does that tell me?...I get a 3.7. Is that good, bad, I don't know. So by doing this probability it's saying, well, if I get a chi-squared that's really large, say, I get an eight, the probability of that happening is less than one percent. [NCP7 168-172]

Beyond this, the teacher needs to understand the reason the chi-square probability chart has the values that it does. For example, the reason that large values of the chi-square statistic have small probabilities is that if the null hypothesis of no association is true, then the differences between observed and expected values are only due to sampling variability and therefore should be minimal. If the differences between observed and expected values are not minimal, resulting in a large chi-square statistic, then there is a small chance that would happen if the null hypothesis were true and thus the p-value is small.

The third and last knowledge component is the chi-square test for independence. This includes its purpose, data types that can be used with this test, hypotheses, assumptions, procedure, and conclusions. The following example illustrates the need for knowledge concerning the designation of the hypotheses for the test.

In a chi-square test of independence, the two options are that the variables are independent or they are not independent. Mr. Tablet needed to explain to his students the necessity of making the null hypothesis the option of independence to respond to a student who asked, “Why do we make that the null hypothesis?” [NT20 107]. Mr. Tablet explained the need for that assumption when calculating the expected values for a test involving the variables physically fit and watching television in the following way:

> Two variables are independent. If I multiply the probability of being physically fit and the probability of watching no TV, I end up, there’s a two percent chance of a person being in that group. Would you agree with that based on the totals that I have? And then if there’s 1200 people total, I would have expected twenty-five; if the two variables were independent of each other, I would have expected 25.48 people to be in that cell if the two variables were independent of each other. [NT20 136-143]

Technology

Graphing Calculator. With categorical data, the participating teachers used the graphing calculator for two purposes: computing proportions and carrying out the chi-square test for independence. Teachers need to know how to perform both of these functions with the graphing calculator, as well as know common reasons for errors and limitations the calculator may impose. In the summary document outline found in Appendix H, section III.H. contains a listing of the relevant knowledge components for teaching association of categorical variables with the graphing calculator.

The following incident involving the graphing calculator reveals the interrelatedness of the knowledge components for teaching statistical association. Mr. Tablet was teaching his class how to use the χ^2cdf command to calculate the p-value for a chi-square test for independence. The syntax for the command is χ^2cdf(lower bound, upper bound, degrees of freedom). Here is the verbal exchange between Mr. Tablet and a student:

Mr. Tablet: How do we calculate the p-value? p is equal to--

Student: χ^2cdf

Mr. Tablet: cdf

Student: Then, you do the--

Mr. Tablet: 37.76 [chi-square value they had already computed for this problem] comma--

Student: comma 1000

Mr. Tablet: Put 1000 in there. Pick a big number because we've got to go way out into that tail, and that tail, you know, it can get pretty high, so let's go to 1000. And then what do I have to include?

Student: Five.

Mr. Tablet: Five degrees of freedom. [NT18 63-67]

Here a student suggested 1,000 as the upper bound for the region under the chi-square distribution used to compute the p-value for the test. The teacher needed to know that the region should go to positive infinity, but that, due to the limitations of the calculator, a value needed to be provided for the upper bound. In addition, he needed to assess whether the student's suggestion of 1,000 was an appropriate upper bound. To do this the teacher used his knowledge of the chi-square distribution with five degrees of freedom, including number sense regarding the chi-square statistic, to determine that 1,000 was far enough to the right on the chi-square distribution that the area to the right of it under the distribution would be insignificant and so could be left out of the calculation of the p-value.

Computer Output. Teachers need to know how to interpret standard computer output for the chi-square test for independence. This involves more than just reading off the value of the test statistic and its p-value, as the next example demonstrates. Mr. Tablet was teaching his class how to read the output from Minitab, a statistical software program, which included a warning stating "6 cells with expected counts less than 1.0. Chi-Square approximation probably invalid. 6

cells with expected counts less than 5.0” [NT22 61-62]. To interpret this portion of the output the teacher needed to know about the assumptions for the chi-square test for independence, including what they are and why they are needed.

Conclusion

The findings of this study were presented in three formats: a concept map, a narrative description including examples from the records of practice, and a summary document outline (contained in Appendix H). Each presentation contained something that the others did not. The concept map gave a visual presentation that showed the relationships between the knowledge components. The narrative description linked the findings to the act of teaching and gave more meaning to the knowledge components used in the concept map. The summary document outline provided a detailed description of the knowledge components needed by teachers. Thus I hope that through all three presentations, the reader was able to come away with a comprehensive understanding of the findings of this study.

CHAPTER V

CONCLUSIONS AND RECOMMENDATIONS

Introduction

This final chapter presents the central findings of this study to determine the subject matter knowledge teachers need to teach statistical association at the secondary level. The chapter begins with a summary of the study and its primary results. The following section describes the study's contributions and implications. Limitations of the study are discussed in the next section, followed by a section providing recommendations for future research related to the present study. The chapter concludes with a description of the significance of the study.

Summary of the Study

The nature of the knowledge needed for teaching is largely under-specified and unresearched (Ball & Bass, 2003; Ball et al., 2001; National Research Council, 2001; RAND Mathematics Study Panel, 2003). While it seems obvious that mathematics teachers need to know mathematics, what is unknown is exactly what aspects of mathematics teachers need to know, how they need to know it, and how and where this mathematics knowledge is used in practice (Ball & Bass).

The need for specification of teacher's subject matter knowledge is particularly acute in the area of teaching statistics. Statistics has become an accepted strand of mainstream school mathematics curricula in the United States (Moore, 2004), but there is concern regarding the knowledge involved in teaching statistics and whether teachers in the United States have this knowledge (CBMS, 2001; Franklin, 2000; Kettenring et al., 2003; Makar & Confrey, 2004). Some have suggested (e.g., Franklin; Kettenring et al.) that the number one priority in statistics education is addressing the shortage of K-12 teachers prepared to teach statistical concepts. In order to prepare teachers to teach statistics, the subject matter knowledge needed for teaching statistics must be defined and described in detail. This study aimed to do just that with a focus on statistical association at the secondary level.

I observed three secondary mathematics teachers as they taught statistical association during the 2005-2006 school year. Following each observation, I also interviewed each teacher interviewed regarding his or her use of subject matter knowledge during the preparation, teaching, and assessment phases of teaching the class that day. This practice-based approach was chosen because it was the best means to research the topic in order to study firsthand the knowledge of content used in teaching (Ball & Bass, 2000a; Fennema & Franke, 1992). A total of fifty class sessions were documented through observation and interview transcripts, fieldnotes, handouts, copies of pages from textbooks and teachers guides, and copies of student work. I created a compilation document for each of the class sessions which brought together all of the relevant records of practice in one file.

In the analysis phase of the study, I studied each of the fifty compilation documents. The records of practice documented in each compilation document served as a catalyst for me to make conjectures regarding the subject matter knowledge needed by the teacher in a particular teaching incident. The purpose was not to document the subject matter knowledge of the participating teacher, but rather to consider what the teacher was doing or called upon to do in teaching and document the subject matter knowledge required for those tasks. Each knowledge component listed in the resulting analysis report was accompanied by a reference to the line numbers in the compilation document which spurred the description of the knowledge component. In this way the results maintained a genuine connection to the teaching practice from which they came.

Twelve compilation documents, representative of the multiple subtopics of statistical association, were selected for study by two additional analysts: a statistician and a statistics education expert. These experts followed the same process I used to create their own analysis reports. Their findings were synthesized with mine to create a summary document that describes the subject matter knowledge needed for teaching statistical association at the secondary level.

In the following sections, highlights of the summary document describing the subject matter knowledge needed by secondary teachers to teach statistical association are presented. In the first section, I detail general knowledge components, while in the following two sections, I describe the knowledge needed to teach association of quantitative and categorical variables respectively.

Statistical Association

Teachers at the secondary level need to know the overall meaning of association in statistics as the correspondence of variation of two statistical variables. To form a complete understanding of this meaning, teachers need to realize the distinction between association and causation; that it describes a general trend; and that variables can be associated yet not be termed dependent.

Fluency in the language of statistics is crucial for teachers of statistical association. Throughout the teaching process teachers must correctly use the vocabulary of the subject. Correct usage of statistical terms such as correlation, independence, and association is important for teachers. The knowledge needed to do so properly is not trivial, as terms such as these have intertwined yet distinct meanings. Teachers need to be comfortable speaking and writing with the vocabulary used in describing various aspects of statistical association. Sleep and Ball (2007) have also noted that fluency and skills with mathematical language are critical parts of mathematical knowledge for teaching.

Knowledge of context or world knowledge is another vital component of the knowledge needed for teaching statistical association. Data sets analyzed for statistical association come with a context, and the teacher needs to know that context in order to do such things as suggest a possible lurking variable or judge the reasonableness of predictions made using a mathematical model. Gal (2004) identified knowledge of context as a necessary piece for statistical literacy, and my findings agreed that it was a necessary knowledge component for teaching statistical association as well. Together, these three knowledge components, meaning, terminology, and context, are essential for teaching statistical association. In the following sections, overviews of the knowledge needed for teaching association of quantitative variables and categorical variables will be presented.

Association of Quantitative Variables

Teachers need to start with an understanding of the data with which they are working. In addition to understanding its context, which has already been described, teachers need to understand the different types of data and how to differentiate between them; the necessity of data; and the inherent variability in quantitative data.

To analyze quantitative variables for association, there are three general methods that teachers need to know. The first is a numerical analysis, which can start from an informal judgment based on a table of values and proceed to the creation of a mathematical model. Included in this knowledge base are the correlation coefficient r, the coefficient of determination r^2, the least-squares regression line (LSRL), and residual.

Graphical methods of analysis comprise the next knowledge category. The scatterplot is the primary graph used to present bivariate quantitative data, and thus that is the focal point of this category. Teachers need to know about such things as scaling effects; reasons and repercussions for the placement of the explanatory variable on the x-axis; and how to visually analyze a scatterplot for association between the variables.

The t-test and confidence interval for the slope of the regression line are the two inference methods taught at the secondary level that deal with association of quantitative variables, and so teachers need to be knowledgeable about these procedures. More than the knowledge needed to carry out these procedures, however, teachers also need to know why they are used; the meanings of the statistics and computations involved in them; and the assumptions necessary for each.

Technological tools, such as the graphing calculator and computer software programs, are used to analyze quantitative variables for association. My research showed that teachers must know how to read relevant statistical software output from standard computer programs. It is not, however, necessary that they know how to use such software to generate the output. The graphing calculator was frequently used by the participating teachers to analyze raw data, and thus teachers need to know how to carry out analysis with graphing calculators, common causes of error messages by the calculator, and the limitations of the calculator.

Association of Categorical Variables

The meaning of statistical association for categorical variables, while still involving covariation, is not exactly the same as that for quantitative variables. Teachers need to know about the nuances of the meaning of association for different types of data variables and the resulting differences in their analysis methods.

As with quantitative variables, association can be analyzed numerically, graphically, and with inferential methods, and teachers at the secondary level need to have knowledge of all three. For numerical analysis, a two-way table is the preferred data representation, and thus teachers need to know how to create and numerically analyze one. This knowledge is dependent upon a solid sense of proportional reasoning, which is needed to understand the analysis of conditional distributions necessary for determination of association. Graphical analysis of association of categorical variables can be done with segmented bar graphs, which teachers need to know how to create and interpret.

The chi-square test for independence is the inferential method secondary teachers must know to teach statistical association of categorical variables. Components of the knowledge needed for this method include comprehension of the chi-square statistic, its distribution, and the reasoning behind the statement of no association as the null hypothesis for the test. Finally, teachers need to know how to use the calculator to analyze categorical variables for association and how to read standard computer output for chi-square tests for independence.

Contributions and Implications

The primary result of this study, the detailed description of the subject matter knowledge needed for teaching statistical association at the secondary level, has uses for several communities: secondary statistics teachers, post-secondary statistics teachers, teacher educators, developers of curriculum for teachers, policy makers, assessment developers, and researchers in statistics education. The uses of the findings of the study will be explained by addressing each of these communities in turn.

Secondary statistics teachers can use the summary document both as a learning tool and an informal assessment tool. Although some recent textbooks (e.g., Bock, Velleman, & De Veaux, 2007) have much of the content included in the summary document, it is unlikely that any textbook would have everything contained in the summary document. Shaughnessy and Chance (2005) wrote a book that answered some of the most common content-related questions that arise in the statistics classroom, a reflection of the fact that common statistics textbooks do not contain all of the content that statistics teachers need. Thus the summary document may be used as a reference from which secondary statistics teachers can learn more about statistical association and its associated prerequisite knowledge. Secondary teachers can also use the summary document to gauge their own knowledge in the field, determine areas of strengths and weaknesses, and determine any follow-up actions that may be needed (if any).

Pre-service secondary mathematics teachers are usually required to take a college statistics course as part of their graduation requirements (Schaeffer, 2005). The summary document can help inform the statistics teachers responsible for designing and implementing the curriculum for such a class as to the knowledge needed by these teachers as they go out into the workforce. In a similar way, the summary document can also aid teacher educators and developers of curriculum for teachers as they design both pre-service and in-service teacher education.

Many policy documents (e.g., CBMS, 2001; NCTM, 2005) have emphasized the importance of a strong subject matter knowledge base for teaching, but none have described this knowledge base at the level of depth or detail contained in the summary document from this study. The findings of this study can contribute towards an expanded description of the subject matter knowledge for teaching statistics and help policy makers make decisions regarding the requirements for teaching certification. In this same vein, the findings of this study can speak to the content that should be contained in assessments for teachers used in the certification process.

The statistical literacy model defined by Gal (2004) included a knowledge component which had three knowledge areas relevant to the current study: statistical knowledge, mathematical knowledge, and context knowledge. While I did not categorize the findings of this study into these areas, it is possible to find multiple examples that would fit into each category from the summary document outline. Moreover I identified additional knowledge components that would not fit under any of these areas, including knowledge regarding technology and relationships among statistics (e.g., slope of the LSRL and the correlation coefficient). My findings could thus be used by statistics education researchers like Gal to inform revisions to their work.

Limitations

In this section, I identify a number of limitations in my study. First, the validity of the results is limited. With only three participating teachers in the study, the observed practice of teaching was not comprehensive. It is certainly possible that teachers in the United States at the secondary level currently teaching statistical association could be using or needing subject matter knowledge that was not identified by my study. There were multiple efforts made to minimize this possibility, such as including teachers of three different courses using three different textbooks and curricula, but it remains a limitation of the study. Therefore the results should be considered a model for discussion and further research, with the potential for future refinement and improvement.

As the only observer and interviewer, I created all of the compilation documents and the transcripts within them. While the intent was for the compilation documents to accurately represent the observed class sessions and interviews, they were subject to my interpretation. All of the analysis was based upon the compilation documents since not all of the analysts could participate in the observations and interviews, and, thus, there is the potential for misinterpretation. The checking of the transcripts by the participating teachers was included in the procedures of the study to minimize the misinterpretation of the observations and interviews.

Another limitation is the restricted number of compilation documents analyzed by all three analysts. Of the fifty total compilation documents, twelve were studied by all three analysts, and the remaining documents were studied only by me. It is possible that the data in the remaining documents would have been interpreted differently by other analysts. Allowing for the fallibility of human beings, it is also possible that there are aspects of the subject matter knowledge that one or all three analysts missed entirely. Errors in the analysis process were minimized through a review of the summary document, which contained the results from all of the analysis reports by all three analysts and all three participating teachers. In addition, the overall agreement of the findings of the multiple analysts for the twelve selected compilation documents indicates that my analysis was on target.

Recommendations for Future Research

There are many directions that future research regarding the mathematical knowledge for teaching statistics may take. One possible direction would involve applying the results of this study to other sources that address the mathematical knowledge needed for teaching. Stylianides and Ball (2004) suggested that the study of mathematical knowledge for teaching be done in an iterative cycle with school mathematics practice at its core. The present study started at the core by taking a practice-based approach to examine the mathematical knowledge for teaching statistical association. Now the results of this study can be compared to the other five sources for studying mathematical knowledge in the framework of Stylianides and Ball: policy documents, teachers' mathematical curricula, studies of teachers' mathematical knowledge done outside the classroom, students' mathematics curricula, and studies of students' mathematical knowledge conducted outside of the normal classroom setting. For example, a curriculum analysis could be done to study the content of the courses designed for prospective teachers to learn about statistical association and see how that compares with the knowledge components needed by teachers as found by this study. The current findings from practice could also be used to examine the utility of recommendations found in policy documents that set standards for teacher education. Again, the results of this study can provide the impetus to reconsider what is recommended for teachers to know, what is promoted in curricula materials for teachers and students, or what aspects of teachers' and students' knowledge are deemed important to examine in the content area of statistical association.

A second area of research could focus upon efforts to help teachers develop the knowledge needed for teaching statistical association as identified by this study. Work needs to be done to determine how best to help teachers not only possess this knowledge but possess it in a way that is readily accessible and ready to be used in the unpredictable arena of the classroom (Ball & Bass, 2000a). This should include learning opportunities for both prospective and in-service teachers through offerings like courses or teacher study groups.

There is also a need for research describing when subject matter knowledge is used by teachers as they teach statistics in order to better support and understand their work. In the analysis process, the analysts of this study were asked to identify the occasions when the teacher needed to call on his or her subject matter knowledge of statistical association. However, due to the large quantity of data to contend with and the time constraints for completing this study, this portion of the analysis was not completed. It is hoped that this addition to my study could provide insight into how the knowledge of statistics interacts with the actual work of teaching it.

The research of Hill et al. (2005) showed that greater knowledge of mathematics by teachers was linked to higher achievement of students. Another line of research could investigate whether this is true for teachers and students in the area of statistical association, using the description of the knowledge needed for teaching statistical association from the present study as the foundation. Research that seeks evidence for the effects of greater teacher knowledge is important to justify the work that is done in teacher education and to attempt to establish empirically the relationship between teacher knowledge and student achievement (Ball & Hill, 2005).

This study used the framework of teacher knowledge developed by the researchers at the University of Michigan's Mathematics Teaching and Learning to Teach Project. In their framework (Ball & Hill, 2005), teacher knowledge is divided into two main parts: pedagogical content knowledge and subject matter knowledge. Pedagogical content knowledge consists of two subcategories: (1) the knowledge of students in interaction with particular content, and (2) knowledge of teaching and the curriculum. Subject matter knowledge also consists of two components: common content knowledge and specialized content knowledge. As a result of this study, a description of teachers' subject matter knowledge for teaching statistical association was created but it did not differentiate between common and specialized knowledge. Research to understand the differences between and separately depict common and specialized subject matter knowledge for teaching statistical association could be an extension of the present study. In addition work to determine what pedagogical content knowledge is needed by teachers of statistical association could be carried out. This would provide a more complete description of the knowledge needed by teachers to teach statistical association, which could be informative for teachers, teacher educators, education researchers, and policy makers.

The current study may also lead to research regarding an overall framework describing the knowledge for teaching statistics. Frameworks describing the knowledge for teaching geometry (Wing, Driscoll, & Heck, 2006) and the knowledge for teaching algebra (Ferrini-Mundy, Floden, McCrory, Burrill, & Sandow, 2005) have been developed by researchers in those fields. A similar general framework for the field of statistics has not yet been developed.

The practice-based approach, the methodology I used in this study, was the most desirable approach for studying the subject matter knowledge for teaching statistics but was also extremely labor intensive. Research could be done to study alternative methods and compare their results to those of this study. For example, how would a description of the subject matter knowledge created by a panel of experts compare my study's description created from studying practice? How would these descriptions compare to the content in the relevant pages of a statistics textbook? Such research would contribute towards a better understanding of the advantages and disadvantages of certain research approaches for studying knowledge for teaching.

Finally, this study was limited to the secondary level of schooling. Similar studies could be done at the elementary, middle school, and college levels to describe the subject matter knowledge for teaching statistical association. This would result in a thorough description of the subject matter knowledge needed for teaching the topic and allow an analysis of how that knowledge differs based on the age of the students participating in the learning process.

Significance of this Study

This study is unique because it is the first to use a practice-based approach to study the subject matter knowledge needed for teaching statistics at the secondary level. In the literature, there has been a call for research on teacher knowledge that studies it in the context where it is used (Ball & Bass, 2000a; Ball et al., 2001; Fennema & Franke, 1992; Sorto, 2004). This study has made a significant contribution to the field by taking a practice-based, direct approach to study the knowledge for high quality teaching as it occurs in the classroom setting. My conclusions regarding teacher knowledge are grounded in the work of high quality teaching and thus better able to improve teacher practice and policy (Ball et al., 2001).

In addition the present study used the practice-based approach to examine teacher knowledge in a novel setting, that of teaching statistics at the secondary level. To my knowledge no research has been done that studied the knowledge needed for teaching statistics at the secondary level through a practice-based approach. Thus this study contributes to the research base regarding the subject matter knowledge for teaching statistics, an area of need identified in the literature (Burgess, 2007; Shaughnessy, 2007), through the most desirable way of studying it, in the context where it occurs.

Research by Ball and her colleagues (Hill et al., 2005) found that higher student achievement occurs with those students whose teachers have higher mathematical knowledge after controlling for significant student and teacher-level covariates. Therefore students' mathematics achievement can improve by improving their teachers' knowledge of mathematics. Such research has not yet been done in the area of knowledge for teaching statistics. However, the first step needed for this type of research is identification of the subject matter knowledge needed for teaching statistics, which this study has done regarding the content area of statistical association. Thus it is hoped that this study can inform the creation of a general description of the knowledge needed for teaching statistics and has the long-term potential to contribute towards improving the achievement of students in statistics.

REFERENCES

Ball, D.L. (1990). The mathematical understandings that prospective teachers bring to teacher education. *Elementary School Journal, 90*(4), 449-466.

Ball, D.L. (1991). Research on teaching mathematics: Making subject matter part of the equation. In J. Brophy (Ed.), *Advances in research on teaching* (Vol. 2, pp. 1-48). Greenwich, CT: JAI Press.

Ball, D.L. (1993). With an eye on the mathematical horizon: Dilemmas of teaching elementary school mathematics. *Elementary School Journal, 93,* 373-397.

Ball, D.L. (2002). Knowing mathematics for teaching: Relations between research and practice. In *Mathematics and Education Reform Newsletter, 14*(3), 1-5.

Ball, D.L., & Bass, H. (2000a). Interweaving content and pedagogy in teaching and learning to teach: Knowing and using mathematics. In J. Boaler (Ed.), *Multiple perspectives on the teaching and learning of mathematics* (pp. 83-104). Westport, CT: Ablex.

Ball, D.L., & Bass, H. (2000b). Making believe: The collective construction of public mathematical knowledge in the elementary classroom. In D. Phillips (Ed.), *Yearbook of the National Society for the Study of Education, Constructivism in Education* (pp. 193-224). Chicago, IL: University of Chicago Press.

Ball, D.L., & Bass, H. (2003). Toward a practice-based theory of mathematical knowledge for teaching. In B. Davis & E. Simmt (Eds.), *Proceedings of the 2002 Annual Meeting of the Canadian Mathematics Education Study Group* (pp. 3-14). Edmonton, AB: Canadian Mathematics Education Study Group.

Ball, D.L., Bass, H., Sleep, L., & Hoover, M. (2004). *Using records of practice to tackle the unsolved problem of teachers' mathematical knowledge.* Presentation at California State University-Northridge, February 23, 2004. See also http://www-personal.umich.edu/~dball/Presentations/RecentPresentations/022304_bbsh1CSUN.pdf

Ball, D.L., & Hill, H.C. (2005). *Knowing mathematics as a teacher*. Presentation at the Annual Meeting of the National Council of Teachers of Mathematics, Anaheim, California, April 8, 2005. See also http://www-personal.umich.edu/~dball/Presentations/RecentPresentations/040805_NCTM_ballhill.pdf

Ball, D.L., Hill, H.C., & Bass, H. (2005). Who knows mathematics well enough to teach third grade, and how can we decide? *American Educator*, pp.

14-46. See also http://www-personal.umich.edu/~dball/Publications /SelecteJournalArticles/BallHillBassAmericanEducator05.pdf

Ball, D.L., Lubienski, S., & Mewborn, D. (2001). Research on teaching mathematics: The unsolved problem of teachers' mathematical knowledge. In V. Richardson (Ed.), *Handbook of research on teaching* (4th ed., pp.433-456). New York: Macmillan.

Batanero, C., Estepa, A., & Godino, J.D. (1997). Evolution of students' understanding of statistical association in a computer-based teaching environment. In J.B. Garfield & G. Burrill (Eds.), *Research on the role of technology in teaching and learning statistics: Proceedings of the 1996 IASE Round Table Conference* (pp. 191-205). Voorburg, Netherlands: International Statistical Institute.

Batanero, C., Estepa, A., Godino, J.D., & Green, D. (1996). Intuitive strategies and preconceptions about association in contingency tables. *Journal for Research in Mathematics Education, 27,* 151-69.

Begle, E.G. (1979). *Critical variables in mathematics education: Findings from a survey of the empirical literature.* Washington, D.C.: Mathematics Association of America and the National Council of Teachers of Mathematics.

Ben-Zvi, D., & Garfield, J. (2004). Statistical literacy, reasoning, and thinking: Goals, definitions, & challenges. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning, and thinking* (pp. 3-15). Dordrecht, Netherlands: Kluwer Academic Publishers.

Bock, D.E., Velleman, P.F., & De Veaux, R.D. (2007). Stats: Modeling the world (2nd ed.). Boston, MA: Pearson Education, Inc.

Burgess, T.A. (2007). *Investigating the nature of teacher knowledge needed and used in teaching statistics.* Unpublished doctoral dissertation, Massey University.

Burrill, G. (1996). *Graphing calculators and their potential for teaching and learning statistics.* Presentation at the International Association for Statistical Education Round Table, Granada, Spain. See also http://www.stat.auckland.ac.nz/~iase/publications/8/2.Burrill.pdf

Casey, S.A. (2005). *Teachers' conceptions of association.* Unpublished manuscript.

College Board (2006). *Advanced Placement Statistics Course Description.* New York, NY: Author.

Conference Board of the Mathematical Sciences (CBMS) (2001). *The Mathematical Education of Teachers.* Washington, D.C.: Mathematical

Association of America; Providence, R.I.: American Mathematical Society.

Creswell, J.W. (1998). *Qualitative inquiry and research design: Choosing among five traditions.* Thousand Oaks, CA: Sage Publications.

delMas, R.C. (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). Dordrecht, Netherlands: Kluwer Academic Publishers.

Denzin, N.K. (1978). *The research act: A theoretical introduction to sociological methods* (2nd ed.). New York: McGraw-Hill.

Estepa, A., & Batanero, C. (1996). Judgments of correlation in scatterplots: Students' intuitive strategies and preconceptions. *Hiroshima Journal of Mathematics Education, 4,* 25-41.

Estepa, A., & Sanchez-Cobo, F.T. (1998). Correlation and regression in secondary school textbooks. In L. Pereira-Mendoza, L. Seu, T. Wee, & W.K. Wong (Eds.), *Proceedings of the Fifth International Conference on the Teaching of Statistics* (vol. 2, pp. 671-676). Voorburg, Netherlands: International Statistical Institute.

Estepa, A., & Sanchez-Cobo, F.T. (2003). Evaluacion de la comprension de la correlacion y regression a partir de la resolucion de problemas. *Statistics Education Research Journal, 2*(1), 54-68.

Fendel, D., Resek, D., Alper, L., & Fraser, S. (1998). *Is There Really a Difference?*. Emeryville, CA: Key Curriculum Press.

Fendel, D., Resek, D., Alper, L., & Fraser, S. (2004). *The Overland Trail* (2nd ed. field test version). Emeryville, CA: Key Curriculum Press.

Fennema, E., & Franke, M.L. (1992). Teachers' knowledge and its impact. In D.A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 147-164). Reston, VA: NCTM.

Ferrini-Mundy, J., Floden, R., McCrory, R., Burrill, G., & Sandow, D. (2005). *A conceptual framework for knowledge for teaching school algebra.* Accessed April 17, 2008 at http://www.msu.edu/~kat/ESMarticle.doc

Ferrini-Mundy, J., & Graham, K. (2003). The education of mathematics teachers in the United States after World War II: Goals, programs, and practices. In G.M. Stanic & J. Kilpatrick (Eds.), *A history of school mathematics* (pp. 1193-1308). Reston, VA: NCTM.

Flanders, J.R. (1994). Textbooks, teachers, and the SIMS test. *Journal for Research in Mathematics Education, 25*(3), 260-278.

Franke, M.L., Kazemi, E., & Battey, D. (2007). Mathematics teaching and classroom practice. In F. Lester, Jr. (Ed.), *Second Handbook of Research on Mathematics Teaching and Learning* (pp. 225-256). Reston, VA: NCTM.

Franklin, C. (2000). Are our teachers prepared to provide instruction in statistics at the K-12 levels? *Mathematics Education Dialogues,* October. Accessed December 7, 2005 at http://www.nctm.org/dialogues/2000-10/areyour.htm

Franklin, C. (2004). *Guidelines for assessment and instruction in statistics education preK-12 report.* Presentation at the Joint Statistical Meetings, Toronto, Canada, August 10, 2004. See also http://it.stlawu.edu/~rlock/gaise

Franklin, C., Kader, G., Mewborn, D. S., Moreno, J., Peck, R., Perry, M., Schaeffer, R. (2005). *A curriculum framework for preK-12 statistics education.* Paper presented to the American Statistical Association Board of Directors, March 17, 2005. Accessed December 12, 2005 at http://it.stlawu.edu/~rlock/gaise/GAISEPreK-12.doc

Gagné, R.M. (1985). *The conditions of learning.* New York, NY: Holt, Rinehart, & Winston.

Gal, I. (2004). Statistical literacy: Meanings, components, responsibilities. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning, and thinking* (pp. 47-78). Dordrecht, Netherlands: Kluwer Academic Publishers.

Garfield, J., & Ben-Zvi, D. (2004). Research on statistical literacy, reasoning, and thinking: Issues, challenges, and implications. In D.

Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning, and thinking* (pp. 397-409). Dordrecht, Netherlands: Kluwer Academic Publishers.

Glaser, B.G., & Strauss, A.L. (1967). *The discovery of grounded theory.* Chicago, IL: Aldine.

Grossman, P.L. (1990). *The making of a teacher: Teacher knowledge and teacher education.* New York: Teachers College Press.

Hall, R. (2000). Videorecording as theory. In A.E. Kelly & R.A. Lesh (Eds.), *Handbook of research design in mathematics and science education* (pp. 647-664). Mahwah, NJ: Erlbaum.

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

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

Kettenring, J., Lindsay, B., & Siegmund, D. (Eds.) (2003). *Statistics: Challenges and opportunities for the 21st century.* National Science Foundation Report. Accessed November 14, 2004 at http://www.stat.psu.edu/%7Ebgl/nsf_report.pdf

Key Curriculum Press (2008). *Interactive Mathematics Program.* [Brochure]. Emeryville, CA: Author. Accessed January 13, 2008 at http://www.keypress.com/Documents/IMP/IMPBrochure.pdf

Konold, C., & Higgins, T. (2003). Reasoning about data. In J. Kilpatrick, W.G. Martin, & D. Schifter (Eds.), *A research companion to principles and standards for* School *Mathematics* (pp.193-215). Reston, VA: NCTM.

Kvale, S. (1996). *InterViews: An introduction to qualitative research interviewing.* Thousand Oaks, CA: Sage Publications.

Lampert, M. (1990). When the problem is not the question and the solution is not the answer: Mathematical knowing and teaching. *American Educational Research Journal, 27*(1), 29-63.

Lesh, R., & Lehrer, R. (2000). Iterative refinement cycles for videotape analyses of conceptual change. In A.E. Kelly & R.A. Lesh (Eds.), *Handbook of research design in mathematics and science education* (pp. 665-708). Mahwah, NJ: Erlbaum.

Ma, L. (1999). Knowing and teaching elementary mathematics: Teachers' understanding of fundamental mathematics in China and the United States. Mahwah, N.J.: Erlbaum.

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

Mathematical Association of America (MAA) (1983). *Recommendations on the mathematical preparation of teachers.* Washington, D.C.: MAA.

Mathematical Association of America Committee on the Undergraduate Program in Mathematics (1961). *Recommendations for the training of teachers of mathematics.* Buffalo, NY: MAA.

Mathematical Association of America Committee on the Undergraduate Program in Mathematics (1971). *Recommendations on course content for the training of teachers of mathematics.* Washington, D.C.: MAA.

Merriam, S.B. (1998). *Qualitative research and case study applications in education.* San Francisco, CA: Jossey-Bass Publishers.

Mickelson, W.T., & Heaton, R.M. (2004). Primary teachers' statistical reasoning about data. In D. Ben-Zvi & J.Garfield (Eds.), *The challenge of developing statistical literacy, reasoning, and thinking* (pp. 327-352). Dordrecht, Netherlands: Kluwer Academic Publishers.

Moore, D. (1990). Uncertainty. n L.A. Steen (Ed.), *On the shoulders of giants: New approaches to numeracy* (pp. 95-137). Washington, D.C.:

National Academy Press.

Moore, D. (2004). Foreword. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning, and thinking* (pp. ix-x). Dordrecht, Netherlands: Kluwer Academic Publishers.

Moritz, D. (2004). Reasoning about covariation. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning, and thinking* (pp. 227-255). Dordrecht, Netherlands: Kluwer Academic Publishers.

Moschkovich, J.N., & Brenner, M.E. (2000). Integrating a naturalistic paradigm into research on mathematics and science cognition and learning. In A.E. Kelly & R.A. Lesh (Eds.), *Handbook of research design in mathematics and science education* (pp. 457-486). Mahwah, NJ: Erlbaum.

National Board for Professional Teaching Standards (2002). *What teachers should know and be able to do.* Arlington, VA: Author. Accessed May 9, 2008 at http://www.nbpts.org/the_standards/the_five_core_propositio

National Commission on Teaching and America's Future. (1996). *What matters most: Teaching for America's future.* New York: Author.

National Council of Teachers of Mathematics (NCTM) (1981). *Guidelines for the preparation of teachers of mathematics.* Reston, VA: Author.

NCTM (1989). *Curriculum and evaluation standards for school mathematics.* Reston, VA: Author.

NCTM (2000). *Principles and standards for school mathematics.* Reston, VA: Author.

NCTM (2005). Highly qualified teachers: A position of the National Council of Teachers of Mathematics. *NCTM News Bulletin, 42*(3), 4.

National Mathematics Advisory Panel (2008). *The final report of the National Mathematics Advisory Panel.* Washington, D.C.: United States Department of Education. Accessed May 9, 2008 at http://www.ed.gov/about/bdscomm/list/mathpanel/report/final-report.pdf

National Research Council (2001). *Knowing and learning mathematics for teaching: Proceedings of a workshop.* Washington, D.C.: National Academy Press.

No Child Left Behind Act of 2001 (2002). Public Law No. 107-110, 115 Stat 1425.

Patton, M.Q. (1990). *Qualitative evaluation and research methods* (2nd ed.). Thousand Oaks, CA: Sage Publications.

Peterson, P.L. (1988). Teachers' and students' cognitional knowledge for classroom teaching and learning. *Educational Researcher, 17*(5), 5-14.

RAND Mathematics Study Panel (2003). *Mathematical proficiency for all students: Toward a strategic research and development program in mathematics education.* Washington, D.C.: Office of Education Research and Improvement, U.S. Department of Education.

Romberg, T.A., & Carpenter, T.P. (1986). Research on teaching and learning mathematics: Two disciplines of scientific inquiry. In M.C. Wittrock (Ed.), *Handbook of Research on Teaching* (3rd ed.) (pp. 850-873). New York: Macmillan Publishing Company.

Romberg, T.A., & Webb, N. (1993). The role of the National Council of Teachers of Mathematics in the current reform movement in school mathematics in the United States of America. In *Science and mathematics education in the United States: Eight innovations* (pp. 143-182). Paris, France: OECD.

Roschelle, J. (2000). Choosing and using video equipment for data collection. In A.E. Kelly & R.A. Lesh (Eds.), *Handbook of research design in mathematics and science education* (pp. 709-729). Mahwah, NJ: Erlbaum.

Schaeffer, R. (2005). *Statistics for all: Nearer our destination or slip sliding away?* Presentation at the United States Conference on Teaching Statistics, Columbus, Ohio, May 20, 2005.

Schmidt, W.H., McKnight, C.C., Cogan, L.S., Jakwerth, P.M., & Houang, R.T. (1999). *Facing the consequences: Using TIMSS for a closer look at U.S. mathematics and science education.* Dordrecht, Netherlands: Kluwer Academic Publishers.

Senk, S., Viktora, S., Usiskin, Z., Ahbel, N., Highstone, V., Witonsky, D., Rubenstein, R., Schultz, J., Hackworth, M., McConnell, J., Aksoy, D., Flanders, J., Kissane, B. (1988). *Functions, Statistics, and Trigonometry* (2nd ed.). Glenview, IL: Scott Foresman Addison Wesley.

Shaughnessy, J.M. (2007). Research on statistics learning and reasoning. In F. Lester, Jr. (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 957-1009). Reston, VA: NCTM.

Shaughnessy, J.M., & Chance, B. (2005). *Statistical questions from the classroom.* Reston, VA: NCTM.

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

Simon, M.A., & Blume, G.W. (1994). Building and understanding multiplicative relationships: A study of prospective elementary teachers. *Journal for Research in Mathematics Education, 25*, 472-494.

Sleep, L., & Ball, D.L. (2007). *How does mathematical language figure in the work of teaching?* Presentation given at the Association of Mathematics

Teacher Educators Annual Meeting, Irvine, CA, January 27, 2007.

Smith, L.M. (1978). An evolving logic of participant observation, educational ethnography, and other case studies. In L. Shulman (Ed.), *Review of Research in Education* (pp. 316-377). Itasca, IL: Peacock.

Sorto, M.A. (2004). *Prospective middle school teachers' knowledge about data analysis and its application to teaching.* Unpublished doctoral dissertation, Michigan State University.

Stake, R.E. (1995). *The art of case study research.* Thousand Oaks, CA: Sage Publications.

Stigler, J.W., Gonzales, P., Kawanaka, T., Knoll, S., & Serrano, A. (1999). *The TIMSS videotape classroom study: Methods and findings from an explanatory research project on eighth-grade mathematics instruction in Germany, Japan, and the United States.* Washington, D.C.: National Center for Education Statistics.

Stylianides, A.J., & Ball, D.L. (2004). *Studying the mathematical knowledge needed for teaching: The case of teachers' knowledge of reasoning and proof.* Paper presented at the 2004 Annual Meeting of the American Educational Research Association, San Diego, CA, April 14, 2004.

Thomas, C.S. (2002). Data analysis knowledge among preservice elementary education teachers. In B. Phillips (Ed.), *Developing a statistically literate* *society: Proceedings of the sixth international conference on the teaching of statistics.* Voorburg, Netherlands: International Statistical Institute. Accessed May 17, 2008 at http://www.stat.auckland.ac.nz/~iase/publications/1/10_23_ct.pdf

Thompson, P., & Thompson, A. (1994). Talking about rates conceptually, Part I: A teacher's struggle. *Journal for Research in Mathematics Education, 25*, 279-303.

Tobin, K. (2000). Interpretive research in science education. In A.E. Kelly & R. A. Lesh (Eds.), *Handbook of research design in mathematics and science education* (pp. 487-512). Mahwah, NJ: Erlbaum.

United States Department of Education. (2002). *Meeting the highly qualified teachers challenge: The secretary's annual report on teacher quality.* Washington, D.C.: Author.

Usiskin, Z. (2000). Teachers' mathematics: A collection of content deserving to be a field. *UCSMP Newsletter,* 28.

Usiskin, Z., Peressini, A., Marchisotto, E.A., & Stanley, D. (2003). *Mathematics for high school teachers: An advanced perspective.* Upper Saddle River, NJ: Prentice Hall.

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

Wilson, P.S. (2004). *GAISE report discussion.* Presentation at the Joint Statistical Meetings, Toronto, Canada, August 10, 2004. See also http://it.stlawu.edu/~rlock/gaise

Wing, R., Driscoll, M., & Heck, D. (2006). *A habits of mind framework supports knowledge for teaching geometry.* Presentation at the Research Pre-session of the Annual Meeting of the National Council of Teachers of Mathematics, St. Louis, Missouri, April 25, 2006.

Yackel, E., & Hanna, G. (2003). Reasoning and proof. In J. Kilpatrick, W.G. Martin, & D. Schifter (Eds.), *A research companion to the principles and standards for school mathematics* (pp. 227-236). Reston, VA: NCTM.

Yates, D.S., Moore, D S., & McCabe, G.P. (1999). *The Practice of Statistics: TI-83 Graphing Calculator Enhanced.* New York: W.H. Freeman.

APPENDIX A

INTRODUCTORY SESSION TEACHER HANDOUT

AN OVERVIEW OF THE STUDY

RESEARCH QUESTION:
Using a practice-based approach, my research will focus on the following question:

> What is the subject matter knowledge for teaching statistical association at the secondary school level?

DESCRIPTIONS OF TERMS IN RESEARCH QUESTION:

Practice-based approach: studies teachers as they work, similar to a job analysis; contrasts with other approaches such as studying policy documents or textbooks

Subject matter: the content one is teaching (here it is statistical association), as opposed to pedagogy; includes mathematics, statistics, and context/world knowledge

Subject matter knowledge for teaching: understanding of the content entailed by the work of teaching; as opposed to pedagogical content knowledge; includes common content knowledge and specialized content knowledge.

Statistical association: correspondence of variation of two statistical variables that are both numerical (typically displayed with scatterplots) or both categorical (typically displayed with two-way tables); its study can include the following:

Numerical: data displays (e.g., table of values, scatterplots), determining if there is an association between the variables, modeling, correlation coefficient r, coefficient of determination r^2, regression line or curve, linear transformation of variables, inference for regression, correlation vs. causation

Categorical: data displays (e.g., two-way tables, ribbon charts), determining if there is an association between the variables, chi-square test of independence

Work of teaching: tasks in which teachers engage, and the responsibilities they have, to teach mathematics, both inside and outside of the classroom

METHODS:

Observation: All class sessions which include teaching and learning about statistical association will be observed; class sessions which only involve non-verbal assessment of students' knowledge of statistical association will not be observed. The class sessions will be audio recorded with a digital voice recorder equipped with a tie-clip microphone you will wear. I will also take fieldnotes, with a focus on incidents involving your subject matter knowledge to discuss at the interview.

Interview: I will conduct an interview of you for each class session observed. Ideally this will occur immediately following the observed class session, but at most it should occur one day after the observed class session. It will be a semi-structured interview, meaning there are some pre-determined questions that I will ask if they are relevant to the observed class session (see attached interview protocol) and there will also be questions formulated by me that relate to the specific class session observed. The time length of each of these will vary based on how much we have to discuss, but I will try to keep them to 30 minutes maximum. These will also be audio recorded.

DATA SOURCES:

- Audiotapes, transcripts, & fieldnotes of observed class sessions
- Handouts given to students
- Copies of relevant student work with names removed; I'll let you know if I am interested in getting this for a particular assignment
- Copies of your documentation of the lesson (plans, notes, or journal reflections); I am not asking that you make anything for me you wouldn't do normally
- Audiotapes & transcripts of interviews

MEMBER CHECKS:

At two points in the analysis of the data I will be asking for you to "check" my work. The first occasion will be your reading of the transcripts I will make from the observed class sessions and interviews. You will have the opportunity to read through the transcripts and let me know if you think I have accurately documented what happened in the class session or interview. Discrepancies between your recollection and the transcription will be discussed between us, with the goal of creating an accurate representation in mind.

The second member check will occur after the data have been analyzed and provisional answers to the research questions have been developed. The results of the analysis will be brought back to you for your review. You can give me feedback, including specific areas of agreement and disagreement.

WHAT DO I NEED TO DO AS A PARTICIPATING TEACHER?

1. Sign your consent form.
2. Provide me with an extra copy of the textbook and other instructional aids.
3. Discuss the study with your class and distribute the consent forms for students & parents. If any student or parent declines involvement, I can not include his or her statements in my transcript and I can't copy his or her work.
4. ALERT ME TO WHAT DAYS YOU PLAN TO TEACH STATISTICAL ASSOCIATION. If you could please let me know at least a week in advance, that would be great! I know things can change, but due to my child care needs I can't

come at a moment's notice, so I really appreciate your help with this. We will plan when to do the interview session too. You can alert me via email sacasey@ilstu.edu or phone.

5. On observation/interview days, please arrange for me to have a seat, copies of student handouts, and clothes with a pocket for the recorder to go in!

THANK YOU SO MUCH FOR YOUR HELP!

APPENDIX B

INTERVIEW PROTOCOL

Verbal assent: I would like to ask you some questions about your use of subject matter knowledge involved in this lesson. Recall that my definition of subject matter knowledge consists of mathematical knowledge, statistical knowledge, and context knowledge. Would that be alright? [If affirmative, proceed; otherwise, end interview].

Planning of the lesson:

Specifically reflecting on your thinking involving subject matter knowledge, talk me through your planning of the lesson.

Did you use any resources, such as textbooks or web pages, in helping your plan your lesson? If so, what? Did you modify the content of these resources? Why? What were you thinking about from a content perspective as you did this?

How did you choose any mathematical representations for use in the lesson? What knowledge of subject matter did you use in choosing them?

How did you choose the definitions of terms and their symbols used in the lesson?

How did your knowledge of the topic(s) covered in the lesson relate to your knowledge of other mathematical or statistical topics?

Teaching of the lesson:

Revisit with me the class session, explaining to me your use of your subject matter knowledge as you proceeded.

As students posed ideas and questions, what were you thinking about from a subject matter perspective as you listened and responded to them?

Was there any point during class where you needed to analyze student(s) solutions? If so, please describe that to me. What subject matter thinking did you do in this analysis?

Were there any points during class where you found students having difficulty? If so, what were you thinking about from a subject matter perspective as you helped them manage those difficulties?

<u>Assessment of the lesson:</u>

What subject matter knowledge did you use to analyze the students' solutions, correct or incorrect?

From a subject matter perspective, what were you thinking about during the design of summative assessments? During their evaluation?

APPENDIX C

SAMPLE COMPILATION DOCUMENT

GBS DAY 11 COMPILATION DOCUMENT

GBS DAY 11 OBSERVATION

Topic: 2.8 Choosing a good model

Setting: Intro activity

Teacher wrote: Read 2.8

1. What is a residual?
2. What is a residual plot?
3. What is a good “pattern” for a residual plot?
4. What is a bad “patterns” for a residual plot?

OK, I want you to open your books to section 2.8. You need to read that, pay careful attention to what they’re doing. This is a section where they’re pulling together a lot of things for you, OK. You have to know what a model is, they take a look at errors and that sort of stuff, and they do stuff with that. And there’s some questions that you need to be able to answer by the time you get done with your reading. While you’re doing that, surprise surprise I’m gonna come around and check your work.

(Skipped section 2.7 step functions homework discussion-irrelevant)

Setting: Discussion of intro activity

What’s a residual? *Error.* Error or…*observed minus predicted.* Deviation. OK.

So what’s a residual plot? *A scatterplot.* It’s a scatterplot that does what? *Uhh, relations of the values of the independent variable.* Yeah, you’re reading it out of the…what does it mean? *The proposed model is good or not…*Yeah, what does it mean? That’s what we’re trying to determine. But a residual plot, a scatterplot, what determines the points in the scatterplot? Where do we get that information from? Yeah. *The errors.* Errors and what? ‘Cause that’s only one coordinate. We need a second coordinate if we’re gonna plot. *The independent variable.* What’s the independent variable? *I don’t know, it depends.* OK. We get the other part of that from the observed x-value. Not the observed y, the observed x-value. So a residual plot you have ordered pairs. The first one is the observed x-value, ‘cause that’s where you calculated the error. You know you plugged the observed x into the predictor, the model, and that was your predicted y. And then you subtracted that from the observed y, OK. And that’s how you got your error. It’s the observed x value where you calculated your error, and then you put the error with it. That's how you get your residual plot.

Teacher wrote: ordered pairs ($x_{observed}$, error)

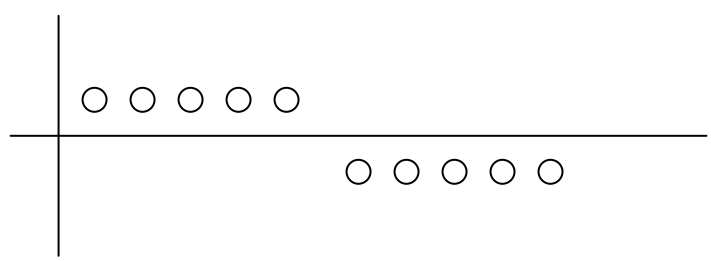
Questions about that? If you were doing this on your calculator, what two lists… OK, lets just.

Teacher wrote: <u>L1 L2 L3</u>

OK. So what would we use? Here's our…*L3 and L4, yeah.* Here's our, yeah, and these are our errors that we calculated or our residuals. (Teacher labels L1 as x_{ob}, L2 as y_{ob}, L3 as residual, error). We would use L1 and L3, we would tell the scatterplot to use L1 and L3 for the residual plot because that's where we get our observed x's and the errors. OK.

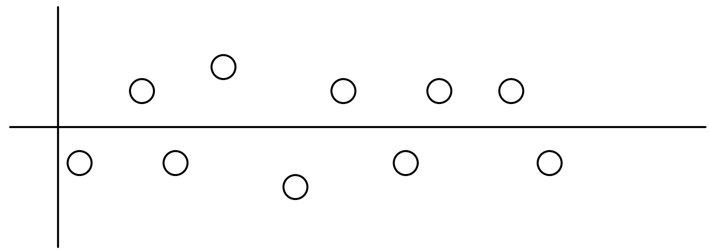
So what's a good pattern for a residual plot? How would you characterize that? *Umm a plot that has a horizontal band centered around the zero, like the x-axis (book definition).* OK, so. *It's like a bunch of dots around the x-axis.* OK.

Teacher wrote: PLOT 1



This is, who knows what the scaling is, OK. We have something there and then we have some more going this way. So if I did this, this would be a good residual plot. *No.* No. What's wrong with that for being a good residual plot? *Oh, it's negative? It has negative?* NO. Around the x-axis, you would expect things positive and negative. *Not on both sides.* Sure there is, there is some up here and some down there. *Not at the same time.* At the same time? OK. So if it looked more like a…*A dot up and then a dot exactly below it.*

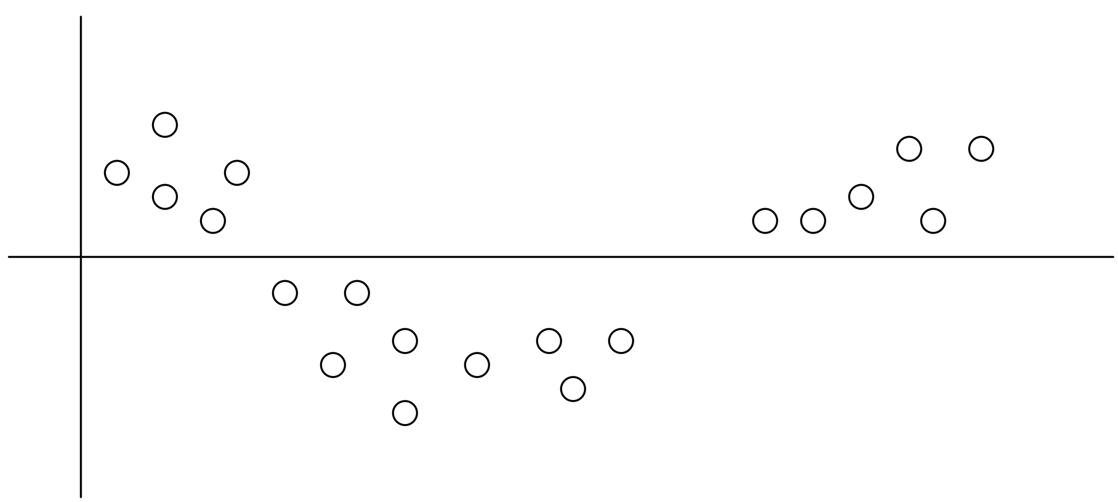
Teacher wrote: PLOT 2



You basically, you could draw a band like this and they would all be inside it. The other thing besides being in a band around zero, around the x-axis, you need some scattering. You have to change sign quite frequently. If you have one clump of them for a long way where the sign is always positive or always negative? That means that your model is missing the data consistently on one side. And you don't want that. You're trying to get as close to the data, or maybe like thread your way through the data. That's what you want for a good model. And the residual plots, if it's not sort of scattered positive and negative in a consistent band, OK. So that's a good pattern for a residual plot.

What are bad patterns for a residual plot? *Cone-like or u-shaped.* Cone-like. *Yeah, or should I say funnel-like?* Funnel-like. *Or a math term.* Well, you're just looking at a picture and telling me. Can you describe in words what we're talking about? *Both ends either positive or negative.* OK, so part of it, if we had our plot like this there would be a bunch of points here (left side positive), and then there would be a bunch of points here (middle points negative), and then there would be more points there (right side positive).

Teacher wrote: PLOT 3

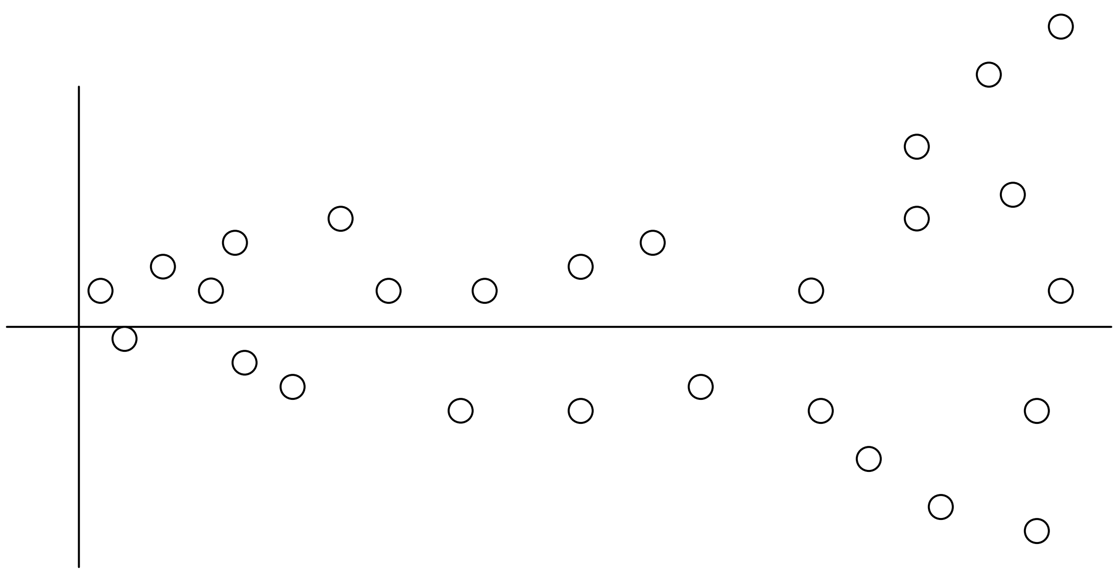


And this is just what I did when (student name) told me it had to be in this sort of narrow band close to the thing. I put a bunch on one side and then a bunch on the other side and there wasn't a whole lot of up and down. Well here, in this section OK, what this means is that all the data points are below where the model is because this is negative, OK? So if you try to make an interpolation here, you're almost guaranteed to be off. Because the model isn't really kind of going with the trend in the data. It seems to be above it consistently. At the ends, it's below it, so you want to know something? You do any kind of extrapolation and you're gonna get an underestimation if they, the, I'm assuming the errors would get even bigger cause the model is getting farther away, cause the errors are bigger at the ends. It's not a good situation either 'cause you can't extrapolate, OK.

So, good pattern, not so good pattern (pointing to PLOTS 2 & 3). What's another pattern? Can you like describe your funnel things for us (student name)? *It's like closer on one end, and like yeah, gets closer.*

What happens if it looks like this? Like near zero and the farther out you got the bigger your errors got in general.

Teacher wrote: PLOT 4

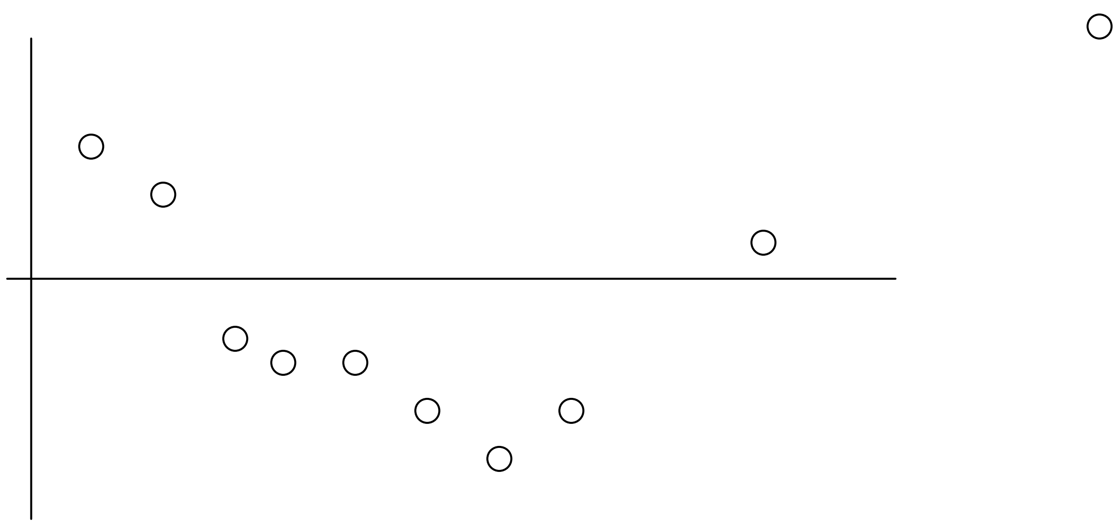


OK. So the errors, the size of the error starts to grow if you go one way or the other. That's not a good pattern. That's not in a consistent horizontal band. OK, that's what we're looking for if we're looking for a residual plot.

Now, why do we call that a good residual plot (PLOT 2)? Because what it tells us is that this model is pretty good because it's sort of staying with the data you know sometimes a little above, sometimes a little below. It’s not consistently one way or the other. *Is this one (PLOT3) better than that one (PLOT 4)? But in this one (PLOT 4) isn’t the model in the middle?* Ummm, not for extrapolation this way cause I think the errors are getting even bigger here than that one, OK. And so it’s just missing the mark, you know, that one’s (PLOT 3) consistently missing it one way, and you might be able to guess, but here (PLOT 4) because the errors are just larger and larger you can’t even guess if it’s gonna be low or high, but you just probably know it’s gonna be way off. You have a good chance of it being way off. So this is, that kind of residual plot (PLOT 2) tells us we have a good model. That one (PLOT 3) or that one (PLOT 4) tells us we need to go looking for another model. OK, we need to go looking for another model if the residuals look like that or like that.

And sometimes you don’t have enough data points for them to look like that, maybe you only have 8 or 10. But even if you only have 8 or 10 and it looks like this

Teacher wrote: PLOT 5



You know if it just looks like that that’s not good because that’s more like that pattern over there (PLOT 3). You just kind of look at it, what you want it, you actually see how you can describe a trend, here (PLOT 4) the errors are getting bigger, here (PLOT 5)the errors are above and below, they're sort of curved. This (PLOT 2) is just sort of like noise (makes static sound). There's not a whole lot of pattern to that, it's just like kind of stuck there and it's just bouncing around, OK, but it's not going bigger or whatever. What you want, this part here to be, this is why this is a good model, OK. It's (PLOT 2) sticking with the data pretty well. Here (PLOT 4) it's not. Whatever's going on, we're getting larger and larger errors. And that (PLOT 3) we're just missing the boat. OK, we're not following the trend very well. We hit a couple of pieces of data we got close to, but not not the rest of it. So when things like this or this happen, we go looking for another model. Say you fit a linear model and you got something like that over there (PLOT 3). So you know what you do? You go and try quadratic or exponential, OK. You get something like this (PLOT 5) and you have an exponential model or something, OK. You go try something else, OK. When you get one of these patterns that aren't very good for your residual plot, you try another model. OK? Even if you have an r^2 or an r, yeah we're gonna start talking about r^2, even if you have an r^2 value like .9, and you get something like that (PLOT 5), you go looking for another model, OK, because it's more than just one number that comes out of a calculator that…*You go looking for another model.* Say that was linear? *Yeah.* You pitch all the linears and you try like quadratic model or another best fit model. That's what I mean by going looking for another model. You go looking for a different family of functions. You don't look at another line, you look at a different kind of shape altogether. Because whatever shape you're using isn't fitting with what's going on.

OK. Now, so, if you look at the problems that you have to do for tomorrow. OK. In 2.8, there's a covering the reading, but the 3 problems where you actually have data, you have to start analyzing is this a good model depending on the residual plot. Do not depend on r or r^2 to answer that question. Develop a residual plot and decide, and make your decision that. Because our criteria that we're gonna use to pick a model is going to be 3. We're gonna look at r^2 and compare those. We're gonna look at residual plots and compare those. And we're gonna look at extrapolation. Do you think that whatever we're physical situation we're talking about really is going to behave like this in the long run? OK? We're gonna be looking at all 3 of those things to try and decide if we have a good model or not. So for um tonight, you're just working on that one piece, can I tell that from residuals, OK. In the writing assignment I'm going to ask you to pull all 3 of those together, OK, but just for tonight just worry about the residuals and try to analyze it depending on the residuals. OK.

GBS DAY 11 INTERVIEW

Yesterday you started off the day by writing on the board 4 questions for the students. (Read questions). What was your thoughts as you crafted those questions and put them forth for the students? I was really aiming at the last two, you know to have them make some, get some kind of feel for that or some kind of whatever and why. But if you just like put that other stuff, you're gonna just do (questions) 3 and 4, they may not know what any of that stuff is you're talking about. So they have to figure out what a residual and what a residual plot is first, and then we'll talk about whether it's a good or bad pattern. So it was just sort of like pre-cursors. *Ok, but your emphasis was on how to interpret the residual plot it sounds like. I was curious, you put pattern in quotes. Why did you do that?* I don't know. Ummm *Oh, just on a whim?* Yeah, it was more on a whim, it was like…it's not really a pattern you know, but there's some characteristic that you want to think about and that's the scattered up and down and pretty much same sized errors all the way across within a certain distance from zero. *OK, and that certain distance is…*Dependent on how big the data is and what you're measuring. They may think it has to be within 1, but if you're, I don't know, *Your units are in millions, off by one is not going to be a problem.* Off by a million is not a problem. *Gotcha.*

Umm, then you did homework on step functions which didn't relate. Then you came back to student comments and questions about stuff. The students were describing what a residual plot was, and they were saying that it's a plot of the errors. And your question was what's the other variable. What prompted you to have that question? Kids easily when they think of function only think of the y, and not the function as represented by the ordered pair, so they're leaving half of it out. And then that causes, if they don't think about that, that causes a problem when they have to figure out what they're supposed to graph. 'Cause they only have one thing, and what else am I supposed to use? It doesn't come out on it if I don't have something else, so it was that kind of thing. *So you wanted to emphasize that you need ordered pairs in order to create a scatterplot. Is that correct?* Yeah. *And that that first coordinate is going to be the observed x-values. Umm, why do you think that, or what's your understanding of the reason they use observed x-values as the variable for the x-axis?* Oh, because that's, I mean, I suppose you could put them like in order but that's, that doesn't give you a feel. You want to know you want to see as you go across the domain where you picked out the data, where you have actual observations. And the other thing, you want that scattered the same way you have your observations one, and two the error that you calculated was on a particular if you want, a vertical line, and it was along that where that x-value was fixed so you would pair it with the x that your plotting so you can get some kind of idea which error is. You know, it gives a better representation I think. *Ok, that's good.*

Umm, So are you going to tell me the real answers to all these questions after you finish? *Oh, you think I know the answers? That's where you're wrong! I ask questions I don't always know the answers to.*

#3 said what's a good pattern for a residual plot, and a student literally read the book definition and said in a horizontal band, close to zero, going across, and you drew this picture (PLOT 1) where you put 5 points up above and 5 points down below. Tell me what you were thinking and why you chose to draw that. Because they ignored the fact that it was supposed to be, there was supposed to be scattered some positive some

negative. *This has some positive some negative?* Yeah, but it has to be interspersed. And so you know they're not, while they're reading it, they're not really doing what I asked when I said what is it about and trying to figure it out and put it in their own words. So I'm gonna try to twist it so it's wrong just to force them to think about the other things. *Right.* 'Cause I'm mean. *And what's your thinking about why it needs to be interspersed. Positive and negative along the way.* Well, this I mean at least for interpolation, this is not gonna, you don't, in this region you don't have a whole lot of hope of this interpolated value being, from your model being any good *Cause you're…*Cause you're *all positive residuals* yeah, you're below the things and then all of a sudden you're above them. You're not gonna have anything there. I suppose that these, you just got lucky and picked them all here and they actually oscillate and come down in your model some place in between, but you can make all kinds of assumptions, but the one here if it's like this is your model is missing the boat. OK, it should somehow be weaving through the data points if it's really doing its job. It should be following the trend a little more closely if it's doing what it's supposed to be doing. *Rather than consistently under predicting then over predicting.* Yeah.

OK, here's a student question. You were talking about, you asked what was a bad residual plot so that you had the like parabola kind of example and you had the funnel kind of example. And you said that the size of the errors grows, that's not good. It's not a consistent horizontal band. And then you said why is that good, why do you want a consistent horizontal band. Because the model stays with the data, it's not consistently above or below. Kind of what we were just saying. But then a student asked you, yeah, but in this plot (PLOT 4) isn't the model in the middle. Isn't the model in the middle? *Yeah. What do you think of that question and what she's saying?* Yeah, I was trying to emphasize that. I can't remember what I said, but the point is that the errors are growing, not that the model's in between them. And so again the whole point is, whatever's going on, it's not mimicking what the data's doing otherwise the errors wouldn't be growing. *So is there a way to get a model that would mimic the data?* I don't know. I'm sure we could put a polynomial model through there…*Get your little twists and turns here and there.* I can make it go through every single point you want I think. Well, just about. I could put it pretty close. You know, when it's scattered like that, whatever's going on there, it almost seems like if the errors are big like that that things aren't, usually, at least with the data sets we do, we don't, unless its oscillating and we're just trying to put a line through there, the errors don't necessarily grow. They just like run off in one direction. For the ones we do. But you would have… *Something like this residual plot, with the funnel, what's that?* I don't know. What you would do to fix it? *Or I mean is it even possible to, I mean…*Well you might try some different things, like for here like this you might try something that oscillates a little. *Ummhmm.* If that would do it, I don't know. But we're not gonna get into that like fixing it completely. We're just gonna go yeah if it's running off, if the errors are running off one way or the other is more what we're gonna deal with. *Yeah, this is a question I have for myself too. Once again a question I don't have an answer to, but, ummm…*

You talked about this plot (PLOT 2) which you had as an example of a good residual plot, what you're looking for. You described it as noise. Tell me where you got that conception from or what you're, why were you thinking that. It's like there's, there's not a real like more like curved pattern to it or anything like that. It just reminded me of snow on a TV screen like (makes static noise). Like that. And if snow was only in the bottom half of the picture it wouldn't be good. So it was just, and the other thing is that umm one of the things, one of the ways that you can think about it is the your error in measurement. Maybe you made an observational error. You can think of that as something *noise* something was extra in your perception or something. There was something that was like messing with the way you were doing it or it was jiggling, I don't know. It just seemed like an appropriate thing to say at the moment. *OK, I'm not saying, I liked it.* Oh, OK, you know where it might have come from? I watch (the TV show) Numbers a lot, and they talk about numbers. They guy is a mathematician, so the guy is trying to filter data and one of the things is we've got to get the noise out of the data. You know, there's these errors in there and so in some way *you picked up the phrase there, ummhmm.* I might of picked up the phrase there. The other thing I might have picked it up from was in umm in fractals. Some of the random stuff they, when you look at data that was a lot of jaggedness in there. That it's hard to tell what's actually going on with the model and what's other variables in there that are just causing trouble. And they call that noise as well, I think. *Do they? OK, interesting.* So those are the places it might have popped out from. I have no idea; I just decided it was noise on the spot up there. *I like it.*

You said that even if you have a r^2 *value like .9 and get a residual plot like this, like a parabola, you need to go look for another model. What.* Because, it may be for the small data, for the bits of data you have, it may be sort of close, 'cause it's somehow following the trend. But it's really being biased in the different pieces of the domain, and so it's not gonna interpolate well. And if you think about it, if these errors are really growing at each end, you don't want to predict outside there either. So it kind of like messes up why you are looking for a model a lot of times. If you're thinking about making an extrapolation, a prediction outside, you don't necessarily want to go there if it's like that. If it's more threaded, then *you feel more comfortable* you feel more comfortable, you should feel more comfortable going outside for some distance, you know. You can't go outside for any distance you want; you ought to get nervous about that, but…

Anything else where you can think about what your knowledge of the subject and planning this, things that link to other mathematical topics that came up. Nope? I don't think there were too many student questions. I think I brought up the couple that I was interested in. I learned this when I was at Penn State and was taking statistics, a graduate statistics kind of thing. *Yeah, I didn't learn about residual plots until some kind of staff development thing I went to, yeah.* Cause they updated it (the textbook) between the 1st edition and the 2nd edition. I think they put the residual plots in the 2nd edition. But I came back and went oh yeah, I know what those are. *Luckily you just learned them huh?* But that was graduate school. What are you doing this to high school kids for? *Well they aren't complicated* No, they aren't. *You know, as a statistical tool, they're not beyond the scope of secondary kids.* No. Well, if they're thinking about it. Otherwise most things are beyond the scope of secondary.

<u>TEXTBOOK PAGES</u>: Section 2.8 Choosing a good model (pp. 134-140)
<u>HOMEWORK:</u> Section 2.8 Problems #1-13 (pp. 139-140)

APPENDIX D

SAMPLE ANALYSIS REPORT

CASEY'S ANALYSIS REPORT OF GBS DAY 11

LINES	KNOWLEDGE DESCRIPTOR
10-14, 161-174	Residual plots: purpose of residual plots to assess appropriateness of that function as a model (big picture of their usefulness); how to graphically analyze residual plot
32-44	Residual plots: scatterplot of the (x, residual) pairs
55-82	Residual plot: need for random scattering of positive and negative residuals throughout plot, otherwise it means the model is over predicting or under predicting for a portion of the domain
84-104	Residual plot: u-shaped plot bad because the model isn't following the trend of the data; will lead to poor interpolation and extrapolation
110-134	Residual plot: funnel shaped plot bad because the data points are systematically getting farther from the model, therefore prediction will be less accurate when x is going toward the direction of the funnel
154-158, 283-303	Model as search for signal in noisy processes (*cross reference* Konold, C. & Pollatsek, A. (2002). Data Analysis as the Search for Signals in Noisy Processes. *Journal for Research in Mathematics Education*, 33, 259-289*)*; Model represents the signal; expect random variation from the model-that is the noise; residual plot shows the noise plotted versus the independent variable; if model is appropriate, the noise/residuals should be randomly scattered because they are the random variation from the model in the process being observed; why we have noise and its sources
181-182	Residual plots: *Can one compare residual plots from different function models to one another to determine which is the best model?* From Rossman: Yes. If one model's residual plot reveals a random scatter, and another model's residual plot has a clear pattern, then the first model is more appropriate than the second.
256-281	Modeling: If the residual plot shows a funnel shape then one should try taking a transformation of the response variable, or perhaps both variables.
Textbook p.136	Modeling: theory-based models are those created based on the scientific theory relating the two variables; impressionistic models are those based on collected data; ideally the two types of models should agree
Textbook p.137	Extrapolation: dangerous especially when no theory to support the impressionistic model
Textbook p.138	Modeling: not everything has a good model (stock market example)

APPENDIX E

DIRECTIONS FOR ADDITIONAL ANALYSTS

Q & A: THE ANALYSIS PROCESS FOR DESCRIBING TEACHERS' SUBJECT MATTER KNOWLEDGE

What is a compilation document?

It is a written data file that documents the instruction for each day that I observed the teaching of statistical association. It usually includes the following items in this order: observation transcript, interview transcript, textbook pages, homework, student work. Not all compilation documents have all of these components. The lines are numbered so that they can be referenced in the analysis process; more is said about that in the analysis report questions below.

In the observation and interview transcripts, regular font is used for the teacher's words. Italic font is used for student words in the observation transcript and for the interviewer's words in the interview transcript.

Each compilation document is named and saved with a school initial reference and a day number. For example, NT Day 9 means it is from school NT and it is the 9^{th} day I observed at that school. The days are numbered consecutively based on the day I observed there, not necessarily consecutive class days for the teacher (I didn't observe days with irrelevant material or when students only took quizzes or tests).

What am I looking for in the compilation documents?

You are looking at the documentation of the teaching of statistical association (for a description of the content of statistical association, see the next question below). This includes not only the face-to-face interaction with the students, but also the designing of quizzes and the use of the textbook. All of these are included in the compilation documents.

I also interviewed each teacher following each observation regarding their thoughts as they planned for the day and as they carried out the lesson. The interviews were used to work towards understanding the subject matter knowledge used in teaching statistical association from the participating teacher's perspective rather than mine, the researcher's. The interviews may provide you important additional information regarding the teachers' thoughts that are not apparent from the observations.

It is impossible to observe knowledge directly as it is not visible. However, it is possible to observe directly the actions, verbalizations, and objects teachers and students use in the classroom. Observation of these things in the natural environment in which they occur provokes discussion of the subject matter knowledge used by the teacher or what might have been useful for the teacher to know in each situation.

What is statistical association?

It is the correspondence of variation of two statistical variables that are both numerical (typically displayed with scatterplots), both categorical (typically displayed with two-way tables), or one categorical and one interval variable (possibly displayed as side-by-side box plots); its study can include the following:

Numerical: data displays (e.g., table of values, scatterplots), determining if there is an association between the variables, modeling, correlation coefficient r, coefficient of determination r2, regression line or curve, linear transformation of variables, inference for regression, correlation vs. causation

Categorical: data displays (e.g., two-way tables, ribbon charts), determining if there is an association between the variables, chi-square test of independence

What am I to think about as I analyze?

I want to start by saying what you are not doing in the analysis: you are not documenting the subject matter knowledge of the participating teachers. Instead, the data sources contained in each compilation document should be used as a catalyst for developing conjectures regarding the subject matter knowledge a particular teaching incident might entail. Considering what the teacher is doing or called upon to do, you should consider what subject matter knowledge is needed by the teacher and where in the practice of teaching it is used.

Therefore, there are two things you are asked to think about as you analyze:1) the subject matter knowledge involved in teaching statistical association; 2) the teaching activities where this subject matter knowledge is utilized.

What is subject matter knowledge?

Referring to Gal's (2004) statistical literacy model, for my study subject matter knowledge includes mathematical knowledge, statistical knowledge, and knowledge of the context. This is a starting place for our work; if we feel that other components need to be included in this description (such as knowledge of technology) then we can change this.

It should be noted that in the theoretical framework for my study from Ball and Bass subject matter knowledge is considered separate from pedagogical content knowledge. Some of the participating teachers had difficulty separating the two when we were talking during the interviews. Do your best to try to separate and isolate subject matter knowledge from pedagogical content knowledge, as this study does not include an analysis of pedagogical content knowledge.

What teaching activities am I noting?

The occasions where the teacher needs to utilize his or her subject matter knowledge for teaching statistical association are the ones we are interested in. In their research, Ball and Bass have included descriptions of such activities, which have included the following:

- Understand and judge student created methods, explanations, claims, or solutions;
- Understand reasons behind methods used in mathematics;
- Use, choose, and judge representations;
- Model operations and concepts;
- Choose examples;
- Choose and develop usable, appropriate definitions;
- Judge and modify treatments of topics in instructional materials;
- Choose and develop usable, appropriate definitions for mathematical terms;
- Determine what counts as mathematical explanation;
- Create mathematical explanations that are accurate, comprehensible, and useful for students;

- Reconcile multiplicities of interpretations, definitions, representations, solution methods, or solutions;
- Explain learning goals and mathematical purposes to others;
- Design assessments;
- Assess students' mathematical learning and take appropriate next steps;
- Determine mathematical questions and problems that are productive for student learning.

This list gives you some examples of the types of activities we may find in our data sources.

<u>What time frame am I to think about in my analysis?</u>

I would ask that you consider two perspectives with respect to time. One perspective considers the knowledge used by the teacher to respond immediately to the current real-time demands of teaching. Instruction over time is also considered, forming the second perspective. This includes knowledge of the larger picture of the statistical topic and its associated practices, and how these connect with other topics. These other connecting topics may be referenced within the compilation documents but they certainly may come from your knowledge of the subject as well.

<u>Is knowledge of technology considered subject matter knowledge?</u>

Yes. I have spent some time reflecting on this question, both individually and with my advisor. I have decided that it is a necessary component due to our reliance upon and use of it in the teaching of statistical association. After all, who finds a least-squares regression line without technology? While knowledge of technology is a component of subject matter knowledge, it would not stand alone without knowledge of the other components.

Some of the teachers use the TI-83 graphing calculator and one teacher uses Minitab. I don't think the type of technology is as important as what it is the teacher needs to know about the technology to use it effectively in instruction.

<u>How far back and forward should my description of subject matter knowledge go?</u>

As a guideline, the description should go no farther than two grade levels/classes before or past the grade level/class described. For example, in thinking about the chi-square test of independence for a two-way table, the teacher needs to know how the chi-square test statistic is calculated. This calculation involves the addition of chi-square terms for each cell in the table, but it isn't necessary to describe that the teacher needs to know the commutative property of addition to understand that the terms can be added in any order.

<u>What level of detail should I use in describing subject matter knowledge?</u>

A description of what the concept is that the teacher needs to know as part of his or her subject matter knowledge is the level of detail that the analysis should contain. It is more than a listing of a topic; it is an explanation of the particular understanding of the topic that is needed by teachers.

For example, teachers probably need to understand slope to effectively teach least-squares regression lines. But what is it about slope that they need to understand? I would conjecture that it is more than knowing the change in y over the change in x formula; it is an understanding of slope as a ratio describing the rate of change of the line. They may also need to understand how changing the location of one point can greatly change the slope of the least-squares regression line.

Keep in mind that the goal is for the results of this study to be at a level of depth and detail to be useful for teachers, teacher educators, and researchers. However, if you feel that you are getting too bogged down in the process, back up the level of detail a little. I can always ask you for further clarification later if it is needed.

What is an analysis report?

A written report prepared by the analyst for each compilation document analyzed. It will consist of two parts: a description of the subject matter knowledge needed for each particular referenced teaching incident, and a list describing where in the practice of teaching subject matter knowledge was salient for this data source.

What is the format for an analysis report?

At the top of each analysis report, please record the name of the compilation document it refers to (for example, NCP Day 3 Analysis Report).

Next comes the **description of the subject matter knowledge with reference to the compilation document's line numbers**. For example, you could write "Lines 113-127: independent and dependent variables: definitions, differences between them, and knowledge of context to determine which statistical variable should be which in a pairing of variables". It is important to tie the conclusions to the data through the line numbers. You can do this in any format that makes sense to you. You can see my example report for one way to do it.

Finally, provide a **list of the teaching activities with reference to the compilation document's line numbers** where the teacher's subject matter knowledge was needed.

Are there any examples of analysis reports?

I have done an analysis report for the NCP Day 3 compilation document. This is included on the CD as an example for you to refer to as you wish, particularly for the level of detail and format of the report.

What if I have questions I want to ask my fellow analysts to respond to?

In starting the analysis process, I know that I have questions that I want to ask all of you to help me think through and respond to. I chose to include these in my analysis reports in italics (you can see that in my example report I provided on the CD). If you have questions you would like to ask the group of analysts to respond to, you can either do as I did and include them in your analysis reports or you can email them to me separately. I will then compile the questions and send them out to the group to respond to via e-mail.

How do I get the analysis reports back to Stephanie?

If the sizes of the files aren't too large, I would prefer to receive them back via e-mail to sacasey@ilstu.edu. If they are too large, you can send them to me on a CD that I can provide you; just let me know if this is needed.

What happens after Stephanie gets my analysis reports?

I will read through them and ask for any additional clarification I may need from you. Then I will compile the results, noting any discrepancies. Next I will share the results with all of you, opening up the findings for discussion. The results will not be linked back to any analyst by name; everyone's results will be compiled together confidentially. Modifications to the results may be made as a result of this process.

APPENDIX F

SAMPLE META-ANALYSIS REPORT

META-ANALYSIS REPORT OF GBS DAY 11

Burgess (statistics education expert)-normal font
Casey (researcher)-**bold font**
Clark (statistician)-*italicized font*

LINES	KNOWLEDGE DESCRIPTOR
27	Terminology: (various words for same concept) – error, residual, deviation.
27	*Residual = observed value – predicted value*
29-53, 209-228	Residuals & Residual plot: Calculation of residuals. How to graph these. Using the graphic calculator to plot residuals. Residual plot as a type of function (linking independent and dependent variables – which is which).
32-44	**Residual plots: scatterplot of the (x, residual) pairs**
29-42	*Residual plot: scatterplot of residuals vs. independent variable; residuals are the dependent variable and are plotted on the vertical axis. (Note: Residuals do not have to be plotted against the independent variable – they can be plotted against the fitted values for example.)*
45-53	*Calculator: How to use a calculator to make a residual plot*
55-174, 194-207, 234-254, 263-281, 305-315	Residual plot: Evaluating a residual plot to determine whether the model is a good one. Variation in positive/negative residuals suggests a better model than some regularity in residuals. Reliability of prediction (through interpolation and extrapolation) affected by regularity/patterns in residuals.
10-14, 161-174	**Residual plots: purpose of residual plots to assess appropriateness of that function as a model (big picture of their usefulness); how to graphically analyze residual plot**
55-82	**Residual plot: need for random scattering of positive and negative residuals throughout plot, otherwise it means the model is over predicting or under predicting for a portion of the domain**
55-82,122-124, 131-132	*Residual plot: Patterns in residual plots – a good linear model has no discernable pattern in the residual plot, and the residuals are 'small'.*
84-104	**Residual plot: u-shaped plot bad because the model isn't following the trend of the data; will lead to poor interpolation and extrapolation**
95-104, 127-131, 182-183	*Residual plots & extrapolation/interpolation*
110-134	**Residual plot: funnel shaped plot bad because the data points are systematically getting farther from the model, therefore prediction will be less accurate when x is going toward the direction of the funnel**

59-65, 84-89, 92, 95-104, 106-108,110-111, 116-120, 127-131	*Residual plot: Recognizing patterns in residual plots; patterns indicate the model is a poor fit*
116-120, 129-131	*Residual plot: Errors growing as independent variable increases in residual plots*
178-182	*Modeling: assessing a model: r / r^2 alone do not determine whether the model is a good fit; need to see the scatterplot and residual plots as well.*
218-228	Graphing: Must use observed x values for independent variable so that deviations between observed y and predicted y can be calculated.
283-303	Variability: Variation conceptualized as noise in data (noise around a 'signal', or trend) – (reference Konold, C., & Pollatsek, A. (2004). Conceptualizing an average as a stable feature of a noisy process. In D. Ben-Zvi & J. B. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning, and thinking* (pp. 169-199). Dordrecht, The Netherlands: Kluwer.
154-158, 283-303	**Modeling: Model as search for signal in noisy processes (*cross reference* Konold, C. & Pollatsek, A. (2002). Data Analysis as the Search for Signals in Noisy Processes. *Journal for Research in Mathematics Education*, 33, 259-289*)*; Model represents the signal; expect random variation from the model-that is the noise; residual plot shows the noise plotted versus the independent variable; if model is appropriate, the noise/residuals should be randomly scattered because they are the random variation from the model in the process being observed; why we have noise and its sources**
181-182	**Residual plots: Can one compare residual plots from different function models to one another to determine which is the best model?** From Rossman: Yes. If one model's residual plot reveals a random scatter, and another model's residual plot has a clear pattern, then the first model is more appropriate than the second.
256-281	**Modeling: If the residual plot shows a funnel shape then one should try taking a transformation of the response variable, or perhaps both variables.**
Textbook p. 136	**Modeling: theory-based models are those created based on the scientific theory relating the two variables; impressionistic models are those based on collected data; ideally the two types of models should agree**
Textbook p. 137	**Extrapolation: dangerous especially when no theory to support the impressionistic model**
Textbook p. 138	**Modeling: not everything has a good model (stock market example)**

APPENDIX G

PREREQUISITE SUBJECT MATTER KNOWLEDGE

DESCRIPTION AND SUMMARY DOCUMENT

OUTLINE

Prerequisite Subject Matter Knowledge

Algebraic Reasoning and Concepts

The results of the analysis process repeatedly indicated that teachers must have a strong knowledge base in some specific algebraic concepts and reasoning skills in order to have the knowledge needed for teaching statistical association. Brief descriptions of the knowledge pieces identified in this area are presented here; a further description can be found in the following prerequisite subject matter knowledge summary document outline.

Proportional Reasoning

Teachers need to have excellent proportional reasoning skills and knowledge and know when to apply them in practice. These skills were called for in the teaching of a multitude of topics, from slope to analyzing two-way tables. Understanding part-to-whole relationships, understanding the difference between raw numbers and ratios, and finding equivalent ratios are some of the necessary components of knowledge involved in proportional reasoning.

Lines

A best fit line is a major topic of statistical association taught at the secondary level. As a result, teachers must have a comprehensive understanding of lines, including their properties, representations, and relationships between multiple lines. The slope of a line was a particularly common and important topic in the observed class sessions and thus a point of emphasis in this knowledge piece.

Function families

As regression models move outside the linear realm to include non-linear models that may be exponential, logarithmic, or quadratic, teachers are called on to use their knowledge of function families. Teachers need to be able to look at a data representation, such as a table of values or scatterplot, and determine what type of function to try as a potential model. Thus they need a firm foundation in the different types of functions and their behavior from numerical, graphical, situation, and algebraic perspectives.

Statistics and Probability

As with algebraic reasoning and concepts, topics in statistics and probability outside that of statistical association were repeatedly identified as necessary knowledge pieces for teachers who teach statistical association. A short narrative of the knowledge identified in this area is described below. A more complete description is found in the following outline.

Graphs

Histograms, boxplots, and side-by-side boxplots were deemed to be important graphs for teachers of statistical association to know. This knowledge must include how to create the graph, interpret the graph in context, and determine which graph is most appropriate to represent a data set(s).

Mean

The mean is a vital statistic to the study of statistical association. It is seen in many different lights, from a way to calculate the expected value to the center of gravity (x-bar, y-bar) as a point that the least-squares regression line (LSRL) always goes through. Teachers need to know multiple ways of conceptualizing the mean, how it is calculated, and its properties.

Normal Distribution and Standardized z-scores

Teachers need to know the properties of a normal distribution, how to determine if a data set's distribution is approximately normal, and how to use the area under the standard normal curve to calculate the probability of an event. This calculation may be done using standardized z-scores, which is another important topic that also relates to the calculation of the correlation coefficient. Standard deviation is another prerequisite knowledge topic that teachers must know in order to understand topics like standardized z-scores or the correlation coefficient.

Probability

Probability is a prerequisite knowledge area for statistical inference and analysis of categorical variables for association. Of particular importance is its definition as the long-run relative frequency of an event. Other concepts essential to the understanding of probability needed for teaching statistical association include conditional probability, independent events, and probability distributions.

Hypothesis test

A t-test for the slope of the regression line is used to determine if there is a significant association between two quantitative variables. The chi-square test of independence is the inferential procedure used to determine if there is a significant association between categorical variables. Thus it is necessary for teachers who teach these topics at the secondary level to understand the general logic and necessary steps for carrying out a hypothesis test.

Confidence Interval

The purpose, general formula, and interpretation of a confidence interval were identified as prerequisite knowledge items a teacher needs to know in order to teach the confidence interval for the slope of a regression line. Specifically a teacher needs to be able to apply these concepts to the context of a problem involving the study of the rate of change between two quantitative variables.

Summary Document Outline

Algebraic Reasoning & Concepts

I. Proportional Reasoning
   - A. Part-to-whole relationships: find a given proportion of the total
   - B. Equivalent ratios
     1. Multiple ways to calculate equivalent ratios, including solving a proportional equation, using rates, and multiplying by a version of the number one (e.g., multiply by 300/300)
     2. Know when proportional reasoning applies to a situation (i.e., constant ratio)
   - C. Ratio vs. raw numbers
     1. Ratio is a better way to compare the numbers in multiple groups when the total possible isn't the same in all of the groups. Changing to a percent is a commonly used method

because it makes all ratios out of 100 so that they can be easily compared.

2. Difference between the meaning of raw numbers and ratios
3. Number sense regarding ratios vs. raw numbers (e.g., if the total is small, one unit change in the numerator of the ratio can result in a large change in the ratio/percentage)

D. Two-way table: The following calculations are equivalent:

1. (row total) x (column total) / (grand total)
2. ((row total) / (grand total)) x (column total)
3. ((column total) / (grand total)) x (row total)
4. ((row total)/(grand total)) x ((column total) /(grand total))x (grand total)

II. Slope of lines

A. Calculate

1. From equation, graph, situation, or in-out table
2. Since a line has a constant rate of change, it doesn't matter which two points on the line are chosen to calculate the slope. If the points chosen have an x-coordinate difference of one, the difference in the y-coordinates will be the slope of the line.
3. Positive when line goes up from left to right; negative when line goes down from left to right; zero when horizontal; undefined when vertical; understand from calculation and from meaning of slope

B. Meaning

1. Each additional increase of 1 unit in the x-variable predicts an increase/decrease of slope units in the y-variable
2. Interpret meaning of slope in context of the problem
3. Rate of change of the line

C. Equivalent ratio: calculate equivalent ratio to make slope easier to interpret or the numbers more meaningful in the context of the problem

D. Lines: Constant rate of change (i.e., the same slope for the entire graph or in-out table) implies linear relationship between the variables

E. Least-Squares Regression Line (LSRL): Its slope represents the average rate of change for the entire data set. It is not necessarily the same as the slope from any given point to another in the data set, nor is it the mean of the slopes between the points.

III. Lines

A. Slope: constant for lines

B. Representations
    1. Equation, graph, situation, and in-out table: knowledge of each representation and the ability to move from one representation to another
    2. Equation
        a. Why the equation $y = mx+b$ works to graph a line and produce the points on it
        b. Multiple ways to write the equation of a line (e.g., slope-intercept equation & point-slope equation)
    3. Graph
        a. Picture of the infinite number of solutions to that equation
        b. As the absolute value of the slope increases, the graph becomes steeper
    4. In-out table: list of ordered pairs on the line

C. Intercepts
    1. X-intercept
        a. Value of the x-variable when the y-variable is zero
        b. Where the graph crosses the x-axis
        c. Meaning in the context of the problem
    2. Y-intercept
        a. Value of the y-variable when the x-variable is zero
        b. Where the graph crosses the y-axis
        c. Represents the starting point (0, b)
        d. Use the context of the problem to determine

D. Find a point on the line
    1. Use any of the representations (equation, graph, in-out table) to find a point on the line, including one not known originally
    2. Proportional equation involving slope can be used for prediction only if y-intercept is zero (can use equation then add/subtract y-intercept to predict in all cases)

E. Intersection of lines
    1. Find from any of the representations (equations, graphs, in-out tables)
    2. Meaning of intersection point in context

F. Parallel lines
    1. Since they never intersect, they must have the same slope so that they have the same steepness.

2. Given a line, a parallel line can be created by adding or subtracting a fixed amount from the y-coordinate of every point on the line; graphically this is a vertical translation.

IV. Algebra
   A. Distance equals rate times time (and its equivalent equations)
   B. Number sense
      1. Rounding of numbers
         a. Effects on subsequent calculations
         b. When it is justified
         c. Appropriate precision of a calculation given the context and magnitude of the numbers
      2. Order of operations
      3. $(a-b)^2 = (b-a)^2$
   C. Function
      1. Definition and application of definition to determine if a graph, equation, situation, or in-out table represents a function
      2. Knowledge of different function families
         a. Including linear, quadratic, exponential, power
         b. Behavior from numerical, graphical, and algebraic perspectives, with specific attention to the differences in their rates of change and the resulting differences in their graphs
         c. Linear, quadratic, and exponential functions can look very similar for a portion of the domain because parts of a quadratic and exponential function can look nearly linear. The differences in their graphs will be more profound for very large or small values of x.
   D. Exponents & logarithms
      1. Relationship as inverse functions
      2. General form (including each parameter's meaning in context), restrictions on parameters, domain, range, and graph of each; includes understanding the reasons for each of these (e.g., in the exponential general form $y = a * b^x$, if $0<b<1$ this is exponential decay because multiplying by a number between 0 and 1 makes a quantity smaller)
      3. Large values of x cause overflow errors in the calculator for exponential models because the base value is being raised to a large value of x creating an enormous value that the calculator can't handle. To prevent overflow errors, the x-

values can be scaled down by subtracting a constant value from each of them. This will not change the nature of the behavior between x and y because it is a horizontal translation by the amount subtracted.

4. Solve a exponential or logarithmic equation
5. Exponential functions have a constant multiplicative rate of change and how to find this common ratio
6. Properties of exponents and logarithms (e.g., $e^{mx+b} = (e^{mx})(e^{b})$)
7. Convert between exponential form and logarithmic form
8. Logarithm base e of x is the natural log, written as ln(x)
9. If *y* vs. *x* is exponential then *log(y)* vs. *x* is linear because in an exponential function the exponent changes linearly
10. If y vs. x has a power relationship, then log (y) vs. log (x) will be linear. Understanding the mathematics behind this involves knowledge of logarithms, converting to exponential form, and exponent rules.

E. Quadratics

1. Understand meaning of parameters a, b, and c in the general form of a quadratic equation $y = ax^2+bx+c$ and relate them to real-world applications (e.g., c would be the starting height of a projectile)
2. For projectile motion the quadratic's general equation is $h(t) = - \frac{1}{2} gt^2 + v_0t + h_0$. The acceleration of gravity,g, is 9.8 meters per second squared if measure in meters and 32 feet per second squared if measure in feet. A limitation of this model is it ignores friction due to air.
3. Numerical (using a table), graphical (graph and find x-intercepts) and algebraic (quadratic formula or factoring) ways to solve a quadratic equation

F. Linear transformations

1. Understand the idea and procedure of linear transformation, including taking logarithms or square roots of the y-variable to slow down the growth of the y-variable so that it has a linear relationship with the x-variable
2. Algebraic manipulation to convert transformed data back to the original data by solving for the variable (e.g., convert from logarithmic form to exponential form)

Statistics & Probability

V. Graphs (See also scatterplot II.L. and segmented bar graph III.D. in Appendix H)

A. Determine appropriate graph to represent the data set(s)

B. Histogram

1. Bar graph for univariate quantitative data
2. The x-axis has the variable of interest; the y-axis has the frequency or relative frequency of the x-variable for a given range of values
3. How to construct and how the graph changes based on the range of values determined for each bar
4. How to interpret the distribution displayed, including its center, spread, and shape

C. Boxplot

1. Graph for univariate quantitative data which displays the five-number summary (minimum, 1st quartile, median, 3rd quartile, and maximum)
2. Modified boxplot shows the outliers as well (if any exist)
3. How to construct and interpret

D. Side-by-side boxplots

1. Graph for visualizing bivariate data when one variable is categorical and the other is quantitative; one boxplot is made for each category and all of the boxplots are graphed next to one another on the same scale
2. Categorical variables should be listed in order if the variable is ordinal (i.e., has an order to it)
3. How to construct and interpret, including a focus upon changes in the centers of the distributions as well as the variation of each distribution (leading towards ANOVA)

VI. Mean

A. Conceptualizations

1. Expected value
2. Balancing point
3. Leveling or equal share

B. Calculation

1. Mean
2. Weighted mean

C. Properties

1. Sum of deviations from the mean is zero
2. Hides variation between individual data points
3. Non-resistant (i.e., value largely affected by outliers)

4. Appropriate measure of center when the data is symmetric and without outliers
5. Central Limit Theorem

VII. Normal distribution & standardized z-score

A. Normal distribution
1. Shape
2. Properties (e.g., symmetric, peak in the middle, 68/95/99.7 rule)
3. Difference between normal and approximately normal
4. Using area under the standard normal curve to calculate probability

B. Standard deviation
1. Statistic that measures the differences of the individual data points from the mean of the data set in order to describe the spread of the data set from its center (as measured by the mean)
2. Calculation and why it is calculated that way, including squaring the deviations and dividing by (n-1) as opposed to n when calculating the sample standard deviation
3. Properties, including that it is appropriate only for symmetric data because it is non-resistant

C. Standardized z-score
1. Calculated for an individual data point, it represents the number of standard deviations a point is from the mean
2. Calculation: z = (data point – mean)/standard deviation
3. Properties
   a. Positive if data point greater than the mean, negative if data point less than the mean
   b. If the data point's value is the same as the mean, its z-score is zero
   c. No units, so it makes possible comparisons of value on different variables, with different scales, with different units, or for different populations

VIII. Probability

A. Definitions
1. Long-run relative frequency
2. Count of outcomes classified as the desired event divided by the total number of outcomes (as long as the outcomes are equally probable)

B. Representations: Fraction, decimal, and percent; ability to move from one representation to another

C. Experimental vs. theoretical
   1. Experimental: observed proportion of times the event occurred in a finite number of trials
   2. Theoretical: limit of the experimental probability
   3. Law of large numbers: long-run relative frequency of repeated independent events settles down to the theoretical probability as the number of trials increases

D. Conditional probability: meaning and calculation

E. Independence
   1. Meaning of independent events
   2. How to determine if events are independent from context and through calculations of probabilities (e.g., $P(A|B) = P(A)$)
   3. Difference between independent events and mutually exclusive events

F. Complement: To find the total in category A, you can subtract the number not in A from the overall total

G. Probability distribution (see also Chi-square distribution III.E.2)
   1. Shape, properties, and how to calculate probabilities for each of the following distributions:
      a. Normal
      b. t-distribution
      c. Chi-square distribution
   2. Calculate the empirical probability of an event from a frequency bar graph

H. Sample size & probability
   1. Probability is a long-run proportion and concerns behavior in the long run (large sample size) not short run (small sample size)
   2. Deviations from the expected are more likely with small sample sizes and become less likely as sample size increases, i.e., small differences can be statistically significant with large samples but tend not to be with small ones.

IX. Hypothesis test

A. Logic: general logic and required steps of a hypothesis test

B. Hypotheses
   1. Test is based on assuming the null hypothesis is true. The ways in which the test uses this assumption shows the necessity of making the null hypothesis a statement of equality or independence.

2. How to determine whether a one-tailed or two-tailed test is appropriate (ask why the test is done and what is the hypothesis that motivated the study)
3. Written in terms of population parameters since those are unknown and what you are trying to infer about

C. P-value
   1. Meaning
   2. Calculation from tables, calculator, or computer
   3. Never zero because there is always some area under the probability distribution as it goes to positive or negative infinity; this means that one is never 100% certain in the decision
   4. Link with the null hypothesis which determines the probability distribution used to calculate the p-value

D. Decision
   1. Alpha level
      a. Represents a pre-set cutoff value at which one will reject the null hypothesis if the p-value is less than it
      b. .05 is industry standard; usually never higher than .1
      c. Probability of Type I error (rejecting a true null hypothesis)
   2. Critical value or p-value approach to making a decision
   3. Decision to reject null hypothesis or fail to reject null hypothesis
      a. Reasons these are the two possible decisions
      b. Decide to reject the null hypothesis if the p-value is less than the alpha level
      c. Decide to fail to reject the null hypothesis if the p-value is greater than the alpha level

E. Conclusion
   1. If null hypothesis is rejected, then the alternative hypothesis is accepted.
   2. If the null hypothesis is not rejected, then there isn't enough evidence to convince one that the null hypothesis is not true.
   3. Link with term "statistically significant"
   4. Conclusion in context of the problem

F. Errors
   1. Type I and type II errors and their consequences
   2. How to minimize the chance of either error
   3. Chance of an error compounds when doing multiple tests

X. Confidence Intervals

A. Purpose: Estimate population parameters
B. General formula: statistic ± (critical value)(standard deviation of statistic)
C. Interpretation: Conceptual (including meaning of confidence) and for specific confidence intervals in context

XI. Ethics & Statistics: what is considered ethical and unethical behavior when conducting statistical studies

APPENDIX H

SUBJECT MATTER KNOWLEDGE SUMMARY

DOCUMENT OUTLINE

I. Statistical Association

A. Meaning of association

1. Correspondence of variation of two statistical variables
2. Not necessarily causation
    a. Possible other reasons for association include a common response relationship or a confounding relationship between the variables: differences between each type of relationship and examples in context
    b. Why and how an experiment can establish causation and observational studies cannot
    c. Lurking variable: meaning and context knowledge to determine possible lurking variables
2. Describes a general trend, i.e., does not have to hold true for every piece of data
3. Independence
    a. If two variables are not associated they are called independent
    b. However if two variables are associated that does not mean they are dependent. Dependency suggests a model or pattern but variables can fail to be independent in many different ways.

B. Terminology: fluent in the language of statistics, including a thorough understanding of the following terms:

1. Population parameter
2. Sample statistic
3. Random
    a. Random sample
    b. Random scatter of points (e.g., quality desired in residual plot)
4. Mathematical model
5. Regression line
6. Predictor function
7. Outlier
8. Influential point
9. Error, deviation, and residual (synonyms)
10. Variable, including random variable
11. Explanatory and response variables
12. Discrete and continuous variables

13. Resistant statistic
14. Association & correlation
    a. Correlation is a specific term referring to the statistic that measures the strength and direction of the linear association between two quantitative variables
    b. Association is a property referring to a relationship between two variables, either quantitative or categorical
15. Independence of two variables
16. Mean
17. Standard deviation
18. Standard error

C. Context
1. Select data set: link statistical and context knowledge to select an appropriate data set to analyze
2. Outlier: context knowledge to explain the existence of an outlier
3. Number sense: context knowledge to evaluate the appropriateness of a numerical answer
4. Slope & lines
    a. Context knowledge to determine if it is reasonable to assume a constant rate of change between two variables, both within the original domain and beyond
    b. Interpret equation of line in context
    c. Use the context to determine the y-intercept of the best fit line
5. Association
    a. Context knowledge to produce two variables that are associated (possibly with a specific direction and/or strength) and understand the reason for their association
    b. Context knowledge to determine a lurking variable
    c. Context knowledge to predict the direction and strength of association between two quantitative variables
    d. Determine which variable is explanatory and which variable is response based on the context of the problem
    e. Relate meaning of data in context to type of association

f. Judge the accuracy of a mathematical model using the context of the variables
g. Consider the context and theoretical relationship between two variables to inform the choice of a mathematical function to model the relationship between the variables
h. Judge the reasonableness of predictions made using mathematical models from the context
i. Use context knowledge to determine whether a statistical significance is practically significant

II. Association between quantitative variables
- A. Data
  1. Understand why data is useful and needed (e.g., to make predictions)
  2. Variability
     - a. Natural and to be expected in data sets
     - b. Ways to look for the signal in noisy data
     - c. Sampling variability
  3. Differentiating types of data
     - a. Difference between quantitative and categorical data and how to classify data as one of the types
     - b. For quantitative data, can be further classified as interval or ratio
     - c. Meaning of bivariate data, including the characteristic that the data is collected in pairs from one source
- B. Association
  1. Analyzing
     - a. Numerically
       - i. Helpful to place explanatory variable values in numerical order and the corresponding response variable values across from them
       - ii. Look down lists together and see if the response variable values are changing in a predictable manner (e.g., always increasing/decreasing, estimate slopes, visualize scatterplot)
       - iii. See whether above-average values of the x-variable tend to accompany above-average values of the y-variable (positive) or below-average values of the y-variable (negative)

and the same for below-average values of the x-variable

b. Graphically
    i. Visually analyze a scatterplot to determine if the variables are associated
    ii. If they are associated, determine what mathematical function(s) would be possible models
c. Inference: t-test on the slope of the regression line is used to determine if there is a statistically significant linear association between two quantitative variables (see also II.M.)
d. Sample size
    i. Larger sample size provides a better chance to detect an association if there really is one
    ii. In t-test for the slope of the regression line, $t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$, so as the sample size n increases so does t, even if r stays fixed; also t-distribution based on (n-2) degrees of freedom

C. Correlation coefficient r
1. Variables
    a. Only applicable to bivariate quantitative data
    b. Correlation coefficient of y on x is the same as the correlation coefficient of x on y; can see why through formula for r and through understanding that association is not a concept that involves order for the two variables involved (if x and y are associated that is the same as y and x are associated)
2. Estimate from graph
    a. Estimate correlation coefficient given a scatterplot
    b. Human error in this process (e.g., human eyes can be fooled to see a stronger relationship between two variables in a scatterplot if they are plotted with bigger scales and with more white space around the points)
3. Rule of thumb ± (1-1/k) to estimate r:

a. If you draw a rectangular box to enclose the data, the greater the number of widths in the length (k) the skinnier the box is which means it's closer to a line
b. 1 is a perfect linear association, so 1/k takes away from 1 an amount relative to how far away the points are from a line
c. Longer boxes are going to have bigger r values because the z-scores are bigger (farther away from the mean)
d. Proportional reasoning: the bigger k gets, the smaller 1/k gets, therefore you are subtracting less from one and will have a value closer to 1 as your estimate

4. Calculation
    a. Formula(s) and their meanings (i.e., why and how the calculation of r measures the direction and strength of a linear association between the variables)
        i. For formula involving product of z-scores, $r = \frac{\sum z_x z_y}{n-1}$,understand its meaning through the analysis of four quadrants created by graphing the x-bar and y-bar lines
        ii. Understand why z-scores are unaffected by linear transformations like subtracting a constant, and therefore r is also unaffected by such a transformation
        iii. Why the sum of the products of the z-scores is divided by (n-1) rather than n
        iv. Since z-scores are without units, so is r
        v. Since r uses the standardized z-scores in its calculation, r does not change when the units of measurement of x, y, or both are changed
        vi. Since z-scores are linear functions of the data values, r is only appropriate for measuring the strength of association for linear relationships
    b. Sensitive to outliers through its use of mean and standard deviation in its calculation
    c. The closer the data points are to a line the bigger r gets, not necessarily the points just getting closer to one another

d. Value of r for averaged data is usually too high when applied to individuals because averaging takes away much of the variation in the original data

5. Interpretation
    a. Measures the strength and direction of the linear association between two variables
    b. Measure of how much help x gives you in predicting y, assuming that you use a straight line to do the predicting
    c. Range of values
        i. Between -1 and 1
        ii. +1 if all points are on a line with a positive slope
        iii. -1 if all points are on a line with a negative slope
        iv. Zero if no linear association between the points (with exception of horizontal line-see II.D.6.b), i.e., the values of one variable provide no help in predicting the other variable
        v. If r is found to be statistically significantly different from zero, then the two variables are considered to have a significant association
        vi. Association considered stronger the closer the absolute value of r gets to 1
        vii. Determine reasonable ranges for r to label as weak, moderate, and strong associations based on the context
    d. Non-linear associations
        i. In scatterplots where the points follow a non-linear model closely, r is not necessarily near zero (e.g., data following an exponential trend usually have a strong r value)
        ii. Data following a quadratic trend have a value of r near zero (see also II.C.6.c.)
    e. Sign
        i. Positive r means that the variables have a positive association (they both increase together)

ii. Negative r means that the variables have a negative association (as one variable increases, the other decreases)

f. Can be misleading
   i. Non-linear variables can have a strong correlation
   ii. Unusual observations can also have a big impact on the correlation
   iii. Important to look at the graph of the data along with the value of r

6. When equals zero
   a. Random scatter of points with no association
      i. There is a range of values of y for a given x-value throughout the data set
      ii. If the x-bar and y-bar lines are graphed to create four quadrants, there will be approximately the same number of points in all four quadrants with the same distances from the lines. Therefore the positive z-score products will be approximately the same value as the negative z-score products, resulting in a sum of the z-score products near zero and therefore an r close to zero.
   b. Horizontal line
      i. If the various values of the x-variable all produce the same value of the y-variable, then you're not learning anything helpful (in terms of predicting the value of the y-variable) by learning the value of the x-variable. Conceptualizing correlation as a measure of how much help x gives you in predicting y (assuming that you use a straight line to do the predicting) then it makes sense that in the case of a horizontal line, r equals zero.
      ii. The slope of a horizontal line is zero and r varies directly with the slope of the LSRL, so if the LSRL has a slope of zero then r also equals zero.
      iii. r^2 can be defined as the percentage reduction in the sum of the squared residuals for the

LSRL compare to the model y = y-bar.
The LSRL and this model are identical if the line is horizontal so there is no reduction and $r^2 = 0$. This implies that r equals zero as well.

iv. In the formula for r, the calculation of the z-scores for the y-variables involves division by the standard deviation of y. The standard deviation of y is zero because all of the y values are equal. However, because of the above reasons, statisticians have defined the correlation coefficient of horizontal lines to be zero.

c. Points follow a quadratic model
    i. The positive and negative products of the z-scores balance each other out in the four quadrants created by graphing the x-bar and y-bar lines and the middle point has an x z-score of zero
    ii. The best-fit line would be approximately horizontal with a slope near zero
d. Points follow a non-linear model

7. Adding and removing points and understanding the effect on r (see also II.G. Outliers & influential points)
    a. In particular, understand the effect of a point along the pattern with a larger x-value, a point away from the pattern with a larger x-value, a point away from the pattern with a smaller x-value, a point in the middle of the domain with a larger y-value
    b. For example, adding a point near the line with a considerably larger or smaller x value than the rest of the points will result in a stronger r value because the product of the z scores will be very large because the point is far away from the means. Thinking of the rule of thumb, it will result in a larger k which means a smaller 1/k to subtract from 1.
8. Relationship with coefficient of determination r^2
    a. $r = \pm\sqrt{r^2}$
    b. Determine whether r is positive or negative based on whether the data have a positive (as one variable increases the other variable also increases) or negative

(as one variable increases the other variable decreases) association as seen in the graph or gleamed from the context

9. Relationship with slope of LSRL (see also II.F.5.b.)
   a. $m = r\frac{s_y}{s_x}$
   b. Results of this relationship (e.g., the slope of the LSRL and r have the same sign)

D. Coefficient of determination r^2
   1. Calculation
      a. r^2 = (SSM-SSE)/SSM where SSM = sum of squares about the mean y-bar and SSE = sum of squares for error about the LSRL
      b. Geometric representation of the above formula
      c. Alternatively can calculate by taking r, the correlation coefficient, and squaring it (algebra can prove these are equivalent)
   2. Numerical interpretation
      a. Percentage reduction in sum of squared residuals comparing the LSRL model to the model y = y-bar
      b. Proportion of variability in the y-variable explained by the LSRL of y on x
      c. Remaining variation in y-variable needs to be explained by other factors
      d. Use to judge model's accuracy
   3. Context interpretation: interpret in context of the problem
   4. Relationship with coefficient of multiple determination R^2
      a. $R^2 = \frac{SSresid(y-bar) - SSresid(\text{model})}{SSresid(y-bar)}$, where the model can be of any type, including non-linear and multiple regression
      b. R^2 and r^2 are equivalent for the linear case

E. Mathematical modeling & regression
   1. Purpose
      a. Variability is natural and to be expected in real data
      b. Variation from a theoretical model can be created by many things, including measurement error and biological variation
      c. Variation from a theoretical model creates scatter or noise

   d. Modeling is looking for the signal in the noise
   e. Useful for making predictions and understanding relationships between variables
2. Mathematical modeling
   a. Process used to find mathematical equation(s) to mimic the behavior of bivariate quantitative data
   b. Theoretical vs. impressionistic
      i. Theory-based models come from the scientific theory relating the two variables
      ii. Impressionistic models are based on collected data
      iii. Differences in their use, purposes, and how to find each (if theoretical is possible)
      iv. Impressionistic may give suggestions for the theoretical
   c. Regression is a tool used in mathematical modeling that takes into account the possibility of chance variation
   d. Not everything has a good model (e.g., stock market)
   e. Recognize insufficiency of a small number of data points for modeling
   f. Appropriate to separate the data into different domains and find different models for each domain if the original scatterplot shows that the data follows different trends for different portions of the domain
3. Best fit models
   a. More than one way to find a model for each function type (e.g., in the linear case there is the LSRL and the median-median line)
   b. For the linear case, knowledge of the multiple models, the differences in their criteria and the advantages/disadvantages of each (see also II.F.)
   c. Different criteria can result in different best fit models; e.g., in the linear case, the line which minimizes the sum of the absolute errors is not always going to produce the same line as the LSRL which minimizes the sum of the squared residuals
   d. Modeling techniques are designed to minimize the errors in the predictions of the y-variable, therefore the decision of which variable is x and which variable is y

is important and would result in a different model if they were switched

4. Notation: y-hat for predicted y-value

F. Line of best fit, including LSRL
   1. Idealized regression line for the population
      a. Using all the values in the population, an idealized regression line could be found which describes the linear relationship between x and y
      b. Places the means of the distributions of the y-variable for each x-variable value on the same line, i.e., it is the conditional expected value function of Y given X=x as a function of x; in this way the line can be thought of as an average
   2. Based on the sample
      a. Rare to have the data for the population, thus the line of best fit is usually based on sample data
      b. Assumed that the errors around the idealized regression line are normally distributed with a center at the line for each value of x
      c. Can find line of best fit even when sample data set is not a function
   3. Y-intercept: can make the choice to have the best fit line go through a given point (usually the origin) based on the context of the problem or prior evidence that it should
   4. Idea of best fit line
      a. Goal to describe data which has a linear association with a line which follows the trend of the data
      b. Line usually tries to fit all of the data, unless the removal of an outlier is justified by the context of the problem
      c. May not go through any of the data points
      d. Criterion for best fit should determine a unique line
      e. Visualize location of a best fit line given a scatterplot
      f. Different best fit lines (e.g., LSRL, median-median), including knowledge of their different criteria regarding how to reach the goal, how they are determined, their advantages and disadvantages, comparisons in their location given a scatterplot
      g. In the statistics community it is agreed upon that the criterion of smallest sum of squared residuals, that used by the LSRL, determines the line of best fit

5. LSRL
   a. Criterion
      i. Minimizes the sum of the squared residuals: The residual is a measure of the error of the line in predicting the y-variable. So to make the best predictions (a common use of a mathematical model) these errors should be minimized. The residuals can not be summed because that sum is zero since the total of the positive residuals is the opposite of the total of the negative residuals. Therefore the residuals are squared first then summed to create a measure for the error of the line.
      ii. Geometric representation involving drawing squares to the line from the points, each having a side length equal to the absolute value of the residual
      iii. Consequences of this criteria (e.g., lacks resistance to outlying points)
      iv. Determines a unique line
      v. LSRL of y on x is different from the LSRL of x on y (nor are they inverses of one another) due to the fact that you are minimizing the errors in a different direction
   b. Slope
      i. Slope of LSRL $b = r (s_y/s_x)$
      ii. Ratio of the standard deviations is similar to change in y over change in x formula for slope of a line
      iii. Since standard deviation is never negative, r and b will always have the same sign and vary directly
      iv. Interpret as predicted or average change in y (response variable) for each unit change in x (explanatory variable)
      v. LSRL model assumes the slope is constant throughout the entire data set
   c. Y-intercept
      i. Y-intercept of LSRL a = y-bar – b*x-bar

ii. Does not necessarily make sense in the context of the problem

d. Properties

i. Center of gravity (x-bar, y-bar) is always on the LSRL: can understand from a calculation perspective using the formula for the y-intercept of the LSRL or from a conceptual perspective of the LSRL as an average

ii. Sum of residuals is zero: property of any line going through the center of gravity

iii. Mean of residuals is zero: property of any line going through the center of gravity and a result of the sum of the residuals equaling zero

e. Visualize location of LSRL given a scatterplot

G. Outliers & influential points

1. Outlier

a. Univariate setting

i. A point well separated (often defined as more than 1.5IQRs below Q_1 or above Q_3) from the majority of the data set

ii. Concerned about in parametric tests because of their influence on the mean and standard deviation which parametric tests use as their most important statistics

b. Bivariate/regression meaning

i. A point with a large error, which means it does not follow the relationship pattern shown in the other points

ii. If there are multiple points with x-values near the outlier, then it will not exert as much influence on the location of the LSRL or r.

iii. Use of context to determine if outlier should be removed from data set; common to remove an influential outlier if have justification from the context

2. Influential point

a. One that greatly influences the location of the LSRL (i.e., when removed the LSRL moves notably)

   b. Influential point often has a x-value that differs greatly from others in the data set and because of LSRL criterion the line needs to minimize the residual squared from that point so it is pulled toward it
   c. Analyze a point's influence on the slope of the LSRL

H. Predictions
   1. Usefulness of data for making predictions
   2. Comparing predictions from different models
      a. Models which use all of the data points (e.g., LSRL) compared to those that use less than all of the data points (e.g., algebraic model from two points)
      b. Numerical models from in-out tables compared to graphical models using lines of best fit
      c. Understand why different models could have (possibly large) differences in their predictions, particularly for larger values of x when a model is exponential
   3. Substitute x-value into a model's equation to find the predicted y-value
   4. Extrapolation
      a. Meaning: make a prediction outside the original domain of the data set
      b. Cautions
         i. Risky assumption that the relationship will continue outside of the original domain
         ii. Less confidence in prediction when there is no theory to support the impressionistic model
         iii. Predictions for values of x farther away from the original domain are less reasonable than those for values close to the original domain
      c. Context knowledge to determine if predictions are reasonable and what the reasonable bounds on the x-values are for making predictions
   5. Interpolation
      a. Meaning: make a prediction inside the original domain of the data set
      b. Usually more confident in interpolation than extrapolation
      c. Context knowledge to judge accuracy
   6. Prediction interval

a. Meaning, formula, and assumptions
b. Wider than a confidence interval for the same confidence level because it's harder to predict one response than to predict a mean response because it is more variable than the mean (link with Central Limit Theorem)

I. Residual
1. Meaning: error in the prediction for a given x-value
2. Calculation
    a. Observed y-value minus the predicted y-value for a given x-value
    b. Only involves y-values because that is the variable the model is trying to predict
3. Geometric meaning
    a. Vertical distance that the model's curve missed an observed point by, with a sign associated to show if the model was below the point (positive residual) or if the model was above the point (negative residual)
    b. Measured as vertical distance not shortest distance because the residual represents the error in the prediction which means the difference in the observed y-value and the predicted y-value
4. Size
    a. No absolute scale on which to judge the magnitude of the residuals
    b. Must analyze with respect to the magnitude of the data values and the context of the data set
5. Standard deviation of residuals
    a. Meaning, formula, and how to interpret value
    b. Used in the calculation of the standard errors of all the regression parameters
6. Sum of residuals for the LSRL is zero, but not true for non-linear regression models
7. Sum of squared residuals
    a. LSRL has as its criterion to minimize the sum of squared residuals
    b. Can also be used to compare different models because it is a measure of the error of the model in its predictions that can be compared across all models (vs. r which can only be calculated for linear-based models)

J. Residual plot
   1. Purpose
      a. Assess the appropriateness of a model
      b. Not used to determine which model is the best fit
      c. Useful because it is easier to see and analyze whether there is a pattern in the residuals with the x-axis representing the model/zero residual than on the scatterplot of the original data and the model graphed on it
   2. How to make
      a. Linear case: Intuitive idea of rotating LSRL to become the x-axis to become the residual plot
      b. Scatterplot of the (x, residual) pairs; alternatively can plot as (y-hat, residual) pairs; either gives the same pattern of points because y-hat = a + bx, so y-hat is a linear function of x which means the only difference in the plots is the horizontal scale
      c. By hand or with aid of technology
   3. Analysis
      a. No pattern in the residual plot means the model has picked up any pattern there was in the original data and the random scatter of the residuals represents the random noise in the process; therefore it is an appropriate model
      b. Approximately half of the residuals should be positive and the other half negative so that the model is going through the middle of the points on the original scatterplot
      c. A pattern in the residual plot means that the model is missing the data points in a predictable way and indicates that another function model should be tried
      d. A funnel-shaped pattern shows that predictions will be less accurate for values of x in the large part of the funnel; also a problem for doing regression inference procedures because it shows that the residuals do not have the same variation throughout the data set which is an assumption needed for those procedures
      e. Outliers can be identified as isolated points in the residual plot

K. Comparing mathematical models

1. Visually analyze a scatterplot and/or numerically analyze an in-out table to determine what mathematical function(s) would be possible models
2. Sum of squared residuals as a statistic used to compare across models: reasoning and meaning of the statistic (See also II.I.7.)
3. R^2 as another statistic used to compare across models: reasoning and meaning of the statistic (See also II.D.4)
4. Important to look at the scatterplot with the graph of the model(s), the values of statistics which measure the strength of the association (such as r, r^2, sum of squared residuals, or R^2) for each model, and the residual plots to assess which model is the best
5. Knowledge of counterexamples of poor decisions based on using only one of the criteria listed above
6. Linear transformation for non-linear data
   a. Advantages of transforming non-linear data to be linear, including that straight lines are easy to understand, you can use previously learned techniques such as LSRL, the mathematics and calculations are simpler, and it often makes the spread around the model more nearly the same everywhere which is needed for inferential procedures like t-test for the slope of line
   b. Exponential functions are linear based because log y vs. x is linear (due to the linear growth of the exponent)
   c. Power functions are linear based because log y vs. log x is linear
7. Multiple regression: general concept of adding more predictor variables

L. Scatterplot
1. Usefulness: It is the preferred graph for bivariate quantitative data because it plots all of the data points and displays the pattern or trend if one exists
2. Explanatory & response variables
   a. Meaning and how to determine which variable to classify as which type (e.g., the variable you are trying to predict should be the response variable)
   b. Explanatory variable graphed on the x-axis and response variable graphed on the y-axis because

modeling methods use x to predict y as accurately as possible

3. Scaling
   a. Need for regular scale
   b. Interpretation of graph and relationship between the variables can change depending on the scale used (Note: this is how linear transformations work).
   c. Number sense to determine the length of an axis and its scaling
4. Direction: meaning and how to determine (if applicable)
5. Strength
   a. Meaning and how to determine
   b. Strong association suggests creation of a model
   c. Measured by r for linear association
6. Form, direction, and strength of an association can change if look at isolated portions of the data set rather than the data set as a whole
7. Rescaling: The explanatory variable is often rescaled, especially when it is years, to make it smaller values by subtracting a constant value. This does not change the association between the variables because it is a shift left or right by a constant.
8. Which variable is plotted on which axis does not matter in determining if there is an association (and therefore it does not matter for r or r^2 either), but it does matter when determining the mathematical model for predicting y from x.
9. Addition of categorical variable
   a. Different colors or symbols can be used to plot points when a categorical variable is added to a scatterplot.
   b. To analyze, look for differences in the patterns for points for the different values the categorical variable takes.
10. Median trace
    a. Purpose: smoothing technique to detect a pattern in a scatterplot with lots of points and no clear pattern initially
    b. How to perform technique and interpret trace
    c. Choice of number of groups changes the trace so you may see different things with different traces
    d. Meaning, calculation, and properties of the median

M. T-test for the slope of the regression line

1. Assumptions
    a. Variables are linearly related so that a constant slope can be assumed
    b. Errors from line are independent so that the sampling distribution is accurate
    c. Calculation of the standard error pools information across all of the individual distributions at each x-value, which is only appropriate when each distribution has the same variance; therefore it is assumed that the variability in the errors is constant for all x-values
    d. In order to use a t-model, it is assumed that the errors at each value of x follow a Normal model; less important as the sample size grows per the Central Limit Theorem
    e. How to check all of the assumptions
2. Hypotheses
    a. Typical null hypothesis is no linear association between the variables, which means that the slope of the population line is zero or equivalently that the population correlation coefficient is zero.
    b. How the assumption of the null hypothesis is used in the rest of the test
3. Test statistic t
    a. Formulas
    b. Degrees of freedom is n-2 because two parameters, slope and y-intercept, are being estimated
    c. Whether or not a value of r is significantly different from zero depends not only on the difference between r and zero but also how variable the data is (i.e., the more scattered the data is the bigger the difference between r and zero needs to be in order to be certain that it couldn't be equal to zero) and on the number of points in the data set since more variation is "allowed" in bigger data sets.
    d. Represents the number of standard errors away from the mean of the sampling distribution, so the bigger the value of t the less likely it is to have occurred due to sampling variability.
4. Conclusion

a. Find t-critical values using t-tables
b. Interpret computer or calculator output to reach a conclusion
c. If the null hypothesis value of the slope is zero, rejection of the null hypothesis means that the slope is significantly different from zero and therefore there is a linear association (which is not horizontal) between the variables; whether the difference is practically important is up to the interpreter of the data analysis

N. Confidence interval for the slope of the regression line
   1. Assumptions: See II.M.1.
   2. Calculation
      a. t* critical value
         i. Represents the number of standard errors needed to extend on either side of the estimate to capture the desired confidence level proportion of the sampling distribution
         ii. How to find using a t-table, calculator, or computer output
         iii. Relationship between value and degrees of freedom
      b. Calculate interval manually, with a calculator, or from reading computer output
   3. Interpretation: Interpretation of confidence interval in context

O. Graphing Calculator
   1. Stat List Editor
      a. Input data into a list
      b. Define a list by a formula
      c. Delete an entry in a list
      d. Under the ListMath menu, standard deviation and variance are sample calculations, not population calculations; differences between the calculations and when each is appropriate
      e. Restore a deleted list or deleted list name
   2. Stat Calc
      a. 1-Var Stats and 2-Var Stats: how to execute the command, what is contained in each of the outputs
      b. Regression
         i. "Diagnostic On" command to include values of r and r^2 in regression output

ii. Syntax for doing a regression command, including the option of having the regression equation copied into the Y= screen
iii. How to read the regression output
iv. Every time a regression command is executed, the calculator automatically puts the residuals in a list called RESID
v. Values placed in RESID are different between the linear regression performed on the transformed data and a non-linear regression done on the original data; for example, if an ExpReg is done the RESID values are the differences between the observed and expected values when the expected values are found by plugging x into the ExpReg equation
vi. Value of r is calculated on the transformed data for LnReg, ExpReg, and PwrReg (e.g., the value for r in an ExpReg is calculated by finding r when plotting log y vs. x)

3. Graph
   a. Make a scatterplot
   b. When the calculator plots a point in a scatterplot, the point's exact location is in the center of the square or plus sign
   c. Determine appropriate viewing window and how to set the window
   d. How to graph a function in the Y= window
   e. How to trace a graph (either a function graph or a scatterplot)
   f. Use the table of values to find a predicted value
   g. Know what an error message means and probable causes of it (e.g., Dimension Mismatch, Invalid Dimension)
4. Distribution: difference between pdf (probability density function) and cdf (cumulative density function) and when each is called for
5. Stat Tests: Use the calculator to perform a t-test for the slope of the regression line
6. Number limitations

    a. Rounding errors with calculations, particularly for exponential functions
    b. Recognize limitations of big numbers in calculations with the calculator and know ways to 'rescale' the numbers to cope

P. Computer: Interpret standard computer output, including that for the following calculations:
   1. Regression inference, including that the p-value given is for the two-sided test
   2. Regression output, including graphs and tables

III. Association between categorical variables

A. Data
   1. Understand why data is useful and needed (e.g., to determine if there is an association between two variables)
   2. Variability: natural and to be expected in data sets
   3. Differentiating types of data
      a. Difference between quantitative and categorical data and how to classify data as one of the types
      b. For categorical data, can be further classified as nominal or ordinal
      c. A quantitative variable can be made categorical by breaking the quantitative variable into categories by ranges of values and tallying the number of data points that fall in each category; drawback is that you lose some information through the categorization process (e.g., a number is classified as high or low based on a cutoff value but that does not tell you how far above or below the cutoff value it is)

B. Association
   1. Analyzing
      a. Numerically
         i. Compare conditional distributions to see if the proportions that fall in each category vary for the different groups of the other categorical variable
         ii. Can be done either direction
         iii. Related to conditional probability and independence
      b. Graphically: Compare the conditional distributions visually displayed in a segmented bar graph
      c. Inference: conduct a Chi-Square test for independence

2. Simpson's Paradox
    a. Extreme form of the fact that observed associations can be misleading when there is a lurking variable
    b. Under what conditions it can occur
    c. Weighted percentages to analyze the two-way tables

C. Two-way tables
1. Purpose: display data from two categorical variables, showing the number of data values that fall in each cross category
2. Number sense
    a. Grand total must be the same for both the row and column variable; this does not mean that the breakdown for each variable has to be in the same proportions
    b. Can not compare raw numbers that fall in each cross category unless the totals for each category of a variable are equal
    c. Percentages
        i. Converting to percentages makes all ratios have the same scale (out of 100), thus allowing comparisons
        ii. How percentages behave (e.g., if the total is small, one unit change in the numerator can result in a large change in the percent)
3. Dimensions: m x n means it has m rows and n columns
4. Marginal distribution
    a. Used to determine the proportion of data points that fall in each category of one of the categorical variables, ignoring the other categorical variable
    b. How to calculate and interpret
5. Conditional distribution
    a. Used to study possible relationships between two categorical variables
    b. Describes the interaction between the variables by looking at distributions of one variable for given categories of the other variable
    c. Can be found in either direction (given the row variable categories or given the column variable categories)
    d. How to calculate and interpret

D. Segmented bar graph: make and interpret

E. Chi-square statistic

1. Calculation

a. $\chi^2 = \sum_{all\ cells} \frac{\text{(observed-expected)}^2}{\text{expected}}$

b. Meaning of the calculation

i. Numerator: subtracting two quantities tells you the difference between their magnitudes. The difference between what actually happened (observed) and what would happen if the null hypothesis were true (expected) needs to be quantified in order to test the null hypothesis. The difference is squared because it does not matter if the observed value is more or less than the expected and we want all of the values positive to sum them and have a meaningful quantity. The statistic also wants to emphasize the difference between observed and expected counts with respect to the sample size because deviations from the expected are more likely with small sample sizes and become less likely as sample size increases.

ii. Denominator: the differences between observed and expected counts usually get larger the more data there is. Therefore to get an idea of the relative sizes of the differences they are divided by the expected counts for their cell. Another reason is that the chi-square statistic is similar to all other test statistic calculations in that it is based on assuming the null hypothesis is true. The expected counts are determined using that assumption, so it's appropriate for the expected counts to be the reference point (i.e., the denominator) in judging how much the observed and expected counts differ.

iii. The ratios are summed because the statistic must create a quantity that measures the total difference between observed and expected for all categories in order to test the null hypothesis.

c. Must use observed and expected counts in the formula; not equivalent calculation if use percents
d. Observed counts must be whole numbers since they represent the number counted in a category
e. Expected values
    i. Calculated assuming null hypothesis is true
    ii. How to calculate for each cell
    iii. May be not be whole numbers since they represent a theoretical amount
    iv. Important to keep as many decimals as possible to have an accurate calculation of the chi-square statistic
f. Ranges from zero (when there is no difference between the observed counts and the expected counts for all cells) to positive infinity
g. Number sense: How calculation is effected by…
    i. Increased number of data values
    ii. Smaller and larger expected counts (e.g., cells with larger expected values make larger contributions to the test statistic)
    iii. The difference between observed and expected counts, given a set number of data values
    iv. Converting percents to raw numbers: differences in raw numbers as total changes
h. Calculate manually or with technology

2. Distribution
    a. Degrees of freedom
        i. Concept
        ii. A two-way table's degrees of freedom (d.f.) is equal to (number of rows-1)(number of columns-1). This is because once one knows the values in d.f. number of cells, the others can automatically be determined by subtracting from the totals of the rows and columns.
        iii. Must be at least one (i.e., must have at least a 2 x 2 table)
    b. Shape

i. Skewed to the right since the probability of getting small discrepancies is greater than the probability of getting large ones if the null hypothesis is true

ii. Becomes less skewed with increasing degrees of freedom because with more cells, increased values of the chi-square statistic are to be expected; therefore the distribution needs to be adjusted to reflect that bigger values of the chi-square statistic are more likely

c. Density curve

i. Chi-square values on the x-axis; corresponding probabilities on the y-axis

ii. Total area beneath the curve is equal to 1; therefore area beneath the curve for a given range of chi-square values can be used to calculate the probability that the chi-square statistic will be in that range

d. Probability chart

i. Understand how and why the numbers change in the table as they do, based on knowledge of the calculation of the chi-square statistic and degrees of freedom

ii. Probability given in chart is the chance that the chi-square statistic is its value or larger assuming the null hypothesis is true

iii. How to read

iv. Local linearity allows one to use linear interpolation to estimate chi-square probabilities not given in the chart

v. Number sense: Given a chi-square test statistic value and degree of freedom, estimate its probability or critical value

vi. Approximations: The distribution of the statistic approaches a chi-square distribution as the number of observations, n, goes to infinity. The rate of convergence is very fast and for reasonably large n (expected counts in each cell are at least 5) the difference between the approximation and the theoretical becomes negligible.

F. Chi-square tests
   1. Non-parametric
   2. Used with categorical variable(s)
   3. Purpose: Differences between observed and expected counts due to sampling variability are to be expected. Test done to determine how likely the differences are if the null hypothesis is true, i.e., is the difference probably from sampling variability or is there an underlying cause.
   4. General procedure
   5. Determine which type of chi-square test is called for based on the study and/or data

G. Chi-square test for independence
   1. Purpose: Determine if there is an association between two categorical variables
   2. Data
      a. From one sample, subjects are classified according to two categorical variables
      b. An entire population, with each subject classified according to two categorical variables
      c. Categories must be mutually exclusive
      d. Determine the data needed to collect to answer a research question
   3. Hypotheses
      a. Null hypothesis: the two variables are not associated, i.e., independent
      b. Alternative hypothesis: the two variables are associated
      c. Null hypothesis must be independence in order to calculate expected counts based on that assumption
      d. Null and alternative hypotheses must be complements (i.e., if the null is false the alternative must be true)
      e. Implication of null hypothesis: conditional distributions of one variable should be the same across all categories of the other variable
   4. Expected counts
      a. Calculation of how many subjects should fall in a cross-category if the null hypothesis that the two variables are independent is true
      b. Multiple methods to calculate and their equivalence

i. ((row total) x (column total)) / grand total
ii. (row total) x ((column total)) / grand total)
iii. ((row total)/grand total) x (column total)
iv. (probability of category a of variable A) x (probability of category b of variable B) x (grand total)

5. Assumptions
   a. If data is from a sample, it should be randomly selected; needed to ensure that the reason for possible differences between observed and expected counts is not from a bias in the sampling procedure
   b. Expected count of at least five in each cell (see III.E.2.d.vi. for explanation). Some textbooks amend this to say all expected counts must be at least 1 and no more than 20% of the expected counts can be less than 5.
   c. If the expected count assumption is not met, a remedy is to collapse some of the categories to get a larger expected count of at least 5 in all of the cells.
6. Procedure
   a. Calculate and interpret chi-square test statistic (see also III.E.) (e.g., larger values of chi-square statistic represent bigger differences between observed and expected values, which results in smaller p-values)
   b. Calculate and interpret p-value (see III.E.2.d)
   c. Compared to 2-proportion z-test
      i. If each variable has two categories, then the chi-square test for independence is equivalent to a 2-proportion z-test
      ii. Chi-square statistic in this case is the square of the z-statistic
7. Conclusion
   a. Link the value of the chi-square statistic, the p-value or critical value, and the hypotheses to make a conclusion in context
   b. If fail to reject the null hypothesis, would conclude there is no evidence of an association between the variables; not the same as concluding no association

c. If reject the null hypothesis, would conclude the variables are associated; not the same as concluding dependence or causality
d. If reject null hypothesis, the terms of the chi-square statistic or the standardized residuals can be analyzed to determine which cells are contributing the most to the value of the chi-square statistic

H. Calculator
1. Stat List Editor
    a. Input data into a list
    b. Define a list by a formula
    c. Delete an entry in a list
    d. Restore a deleted list or a deleted list name
2. Distribution: Use chi-square cdf command to calculate the p-value for a chi-square test
3. Stat Tests: Use the graphing calculator to perform a chi-square test for independence and interpret its output
4. Matrix: How to enter and view a matrix which contains the data from a two-way table

I. Computer: Interpret standard computer output for the chi-square test for independence