An Examination of the Statistical Problem-Solving Process as a Potential Means for Developing an Understanding of Argumentation

Brittany DeShae Smith Baum

A Dissertation Submitted to the Faculty of the College of Graduate Studies in Partial Fulfillment of the Requirements for the

Degree of Doctorate of Philosophy in Mathematics and Science Education

Middle Tennessee State University

May 2017

Dissertation Committee:

Dr. Angela Barlow, Chair

Dr. Michaele Chappell

Dr. Ginger H. Rowell

Dr. Jeremy Strayer

This work is dedicated to my parents, Troy and Cindy Smith, and my husband, Jacob. Without your constant love, support, and encouragement, none of this would have been possible.

ACKNOWLEDGEMENTS

First, I would like to thank my entire family for their constant support and encouragement. My mom and dad, Cindy and Troy, my sister, Kelsie, my grandparents, my aunts and uncles, and my many cousins. Thank you all for believing in me and always standing by my side.

To my husband, Jacob, I cannot thank you enough for always being my number one fan.

To my aunt, Cheryl, thank you for being the best math teacher and for helping me develop a love of math and teaching others, a love we now share.

To my dissertation chair, Dr. Angela Barlow, thank you for your guidance, direction, and constant feedback throughout this entire process.

A special thank you to my committee members, Dr. Michaele Chappell, Dr. Ginger Rowell, and Dr. Jeremy Strayer, as well as the many faculty and peers at MTSU who have helped me reach this milestone. Your advice, guidance, and friendships over the years have meant so much to me.

Finally, I would like to thank God for providing me with the strength and perseverance I needed to get to this point. Through Christ all things are possible.

ABSTRACT

As part of the recent history of the mathematics curriculum, reasoning and argument have been emphasized throughout mathematics curriculum standards. Specifically, as part of the Common Core State Standards for Mathematics, the Standards for Mathematical Practice were presented, which included the expectation that students develop arguments and carefully consider others' arguments. Due to its emphasis on reasoning and argument, argumentation is one possible way students can meet the expectations of these standards. However, when used in mathematics, argumentation is commonly limited to proofs. Therefore, the use of argumentation in mathematics in ways apart from proofs is needed. Through an examination of students in a college-level introductory statistics classroom, this study examined how engaging in the statistical problem-solving process served as an avenue for developing students' understanding of argumentation.

This study was a holistic multiple case study with three cases. This study utilized data in the form of lesson documents, writing prompts, interviews, field notes, observation video, and the researcher. The data were analyzed and coded using Creswell's five general steps for analyzing qualitative data, Toulmin's Model of Argument, and levels of understanding of argumentation, which emerged from the data. Results from this study indicated that through participating in argumentation training and engaging in the statistical problem-solving process, the participants displayed an increased understanding of argumentation. However, participants did not provide evidence of reaching a deep level of understanding of argumentation. Participants'

tendencies to limit their engagement in and recognition of the argumentation process to the statistical problem-solving process led the researcher to question the role of authority in the classroom and its potential influence on the argumentation process. The findings from this study suggest a need for future research examining processes for developing a deeper understanding of argumentation in students.

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

CHAPTER ONE: INTRODUCTION

Introduction

Ben-Zvi (2006) defined argumentation as "discourse for persuasion, logical proof, and evidence-based belief, and more generally, discussion in which disagreements and reasoning are presented" (p. 2). Students who are involved in argumentation must externalize their thinking (Erduran, Simon, & Osborne, 2004) and articulate their reasoning to one another as they move through the argument (Billig, 1987). In order to engage students in the argumentation process, argumentation itself should be taught explicitly to students (Mason, 1996). One model for teaching argumentation is Toulmin's Model of Argument, which involves the use of claims, data, warrants, backing, rebuttals, and qualifiers (Toulmin, 1958).

Many scientists and science educators promote the use of argumentation in the classroom (e.g., Aydeniz, Pabuccu, Cetin, & Kaya 2012; Driver, Newton, & Osborne, 2000). These science educators believe there are several benefits to using argumentation in the science classroom. One benefit is that argumentation engages students in practicing reasoning skills verbally with their peers (Aydeniz et al., 2012; Bricker & Bell, 2008; Erduran & Jimenez-Aleixandre, 2008). Another benefit is that the use of argumentation allows students to notice their own inconsistencies in their reasoning, as well as discover inconsistencies in others' reasoning and logic (Ayendiz et al., 2012; Bricker & Bell, 2008; Kuhn, 2010). In doing so, students are also able to gain better reasoning skills.

As with science, reasoning skills are important in mathematics. As part of the recent history of the mathematics curriculum, reasoning and argument have been emphasized throughout mathematics curriculum standards (cf. Common Core State Standards Initiative [CCSSI], 2010; National Council of Teachers of Mathematics [NCTM], 1989, 2000). In 1989, as part of the *Curriculum and Evaluation Standards for School Mathematics*, the NCTM included "learn to reason mathematically" (p. 5) as one of the five general goals for all students. The document further explained that "making conjectures, gathering evidence, and building an argument to support such notions are fundamental to doing mathematics" (p. 6). Similarly, in *Principles and Standards for School Mathematics* (NCTM, 2000), reasoning and proof represented one of the five process standards for acquiring and applying mathematics content knowledge. As part of the reasoning and proof standard, students should both "make and investigate mathematical conjectures . . . [and] develop and evaluate mathematical arguments and proofs" (NCTM, 2000, p. 402).

More recently, the Common Core State Standards for Mathematics (CCSSM) (CCSSI, 2010) presented the Standards for Mathematical Practice. These standards "describe ways in which developing student practitioners of the discipline of mathematics increasingly ought to engage with the subject matter as they grow in mathematical maturity and expertise throughout the elementary, middle and high school years" (p. 8). One of the Standards for Mathematical Practice is "construct viable arguments and critique the reasoning of others" (p. 6). This standard calls for students to "justify their conclusions, communicate them to others, and respond to the arguments of others . . . [and] listen or read the arguments of others, decide whether they make sense, and ask useful questions to clarify or improve the arguments" (CCSSI, 2010, p. 6). One way students can accomplish the tasks described in this standard is through the use of argumentation. However, when used in mathematics, argumentation is commonly limited to proofs (e.g., Boero, 1999; Conner; 2007; Smith, 2010).

One possible venue for building students' understanding of argumentation for use in mathematics is through statistics. After many years of being addressed within mathematics education standards, the Guidelines for Assessment and Instruction in Statistics Education (GAISE) were created. As part of these guidelines, the ultimate goal intended for all students at all levels is to develop statistical literacy and the ability to think statistically (American Statistical Association, 2005; Franklin et al., 2007). A key component for students to gain both statistical literacy and reasoning is through the engagement in the statistical problem-solving process (Franklin et al., 2007). Franklin and colleagues (2007) described four components of statistical problem solving: formulating a question (or questions) that can be addressed with data; designing and employing a plan for collecting data; selecting and using appropriate methods to analyze the data; and interpreting the results in relation to the original question(s) (Franklin et al., 2007). These four components of the statistical problem-solving process align with the components of Toulmin's Model of Argument. Therefore, engaging students in the statistical problem-solving process is one possible way to develop students' understanding of argumentation.

Theoretical Framework

For the purposes of this study, the theoretical framework for argumentation was Toulmin's Model of Argument (Toulmin, 1958). Toulmin's Model of Argument has three basic components and three additional components. The first component is the claim. Toulmin (1958) defined the claim as the "conclusion whose merits we are seeking to establish" (p. 90). This component aligns with the formulating-a-question component from the statistical problem-solving process (Franklin et al., 2007). In a statistics context, the claim could be the hypothesis being tested. The second component of the argument model is the data. The data are the facts or evidence that are appealed to as the foundation of the claim. This component is similar to the designing and employing a plan for collecting data component of the statistical problem-solving process (Franklin et al., 2007). In a statistics context, the data are often collected as a means for gathering evidence. Toulmin's (1958) third component is the warrant. The warrant is the justification or logical statement one makes to connect the data and the claim. This component is similar to the selecting and using appropriate methods to analyze the data component from the statistical problem-solving process (Franklin et al., 2007). In a statistics context, this could come in the form of the selection of a proper statistical test to analyze the data. These three components form the simplest structure of an argument (Toulmin, 1958). However, further questions will often arise that require attention (Toulmin, 1958). Therefore, more complex arguments will also include qualifiers, backings, and rebuttals. A qualifier is a statement or condition that limits when the claim is true. This component is sometimes seen in the interpreting of the results in relation to the original question from the statistical problem-solving process (Franklin et al., 2007). In a statistics context, the qualifier could be the level of confidence for your results from a statistical test. Backing is used to validate and support the warrant made. Finally, a rebuttal is a counter-argument or claim that negates any part of the argument. For the purposes of this study, the use of data, warrant, and a claim constituted an argument.

However, the use of backing, rebuttals, and qualifiers was considered part of the argument, if present.

The components of Toulmin's Model of Argument aligned with the components of the statistical problem-solving process are presented in Figure 1. In this figure, the dotted line represents the cycle of the statistical problem-solving process and the solid lines represent the connectedness of the components of Toulmin's Model of Argument.

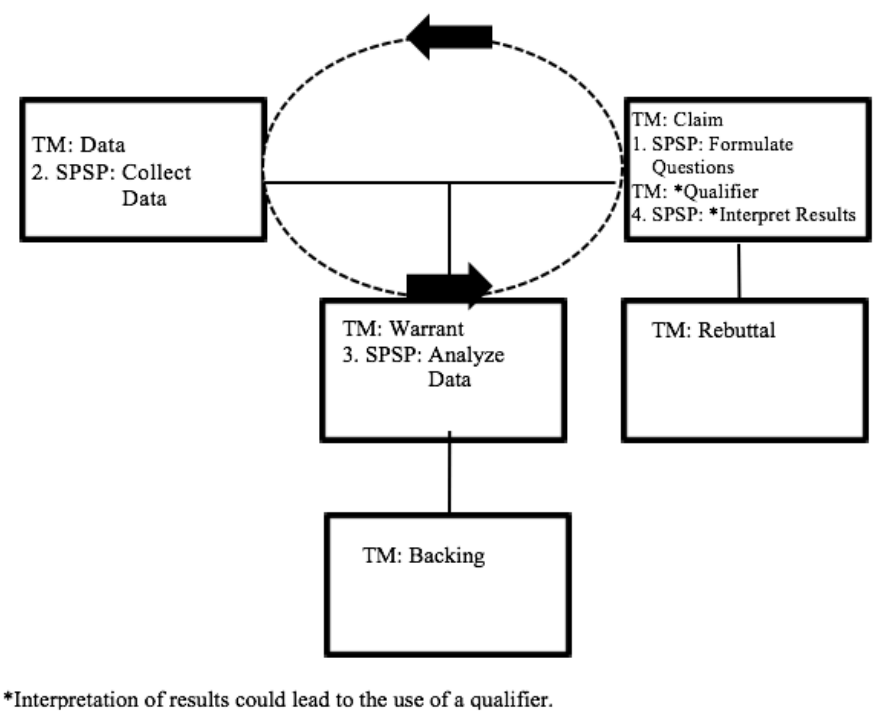

Figure 1. Model aligning Toulmin's Model of Argument and the statistical problem-solving process.

Problem Statement

Recent curriculum standards have emphasized reasoning and argument in mathematics (e.g., CCSSI, 2010; NCTM, 1989, 2000). However, argumentation is commonly limited to proof when used in mathematics (e.g., Boero, 1999; Conner; 2007;

Smith, 2010). One possible venue for building students' understanding of argumentation in mathematics is through engaging students in the statistical problem-solving process (Franklin et al., 2007). The statistical problem-solving process aligns with the components of Toulmin's Model of Argument. However, engaging students in the statistical problem-solving process as a means for developing their understanding of argumentation has not been examined.

Statement of the Purpose

The purpose of the study was to investigate students' engagement in the statistical problem-solving process as an avenue for developing students' understanding of argumentation. Specifically, this study sought to examine students' understanding of argumentation while being engaged in the four components of the statistical problem-solving process. In order to do this, an exploratory case study was used. The specific research question was: How does students' engagement in the statistical problem-solving process support students' understanding of argumentation, if at all?

Significance of the Study

The significance of this study lies in its ability to provide insight into the possible use of students' engagement in the statistical problem-solving process as an avenue for supporting students' understanding of argumentation. In this way, the results of the study will allow for future research investigating argumentation in an introductory statistics classroom. Similarly, results from this study provide insight for a possible way to enhance students' understanding and use of argumentation, which is an important component of the Standards for Mathematical Practice (CCSSI, 2010).

Definitions

To allow the reader to better understand the language of the study, definitions for the following terms are provided.

Argumentation

In this study, argumentation is defined as the engagement or use of one or more components of Toulmin's Model of Argument (Toulmin, 1958). Specifically, these components include: claim, data, warrant, qualifier, backing, and rebuttal. Definitions of these follow and are based on the work of Toulmin (1958).

Claim. Toulmin (1958) defined the claim as the "conclusion whose merits we are seeking to establish" (p. 90).

Data. The data are the facts or evidence that are appealed to as the foundation of the claim.

Warrant. The warrant is the justification or logical statements one makes to connect the data and the claim.

Qualifier. A qualifier is a statement or condition that limits when the claim is true.

Backing. Backing is used to validate and support the warrant made.

Rebuttal. A rebuttal is a counter-argument or claim that negates any part of the argument.

Statistical Literacy

Statistical literacy will be defined as "understanding and using the basic language and tools of statistics: knowing what basic statistical terms mean, understanding the use of simple statistical symbols, and recognizing and being able to interpret different representations of data" (Garfield & Ben-Zvi, 2007, p. 380).

Statistical Problem Solving

Statistical problem solving will be defined by the following four components: formulating a question (or questions) that can be addressed with data; designing and employing a plan for collecting data; selecting and using appropriate methods to analyze the data; and interpreting the results in relation to the original question(s) (Franklin et al., 2007).

Chapter Summary

Recent mathematics curriculum standards have placed an emphasis on reasoning and argument in mathematics (NCTM, 1989, 2000; CCSSI, 2010). As part of the CCSSM, one of the Standards of Mathematical Practice is "construct viable arguments and critique the reasoning of others" (CCSSI, 2010, p. 6). One way to support students in this standard is through the use of argumentation. Since the use of argumentation in mathematics education has previously been limited to mostly proofs, a possible avenue that has not been studied to support students in developing their understanding of argumentation in mathematics is through statistics. As part of the guidelines for statistics outlined in the GAISE report, the ultimate goal intended for all students is statistical literacy and the ability to think statistically, which are apparent through engagement in the statistical problem-solving process (Franklin et al., 2007).

The study investigated students' engagement in the statistical problem-solving process as an avenue for developing students' understanding of argumentation. To this end, this chapter has presented the background, significance, and purpose of the study. The following chapter will review the related literature.

CHAPTER TWO: REVIEW OF THE LITERATURE

Introduction

Argumentation provides a means for students to externalize their thinking (Erduran et al., 2004) and articulate their reasoning to one another as they move through the argument (Billig, 1987). Many scientists and science educators promote the use of argumentation in the classroom (e.g., Aydeniz et al., 2012; Driver et al., 2000) as a way to improve students' reasoning skills (Ayendiz et al., 2012; Bricker & Bell, 2008; Kuhn, 2010). As with science, reasoning skills are important in mathematics. As part of the recent history of the mathematics curriculum, reasoning and argument have been emphasized throughout mathematics curriculum standards (cf. CCSSI, 2010; NCTM, 1989, 2000). One possible venue for building students' understanding of argumentation for use in mathematics is through statistics.

As part of the GAISE (Franklin et al., 2007), a key component for students to gain both statistical literacy and reasoning is through the engagement in the statistical problem-solving process, which includes: formulating a question (or questions) that can be addressed with data; designing and employing a plan for collecting data; selecting and using appropriate methods to analyze the data; and interpreting the results in relation to the original question(s) (Franklin et al., 2007). These four components of the statistical problem-solving process align with the components of Toulmin's Model of Argument (Toulmin, 1958). Therefore, engaging students in the statistical problem-solving process seems to hold potential for developing students' understanding of argumentation.

The purpose of the study was to investigate students' engagement in the statistical problem-solving process as an avenue for developing students' understanding of argumentation. Relevant literature to this study was examined and is described in this chapter. The first area of literature will focus on defining understanding. Next, Vygotsky's social constructivist theory will be described. Then, the influence this theory had on shaping mathematics education reform efforts in the classroom will be described. Following this, a description of mathematical discourse and argumentation will be given. Finally, literature related to argumentation in both science and mathematics education will be described.

Defining Understanding

Understanding of mathematics is emphasized throughout many mathematics education curriculum and publications (e.g., CCSSI, 2010; NCTM, 2000; National Research Council [NRC], 2001). Brownell (1947) defined mathematical understanding as having "meanings of" (p. 9) mathematics, which emphasized students' ability to make connections and see relationships among mathematics concepts. Brownell stressed the importance of creating these connections in mathematics as a means for establishing the meaning of mathematics and, hence, the understanding of mathematics. Further, he cautioned that memorization does not yield an understanding of mathematics.

> To classify arithmetic as a tool subject, or as a skill subject, or as a drill subject is to court disaster. Such characterizations virtually set mechanical skills and isolated facts as the major learning outcomes, prescribe drill as the method of teaching, and encourage memorization through repetitive practice as the chief or sole learning process. In such programs, arithmetical meanings of the kinds mentioned above have little or no place. Without these meanings to hold skills and ideas together in an intellectual, unified system, pupils in our schools for too

long a time have "mastered" skills which they do not understand. (Brownell, 1947, p. 11)

In other words, Brownell indicated memorization of skills and procedures does not produce understanding of the mathematics. Understanding of mathematics requires an ability to see relationships and connections within mathematics and the real world (Brownell, 1947).

In 1976, Skemp further described two different types of understanding, instrumental and relational. Skemp defined instrumental understanding as knowing and doing "rules without reason" (p. 2). This type of understanding involves memorization of procedures where there is "no awareness of the overall relationship between successive stages, and the final goal" (Skemp, 1976, p. 14). The second type of understanding described by Skemp was relational understanding. Similar to Brownell's (1947) definition of understanding, Skemp (1976) defined relational understanding as "knowing both what to do and why" (p. 2). He described four advantages to gaining a relational understanding of mathematics. The first advantage is that a relational understanding is more adaptable to new tasks. In this sense, relational understanding is generalizable and, therefore, can be related and adapted to new problems as they occur. The second advantage is having a relational understanding allows students to more easily retain information learned. Relational understanding allows students to know "how [parts of mathematics] are inter-related [which] enables one to remember them as parts of a connected whole, which is easier" (Skemp, 1976, p. 9). This allows for a learning that is more lasting instead of temporarily memorized. The third advantage is gaining relational knowledge is self-rewarding and should be a goal for students to gain. This means that gaining a relational understanding is motivating for students and a reward within itself. The final advantage to gaining a relational understanding of mathematics is “relational schemas are organic in quality” (p. 10). In other words, once students gain a relational understanding of one thing, they “actively seek out new material and explore new areas” (p. 10) to gain a relational understanding of other areas as well. Overall, Skemp emphasized a relational understanding as one in which students “know both what to do and why” (p. 2) and its importance for students to obtain.

Similar to the relational understanding described by Skemp (1976), the NRC (2001) emphasized conceptual understanding as one of the five interconnected strands of mathematical proficiency. They defined conceptual understanding as “comprehension of mathematical concepts, operations, and relations” (NRC, 2001, p. 116). Further, the NRC explained, “A significant indicator of conceptual understanding is being able to represent mathematical situations in different ways and knowing how different representations can be useful for different purposes” (p. 119). The NRC’s (2001) definition of conceptual understanding mirrors both Brownell (1947) and Skemp’s (1976) definitions of understanding. For example, NRC (2001) suggested that having a conceptual understanding of a mathematical topic indicates that one knows the importance of that topic and how it relates to other topics and concepts in mathematics. This allows for a retention of mathematics learned “because facts and methods learned with understanding are connected [so] they are easier to remember and use, and they can be reconstructed when forgotten” (NRC, 2001, p. 118). Similarly, they indicated a conceptual understanding extends far beyond a memorization of facts or formulas (NRC, 2001), just as both Brownell (1947) and Skemp (1976) explained. Finally, evidence of students' conceptual understanding can be seen in their ability to "verbalize connections among concepts and representations" (NRC, 2001, p. 118).

More recently, the CCSSM (2010) included the term understanding throughout the standards. The authors of this document explained, "One hallmark of mathematical understanding is the ability to justify, in a way appropriate to the student's mathematical maturity, why a particular mathematical statement is true or where a mathematical rule comes from" (p. 4). Throughout the CCSSM, a form of the word understand appears 263 times (Usiskin, 2012). In order to analyze understanding in the CCSSM (2010), Usiskin (2012) divided the use of the term understanding into four dimensions for understanding concepts, which he believed were independent of one another and defined understanding.

> It is because of the relative independence of skill-algorithm understanding, property-proof understanding, use-application understanding, and representation-metaphor understanding from each other that I believe that the understanding of mathematics is a multi-dimensional entity, in the sense that there are independent components that constitute what might be called "real true," "complete," or "full" understanding (Usiskin, 2012, p. 2016).

Usiskin (2012) found that at all grade levels, the term understanding was used most often as property-proof understanding.

Overall, understanding in mathematics extends beyond memorization of formulas or facts (Brownell, 1947; Skemp, 1976; NRC, 2001). It is an ability to know why a mathematical concept or idea is important and how it relates and connects to other mathematical concepts or ideas (Brownell, 1947; Skemp, 1976; NRC, 2001). This type of understanding is often referred to as conceptual understanding. The ability to build connections and see relationships with other ideas is an important component of understanding, not just in mathematics, but in all aspects of understanding. This description of understanding was the goal for students' understanding of argumentation in this study. Understanding argumentation was viewed as students not only correctly using argumentation, but also as students' ability to know when other students are using argumentation and to see connections between the argumentation framework and the statistical problem-solving process.

Social Constructivist Theory

Vygotsky's social constructivist theory is part of the cognitive approaches to learning. The key component of Vygotsky's social constructivist theory is that learning is constructed socially (Vygotsky, 1978). Since argumentation is social by nature (Toulmin, 1958), Vygotsky's social constructivist theory is an important motive for the use of argumentation in the classroom. This section provides a description of Vygotsky's social constructivist theory.

Vygotsky's social constructivist theory explains that learning occurs through social interactions (Vygotsky, 1978). He believed cooperative learning is an essential component of developing a deeper understanding. Vygotsky explained, "The actual movement in the development of the child's thinking occurs not from the individual to some state of socialization but from the social to the individual" (Vygotsky, 1987, p. 76). He believed the development and learning of an individual, which originated from social interactions, is dependent upon signs and tools, specifically thinking and speech, to mediate cognition and mental processes (Smagorinsky, 1995). As described by Toy (2013):

> In learning situations, speech and language are fundamental to mediation. Language as a symbolic system and cultural tool is a central theme in Vygotsky's sociocultural theory as language provides the primary means through which dialogue and co-construction of knowledge and understanding is enabled. (p. 24)

Wertsh (1991) identified three themes of Vygotsky's social constructivist theory that describe the relationship between individual and social processes in learning. The first theme was that individual development begins in social sources. In describing this theme, Palincsar (1998) stated, "From this perspective, as learners participate in a broad range of joint activities and internalize the effects of working together, they acquire new strategies and knowledge of the world and culture" (p. 352).

Wertsh (1991) identified the second theme of Vygotsky's social constructivist theory as one in which human action is facilitated through semiotic means. Semiotic means were described as the tools learned during the social construction of knowledge that are then internalized for future use during independent problem solving (Palincsar, 1998). As part of this theme, logical reasoning is created through social processes (Forman, 2003).

The final theme identified by Wertsch (1991) was that the previous two themes should be viewed through a developmental analysis, which observes the changes that occur through different phases of development. As described by Vygotsky (1978):

> To study something historically means to study it in the process of change; that is the dialectical method's basic demand. To encompass in research the process of a given thing's development in all its phases and changes – from birth to death –

> fundamentally means to discover its nature, its essence, for it is only in movement that a body shows what it is. (pp. 64-65)

Together, these three themes describe Vygotsky's social constructivist theory, which focuses on how learning is constructed socially.

In reference to how Vygotsky's social constructivist theory should be applied to the teaching of mathematics, Forman (2003) indicated that teachers should recognize learning as a dynamic process. Additionally, she suggested that teachers should allow learning for students to occur through communities of practice, as well as through situated practices in which the students and teachers work together. Finally, Forman suggested that mathematical communication, in terms of the meaning of symbolic objects, should involve argumentation. Forman further described that since Vygotsky's sociocultural theory concerns both the social and cultural factors that affect children's learning and development, "teachers are supposed to recognize these students' distinct learning histories, attitudes, motivations, and beliefs about mathematics and to encourage all of them to engage in age-appropriate but also high-level mathematical reasoning and problem-solving" (p. 335).

The ideas of Vygotsky's sociocultural theory can also be applied through the use of argumentation in the classroom. Argumentation involves the interaction between or among students to support or rebut claims made by one another. This interaction among students that is encouraged through the use of argumentation is supported by Vygotsky's social constructivist theory.

Mathematics Education Reform History

The introduction of Vygotsky's social constructivist theory influenced a change in mathematics education reform history (Lambdin & Walcott, 2007). This change led to student communication playing a significant role in students' learning of mathematics (NCTM, 2000). Since the early 1900s, mathematics education reform has progressed through six distinct phases: drill-and-practice, meaningful arithmetic, new math, back to basics, problem solving, and standards and accountability (Lambdin & Walcott, 2007). Recognizing the impact of Vygotsky's social constructivist theory, this section examines these phases and the changes that occurred within each phase.

Drill and Practice

In the 1920s, the drill-and-practice phase, which was based upon Thorndike's theories of connectionism, emphasized an instrumental understanding of mathematics for students (Klein, 2003; Lambdin & Walcott, 2007; Lefrancois, 2006). For example, as part of this phase, the teacher needed to both "recognize and make explicit the essential bonds that constitute the subjects they teach . . . [and] arrange for students to receive the right type of drill and practice on each of the right bonds for the right amount of time" (Lambdin & Walcott, 2007, p. 4). During this phase of mathematics education, interactions among students were not present as part of the learning process, and, therefore, argumentation was not used during this phase.

Meaningful Arithmetic

Following the drill-and-practice phase, the meaningful arithmetic phase focused on developing concepts in a meaningful way in order for students to use mathematics as a life skill (Lambdin & Walcott, 2007). This phase occurred from approximately 1930 to the 1950s and was based on the Gestalt Theory, “which means an organized whole in contrast to a collection of parts” (Lambdin & Walcott, 2007, p. 8). The type of instruction used for meaningful mathematics emphasized activity-based discovery to help students see connections (Lambdin & Walcott, 2007). Argumentation could have occurred during this phase through activity-based discovery among students. However, it is not documented that argumentation was present during this phase.

New Math

The new math phase, which occurred from approximately 1960 to the early 1970s, focused on the properties and fundamental structures of mathematics (Lambdin & Walcott, 2007). During this phase, Bruner recommended a spiral curriculum and discovery learning (Lambdin & Walcott, 2007). As part of the spiral curriculum, “teachers should first promote discovery of mathematical concepts through manipulation of blocks, sticks, chips, or other objects; then present these concepts pictorially; and finally introduce the appropriate mathematical symbols” (p. 11). As with the previous phase of mathematics, argumentation could have occurred during this phase through discovery learning. However, it is not documented that argumentation was used in mathematics classrooms during this phase.

Back to Basics

Many parents and teachers questioned the usefulness of the content being taught in the new math phase, which then led to the back-to-basics phase of mathematics education during the 1970s (Klein, 2003; Lambdin & Walcott, 2007). This fourth phase returned to connectionism and skill development similar to the drill-and-practice phase (Klein, 2003; Lambdin & Walcott, 2007). Therefore, the influence of Vygotsky’s social constructivism was again not present during this phase. Despite the differences among these phases, one common feature they shared was a lack of an emphasis on mathematical discourse and argumentation.

Problem Solving

At the beginning of the 1980s, mathematics educators voiced concern about students' ability to use mathematics in their life beyond high school and America's ability to be internationally competitive in mathematics (Lambdin & Walcott, 2007). This concern led mathematics education into the problem-solving phase in which problem solving and mathematical-thinking processes were the focus of mathematics education reform efforts. In this phase, genuine problem solving was used to "help students develop a deeper understanding of mathematical concepts and skills" (Lambdin & Walcott, 2007, p. 15). During this phase, "considerable attention was given to having students work in cooperative groups and having them verbalize their thinking" (p. 15). Therefore, the influence of Vygotsky's social constructivism was present for the first time in mathematics education. Further, evidence of the influence of social constructivism on the problem-solving phase of mathematics was hinted in NCTM's *Agenda for Action* (1980) through the inclusion of communication of data and results in the recommendation of problem solving as a focus of school mathematics. Also hinted in the *Agenda for Action* (NCTM, 1980) was the use of argumentation in mathematics. In addition to the components of problem solving listed in the recommendations that relate to argumentation, such as formulate questions and seek out appropriate data, developing informal proofs including counterexamples was also a recommended action for mathematics students (NCTM, 1980).

Standards and Accountability

The problem-solving phase led to the current phase of mathematics education, which is known as the standards and accountability phase. This phase began with the release of NCTM's (1989) *Curriculum and Evaluation Standards for School Mathematics*, which focused on problem solving, meanings, communication, connections, and patterns. These standards were updated in 2000 as NCTM released the *Principles and Standards for School Mathematics*. These updated standards included the five process standards, which describe processes in which students should engage in order to learn mathematics effectively. The process standards include: problem solving, reasoning and proof, communication, connections, and representation (NCTM, 2000). As before, evidence of the influence of the Vygotsky's social constructivist theory can be seen by the inclusion and emphasis of communication in the standards. As one of the five process standards, the communication standard calls for students to "organize and consolidate their mathematical thinking through communication; communicate their mathematical thinking coherently and clearly to peers, teachers, and others; analyze and evaluate the mathematical thinking and strategies of others; and use the language of mathematics to express mathematical ideas precisely" (NCTM, 2000, p. 60). Further, the *Principles and Standards of School Mathematics* describes communication as "an essential part of mathematics and mathematics education" (NCTM, 2000, p. 60). Argumentation can also be seen in the communication standard through the analysis and evaluation of other students' mathematical thinking. However, the presence of argumentation in mathematics was mentioned for the first time in the reasoning and proof standard. As part of this standard, students are to "develop and evaluate mathematical arguments and proofs" (NCTM, 2000, p. 56).

More recently, a new set of standards, referred to as the CCSSM (CCSSI, 2010), has been released. These standards have been adopted by 43 states (CCSSI, 2015). The CCSSM includes the Standards for Mathematical Practice, which describe "varieties of expertise that mathematics educators at all levels should seek to develop in their students" (CCSSI, 2010, p. 6). Evidence of the influence of the social constructivist theory and the inclusion of argumentation in mathematics can continue to be seen in the standard that calls for students to "construct viable arguments and critique the reasoning of others" (CCSSI, 2010, p. 6). Similar to the communication standard, this standard calls for students to "justify their conclusions, communicate them to others, and respond to the arguments of others" (CCSSI, 2010 p. 6).

After the release of the CCSSM, NCTM published *Principles to Actions: Ensuring Mathematical Success for All* (2014). The importance and influence of the social constructivist theory and the inclusion of argumentation in mathematics education is again found throughout this publication. For example, as part of the eight Mathematics Teaching Practices, which "provide a framework for strengthening the teaching and learning of mathematics" (NCTM, 2014, p. 9), one practice included is to facilitate meaningful mathematics discourse. This practice describes "effective teaching of mathematics facilitates discourse among students to build shared understanding of mathematical ideas by analyzing and comparing student approaches and arguments" (NCTM, 2014, p. 10). Analyzing and comparing one another's arguments is an important component of argumentation.

Summary

Throughout the first four phases of the mathematics education reform history, argumentation was not present and did not play a role in mathematics classrooms. However, the introduction of Vygotsky's social constructivist theory to the United States through his 1978 publication signaled a turning point in mathematics education history. Through this introduction of Vygotsky's (1978) research, argumentation became an important aspect of the mathematics classroom, which is seen through its inclusion in mathematics education standards (e.g., CCSSI, 2010; NCTM, 1989, 2000, 2014).

Argumentation in the Classroom

As part of the recent history of the mathematics education curriculum, mathematical discourse, reasoning, and arguments have been emphasized in mathematics curriculum standards and publications (c.f. CCSSI, 2010; NCTM, 1989, 2000, 2014). In the sections that follow, a description of mathematical discourse and argumentation will be given.

Definition of Mathematical Discourse

Throughout the different phases of mathematics education, the role of communication and mathematical discourse in the classroom has varied. With the current emphasis on standards, engaging students in discourse has become a key component to advancing students' learning of mathematics (NCTM, 2014). Mathematical discourse is defined as the "purposeful exchange of ideas through classroom discussion, as well as through other forms of verbal, visual, and written communication" (NCTM, 2014, p. 29). Once mathematical discourse is established in the classroom, students have the opportunity to "share ideas and clarify understandings, construct convincing arguments regarding why and how things work, develop a language for expressing mathematical ideas, and learn to see things from other perspectives" (NCTM, 2014, p. 29). To support classroom discourse effectively, however, teachers are responsible for building a safe community for students to express their ideas (NCTM, 2000).

Mathematical discourse and its effect on student learning in mathematics classrooms has been researched throughout the years. Webb (1991) conducted a literature review on the use of small groups in mathematics and its impact on mathematics achievement. Specifically, her paper focused on empirical research studies that "systematically analyzed task-related interaction among students working on mathematics material in peer-directed small groups in the classroom [and] examined the relationship between interactions among students and student achievement" (Webb, 1991, p. 370). There were 17 studies total that were reviewed, and they all occurred within the time span 1979 to 1985. All of the studies examined in her literature review used the same procedures for the small group work that occurred. In these studies, students were first introduced to the topic by the teacher. The students then worked in small groups of typically four students each on a set of problems and were encouraged to work together. The students were also encouraged to make sure all students within the group were learning and told not to divide work up among group members. In the majority of these studies, small groups were videotaped or audiotaped, and the frequency of different kinds of student interaction within the small groups were coded. Students in all of these studies were given an achievement test at the end of the unit, which they completed individually. Webb (1991) found that among these studies, giving elaborate explanations within small groups was positively correlated to mathematics achievement.

She also provided a list of suggestions for discourse that emerged from the research that should occur among small groups. Webb (1991) concluded:

> The optimal small group setting is one in which students freely admit what they do and do not understand, consistently give each other detailed explanations about how to solve the problems, and give each other opportunities to demonstrate their level of understanding. (p. 386)

Webb's (1991) research provides suggestions based on empirical research on how to use small groups and what should be occurring within these small groups to optimize student learning and achievement in mathematics. These suggestions also support the use of argumentation in mathematics classrooms through the use of small groups because students engaged in argumentation with one another will continually communicate and provide explanations for their claims to each other.

Students' beliefs regarding how discourse affects their achievement in mathematics is important to consider when using argumentation in mathematics because argumentation involves communication and interaction among students. Moore (2000) conducted a qualitative ethnographic research study that investigated the role of student discourse in mathematics achievement. The researchers defined discourse as including "verbalizations, behaviors, and social interactions that occur as the students participate in the community of mathematics classrooms" (Moore, 2000, p. 5). The participants in her study were five male African-American high school students. Three of the students were enrolled in Algebra II and the remaining two were enrolled in Precalculus. Data for the study included participant observations, audiocassette recordings, written-document collection, and in-depth interviews. The participant observations occurred three to four times a week for four months. The observations and audiocassette recordings were transcribed, and discourse with the teacher, other students, and among each other was identified. Interview data were also coded into three main themes: classroom environments, curriculum, and external factors for outside of the classroom or school that influenced the students' learning experiences. In addition to looking at each student's data individually, a cross-case analysis was performed.

In reference to the role student discourse has on mathematics achievement, Moore (2000) found various results among her participants. In reference to whether they believed their experiences using discourse in mathematics were important contributing factors to their mathematics achievement, four of the five students indicated yes. Limitations to this study included the use of only African-American male students and the lack of information provided on the types of tasks on which the students were working when discourse occurred throughout the study. Since the participants in this study believed discourse in their mathematics classroom was important for their mathematics achievement, students could possibly view discourse through the use of argumentation in mathematics as having a positive influence on their mathematics achievement.

An important aspect to consider for discourse within the small group setting is how the students will interact with one another. Larriva (1998) conducted a case study that focused on understanding how students communicate within a small group setting in a high school mathematics classroom and the effects of these forms of communication. The participants of the study were two female high school students. One student was Latina and the other white, but both students were native-English speakers. The participants were both enrolled in a third-year college preparatory mathematics class. However, one participant was considered to be the top academic student in the class and the other had a D average. The participants worked with two other students in the class to form a small group of four students. Three lessons were used for this study, all of which occurred within the same academic year. The students worked on problem-solving tasks within their small groups during these three lessons. The curriculum within the mathematics class was “intended to promote numerous opportunities for engagement in verbal and written discourse. . . [which] provided the opportunity to examine the impact of an NCTM-minded reform curriculum on the functioning of a mathematics classroom” (Larriva, 1998, p. 26). Video and audiotape recordings of the small group interactions from the three lessons, videotape from the whole class discussion of the three lessons, student interviews, teacher interviews, and participant observations were the data sources from this study. The video and audio recordings were transcribed and coded for the type of interactions that occurred for both participants.

The participant with the lower average in the mathematics course was found to be less assertive in her communication within the group. Similarly, she had a difficult time gaining the group’s acceptance of her contributions when she provided them to the group. In contrast, the student who was at the top of the mathematics course demonstrated authority within the group and influenced the direction of the tasks in accordance to her beliefs about the problems. Therefore, Larriva (1998) found that students with more assertive behavioral characteristics tended to lead group discussions in mathematics and even overruled mathematical ideas and suggestions from other group members. Larriva also found that a student’s perception of herself and of her other group members tended to influence the roles within the group in reference to who was in charge. However, this study only focused on two members within the group and did not account for the other two students in the same group, which was a limitation of the study. This study is significant, though, because it shows that not all students are given the same respect in terms of mathematical contributions within small group work. Therefore, it is important to be aware of how students are grouped within small groups, especially when students are using argumentation, since argumentation involves analyzing and engaging in one another's arguments. It is also important to create social norms within the classroom that teach students to value the mathematical contributions of all students.

Another important aspect to consider when students are working in small groups is the teacher impact on communication within the group. Kysh (1999) conducted a case study that investigated how students talk about mathematics, both with other students and with the teacher, when working in small groups. The participants in the study were the students in an Algebra I course, which was taught by the researcher the year in which the data were collected. Data for the study included: audio recordings of teacher and student interactions, audio recordings of small group interactions, field notes, copies of students' written work, and audio recordings of follow-up interviews with the students. All of the audio recordings were transcribed and coded based on the type of talk that was occurring. The categories for coding included: about mathematics, in mathematics, with mathematics, beyond mathematics, and attention directors. The *about mathematics* category included talk that was about mathematics but did not include the use of mathematics vocabulary. The *in mathematics* category included talk that described the mathematics a student was in the process of doing and used mathematics terminology.

The *with mathematics* category included talk, both using mathematics terminology and using non-mathematics terminology, during the problem-solving process. The *beyond mathematics* category included talk that does not involve mathematics vocabulary but was used to generalize ideas that emerge from the mathematics. The final category, *attention directors*, included talk used to get another student's or the teacher's attention.

Kysh (1999) found that student talk within the categories of in, with, and beyond mathematics occurred a higher percentage of the time when students talked with each other in their small groups versus when they talked with the teacher. Further, she found that small groups provided an increased number of opportunities for student participation in the classroom, and students used more mathematical language and provided more explanations when in small groups. This study demonstrated the benefits of students working in small groups and communicating with one another versus with the teacher as a means for developing mathematical discourse.

In addition to students' beliefs about small groups and how students interact in small groups, student learning that occurs through the use of small groups is also an important component to consider. In another study, Bradford (2007) conducted a mixed-methods case study to explore the use of student discourse through mathematics dialogue activities as an instructional strategy to improve student learning and achievement in mathematics for low-achieving mathematics students. The participants in the study were 160 students in four different prealgebra courses. Two classes were the experimental classes for the study and the remaining two classes were the control classes. The experimental classes used dialogue activities to promote student discourse, and the control classes used the teachers' regular instructional methods, which did not include the student dialogue activities. The four classes were taught by two different teachers. Each teacher taught one experimental class and one control class. Data for this study included: observations, audio recordings of both experimental and control classes, field notes, student interviews, pre- and post-assessment, and pre- and post-attitudes survey. The qualitative component of the analysis was used to describe characteristics of student discourse and teacher facilitation in both the experimental and control classes. The quantitative component of the analysis was used to evaluate student achievement, student problem solving, student attitudes towards mathematics, and student perception of the dialogue activities.

In this study, Bradford (2007) found that students working in small groups displayed more indicators of student learning and better attitudes. Additionally, students in the experimental classes outperformed students in the control classes in both mathematics achievement and problem solving. Further, the experimental classes also provided more opportunities for student-led questions and explanations. The findings from this study are important because they support the use of student discourse through small groups as a means of improving students' mathematics achievement.

Overall, research indicates the use of small groups in mathematics classrooms can have a positive impact on student learning and achievement. There are many factors to consider when using small groups, such as behavior characteristics of the students within a group and the type of communication occurring within the group. Argumentation is one way to support students in communication of their thinking and ideas to one another within a small group setting.

Argumentation

One form of discourse is argumentation (Mason, 1996). This relationship is seen through a descriptor of the evidence of student engagement in mathematical discourse, which indicates that students should “critique the reasoning of peers, using examples and counterexamples to refute arguments” (NCTM, 2014, p. 35). Similarly, students who are involved in argumentation must externalize their thinking (Erduran, Simon, & Osborne, 2004) and articulate their reasoning to one another as they move through the argument (Billig, 1987). In order to engage students in these argumentation processes, argumentation itself should be taught explicitly to students (Mason, 1996). One such model for engaging in argumentation is Toulmin’s Model of Argument (Toulmin, 1958).

In reviewing the literature regarding studies of argumentation that utilized Toulmin’s Model of Argument (Toulmin, 1958), the majority of argumentation studies have been conducted in science education (e.g., Aydeniz et al., 2012; Driver, Newton, & Osborne, 2000). Therefore, in the following sections, studies of the use of argumentation and Toulmin’s Model of Argument in science education will be discussed first. This will provide an initial perspective of the implementation and effectiveness of argumentation in the classroom in order to better understand the use of argumentation in mathematics. Following the discussion of argumentation in science education, studies involving argumentation and Toulmin’s Model of Argument in mathematics education will be discussed.

Argumentation in science. Many scientists and science educators promote the use of argumentation in the classroom (e.g., Aydeniz et al., 2012; Driver et al., 2000). These science educators believe there are several benefits to using argumentation in the science classroom. One benefit is that argumentation engages students in practicing reasoning skills verbally with their peers (Aydeniz et al., 2012; Bricker & Bell, 2008; Erduran & Jimenez-Aleixandre, 2008). Another benefit is that the use of argumentation allows students to notice their own inconsistencies in their reasoning, as well as discover inconsistencies in others' reasoning and logic (Ayendiz et al., 2012; Bricker & Bell, 2008; Kuhn, 2010). In doing so, students are also able to gain better reasoning skills.

Argumentation has been examined in several empirical studies (e.g., Aydeniz et al., 2012; Cross, Taasoobshirazi, Hendricks, & Hickey, 2008; Venville & Dawson; 2010). Aydeniz et al. (2012) investigated the effect of argumentation-based instruction on college students' conceptual understanding of the properties and behaviors of gases. His study took place at a university in Turkey and involved two general chemistry courses taught by the same professor. One course, containing 52 students, was the control group. The second group, containing 56 students, was the experimental group. Both the control and experimental groups received six hours of lecture on the behavior and properties of gases over a two-week period. Following this, the control group participated in two hours of problem solving in small groups of four. At the end of this problem-solving activity, the professor solved and discussed problems that "emphasize[d] the most important ideas on the board" (Aydeniz et al., 2012, p. 1308). In contrast, following the two-week lecture period, the experimental group participated in two hours of argumentation training based on Toulmin's Model of Argument. During this training, the students participated in one hour of written arguments followed by one hour of verbal arguments within small groups of three students.

The data in this study consisted of pre- and post-test scores and a post-argumentation questionnaire. The researchers conducted several statistical tests on the data and found that engaging in argumentation-based learning improved students' conceptual understanding of the properties and behaviors of gases in a college science course. A limitation to consider for this study was the lack of a thorough description of how argumentation was taught to the students.

Just as the effects of argumentation on students' conceptual understanding is important, the effects of argumentation on student learning is also important to consider. Cross et al. (2008) examined the effect of argumentation on student learning during the implementation of a NASA-funded biology program through a case study. The sample for the study included three African-American male high school biology students in the 10^{th} grade. These students participated in four activity-oriented quizzes over a two-week period. After completing the quiz, the three students formed a group and participated in an "argumentation review routine" (Cross et al., 2008, p. 842). The argumentation review routine was described by the authors.

> Students were told that for each item they should state their answers and then, using data, warrant or explain those individual answers, and – using backings where appropriate – come to an agreement on the most 'sensible solution' . . . Students were also encouraged to use more complex forms of argumentation including rebuttals and qualifiers while discussing their answers. (Cross et al., 2008, p. 842)

As part of the argumentation training, the students were shown an animated cartoon video prior to their first quiz. This video contained characters that demonstrated both a good argumentation review routine and a poor argumentation review routine. The data for this study included the video recordings from the group's argumentation review routine and a pre- and post-exam. The video recordings were transcribed and coded in reference to Toulmin's Model of Argument. These arguments were labeled as lower quality argumentation or high quality argumentation. Lower quality argumentation consisted of arguments that included claims, data, warrants, and backings (Cross et al., 2008). High-quality arguments included rebuttals and qualifiers (Cross et al., 2008).

Overall, Cross et al. (2008) found argumentation had a positive impact on student learning and achievement in science over the two-week period. The authors also noted that "students tend to feel more comfortable and be more competent in arguing about concepts when they are sufficiently knowledgeable about that subject" (Cross et al., 2008, p. 856). One limitation of this study, however, was the lack of information and consideration of teacher facilitation in reference to both the argumentation training and argumentation review routine.

Conceptual understanding is important for student learning. Similar to Aydeniz et al.'s (2012) study that examined the impact of argumentation on students' conceptual understanding, Venville and Dawson (2010) explored the impact of argumentation on students' conceptual understanding of science. Their study was a quasi-experimental embedded within a case and took place in 10^{th} grade genetics classes in Australia. The genetics course was a 10-week course. There were four classes used in the study: two control classes with 46 students total and two experimental classes with 46 students total. The two experimental classes participated in three explicit argumentation lessons during the genetics course. These lessons were 50 minutes long each and were inserted during the sections of genetic diseases, genetic engineering, and cloning.

The first of the three argumentation lessons involved the explicit teaching of argumentation skills. The remaining two argumentation lessons incorporated whole-class argumentation into the genetics lessons. The instructor of the two experimental classes participated in a two-hour, one-on-one argumentation training session prior to the first argumentation lesson. The two control classes did not participate in any of the planned argumentation. During the course time when the experimental groups were participating in an argumentation lesson, the control classes conducted library research on genetic diseases, genetic engineering, and cloning.

The instrument used in this study was a student survey. The survey was analyzed based on the components of Toulmin's Model of Argument, a Mann-Whitney U-test, and the Wilcoxen Signed Rank Test. Overall, Venville and Dawson (2010) discovered a significant improvement in high school students' conceptual understanding of genetics topics after implementing argumentation. Further, they noted that "the explicit teaching and practice of argumentation skills does improve the complexity of students' argumentation" (Venville & Dawson, 2010, p. 969). However, the researchers only analyzed written arguments and did not consider or implement verbal arguments among students, which was a limitation of this study.

Collectively, these three studies (Aydeniz et al., 2012; Cross et al., 2008; Venville & Dawson, 2010) indicate that argumentation can have a positive impact on students' conceptual understanding for certain science concepts. These results are important because they potentially suggest the role of argumentation in developing a conceptual understanding of mathematics, which is a main focus throughout mathematics curriculum standards (e.g., CCSSI, 2010; NCTM, 1989, 2000). Therefore, the use of argumentation in a mathematics classroom could possibly help develop students' conceptual understanding of mathematics.

Several studies have specifically used Toulmin's Model of Argument for implementing argumentation or describing students' arguments (e.g., Bell & Linn, 2000; Erduran et al., 2004; Jimenez-Aleixandre, Bugallo-Rodriguez, & Duschl, 2000). Bell and Linn (2000) conducted a study involving 172 middle school students. These students were dispersed in six class periods and worked in pairs in the study. The study investigated the types of arguments made by middle school students and whether argumentation allowed for knowledge integration (Bell & Linn, 2000). The study was conducted as part of the Knowledge Integration Environment design and used a software named SenseMaker argument editor. As described by Bell and Linn (2000), "SenseMaker allows students to construct and edit their arguments using a graphical representation" (p. 799). This software helps students build arguments and search the world-wide web to find evidence. As part of SenseMaker, students can use the Mildred guide component, which provides prompts and hints for students when asked.

The data for the study were the arguments students created in the SenseMaker argument editor. These arguments were analyzed based on the components of Toulmin's Model of Argument. Bell and Linn (2000) found that, overall, creating arguments can promote knowledge integration. Further, the researchers found that students often used data to support claims, but rarely included warrants and backing in their arguments. This study, however, used a software to help students create arguments. It did not examine how students can create arguments outside of the template of the software, which was a limitation of the study.

As Bell and Linn (2000) discovered, students use of the components of Toulmin's Model of Argument was not uniform across all components. In another study that examined how students use these components, Jimenez-Aleixandre et al. (2000) conducted a study that focused on high school students' ability to develop arguments and the type of arguments they develop without prior argumentation training. This case study occurred in a high school genetics class in Spain. The participants were the 9^{th} grade students in the class; however, for the purposes of argumentation occurring within a small group, the researchers focused on a group of four females for the study (Jimenez-Aleixandre et al., 2000). The participants participated in six one-hour sessions over a period of two weeks. The first four sessions were lecture-style with the dialogue being initiated by the teacher and under her control. Since Jimenez-Aleixandre et al. (2000) believed that problem-solving tasks that asked students to provide reasons for their answers and choices provided an opportunity for students to engage in argumentation, the last two sessions focused on students working in small groups to solve genetics problems.

The data for the study included audiotape recordings from the small group and observations of the small group. The audiotape recordings were transcribed and coded. Instances where students were considered to be "talking science" (Jimenez-Aleixandre et al., 2000, p. 767) were then analyzed using Toulmin's Model of Argument. Jimenez-Aleixandre et al. (2000) found that high school students studying genetics focused mainly on their claims, but rarely supported these claims with data or warrants. Further, the researcher noted that within the small group, two students shared a leadership role and attributed for the majority of the argumentation that occurred within the group. One limitation to the study, however, was that the researchers indicated the teacher lacked the skills to shift the classroom to student-centered even when she wanted a student-centered environment. Therefore, the students in the study were likely not accustomed to student discourse within the classroom, which could impact the amount of argumentation that occurred.

As demonstrated by Bell and Linn (2000) and Jimenez-Aleixandre et al. (2000), the ways in which students use the components of Toulmin's Model of Argument vary, but consist mostly of claims. However, there is a need to analyze these arguments. Erduran et al. (2004) conducted a study that investigated the use of Toulmin's Model of Argument as an instrument for analyzing argumentation discourse both qualitatively and quantitatively. Their study was part of a larger project that took place in the United Kingdom. Twelve middle school science classes were the focus of this study. Each class had a different teacher and all students were in grade eight. The teachers were trained in a total of nine workshops where they were "familiarized with the overall objectives and research design of the study" (Erduran et al., 2004, p. 921). As part of these workshops, teachers were introduced to Toulmin's Model of Argument and allowed to explore applications of argumentation in the classroom. They were also taught to encourage students to use evidence to support claims. One lesson that tasked students to present an argument for and against the funding of a zoo was used in the study. This lesson was an hour long and repeated a second year by the same teachers with different students. Audiotape recorders were placed on the table of two groups of students as well as wired to the teachers each year (Erduran et al., 2004). These recordings were then transcribed and coded.

When presenting the results of the study, the researchers were not clear on whether the data analyzed were primarily drawn from the teacher recordings or the student recordings, which was a limitation of the study. For the quantitative analysis, the arguments were grouped by the number of components from Toulmin's Model of Argument that were included in the argument. Arguments that consisted of fewer components were considered less sophisticated arguments. As the number of components found within the argument increased, the researchers considered the argument to be more sophisticated.

For the qualitative analysis, Erduran et al. (2004) found that argumentation was used in the classroom in both years and used in similar ways in both years. However, there were no patterns found across the teachers, so the use of argument in the class was dependent upon the teacher. In reference to the type of arguments made for the quantitative analysis, Erduran et al. (2004) found that most arguments contained fewer components of Toulmin's Model of Argument. However, there was a statistically significant difference in the more sophisticated arguments used in the second year versus the first. Overall, Erduran et al. (2004) found that Toulmin's Model of Argument could be used as an instrument for both qualitatively and quantitatively measuring arguments in the classroom discourse of a middle school science classroom.

In general, argumentation has shown to have a positive impact on student learning and understanding of some scientific concepts (Aydeniz et al., 2012; Cross et al., 2008; Venville & Dawson, 2010). As part of introducing argumentation to students and analyzing student arguments, Toulmin's Model of Argument has played a key role. However, results from the use of this model have varied (Bell & Linn, 2000; Jimenez-Aleixandre et al., 2000). Therefore, further research on the use of Toulmin's Model of Argument as a means of analyzing students' arguments in science education is needed.

Argumentation in mathematics. Similar to the role of argumentation in science, argumentation plays a fundamental role in mathematics (Banegas, 2013). Although empirical research using Toulmin's Model of Argument is not as prevalent in mathematics education research as in science education research, there are some studies that have used the model (e.g., Conner, 2007; Smith, 2010; Yackel, 1997, 2001). These studies used Toulmin's Model of Argument in a similar way as many science education studies, in that most studies using Toulmin's Model of Argument in mathematics education use the model for evaluation purposes (e.g., Conner, 2007; Smith, 2010; Yackel, 1997, 2001).

One such study was conducted by Yackel (1997). In her study, she found that Toulmin's Model of Argument can be used as a tool to demonstrate changes in learning that take place over time. This led to the use of Toulmin's Model of Argument in three differential equations classrooms taught by the same professor at the same university (Yackel, 2001). However, the last class of the three was the focus of her study. During this study, teaching experiments were conducted anywhere from six weeks to the entire semester (Yackel, 2001). The length of the teaching experiment discussed in this study was not specified, which was a limitation of the study. Throughout the teaching experiments, data were collected in the form of: videotapes of each class session; videotapes of two different small groups; field notes from both the observer and the instructor; records of instructional activity and instructional decisions; student work in the form of in-class work, homework assignments, weekly electronic journal entries, and reflective portfolios submitted twice in the semester; and videotape and artifacts from individual student interviews. These interviews were conducted with selected students in order to “gain initial information about their concepts related to functions and rates of change, to assess their understanding of key concepts of the course, and to inquire about their beliefs about mathematics activity and mathematics learning” (Yackel, 2001, p. 10).

As part of the teaching experiment, the instructor emphasized justification and argumentation through collaborative learning. The instructional design of the differential equations class consisted of students working in small groups for two to four minutes followed by whole class discussions for 15 minutes. Students then returned to their small group discussions for another two to four minutes. This cycle continued throughout the 80-minute class period. During the small group discussions, students were told to “think about some question or issue rather than to solve a specific problem” (Yackel, 2001, p. 12). Because of the emphasis on reasoning and argumentation in the small group discussions, whole class discussions resulted in the emergence of key concepts, rather than the telling of key concepts by the instructor. For example, as part of one class period during the teaching experiments of the third differential equations class, four small group sessions and five whole class discussions occurred during the 80-minute class period. During this time, the instructor repeatedly asked students for reasons and justification of their claim. Similarly, the instructor asked other students in the class what they thought about another student’s argument in order to encourage further development and use of arguments. As described by Yackel (2001):

> The instructor was initiating the expectations that students are to provide arguments to justify their claims, that they are to attempt to make sense of others' arguments, and that there might be more than one argument to support a given claim. From the point of view of argumentation, we would say that the discussion focused on data, warrants, and backing to support conclusions but not on the conclusions themselves. (p. 13)

Overall, as with her previous research, Yackel (2001) found that Toulmin's Model of Argument was a useful tool for demonstrating how students' learning progressed. She described Toulmin's Model of Argument as placing an "emphasis on explanation and justification in a mathematics classroom [which] leads to mathematics learning that emphasizes reasoning" (Yackel, 2001, p. 15).

Similar to Yackel's (2001) use of Toulmin's Model of Argument as a tool for demonstrating the progression of student learning, Conner (2007) used the model as a tool for analyzing arguments made within a mathematics classroom. In her study, Conner (2007) conducted a multi-case study involving three cases that investigated student teachers' conception of proof and how that related to their facilitation of argumentation in secondary mathematics classrooms. The three participants in this study were secondary mathematics student teachers. Several data sources were used in order to "discern the student teacher's conception of proof and justification or describe the argumentation in his or her classroom" (Conner, 2007, p. 41). These data sources included: video and audio recordings of two semi-structured interviews with each of the three student teachers in reference to their beliefs about mathematics and proof within mathematics; audio recordings and field notes from an interview with each student teacher's mentor teacher; field notes from class observations; audio recordings of classroom discourse; and lesson plans, tests, quizzes, and worksheets created by the student teachers.

After data collection, a classroom argumentation analysis was conducted for each student teacher. This analysis consisted of identifying episodes of argumentation that consisted of a claim, data, and warrant; writing a description of the episodes of argumentation; and assigning who was contributing the components of the argument. Additionally, an analysis for each participant's conception of proof was performed using each student teacher's interviews. Patterns were then examined for both analyses for each student teacher to see if there were common themes between the two. Last, a cross-case synthesis was performed where "similarities and differences between the three cases were recorded and analyzed until [the researcher] obtained a unified characterization of a possible relationship between student teachers' conceptions of proof and how they support argumentation in secondary mathematics classrooms" (Conner, 2007, p. 70).

Conner (2007) found that Toulmin's Model of Argument enabled the researcher to analyze argumentation in the classroom. Also, Conner (2007) found that most of the arguments that occurred in the classroom were initiated and led by the student teachers. The student teachers had a tendency to provide many of the warrants or provisions to student warrants when made by a student. However, argumentation did occur in all three classes without it being an intentional goal of any of the lessons, which "gives reason to believe that argumentation may occur naturally in mathematics classrooms and may provide a base upon which to build more robust argumentation practices" (Conner, 2007, p. 238). Additionally, how the student teacher's supported argumentation in their classroom aligned with their conception of proof, which implies a relationship between proof and argumentation. However, the study only accounted for the student teacher's conception of proof as an influence on the use of argumentation in the classroom, which was a limitation of the study since it is likely more factors existed.

The use of argument with proof can be seen in the CCSSM (CCSSI, 2010). Technology can play a role in the development of proofs, especially in geometry. Therefore the effects technology has on argumentation is important to examine. Smith (2010) investigated the arguments that middle school students developed both when using a technological tool and when using non-technological tools. In his qualitative multi-case study, three pairs of students from both the classroom using technological tools and the classroom using non-technological tools were purposefully selected and were the focus of this study. The two classrooms were both considered teaching experiments in eighth grade mathematics classes and both taught by the researcher. The activities and tasks for the class using a technological tool were of a geometric nature, so students could use the technology to explore and investigate theorems, properties, and definitions related to triangles. The activities and tasks for the class using non-technological tools utilized tools such as snap cubes, protractors, rulers, and pre-cut triangles. The tasks, which were inquiry-oriented, for both classes were almost the same, with the exception being the type of tools used by the students. Both the instructional sequence and objectives for the tasks were the same for both classes.

Data collected for the study included: video and audio recordings from both classes; video and audio recordings from the small groups selected within each class; video recordings of the computer screen while students used the technology for the three pairs of students in the technological tools class; and artifacts including copies of students' written work, homework, quizzes, and exams. Smith (2010) transcribed the video recordings and identified claims and then developed a description of the argumentation for each claim where components of Toulmin's Model of Argument were identified. He then identified ways in which the students used tools with each argument. Next, arguments were analyzed based on the explicitness of the warrants in the argument and how tools were used during the argument.

In general, the three pairs of students in the class that utilized technological tools "create[d] arguments of various structures and content including the ways in which the students use[d] the technology" (Smith, 2010, p. 408). Four major themes were identified from these three pairs of students: the explicitness of the warrant with the use of technology; the complexity of the argument with the use of technology; the use of technology as a means for collecting additional data; and the increase in the use of explicit warrants through the unit. The three pairs of students in the class that utilized non-technological tools also created arguments of various structures and content. Three themes that were identified from these three pairs of students included: the majority of the number of arguments made throughout the unit occurred in one task; the tendency to use explicit warrants when not using tools; and the use of tools to collect data. When looking across the cases, the students from both classes created arguments of various complexity and structures. Three themes that were identified from the cross-case analysis included: the technological tool students created more arguments than the non-technological tools students (179 arguments to 88 arguments); students from both classes were more likely to provide an explicit warrant when not using the tools; and students in the technological tools class created more arguments with a complex structure than students in the non-technological tools class. This study indicated that technology can be used as a tool for engaging student in argumentation and enhancing the type of arguments they create versus when not using technology. Therefore, technology can play an important role in the use of argumentation in small groups in the mathematics classroom.

In conclusion, although empirical research using Toulmin's Model of Argument is not as prevalent in mathematics education, there are some studies that have used the model (e.g., Conner, 2007; Smith, 2010; Yackel, 1997, 2001). These studies used Toulmin's Model of Argument for evaluation purposes (Conner, 2007; Smith, 2010; Yackel, 1997, 2001). Overall, Toulmin's Model of Argument has been shown to be an effective way of analyzing student arguments (Conner, 2007; Smith, 2010; Yackel, 1997, 2001). However, as evidenced by the lack of research on argumentation in mathematics education as compared to science education, there is a need for further research of argumentation in mathematics. This need for research of argumentation in mathematics education further extends for a need for research of argumentation in statistics education, which is nearly non-existent in the literature.

Chapter Summary

This chapter provided an overview of the literature related to the study. An examination of Vygotsky's social constructivist theory as well as how this theory influenced the current mathematics education reform efforts were described. Similarly, mathematical discourse and argumentation, along with research of argumentation in both science education and mathematics education, were described.

CHAPTER THREE: METHODOLOGY

Introduction

Argumentation provides a means for students to gain better reasoning skills (Billig, 1987; Aydeniz et al., 2012; Driver et al., 2000). As part of the recent history of the mathematics curriculum, reasoning and argument have been emphasized throughout mathematics curriculum standards (c.f. CCSSI, 2010; NCTM, 1989, 2000). One possible method for building students' understanding of argumentation for use in mathematics is through engaging students in the statistical problem-solving process.

The purpose of this qualitative study was to investigate students' engagement in the statistical problem-solving process as an avenue for developing students' understanding of argumentation. Specifically, this study sought to examine students' understanding of argumentation while being engaged in the four components of the statistical problem-solving process. In order to do this, an exploratory case study was used. The following research question guided the study: How does students' engagement in the statistical problem-solving process support students' understanding of argumentation, if at all?

This chapter will begin with an overview of the research design, which includes a description of the case study selection. Next, the chapter will include a description of both the context in which the study occurred and the participants. A report of the instruments used in the study will follow, along with the procedures for data collection and data analysis. Finally, a discussion of the limitations and delimitations of the study is provided.

Research Overview

The study utilized an exploratory qualitative case study, as defined by Yin (2014). In his text, Yin described a case study as "a study that investigates a contemporary phenomenon in depth and in its real-world context" (p. 237). He further explained that case studies should meet three conditions: involve how and why research questions; not require the control of behavioral events; and focus on contemporary events (Yin, 2014). Creswell (2013) described the type of problem best suited for a case study design as being one that seeks to provide an "in-depth understanding of a case or cases" (p. 104). In addition, one key feature of a case study, as noted by Eisenhardt (1989), is that case studies are "particularly well suited to new research areas or research areas for which existing theory seems inadequate" (p. 548). This case study was used to explore and understand how engaging students in the statistical problem-solving process in their small groups in an introductory statistics classroom could enhance their understanding of argumentation. Due to the limited research of argumentation in a statistics classroom, the case study methodology was appropriate. In this instance, an exploratory case study provided insight to inform future research investigating argumentation in an introductory statistics classroom.

Due to the desire to examine different types of students' understanding of argumentation while engaging in the statistical problem-solving process, a multiple case study approach was used. Of particular interest were students' levels of engagement within the small group structure (e.g., students who are normally active in group discussion, students who are normally moderately active in group discussions, and students who are normally rarely active in group discussions). In general, a multiple case study design is preferred over a single case design (Yin, 2014). The argument for multiple cases recognizes the replication logic in which replications of the original finding are then considered robust (Yin, 2014). For this case study, the units of analysis were the three individual participants identified based upon the frequency of their interactions within a small group. Since these cases were based upon a single unit of analysis, the multiple case design is of a holistic nature (Yin, 2014). Therefore, this methodological design of this study is best described as a holistic multiple case design with three cases.

Case studies do not always involve clean beginning and ending points, and sometimes these studies involve the evolving of ideas throughout the study, including research questions (Creswell, 2013; Patton, 2002). In this study, the original research question was: How does training in argumentation using Toulmin's Model of Argument support students' engagement in statistical problem solving, if at all? Prior to data collection, it was uncertain how argumentation would be used by the students during the Coke/Pepsi lesson. After examining the data, the researcher realized the design of the Coke/Pepsi lesson was presented in a way that engaged the students in the statistical problem-solving process as they worked through the lesson. Therefore, the research question needed to be reshaped to reflect this occurrence. Patton (2002) described how qualitative research can involve changes within the design of the study and after fieldwork is conducted. Therefore, the research question for the study was reshaped to reflect those changes that occurred. The revised research question was: How does students' engagement in the statistical problem-solving process support students' understanding of argumentation, if at all?

Research Context

This study took place at a public university in the Southeastern portion of the United States. According to statistics from the Spring 2015 semester, the semester in which this study occurred, the university hosted more than 21,283 students with approximately 33.5% minority students. Approximately 45.8% of the students at the school were male and 54.2% were female.

The course relevant to this study was the applied introductory statistics course. This course was described as consisting of descriptive statistics, probability, and statistical inference. Prerequisites of the course included both two years of high school algebra and a minimum mathematics ACT score of 19.

The university had a total of 19 sections of the introductory statistics course taught during the Spring 2015 semester. Of the 19 sections, four were unique in that they utilized statistics modules developed by statisticians and a statistics educator to promote active learning in the introductory statistics classroom. The introductory statistics course selected for this study was one of these unique sections.

The course instructor, referred to as Dr. Anderson (a pseudonym), was a professor of mathematical sciences at the university. Dr. Anderson was a statistician who taught primarily statistics classes. She had taught the introductory statistics course many times throughout her career. Just prior to this study, Dr. Anderson received external funding to support the development of active-learning modules for use in introductory statistics classrooms. As a result, she taught selective sections of introductory statistics that utilized these statistics modules.

Participants

Purposeful sampling was used to select the participants in this study. Patton (2002) defined the importance of purposeful sampling when he stated, “The logic and power of purposeful sampling lie in selecting information-rich cases for study in depth. Studying information-rich cases yields insights and in-depth understanding rather than empirical generalizations” (p. 230). Participants in the study were selected from students who were enrolled in the chosen section of the introductory statistics course and who had agreed to participate in the study. Three participants were selected to form one group for the study. These participants were selected based on prior observations in order to gain a sample of three participants as follows: one participant who was very verbal and active in group discussions; one participant who was moderately active in group discussions; and one student who was rarely active in group discussions. These contrasting cases were selected to show different perspectives of the process (Creswell, 2013; Yin, 2014). The three participants taking this introductory statistics course were completing the course as part of their general education requirement in mathematics or as a requirement of their specific program. They are described in Table 1 below.

Table 1

Participant Background Information

Name	Age	Gender	Ethnicity/Race	Major
Chelsea	20	Female	White	Pre-nursing
Bailey	18	Female	White	Biology with concentration in physiology
Megan	*	Female	*	*

Note. * indicates information was not provided by the participant.

Instruments and Data Sources

Based on Yin (2014) and Creswell's (2013) suggested sources of evidence for data collection, the following sources were used in this case study research: lesson documents, writing prompts, interviews, observations, and the researcher. In what follows, each of these sources of evidence will be further explained and connected to the study.

Lesson Document

The participants participated in a lesson titled "Coke/Pepsi Taste Test" (Garfield, delMas, & Zieffler, 2007) in which they designed and conducted an experiment (see Appendix A). The intent of the Coke/Pepsi Taste Test document was to engage participants in the statistical problem-solving process in order to observe their use of argumentation during this process. The participants' responses on the lesson document were in a narrative, short answer form. These lesson documents were collected.

Writing Prompts

Participants were asked to respond to reflective writing prompts three times during the semester (see Appendix B). Responses to these writing prompts were used to gain further insight into students' thoughts about the use of argumentation and their understanding of argumentation.

Interviews

An additional source of data for this qualitative case study were interviews with the participants (see Appendix C). After answering the interview questions, the participants then were shown three video segments of their group from the Coke/Pepsi Taste Test lesson. These three segments were selected based on their alignment with the statistical problem-solving process. The first segment was selected because it included the participants' conversation about the design of their experiment, which is related to the second step, designing and employing a plan for collecting data, in the statistical problem-solving process (Franklin et al., 2007). The second segment was selected because it included the participants' conversation about the selection of the test to use to analyze the data in their study they designed as a group. This was related to the third step in the statistical problem-solving process, select and use appropriate methods to analyze the data (Franklin et al., 2007). The third segment was selected because it included the participants' conversation about interpreting the *p*-value in relation to accepting or rejecting this null. This was related to part of the fourth step in the statistical problem-solving process, interpret the results in relation to the original question (Franklin et al., 2007). From these interviews, the researcher gained insight into the participants' own perspectives about their use of argumentation during the statistical problem-solving process and their understanding of argumentation. The researcher audio recorded and transcribed the interviews with the participants.

Field Notes

Observations occurred throughout the study. The researcher took field notes during these observations. Since the purpose of these observations was to become familiar with the classroom and participants' mannerisms, these observations were not video or audio recorded.

Observation Video

Since not all communication is verbal, the observation during the Coke/Pepsi lesson was video recorded. The purpose of the video recorded observation was to determine if participants used argumentation as part of their statistical problem-solving process when designing and conducting their experiment.

Researcher

In addition to the data collection methods described above, the researcher was also a key instrument for the study (Creswell, 2013). Patton (2002) described the qualitative researcher as an "instrument of both data collection and data interpretation" (p. 50); therefore, qualitative researchers must describe their background and experiences as related to the study. In this study, the researcher graduated with a Bachelor of Science in Mathematics with a Mathematics Education concentration and Secondary Education minor. After completing her degree, she continued her education and earned a Master of Science in Mathematics. During this time, she also began a doctoral program concentrating in mathematics education. As part of her doctoral program, the researcher completed coursework in qualitative research methods. She was also involved with a research group for two years in which she used both qualitative and quantitative forms of research. As a result, she was qualified to serve as an instrument in this study.

Procedures

Once IRB approval was obtained (see Appendix D), the researcher observed the participants in order to become familiar with the classroom and potential participants' mannerisms. The researcher then selected three participants for the study based on the observations. After the participants were selected, the participants responded to the first of three writing prompts (see Appendix B).

The researcher prepared the instructor of the introductory statistics class, Dr. Anderson, for the argumentation training during the first half of the semester. The argumentation training was informal and designed by the researcher based upon examples provided by Toulmin (1958). During the training, the researcher and Dr. Anderson looked at examples provided by Toulmin and identified the components of Toulmin's Model of Argument present. The researcher and Dr. Anderson then created their own examples and identified and discussed components of Toulmin's Model of Argument. Also during this training, the Coke/Pepsi Taste Test lesson was discussed. The researcher and Dr. Anderson talked through the lesson and discussed possible student responses in order to better prepare for the lesson.

Dr. Anderson then implemented the argumentation training and Coke/Pepsi Taste Test lesson in a different class for practice purposes. The researcher and Dr. Anderson met and discussed any adjustments that should be made to the training prior to implementing it in the introductory statistics classroom. For example, during the practice argumentation training, it was clear none of the students were familiar with argumentation. Although most students easily understood claims and data in this training, warrants were difficult for the students to identify and use. Therefore, the researcher and Dr. Anderson decided to spend more time having students identify warrants and create warrants in their own examples during the argumentation training. Similarly, during the practice lesson, it was discovered that time constraints would not allow for each group within the class to perform the experiment their group had designed. Therefore, only one experiment would be conducted during the class period. However, all groups would still design their own experiment as planned.

Dr. Anderson then engaged the class, including the three participants, in the argumentation training during a class period of their introductory statistics course (see Appendix E). During this training, the teacher provided participants with an example of an argument and explained the different components of the argument. The participants then created their own mock arguments where they identified each component. Participants then participated in live mock arguments with their group members in which they practiced using each component of Toulmin's Model of Argument. Although Dr. Anderson had focused on explaining and identifying warrants during this training, participants still appeared to have a difficult time identifying warrants during their mock arguments with their group members.

The class, which included the three participants, then responded to the second writing prompt immediately following the argumentation training. The researcher and Dr. Anderson met to discuss the implementation of the Coke/Pepsi Taste Test lesson during the week in which the lesson was implemented. During this meeting, a review of the previous practice lesson was again discussed. The researcher discussed the materials she would bring for the lesson implementation, including both Coca-Cola and Pepsi, cups with numbers on the bottom for easy identification, and an Excel spreadsheet for recording of student data. Dr. Anderson and the researcher also discussed the best location of the researcher for observing the participants during the lesson.

The class, including the three participants, then completed the Coke/Pepsi Taste Test lesson during a single 85-minute class period. Due to the knowledge of various statistical tests needed to complete the lesson, the Coke/Pepsi Taste Test lesson occurred near the end of the semester. While the participants were engaged in the lesson, the researcher both video recorded the participants as well as took observational field notes. A complete transcription of the lesson can be found in Chapter 4. At the end of the lesson, the participants completed the third writing prompt.

Following the Coke/Pepsi Taste Test lesson, the researcher interviewed each of the three participants separately using the interview protocol (see Appendix C). After answering the interview questions, the participants then were shown three video segments of their group from the Coke/Pepsi Taste Test lesson. These three segments were selected based on their alignment with the statistical problem-solving process. The participants were asked to watch and identify any components of argumentation present in these segments. Figure 2 presents a timeline of these procedures for the study.

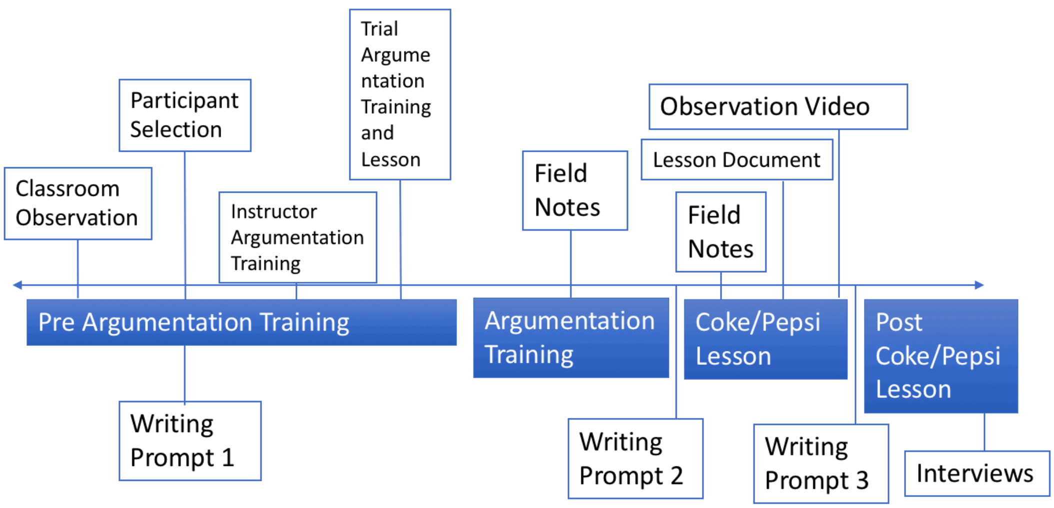

Figure 2. Timeline of the procedures for the study.

Data Analysis

Creswell (2013) described data analysis in qualitative research as consisting of five general steps. These steps are as follows: organizing the data; reading and memoing; describing, classifying, and interpreting data into codes and themes; interpreting the data; and representing and visualizing the data (Creswell, 2013). He further described the data analysis process involving these steps as occurring in a spiral form in which "the researcher engages in the process of moving in analytic circles rather than using a fixed linear approach" (Creswell, 2013, p. 182). Throughout the study, the researcher followed this approach for data analysis in which she moved continuously through each of the five steps.

The researcher prepared a case description for each case, which described how each participant interacted with the other participants and used argumentation when engaged in statistical problem solving. A cross-case synthesis was used to analyze the case study data. Yin (2014) defined cross-case synthesis as "a compiling of data for a multiple-case study, by examining the results for each individual case and then observing the patterns of results across the cases" (p. 238). He further suggested using a cross-case synthesis when "a case study tries to explore whether the cases being studied had replicated or contrasted with each other" (Yin, 2014, p. 167). Therefore, data collected from each of the cases was examined individually and then compared.

Initially, Toulmin's Model of Argument was used to analyze the discussion among the participants during the Coke/Pepsi Taste Test lesson. Coding of the participant interactions occurred while watching the video recordings of the Coke/Pepsi Taste Test lesson. The researcher watched the video and identified any components of Toulmin's Model of Argument that occurred, such as a claim, data, or warrant. The researcher then identified which participants contributed each component identified. Since data were collected from multiple sources for triangulation purposes and in order to determine the consistency of the analysis, the researcher then compared the results of the analysis to the observation notes from the lesson (Creswell, 2013; Yin, 2014).

An analysis of when argumentation was used in the lesson by each participant was then compared to the participant answers from the interview questions and their video analysis for consistency with the results. However, the interview was also used to determine each participant's view of how argumentation was used to see how their view of how they used argumentation compared to how argumentation was actually used during the lesson and to help determine each participant's understanding of argumentation. This allowed the researcher to better understand how the participants interpreted argumentation. The data from this were then compared to the participants' writing prompts from both before and after argumentation training and the lesson. This allowed the researcher to observe any growth in the understanding of argumentation that occurred.

After reviewing the detailed description of each participant that was formed from the analysis, participants' responses in which they described argumentation were then categorized and sequenced in a way to demonstrate advances in their thinking. These sequenced categories were then used to form the levels of understanding of argumentation (see Table 2).

Table 2

Levels of Understanding of Argumentation

Level of Understanding	Description of Understanding
0	Argumentation is viewed as an argument between or among one another. Argument would be defined as "an angry quarrel or disagreement" (Merriam-Webster, 2017, p. 1).
1	The components of argument are described, but the terminology of argumentation is not included in these descriptions. The participant has moved beyond seeing argumentation as arguing and moved toward viewing argumentation in terms of Toulmin's Model of Argument.
2	The components of argumentation are described using proper terminology. However, the participant still inconsistently or inaccurately recognizes the components in action.
3	The components of argumentation are described using proper terminology. The participant can consistently and accurately recognize the components of argumentation in action.

The cases were then reviewed, and each participant's understandings of argumentation were coded based upon these levels of understanding argumentation. Codes were assigned at five instances in the study: prior to argumentation training; after argumentation training; during the Coke/Pepsi Taste Test lesson; after the Coke/Pepsi Taste Test lesson; and after the interview.

If a participant's understanding of argumentation did not completely meet all of the requirements of one level or met requirements of understanding within two levels, the understanding was coded accordingly. For example, if a participant described argumentation as a disagreement, but also related argumentation to the statistical problem-solving process, the participant was coded as a Level 0.5. This coding is reflective of an understanding of argumentation that is similar to a Level 0 in that argumentation is still viewed as a disagreement between two people, but also shows signs of viewing argumentation as a model, a part of a Level 1 understanding of argumentation, as evidenced by the participant's relation of argumentation to the statistical problem-solving process. Similarly, a participant who described the components of argumentation and used proper terminology for most, but not all, of the components in her description would be coded as a Level 1.5. This coding is reflective of an understanding that is similar to a Level 2 in that the components of argumentation were described and proper terminology was used in the description. However, it is also similar to a Level 1 understanding because some of the description of components did not include proper argumentation terminology. Therefore, a coding of 1.5 would be used to reflect this type of understanding.

Limitations/Delimitations

There were five limitations and one delimitation present in this case study research. The first limitation was the decision to limit the participants' engagement in the statistical problem-solving process by not having the participants conduct the study they designed themselves and, instead, conduct a study designed by a different group along with the rest of the class. This decision was made after implementing the practice lesson. The instructor and researcher realized there would not be enough time during the 85-minute class period to have all students conduct the study their group designed. Not having the participants conduct the study they designed during the Coke/Pepsi Taste Test lesson could have eliminated potential use of argumentation by the participants.

The second limitation of the study was also related to the time constraints of the class period. In order to ensure the participants had the opportunity to engage in all four steps of the statistical problem-solving process, the instructor reviewed various statistical tests during the Coke/Pepsi lesson. The review of these tests could have also eliminated potential use of argumentation by the participants that could have occurred when researching statistical tests.

The third limitation of the study was related to the length and detail of the interviews. All three participants had busy schedules during the semester of the study. Each participant could only commit approximately 15 minutes to the interview with the researcher. Therefore, the interviews were rushed and did not contain as much detail and questioning that could have occurred had more time been available.

The fourth limitation of the study was the delay in Megan's interview with the researcher. Both Chelsea and Bailey participated in their interview with the researcher two days following the Coke/Pepsi lesson. Although Megan was scheduled to also have her interview that day, she was absent. Due to other obligations that Megan had, she was not able to participate in her interview until two weeks after the Coke/Pepsi lesson. This delay in time caused her to have a harder time recalling events from the Coke/Pepsi lesson.

The final limitation of the study was the lack of demographic information provided for Megan. The researcher repeatedly sent emails to Megan for this information. However, no response was ever received from this participant.

There was one delimitation of this case study. This delimitation was the selection of a college-level introductory statistics class for the study. This selection, along with the use of a case study, does not allow the results to be generalized. However, the study is generalizable to theoretical propositions (Yin, 2014). Since this study focused on engagement in the statistical problem-solving process as a possible venue to enhance students' understanding of argumentation, the final report of the study is a description of how the participants used argumentation during the statistical problem-solving process.

Chapter Summary

To determine how engaging students in the statistical problem-solving process supports students' understanding of argumentation, an exploratory qualitative case study approach was used. This study involved a holistic multiple case study design with three cases. The three participants served as the three cases in the study. These participants were part of an introductory statistics class. Data collection methods, which included a lesson document, writing prompts, observations, and interviews, were discussed in relation to the study. The Toulmin Model of Argumentation was described and coding and data analysis methods were further discussed. The chapter concluded with a description of limitations and delimitations of the study.

CHAPTER FOUR: RESULTS

Introduction

The recent history of mathematics education curriculum standards has placed an emphasis on both reasoning and argument in mathematics classrooms (c.f. CCSSI, 2010; NCTM, 1989, 2000). One way to help students gain better reasoning skills and develop arguments is through the use of argumentation (Billig, 1987; Aydeniz et al., 2012; Driver et al., 2000). In order to properly use argumentation, students must first be introduced to argumentation. Therefore, a possible way to develop students' understanding of argumentation in mathematics is through engaging in the statistical problem-solving process.

This qualitative case study sought to investigate students' engagement in the statistical problem-solving process as a possible way to develop students' understanding of argumentation. The research question for this study was: How does students' engagement in the statistical problem-solving process support students' understanding of argumentation, if at all? The study utilized an exploratory qualitative multiple case study, as defined by Yin (2014) and included three participants. Instruments and data sources for the study included: lesson documents, writing prompts, interviews, observations, and the researcher. In order to analyze the data, Toulmin's Model of Argument and codes for the levels of understanding of argumentation, which emerged from the individual cases, were used.

This chapter will begin with a description of the Coke/Pepsi lesson. Next, data from each of the three individual cases will be presented. Last, a cross-case analysis of the three cases will be described.

Coke/Pepsi Lesson Description

The argumentation lesson occurred in an introductory statistics classroom. The participants in this study sat in a group of three with their desks in an L shape. The remaining students in the classroom also sat in groups of three or four with their desks pushed together. The camera focused solely on the group of three participants as seen in Figure 3. The researcher sat near the participants.

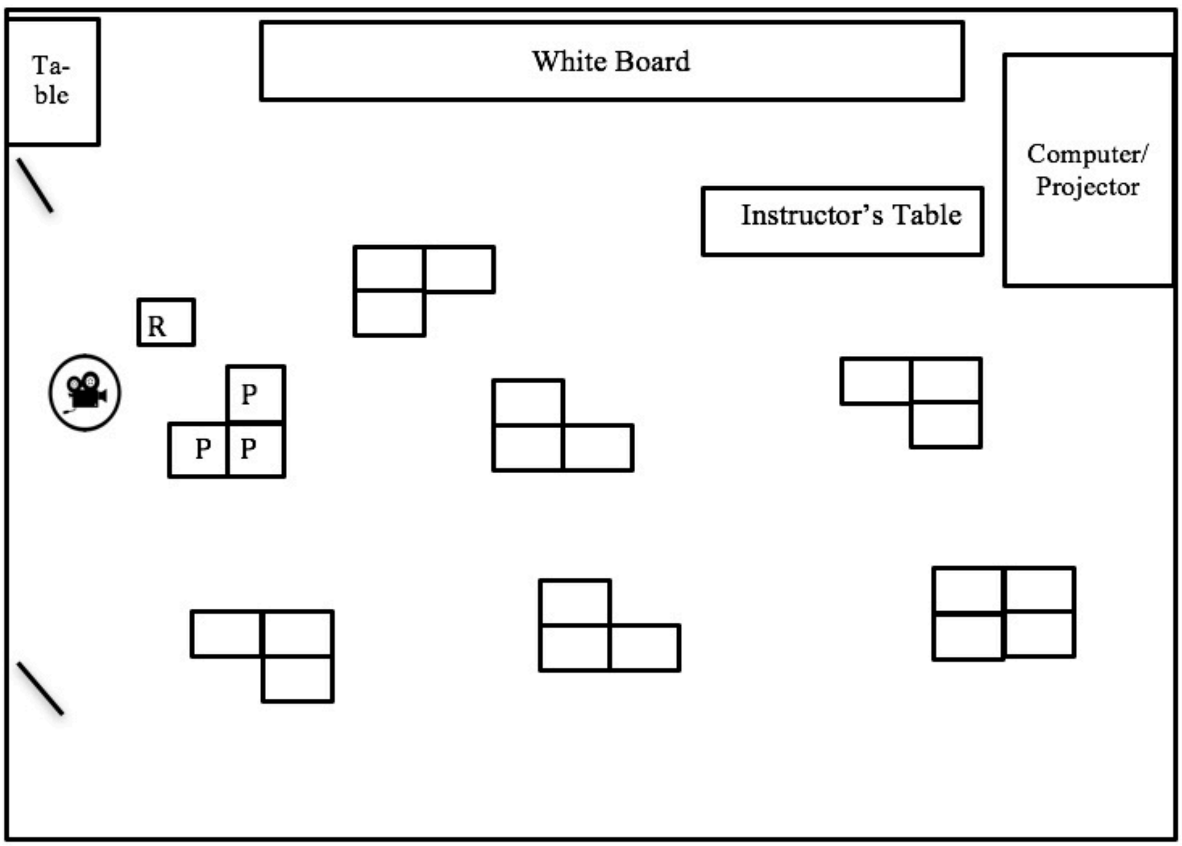

Figure 3. Classroom layout during the Coke/Pepsi lesson.

In what follows, a detailed narrative of the Coke/Pepsi lesson focused on the dialogue within the group of participants is provided. Within this description, the beginning and end of each of the three selected video excerpts viewed by the participants in the individual interviews are marked.

At the onset of the lesson, the instructor, Dr. Anderson, expressed enthusiasm for the lesson as she provided an overview of the day's activities. She introduced the lesson by saying,

> Today is really going to – really be a good day because you have been learning lots of statistical techniques throughout the semester and today you get to do what a statistician does. One of the things is, we are going to design a study. You are going to collect the data yourself, you are going to analyze the data, and you are going to write your conclusions. So, you are going to do the full process. So it's really a good day. And you make the decisions about what gets done. So, umm, part of our goal is that engaging you in that full statistical problem-solving process. So you get to be the statistician today and make the decisions. So to warm you up for that on being the statistician, I would like for you to talk in your groups about what is important to do when solving a statistical problem. Ok, so you talk to your peers and then we will report out.

Within the group of participants, Chelsea initiated the conversation by saying, "Collect a reasonable amount of data." Then, Chelsea and Bailey shared their ideas.

> Chelsea: From a random sample.
>
> Bailey: Yea. And then you have to figure out the right test. That's the hard part.
>
> Chelsea: The significance level.
>
> Bailey: And what kind of test you are going to take. Remember when she made that chart.

Chelsea: Yep. So what kind of test to use.

Bailey: Uh huh.

Chelsea: The sample.

Bailey: [inaudible]

Chelsea: And then at the end it's important to decide if you reject the null.

Bailey: And then your interpretation is important too.

During this conversation, Megan leaned back in her chair and looked at the other two participants. The two participants finished their conversation, mentioning the audio recorder sitting on their desks. After a brief pause in conversation, Dr. Anderson walked over to their group. As she listened, Chelsea and Bailey continued with their conversation about what is important to do when solving a statistical problem.

Chelsea: Oh you have to make sure you don't make a Type I or Type II error.

Bailey: Yea and then figuring out the kind of type – the type of error you got is confusing though.

Chelsea: You mean whether you rejected the null and it was true or accepted the null and it was false.

Bailey: Yea. Uh huh. And especially when it's things like with the coins, you can tell, like, with the test if it's equal the null. But with some tests, it's just like we don't even know if we got it right or wrong. Which is weird.

Chelsea: Right.

Bailey: What are we going to be doing today, like, with the Coke and Pepsi stuff because I guess the most important stuff in doing a statistical study is having a clear question.

Chelsea: Yea.

Bailey: Like what are you (pause).

Chelsea: What are you looking for? Or testing.

Bailey: And making sure that a question is not, umm, worded in like a [inaudible] manner.

Chelsea: Yea.

Dr. Anderson called the class back together. “Ok. It looks like everyone has come up with lots of really good things here. So, umm, what was the first thing that – this group, where you said you had to start with?” A student from one of the other groups in the classroom answered quietly. Dr. Anderson replied, “She said it quietly. She said knowing what you are looking for. Is there a different way to word that? I mean it’s perfect. I mean I understand exactly what she means.” A different student from the class answered to which Dr. Anderson replied, “You might need to know about the population you are talking about. I was thinking about something else. Knowing what you are looking for. Understanding your research question?” A different student in the class responded. The instructor replied,

> Ahh, which is the alternative hypothesis. Very nice, very nice. [Student] what was your first thing? It follows very nicely. Knowing what you are looking for and then, (slight pause) collecting your data to back it up. She was listening last class period, don’t you think? I think so. Ok, let me try this group here. What did you have?

A student responded and Dr. Anderson continued.

> So you moved to the statistical test. So you collect your data, you've done a statistical analysis and you have got some information out from the statistical analysis, and you are going to use that to make a decision. What would you all add after that?

A student again responded.

After this student's response, Dr. Anderson then specifically called on Chelsea, Bailey, and Megan.

> Oh my goodness, the statement that connects the data to the claim. Oh plus one. What do you think? Was she listening last class period? I'm going to leave it at that. Ok, this group [Chelsea, Bailey, and Megan]. So let's say that you have already done all of those things. What else might you think about?

Chelsea and Megan both immediately looked at Bailey. After a brief pause, Bailey leaned back in her chair and responded, "Interpreting our results." Dr. Anderson replied, "Well, of course. Interpreting the results. That is very good. I didn't hear you say that one, but that is very good because that would be the perfect next step." Bailey added, "Oh and making sure you didn't make an error. A Type I or II."

Dr. Anderson continued with the class conversation.

> Analyzing your Type I and your Type II error. Very nice. Congratulations to all of you as a class. You have completed that full cycle, and that is what we will do today. And since you brought it up, let's go ahead and just remind you about the argumentation model we talked about last time, which was exactly what you said. It started with the data (begins drawing model on board) you wanted to use to

make your claim. And this is that simple model. We will just keep that. And then who remembers what it's called when you link those two together?

The class responded "warrant" as a whole. Dr. Anderson continued, "Oh everybody. So we said it's pretty easy to understand the data in statistics, right? And the claim is the research hypothesis just like you said. What is the warrant in there?" There were mumbles among the students about connecting data and claim, so Dr. Anderson continued.

It does connect them, but what is the statistical step for it? The test that's right. And more so than just the test, you have to make sure you get the correct test. So selecting the correct test, and then of course that whole process of using it correctly to analyze. Just like you all said. And then interpreting it so that it makes sense. It's very nice. Very nice job. One of the things you will have to do today that is really important is selecting the correct test. So what tests have we done, statistical tests, for inference we should say? Statistical inference [tests] we have done so far where you want to make an inference about a population, but you use your sample. So if you could quickly name them. There are sort of four approaches for them, and two different types of test that you have seen so far. Ok.

There was a brief pause and Megan searched through her bag. Dr. Anderson continued by saying, "Ah, that one was harder. You had the other process perfect. So look some things up. What are the different tests you have done so far. Could be more, but actually I wrote a bunch on the board." Megan and Chelsea both took out their notes and read them. Bailey then initiated the group conversation.

Bailey: So like one-sample, two-sample.

Chelsea: [She reads from notes.] Mean tests, proportions or, ah, *t*-test, *z*-test for proportions. Umm, a two-sample *t*-test.

Dr. Anderson: And while you are naming them the important thing is when do you use each one. So test and when do you use it.

Chelsea: You use the *t*-test for mean.

Bailey: For one-sample.

Chelsea: Um hmm. You use a one-proportion *t*-test for proportions of one sample. Use a two-sample *t*-test for a two-sample mean. Umm, and then you've got - you do two proportion *z*-test for proportions of two samples.

Bailey: And then more than one-sample you use chi-square for proportions.

Chelsea: Do what?

Bailey: Chi-square for proportions. (She points to it in Chelsea's notes.)

Chelsea: Chi-square? That's x-squared.

Bailey: Yea, but it's called chi.

Chelsea: Oh, it's called chi. Oh. Sorry. I though you said pie. (laughs).

Dr. Anderson interrupted to announce to the class, "You have permission to pull out your phone to check your pictures. That was a board that a lot of people took pictures of. The other way is looking in your calculator for the list of all those tests and scroll way down." She then walked over to Chelsea, Bailey, and Megan.

> What did you all come up with? Oh, very nice notes. So everyone is sort of mentioning tests out loud. What corresponds to the test? There is hypothesis testing, and then there is something else that is a partner with it.

She then walked away from this group to question another group. Bailey continued the conversation by saying, “Umm, confidence intervals?” Chelsea responded, “Maybe.” Bailey and Chelsea looked through their calculators, and Megan read her notes. After a brief pause, Megan looked at Bailey.

Megan: You were right.

Bailey: Hmm?

Megan: Confidence intervals.

Bailey: Oh.

Dr. Anderson brought the attention of the class together again and asked, “So what are the statistical tests that we have done so far?” A student from the class responded.

Dr. Anderson: Very nice. A one-proportion z-test. What in the world is that for?

Chelsea: It was one-sample mean?

Dr. Anderson: Oh, and there is a one-sample mean. Does that go under the list of statistical tests?

Chelsea: I thought you were asking what the one-proportion z-test was for.

Dr. Anderson: I am. I am sorry. So could you answer that question?

Bailey: [speaking quietly to Chelsea] One-sample proportion.

Chelsea: Oh, it’s z-test. I am sorry. So it’s one-sample proportions. I looked at the wrong thing. My fault.

Bailey: [inaudible] (to Chelsea)

Chelsea: I thought she said t.

> Dr. Anderson: This class is a great team-working class aren't they. Everybody helps everybody out, and as long as you don't do that in the middle of a test we are ok.

Dr. Anderson continued with the group conversation and asked, "So this is for one-sample proportions. So what does the null look like?" A student from the class responded. The instructor continued,

> *p*. Oh, well. You have to be careful. That would be the case if there were no preference, but *p* equals some value, *p*-naught, which is given somehow in the problem. It's the baseline measure. I believe that 75% of you want to make an A in this class. So the proportion of people that want to make an A in this class is 75%. I'm guessing that you might tell me that the alternative to that is what? Greater than, less than, or not equal to?

A student from the class responded, "not equal to." Dr. Anderson replied, "Well you could, but I figured you might say greater than 75% of you might want to make an A [students laugh]. But you are right. If I don't tell you, which one should you pick?" A different student from the class answered, "not equal to." Dr. Anderson continued, "Not equal to. So this is some specified value that we call the hypothesized value. Ok, great. So what's the next one? There is a plain old *t*-test. And what is it for?" A student answered "one-sample means." Dr. Anderson replied, "One-sample means. Ok. We know two more. Oh, so each one of these tests has a corresponding what? A partner. Yea I heard it." Chelsea answered, "Confidence interval," and Dr. Anderson replied, "Right. So you could have a test, or on the calculator it's called like the *t*-int or interval depending on how long the word is. But the confidence interval is the corresponding one.

So then what's the null for means?" There was a brief pause. Dr. Anderson continued, "We have done it on the homework. We did it in class. We entered the data for one last time. Megan, what was it?"

Megan: I didn't say anything.

Bailey: It could be equal to.

Dr. Anderson: It's going to be equal to. Bailey says it's going to be equal to. Very good. She took the easy part.

Bailey: Mu.

Dr. Anderson: There is a mu. Means. Population mean. Equal some hypothesized value. All right. Now, Megan, how many choices do I have for my alternative?

Megan: Uhhh, three.

Dr. Anderson: Yea. And so it would be greater than, less than, and not equal to. Pick one. You only get to pick one. All right. There are some more. What are they? These were one-sample. What are the others? [Student from class responds two-sample.]

Dr. Anderson: There is a two-prop, thank you, *t*-test. And they call it on the calculator a two what with a *t*? (Pause). I think they call it two-sample *t*-test. And each of those go with confidence intervals, and they are basically saying that you use them instead of when there is one-sample when there are two. (Pause). [Instructor wrote each of these tests and when to use them on the board at the front of the classroom.] Ok, so we are ready to develop a test. You should have those in your mind. And this, if you are guessing, there is some Coke and there is some Pepsi over there. You are writing. Let me give you a second. [Brief pause]

Ok, so did everyone have a chance to do that now? All right. So what I want to look at for today, in other words, what is the claim or research question? We are going to state it in the form of a question instead of directly as a claim. So what I want you to do is design a study in your groups. And you may all come up with different answers so don't, umm, just focus in your group as you design a study. And what you want to do is design a study to determine if people could correctly identify a drink as either a Coke or Pepsi [she wrote this statement on the board]. So you want to design a study to determine if people could correctly identify a drink as either a Coke or Pepsi. So let me let you think about that along with everything we have thought about for the semester to kind of pull it together and what would be a good way to design a study that would be effective. And that's the key at answering this question. You will get a little bit longer to talk as you need it. Ok, so get started. You probably need to take notes on this because I will call on you for that. But they may not be perfect the first time, so you may have to modify them as you go.

Each of the groups worked individually again. Chelsea initiated the conversation for her group's discussion.

Chelsea: So the claim is people can correctly identify a drink as either a Coke or a Pepsi? And then you have to decide your hypothesis and your null?

Megan: Yea, so right now what we are doing is deciding how we would go about designing a study.

Chelsea: Oh ok.

Bailey: So [inaudible].

Megan: You have to have people. Your sample.

Bailey: Yea, and I'm assuming everybody in this class. Like 1. So our A would be [inaudible]. So we would need two samples. (She counts students in classroom.) Sixteen. So 16. Then our sample is 16. That's, like, very small.

Chelsea: Yea.

Bailey: Given our circumstances, 16 is probably going to be our sample.

Chelsea: Mm hmm.

There was a brief pause. The instructor then walked over to the group.

Bailey: (to Dr. Anderson) So we are trying to figure out how we are going to (pause).

Dr. Anderson: Yea, so you are going to need more time to talk.

The instructor then walked away. The group conversation among Chelsea, Bailey, and Megan continued.

<Interview Video Clip 1 Begin>

Bailey: I was going to ask one question.

Chelsea: Do we need to change the significance level?

Bailey: I don't even know what we are really doing. Like, what is our specific question. Like, it could be do people prefer Coke or Pepsi.

Megan: If they can identify.

Chelsea: If they can tell the difference between them. Like, if you don't have a label, can you tell if this is Coke or Pepsi.

Bailey: Oh ok. So we are assuming the null would be that they can or they can't?

Chelsea: I would say that they can't.

Bailey: I mean because there is like a 50-50 chance that they would get it right or they would get it wrong if they can't tell the difference.

Chelsea: But I think you would say equals.

Bailey: Yea.

Chelsea: And if it's equal, they can't identify. And then, greater than, your research hypothesis would be greater than 0.5.

Bailey: But what did you say earlier that we need to have?

Chelsea: Significance level.

Bailey: Yea. So.

Megan: But aren't we just supposed to be deciding how we would do a test though? We aren't supposed to be finding all of this, we are literally just.

Chelsea: But what we are doing - we are just kind of saying to design the study we would have (brief pause). Are you saying, like, how we would test it, or are we supposed to do that?

Megan: Yea. Yea.

Chelsea: Ok. I thought we were supposed to.

Megan: I thought we were going to be conducting a test. Is that what you were trying to say?

Dr. Anderson walked over to the group during their discussion. Megan's question was directed toward her. Dr. Anderson responded, "You are doing it. Yes, you are designing an experiment. And you need all of those details or else you won't be able to do it. Excellent." She then walked away from the group. Bailey continued the group conversation.

Bailey: So each individual will need two cups.

Chelsea: Ok my fault. I was confused.

Megan: It’s ok.

Bailey: So you randomly decide if you drink the Coke first or the Pepsi first?

Megan: Like, we could say we go on campus and have, like, a booth for people to come up and try it and see if they could.

Chelsea: Tell if it’s Coke or Pepsi.

Megan: Yea.

Bailey: So set up, like, a booth.

Megan: Set up the booth on campus or something like that.

Chelsea: Like a lemonade stand.

Megan: Yea.

Bailey: We could flip a coin to decide if they will be trying Coke or Pepsi.

Megan: Or do you make them try both?

Bailey: Both?

Megan: Yea.

Chelsea: Yea I would make them try both and label this one is Coke or this one is Pepsi.

Megan: You just put numbers, like one or two, and then you know which number is which, but they don’t.

Chelsea: Assign a number to a product.

Megan: Yea.

<Interview Video Clip 1 End>

There was a brief pause in the conversation while Bailey, Chelsea, and Megan all three recorded their ideas. Dr. Anderson walked over to the group and said, "I came at the wrong moment. There's no conversation." Bailey responded.

Bailey: Oh, because we were ready [inaudible].

Chelsea: We were writing.

Dr. Anderson: I know.

Megan: We said that we would set up a booth on campus and then assign a number to a product. And then let them taste test it and tell us which is which.

Chelsea: Taste both.

Dr. Anderson: Ok. So what are you giving them and… (pause)

Megan: We would give them one of both. And one would be number one and one would be number two. And they won't know, but we will know what it is obviously. And they would tell us which is which.

Dr. Anderson: So while you are thinking about this, you might think about what you are testing, the null, the alternative, just going through that whole process.

Dr. Anderson then walked away from the group and their conversation continued.

Megan: So now back to what you were saying.

Chelsea: So the claim is that people can tell a difference?

Bailey: If it is we can, then it would be greater than?

Chelsea: Umm hmm. Your research would be.

Bailey: Yea, you can tell a difference because [inaudible].

Megan: Wouldn't the claim be that is there a difference?

Chelsea: Mmm hmm.

Megan: I mean I guess that makes sense. I just would word it different than it is right now.

Chelsea: How would you put it then?

Megan: I would word it like you're saying if people can tell a difference, but you wouldn't know that unless you have already done it.

Chelsea: But you are researching. So, like, your research hypothesis.

Megan: Ok.

Chelsea: No, I'm just confused what you are saying [laugh]. So if you are saying people can tell a difference, your null is equal to 0.5. And your research is greater than 0.5.

Megan: Ok.

Chelsea: If you wanted to word it where people couldn't tell a difference, your null would be still equal to and your research would be less than 0.5.

Megan: Correct.

Chelsea: Or would it be?

Bailey: If they can't tell, then would it be (pause). This is confusing me. I don't know why.

Megan: I need to go back and read my notes.

Bailey: How did you get number values for yes and no?

Megan: Fifty-fifty chance.

Bailey: So, if like somebody got it right, you would put theirs down as one? And if you got it wrong you would put no as a zero?

Megan: No idea.

Chelsea: Wait, what?

Bailey: Like, if, ok. So let's say you gave somebody two cups and you are like is it one or two, which one is Coke. And then they take the second one and they say this one is Coke and they were wrong. How would we put that in numbers?

Chelsea: You just have columns of yes and no, and if you get it wrong you put no.

Bailey: Oh. So, like, the proportion of the ones that got it right, and then proportions of ones that got it wrong.

Chelsea: Yea. I mean, that's all I would do because I don't know what else.

Megan: Yea, that's the best way to do it.

Chelsea: So what was your claim?

Megan: I would say, like, is there a difference between Coke and Pepsi instead of people can tell a difference.

Chelsea: But how would you (pause).

Bailey: Because they are not the same product. So we already know there is a difference. The type that she has says to determine if people can identify a drink as either Coke or Pepsi.

Chelsea: Ok.

Bailey: So I think we have to make it center around people and not about the actual products.

The instructor walked over to the group while Bailey was speaking.

Chelsea: Ok.

Bailey: And so do you [inaudible]. And so our null would be that (pause).

Chelsea: p equals 0.5.

Dr. Anderson: And what does *p* stand for?

Chelsea: The proportion of people that can identify, or the proportion of people that… (pause)

Bailey: Correctly identify.

Chelsea: Ok.

Bailey: Is that what *p* would be in this situation?

Dr. Anderson: That's what I am asking you.

Bailey: So our claim was people can tell a difference between Coke and Pepsi. So our null hypothesis would be that the proportion of people that can tell a difference would be greater than 0.5.

Chelsea: But I thought that was the research.

Bailey: Oh yea.

Chelsea: So the proportion of people correctly can (pause).

Bailey: Oh yea. So would it be that (pause).

The instructor walked away while Bailey was speaking. The conversation among the group members continued.

Chelsea: So you are saying that people can tell a difference between Coke and Pepsi. So that's your research - so that's. The proportion of people that correctly identify would be greater than 0.5. Yea. Because that means people can definitely tell a difference. If it's significantly greater than 0.5.

Bailey: So then our research would be greater than 0.5 for people who got it. The proportion of people who can tell a difference is greater than 0.5, because if it was

0.5 then it means they just blindly chose. They didn't know what they were talking about.

Megan: So it's greater than why? I'm not saying it's not, I'm just wondering because I'm confused.

Bailey: Because instead of it being - like 0.5 is no preference. So if they just randomly chose one, and it could have been either one. They didn't know. But then if you can tell a difference, then that means there is something guiding your response. So it's not like you just chose something, like heads or tails [inaudible].

Chelsea: I think that would be so much, or this would be so much easier, if the claim or the question was do people prefer Coke or Pepsi. I thought the wording is confusing.

Bailey: Oh, and then part of our study we didn't clarify is what we are going to actually be asking them. Like, we are going to give them two cups and then (pause).

Chelsea: Can you identify which is Coke?

Bailey: So we are actually asking them (pause). Ok. But I'm assuming we are not going to make the, like, booth [inaudible]. Like, if we actually left the classroom and made a little booth.

Chelsea: So now we need to pick an alpha level. So for a 95% confidence interval, you pick a 5% significance.

Bailey: So we are choosing 95? I mean, obviously, because that's what we usually do.

Chelsea: Well, yea, that means that alpha would be .05.

Bailey: I think that's good.

Chelsea: (to Megan and pointing at a place in her notes) So we did this, and then that's this.

Megan: Ok.

Chelsea: So you just then say .05.

There was a brief pause while Megan, Bailey, and Chelsea wrote. Dr. Anderson returned to the group members. Bailey continued the group conversation when she finished writing.

Bailey: So our confidence levels are [inaudible].

Chelsea: Ok. Now we have to collect data, which we can't do.

Dr. Anderson: What's your experiment about again? There's a booth.

Chelsea: Set up a booth, and you give people both Coke and Pepsi that's not labeled. It's just in a cup numbered one and two. And we tell them to identify.

The instructor walked away. Megan asked a question to the group.

<Interview Video Clip 2 Begin>

Megan: Did we pick 95 because that's what we always do?

Chelsea and Bailey: Yea.

Bailey: I mean we could pick 90 because it's not like a life or death situation, so who cares.

Chelsea: I say just go with 0.5, or wait 0.05.

Bailey: Collect data, so what test do we do.

Chelsea: Umm. I don't think it's two-sample.

Megan: Yea I think it would be one-prop *z*.

Chelsea: Yea.

Bailey: Test or interval?

Chelsea: Test.

Bailey: So what's the difference between *t*-test and *t*-interval?

Chelsea: If you are trying to find the interval you would do the (pause).

Megan: Like when we were doing the height stuff, and when you are getting the range.

Bailey: And so wouldn't the test do (pause). I feel like I should probably know this, but I don't.

Chelsea: That should be the *p*-value.

Bailey: Ohh. Ok.

<Interview Video Clip 2 End>

Dr. Anderson walked to the front of the classroom and called the class together again for a class discussion.

> I think you have done a great job. I have heard lots of incredible things that I need for each other to hear, for you to hear each others', so we can sort of see what we are going to do as a class. So let me start with the group in the back. Who is the spokesperson back there?

Even though they were not in the group in which Dr. Anderson called, Chelsea turned to Bailey and said, "You be the spokesperson." A student from the group in the back responded. The instructor called on a different group, and they responded as well. Dr. Anderson then continued in reference to the group's response.

And tell them why you are going to have two trials for it. [A student responded.] We are a small class today and we are going to have some people that are going to help pass out. And they counted and n is 16, so she knew the sample size was going to be too small. So she was going to double it.

Dr. Anderson then called on the next group, and they responded as well. The next group on which she called was Bailey, Chelsea, and Megan's group.

Dr. Anderson: Ok, so what are you going to do?

Bailey: We said that we are going to set up a booth somewhere on campus, and we are going to hand each person two cups, one that has the number one and one that has the number two. And we are going to ask them which one was Coke and which one was Pepsi.

Dr. Anderson: Ok, so they brought up something really important with that booth on campus. She has helped me identify, whether she knows it or not, the population. So what is our population?

Several students in class: [University] students.

Dr. Anderson: Ok, so the population is going to be [University] students. And this group over here, they have helped me identify the sample. What's our sample? (pause) The sample that we are going to actually do the test with? She thought maybe to get it larger, we could go down the hall and start asking people. And we could, but we don't have IRB approval for that. So the sample is our class. All right. Very nice. That's really important. Ok, this group. What was your experiment?

The next group described their study to the class. Dr. Anderson responded, "So the two cups again. And there are lots of great ideas. I am really pleased that you thought of the details about making sure that people don't know which is which one way or another." A student raised their hand and asked Dr. Anderson a question. Dr. Anderson answered,

> Very good. So because we are doubling the sample size, she said ideally we need to use crackers to cleanse your pallet. And, in fact, every one of you who had two cups needs the same thing, the crackers to cleanse your pallet, right? Because you need to make sure you don't get the taste all mixed together in your mouth and then you can't tell. That's exactly right. Very good.

A different student asked Dr. Anderson a question. She responded, "There is multiple ways to do this, but the random is really nice because why?" Chelsea's hand was raised, so Dr. Anderson called on her to answer the question. Chelsea responded, "I have another question." Dr. Anderson replied, "We'll, let her finish up and then you go. Ok, go ahead. Why would you do random instead?" The student responded. Dr. Anderson replied,

> You were right. Oh, so don't doubt yourself. Randomization is critical to taking care of the other factors, but also that way - somebody else tell me why. Just to increase our sample size today we are going to do two, right? That was the only reason she was doing it. Then, how is one cup better than two in the experiment? How could it be better than giving them two?

The student again replied. Dr. Anderson continued, "All right, and so they can get to taste both of them in your head and discern, where with one you don't even have any comparison." Another student asked a question, and Dr. Anderson continued.

You could do that. So yea, you have done a great job. There are lots of different experiments. We are going to need to pick one so that we are all on the same page. And, umm, I think, yes, that we are ready for this. So you can sort of transfer some of the information that you have already written right onto this sheet, but let's pick one together. And I actually get to help. Because of the timing and getting everybody to be able to participate, we sort of have to decide a little bit of something in advance. And so we are going with this group's model of one cup each, but we will replicate because your sample 16 is so small that we will do that. Now, Chelsea, you had another question.

Chelsea: Would you, like, in a test like this, would you want to I guess ask - sample people individually so that don't have the influence of the person. Would that make a difference in your data?

Dr. Anderson: Yes, it could, but hopefully you are going to, if you do the two cups, what are you going to have to do? Can you always put the Coke on the right?

Chelsea: No.

Dr. Anderson: What do we call it what you are going to do?

Bailey: Alternating.

Dr. Anderson: Well, would you alternate it or would you randomize?

Chelsea: Randomize.

Dr. Anderson: You would randomize. Yes. So, and yes, you want to make sure, they blindfolded to make sure no one would see what was going on with anybody else. Same kind of idea. Very nice. So go ahead and specify your null and alternative up here, and we want to erase the board. And we ask you what statistical test you will use, and then we will get ready to collect the data. Umm, and I will pull the screen down, so I will let you have a moment to write all of that. (Pause). You have it in your notes which test to use when. (Pause). So just one more time to make sure that we are all on the same page. [Student] can you tell them one more time what we are going to do?

Dr. Anderson began a review of the study the students conducted based on one of the group's study design.

Student: We are going to have two trials and within the trials you will be given a cup at random.

Dr. Anderson: So we are going to give each person one cup because we are just pretending like there are two of you. So we are not going to mention the second trial at the moment. So give each person one cup and that cup has what?

Student: Either Coke or Pepsi.

Dr. Anderson: Ok. And then what are we going to do? What are we going to ask them?

Student: We are going to ask them which they thought it was.

Dr. Anderson: Ok and then? Ok. Coke or Pepsi. All right and then what?

Student: We need to record.

Dr. Anderson: Record the answer. Ok. Guess what? We have a sheet for that already prepared. And then what? And then she said, because we are doubled, we are going to need crackers. Do we have any crackers? No, we don't have any crackers, so we will pretend you have crackers. So you will record your answers and then you will analyze the data, but we are going to repeat this twice just so we can get a little more data. So analyze and write your conclusion. All right, which is what you told me. Oh, so you have to, in analyzing that data, you have to run a test. So right now you are determining, in your groups, which test. And when you write your null and your alternative, I want you to specify exactly what your symbol means. Ok, go. You have about one minute for that. Ok, two.

There was a brief pause. Megan began the conversation when the instructor walked over to the group's table.

Megan: So p stands for the (pause).

Bailey: Proportion of people who believe, who got it right.

Chelsea: Who identified the content correctly.

Dr. Anderson: There's, umm, a keyword about what p stands for versus like p-hat.

Bailey: Oh. Population proportion.

Dr. Anderson: And you all helped to specify that. [Instructor walked away.]

Bailey: Oh. The population proportion of [University] student who correctly.

Chelsea: Yep.

There was another brief pause while all three group members wrote. Megan stopped writing and looked around. When Chelsea finished writing, Megan looked at Chelsea's paper and then continued writing. Chelsea said to her, "That's what I thought. I'm not completely positive." Chelsea then turned and asked Bailey, "What did you put?" Bailey showed Chelsea what she had written on her page. Chelsea read it out loud. "Population proportion of [University] students who correctly identify the drink as Coke or Pepsi. Ok." There was a brief pause. Dr. Anderson walked over to the group and asked, "So what did you decide, Megan, that *p* stands for?" Megan responded, "Umm the population proportion of [University] students who correctly identify." Dr. Anderson walked away, and Megan said, "We worked hard on it. Just kidding." Megan, Bailey, and Chelsea then had an off-topic conversation while the rest of the groups finished.

Dr. Anderson walked to the front of the class and began the group discussion again.

> Ok so I have had three possible null hypotheses so far. Three of them. I have [the null hypothesis] is that *p* equals 0.75. I have that *p* equals 0.5, and I have that *p*-Coke equals *p*-Pepsi. You all knew that the null is always what?

Dr. Anderson wrote this information on the board as she spoke it. Bailey responded, "Equals." Dr. Anderson continued.

> Yea, you can do better than that. [More students say equals.] Equals. Yes, the null [hypothesis] is always equals. That's how we do it in this class. There are other ways that are equivalent, but this is the simplest way, and that's what we do. Now, being respectful to your colleagues because there are multiple ways to do this and none of these are wrong, they are just all different. Given what we said to keep it as simple as possible and require the least amount of evidence that people can [tell a difference between Coke and Pepsi], well, what's *p*? What's *p*?

A student responded, "Proportion." Dr. Anderson replied, "It is proportion, but give me more." The student continued, "Population proportion." Dr. Anderson then responded, "It is population proportion. So it is the population proportion of what?" This time Bailey answered, "[University] students." Dr. Anderson continued.

> And we said [University] students. Go ahead. Who can (pause) correctly identify one cup as Coke or Pepsi. So that one cup that they get, they can correctly identify it as Coke or Pepsi. Ok. So that's what *p* is. Which one of these should we choose to be our null [hypothesis]?

A student from a different group answered. Dr. Anderson replied, "Oh, so this would be, she said rater preference, but what would be another way of saying that." Bailey once again answered, this time stating, "No preference." Dr. Anderson continued, "No preference. What would they be doing?" A student from a different group responded again. Dr. Anderson said, "Anybody else want to add some words to that? (pause) Would that be the same thing as they were just guessing?" Bailey answered, "Yes." Dr. Anderson replied, "Just by random chance. She said that it is the same as having a random chance. And why is it a half instead of a third?" A student from another group responded. Dr. Anderson was pleased and responded, "Very nice. Because there are only two options. Very nice. Ok, so that would be just guessing, and that would mean that they can or cannot tell a difference?" This time Chelsea responded and said, "Cannot." Dr. Anderson continued.

> They can't correctly identify because that would be the same as just guessing. So they can't correctly identify. Ok, 0.75. If they got 0.75 that actually makes you think that they can identify it correctly. In fact, it's very specific. They can identify it correctly exactly 75% of the time. Unfortunately, if they identified it 90% of the time that would be even better, but it might not be detecting the same thing. Do we need to require that they make it 75% of the time? That's a little strong. Nothing wrong with it, it's just a little strong. And then this one, what

> would this be? Population proportion who can identify Coke correctly is the same as the population proportion who can identify Pepsi correctly. Umm, this one is a little tricky. This implies how many samples? Two. How many samples are we going to be able to have in our class? One. Only way we would have two is if I divided you by boys and girls and that's not going to work because [one student] is not testing, so we can't right? I mean there are other things. Freshman and Seniors. But why would that have any preference, right? So, umm, in this case given what we have, this will be our null [hypothesis]. (Instructor points to the board). And what's the alternative so that we are all on the same page? Bailey, how about your group. What did you all say the alternative was?

Bailey responded, "The population proportion was greater than 0.5." Dr. Anderson asked, "Why do we choose greater than 0.5 instead of not equals to?" This time Chelsea responded, "Greater shows that, well, greater shows a preference where they can tell the difference." Dr. Anderson continued.

> They can tell the difference. If you had a 25% percent that's even less, so they are guessing right? They can tell the difference. Or can correctly identify is what I should say. Ok, well, we are ready to collect our data. So if [student], [student], and [student], if you three can come up and what they are going to do is they are going to bring you one cup. And what I need you to do is take your taste and write either Coke or Pepsi, whichever one you think it is, and then when you finish that look on the bottom and you will have a number. And I will need that number and then either Coke or Pepsi is what we will need for the data collection. So if the three of you will come up here and be our servers and take them to the groups.

And then each of you for the data collectors, write your number and then either Coke or Pepsi, and we will go from there. So you collect your data, and I will be ready to take your data.

The three group members had a conversation while tasting their drink.

Megan: I can tell the difference easy. I think they taste totally different.

Chelsea: Pepsi is so much sweeter.

Megan: Oh yea.

Chelsea: But I like Pepsi better than Coke.

Megan: I like Coke better.

Chelsea: See everyone in my family likes.

Megan: But I drink diet. I like Diet Coke a lot better.

Chelsea: Do we just go?

Megan: I guess.

They all three tasted their drink.

Megan: Is this Coke? Definitely.

Chelsea: I don't know what you have. We have different ones.

Megan: I think so because they are all numbered. They must be different. That's definitely Coke.

Bailey: I wish I had something to compare it to.

Chelsea: Mine is Coke.

Megan: I think it makes a difference too that they are not, like, super cold.

Chelsea: Yea. When they are cold and on ice it tastes different.

Dr. Anderson: (to whole class) Ok, it looks like they are probably ready for their second one. We are going to give you two just to increase the sample size. Or maybe three. That's not a bad idea. Ok, so I am going to collect data so we don't accidentally get them mixed up.

Megan: I think I have Coke again.

Chelsea: This is so frustrating.

Megan: You are right. When they are cold you can tell.

Chelsea: When it's cold there is a difference.

Megan: I think I have Coke though.

Chelsea: I thought I knew. I thought I knew the difference. I think this is Coke though.

Dr. Anderson: (to whole class) So do you think the idea of randomizing them where you could have two of the same, two different, or one of each, is good?

Chelsea: Yes.

Dr. Anderson: Did that make it trickier?

Chelsea: Yes.

Dr. Anderson: As opposed to if we were just discerning between the two. Ok. Well, let's see what we have. There is still some left. Let's go another round really quick.

Megan: There is no way we are [inaudible].

Chelsea: I think it's random.

Bailey: I put [inaudible].

Chelsea: Do you think yours is Pepsi?

Bailey: I feel like Pepsi has a faulty taste.

Chelsea: I think this is a good experiment too because, like, I can't stand Pepsi, so if one of these is Pepsi, you're not gonna [inaudible].

Megan: This tastes like Coke. Everything is Coke.

Bailey: I know.

Chelsea: I don't even know. I have no idea.

Bailey: (Question asked to instructor). Are we allowed to change our answer for old ones after we taste.

Dr. Anderson: No, because you are technically a new person now.

Bailey: Oh ok.

Dr. Anderson then spoke to the whole class again.

> Bailey asked a great question, and what I want you to keep in mind is that we are really pretending that there are three of you. You are three people. So you can't go back and change the first one because you have a new taste in your mouth. That changes the experiment. We are pretending there are three of you. All right. So are we ready to get the data? Has everyone made their decision? (pause) Ok, I think we are ready. [She opened up an Excel spreadsheet on the computer on the board in front of the class.] Ok, so listen for your number and you are just supposed to tell me if it's Coke or Pepsi.

Dr. Anderson then called out numbers to represent each sample. As their number was called, the student responded either Coke or Pepsi. After all students responded, Dr. Anderson went through the spreadsheet and marked each answer as correct or incorrect. She then continued.

> Ok. So this is - the answers are under here. So this is what we had. So while we are on this page, I am sure you can do it better than I can, so count the corrects and wrongs. Ok. So we need to go ahead and count on here how many are correct and how many are wrong. So we need a total on each. Double check me.

The students corrected a few mistakes as they saw them. Dr. Anderson again continued, “So how many wrongs were there? Only seven. Out of how many? Twenty-five minus one, two, three. Ok, so how many got it correct? Ok, 15 out of 22 on that page. Let’s go to the next page.” She then continued to call numbers, and students responded either Coke or Pepsi as their number was called. Once she finished this she continued.

> So there are seven wrong on this one out of 25, so how many are correct? Ok, so now do you have enough information? I will put the two of them on there. You will need both. So what is the total *n* [sample size]? Yea, it is 47. So *n* is 47. So that is what we will be using. *N* is 47. So if you would go ahead and analyze the data, and take some time and see if you can write a good conclusion. Ok. So analyze your data. You will need your calculators I’m sure. Write your conclusion. All of that good stuff.

A student in the class asked, “Do you plan to tell us which test to use?” Dr. Anderson responded, “No. You have already decided that. Well, I say that, but I want you to decide for this right here. I want you to decide.”

Bailey, Chelsea, and Megan had a discussion.

<Interview Video Clip 3 Begin>

> Bailey: So is our x the number of corrects? (She is referring to the x in the calculator where you input the information for the chosen statistical test.)

Chelsea: Yes.

Bailey: One-proportion *z*-test. Can I take intervals too? It's easy to interpret the one-proportion *z* because the interval doesn't even matter for the *p*-test.

Chelsea: The significance level?

Bailey: Yes.

Chelsea: Yes, it is. Yea, because it is lower. If your *p*-value is lower than the 0.05 then you reject the null [hypothesis] and take the alternative [hypothesis].

Bailey: Oh ok. So this *p*-value it gave us is way lower than our significance level, so we are going to reject the null [hypothesis].

Chelsea: Yea, because it's lower than alpha, you reject null [hypothesis] and take alternative [hypothesis]. We didn't say what alpha was did we? (Reading from notes.)

Dr. Anderson walked up to the group.

Dr. Anderson: We didn't say what alpha was did we?

Chelsea: Do we not assume it's 0.05?

Dr. Anderson: You get to make that choice.

Megan: And we look at *p*, not *p*-hat right?

Chelsea: Right. That's why I kept messing up on homework last night.

Megan: Ok.

<Interview Video Clip 3 End>

There was a brief pause while all three wrote on their paper.

Megan: Do we say confidence level? Ninety-five percent confidence level?

Chelsea: You just use the 5%, at the 5% level. That's what I wrote down in my notes.

Megan: I am always bad at interpreting these.

There was another brief pause while all three wrote. Megan looked over at Chelsea's paper. Chelsea erased something on her paper and noticed Megan looked at her paper.

Chelsea: Oh. Did you copy me? Sorry. I did it like if you accept. (Looked back at notes.) Oh, I was just kidding. Oh, I did it right. Crap. Because I thought. Well I looked at these two examples. So because the *p*-value was smaller, we reject the null [hypothesis] and take the alternative [hypothesis]. So it's this one (points to notes). I am so sorry.

Megan: I mean I could not be lazy and look at my own notes so that's ok.

There was another pause while all three wrote. Dr. Anderson walked over, but all three were writing. She walked away. After a few minutes, Bailey began the conversation again.

Bailey: So I want to compare my interpretation to yours because mine is different. So I said at the 5% confidence level, we have strong evidence to suggest that the population proportion of [University] students who can correctly identify a drink as Coke or Pepsi is higher than 50%.

Chelsea: I said at the 5% level, the true population proportion of [University] students who can correctly identify a drink as either Coke or Pepsi is significantly greater than 50%.

Megan: So if it would have been – if the *p*-value would have been greater than, then we would have put it significantly, we would say (pause). Well, how would you word it?

Chelsea: This one that I wrote the other day says at the 5% level there is not enough evidence to say that the true population proportion of [University] students who can correctly identify Coke or Pepsi is significantly different from 50%. Or you could say there is not enough evidence to say all [University] students can.

Megan: It's weird that, like, if the *p*-value is bigger or smaller then it's greater. It's kind of, like, opposite of how my mind thinks.

Chelsea: The only reason that I remembered that is [because] that homework last night had like eight questions of would you reject the null given this *p*-value.

Bailey: I missed that a lot.

Dr. Anderson walked over to the group.

Dr. Anderson: Umm, so I had a question. I heard you say that you were going to choose alpha last time. What did you all choose? You chose 5% right?

Chelsea: Yea.

Dr. Anderson: What does alpha stand for? What does it mean? Is 5% like a choice?

Dr. Anderson then walked away from the group.

Chelsea: Alpha is your chosen significance level.

Bailey: And I think 5% was a good choice because if it was 99% that would make, like, what is the word?

Chelsea: Interval?

Bailey: Yea. Our interval would be larger.

Chelsea: Our interval larger, so error. Like it gives you a little bit of leeway. Is that the word?

Bailey: Yea. And [inaudible].

Megan: I get this better after today. Like, I didn't understand it at all before. I had a lot of trouble with the homework. I just kind of guessed. Like yes or no. I did it late so I forgot.

Dr. Anderson began the whole group conversation again. She asked, "Let's see, Chelsea, could you tell me about what you guys concluded?" Chelsea responded, "Umm, at the 5% level, the true population proportion of [University] students who can correctly identify a drink as Coke or Pepsi is significantly greater than 50%." Dr. Anderson then asked Megan, "Do you agree? So we agree? How did you know, let's see, Megan, how did you know to do that?" Megan replied, "Umm, because the, umm, alpha, the *p*-value was smaller than the alpha, umm, sigma, sorry." Dr. Anderson responded, "Alpha. You were right. And what does alpha stand for?" Megan hesitated and said, "Umm the, we just said it." Chelsea replied, "Significance level." Dr. Anderson asked, "And who got to choose that today." Chelsea responded, "We did." Dr. Anderson replied, "You all did because you are the researchers and what did you all choose?" Chelsea again responded and said, "Five percent." Dr. Anderson replied, "Five percent is a common value. And what does alpha stand for?" This time Bailey responded and said, "The one minus." Dr. Anderson asked, "Yea, but what is it? It's probability of what?" A student from a different group answered, "Rejecting the null [hypothesis]." Dr. Anderson continued, "Yea, when the null [hypothesis] is true. So is that getting an answer right or wrong? Is alpha error or correctness?" Chelsea replied, "Error." Dr. Anderson responded,

> Error. So it's the maximum amount of Type I error you are willing to tolerate. Be sure to study that. That is what that homework that is due on Thursday is on. Ok, so umm, (pause). So back group could you ask a question to help clarify the process? What still needs clarifying in your mind about this? Anything?

A student asked, "Is it always going to have to be lower or greater than 5%?" Dr. Anderson answered, "That's a good question. So are you asking me if you are always using 5%?" The student answered, "Yea." Bailey then answered, "Depends on the situation. It might change." Dr. Anderson continued, "Depends on the researcher. The most common value is 5%. Other common values, you could choose are 1% or 10%. Depending on the setting, you might choose 20% if you have a really small sample size. That's a great question." Megan then asked, "What would be a situation where you would have to do less than? Because every time I've pretty much done it, it has been greater than." Dr. Anderson responded,

> Yea, so many of them have been greater than because we want them correctly identifying so, umm, so if I said that it turned out in the research that 75% of the people who took this new drug got really sick, I would want the alternative to be, well, that the benchmark of an old drug is 75%. So it made 75% of people get sick. It's a cancer drug right. So I have a new drug. I want to show that less people get sick. The proportion of people that get sick is less, so that's what I want to show. So it depends all on what you want to show. That was a great

question. So when you were coming up with this process – well, let me ask this. What if there were three colas involved?

Bailey answered, “It would be 0.66, I mean 0.33. We would make our null [hypothesis] equal to 0.33 and then our alternative [hypothesis] greater than.” Dr. Anderson continued.

> So she said that - why would she pick, somebody else, why would she say that the null [hypothesis] would be a third, that the proportion would be a third, if we had three colas? So I should have asked for the four colas. Let me ask that one too. So this group said we are going to give them four different colas to taste from. Why, what would their null be?

The class responded 0.25. Dr. Anderson asked why, and a student answered, “Because there are four outcomes.” Dr. Anderson continued.

> And if you are just guessing? It would be one out of four. Right. Very nice. Good job. And what about, what would be the next step in terms of your analysis? [Pause.] So remember when we talked about that model when you asked some more questions from it? So what might be the next step in your analysis? Just, I mean, that’s the next question you should ask yourself. So I will ask you one last question, and then we will have a writing prompt. So tell me, talk to the members of your group about what is most important when designing or conducting a study.

Bailey, Chelsea, and Megan discussed different components of a study they felt were important. Chelsea began the conversation.

> Chelsea: Choosing your correct test.
>
> Megan: Yea, because then all of your stuff would be wrong.

Bailey: Or choosing or getting a good sample because, like, the way you got your sample, because of the sample stuff.

Chelsea: I feel like it is all pretty much important because if you mess up your null [hypothesis] and your alternative [hypothesis], then you are wrong. If you mess up the design, you are wrong.

Bailey: If you mess up the interpretation, it's wrong.

Chelsea: Yea, if you choose the wrong test, then you are definitely wrong.

Bailey: Yea.

Dr. Anderson said to the class, "I have to tell you that you have some great answers. Let me hear answers from this table. One word." A student answered, "Data." Dr. Anderson called on another group, "The data. Let me hear an answer from this one back hear. I think I hear it." A student answered, "Random." Dr. Anderson continued, "You need to make sure you have got randomization. I heard something from this group what did you say?" A student answered, "The warrant." She again continued, "The warrant. Choosing the correct test. Exactly. What about this group?" Bailey responded, "We said all parts. Like all parts." Dr. Anderson replied,

> All parts are important. I certainly agree. If you mess up on one you are in trouble. Ok, I will take up both [sheets of paper]. I am going to take up the writing prompt and the work that you did in class. Ok, writing prompt and your work. So this writing prompt is about argumentation. The writing prompt is about argumentation. Please make a good link. The writing prompt is about argumentation. Make sure that I know what you are thinking in your head because it may have been so clear to you, right? That whole thing. So make sure I

> understand the link you are making, so make that clear for me. Oh, and those are individual. I know you have done everything together in your groups, but make the writing prompts individual.

Megan raised her hand to ask Dr. Anderson a question. She walked over to Megan, and Megan asked, “Am I allowed to ask a question about this [the writing prompt]? I know, like, the process of like, what argumentation is but, I still don’t understand what this is?” Dr. Anderson answered, “That’s difficult for me, too. That’s the hardest part I think. So maybe in this case if you spell out what you think it might be that would be great since it’s something you are questioning, and then we will look at it later.”

The students all responded in writing to their writing prompts. When they were finished, they handed them to Dr. Anderson and left class.

Participant One – Chelsea

In this section, a description of Chelsea’s data from prior to argumentation training, during argumentation training, post argumentation training, Coke/Pepsi lesson, and post Coke/Pepsi lesson is provided.

Prior to Argumentation Training

During an observation of students working in small groups prior to argumentation training, the researcher identified Chelsea as an active member of group work and discussions. Chelsea either took part in or initiated almost every discussion that occurred within the group by either asking a question or providing her opinion to the question posed by the instructor (Field Notes, 3/26/15). After being identified as an active participant in group discussions by the researcher, Chelsea was given a writing prompt that asked, “In your opinion, what is argumentation?” (Writing Prompt 1, 4/7/15). Her response to this writing prompt stated, “Argumentation, in my opinion, is giving sufficient evidence to support your ideas, views, or opinions in response to someone else’s views, ideas, or opinions” (Writing Prompt 1, 4/7/15). In this statement, Chelsea mentioned none of the components of argumentation by name. However, she hinted at the components of data, claim, and rebuttal in her description. Therefore, Chelsea’s response was coded as a Level 1 on the levels of understanding of argumentation.

Argumentation Training

After the first writing prompt was given, Chelsea participated in argumentation training with the rest of her classmates. During this training, she was active in her group discussions by leading or initiating many of the conversations that occurred, as well as by participating in every group discussion that occurred (Field Notes, 4/9/15). Chelsea also participated in the whole class discussions by answering the instructor’s class questions on four different occasions throughout the lesson (Field Notes, 4/9/15).

Post Argumentation Training

At the conclusion of the argumentation training, Chelsea was given a second writing prompt that asked, “What do you think argumentation is now?” (Writing Prompt 2, 4/9/15). Her response to this writing prompt stated, “Providing a claim or statement with data and linking that data to your claim to reach the goal of in fact proving that claim” (Writing Prompt 2, 4/9/15). In her response to this writing prompt, Chelsea used the terms claim and data to describe argumentation, neither of which had she used in Writing Prompt 1 when asked to define argumentation (Writing Prompt 2, 4/9/15). However, she still included only a description of a warrant without using the proper terminology. Because of her inclusion of the proper terminology for some argumentation terms, but inclusion of a description of a term instead of using the proper terminology in her response, Chelsea's response was coded as a Level 1.5 on the levels of understanding of argumentation.

Engagement in the Coke/Pepsi Lesson

During the next class meeting following the argumentation training, Chelsea engaged in the Coke/Pepsi lesson. After being asked by Dr. Anderson to discuss within each group what is important to do when solving a statistical problem, Chelsea was the first to initiate conversation by suggesting you need to "collect a reasonable amount of data" (Video Transcript, 4/14/15). After a response from Bailey, she added to her suggestion, "from a random sample" and "the significance level" (Video Transcript, 4/14/15). Chelsea continued to provide suggestions to her group, such as "the sample" and "at the end it's important to decide if you reject the null" (Video Transcript, 4/14/15). Her suggestions were supplemented at times by Bailey. For example, after her suggestion about the importance of deciding whether to reject the null, Bailey added, "And then your interpretation is important, too" (Video Transcript, 4/14/15). However, Chelsea's suggestions were a main source of information within her group conversations.

When the instructor approached, Chelsea quickly continued with more suggestions of important components of statistical problem solving, such as "make sure you don't make a Type I or Type II error" (Video Transcript, 4/14/15). After the instructor left, Bailey took a more active role in the conversation and added suggestions as well. Chelsea responded by asking a clarifying question of, "You mean whether you reject the null and it was true or accepted the null and it was false?" and agreed with a suggestion made by Bailey (Video Transcript, 4/14/15).

Throughout the dialogue that occurred among the group members on what they believed is important when solving a statistical problem, Chelsea did not search her notes and was quick and calm with each response, which seemed to indicate she was confident in her knowledge and answers. Through her confidence and initiative to begin group conversation, she appeared to establish herself as a leader of the group.

When Dr. Anderson brought the class back together for a whole-group conversation, Chelsea kept her eyes on the instructor. When Dr. Anderson called on Chelsea's group to answer a question, both Chelsea and Megan turned to Bailey to respond (Field Notes, 4/14/15). Bailey responded to the instructor, and she continued with the class discussion. After a few moments, Dr. Anderson asked students to discuss with their groups different tests they had learned about in the semester and when they used them. Chelsea immediately took out her notes (Field Notes, 4/14/15). She then named several different statistical tests to her group members as she read from her notes, "mean tests, proportions or *t*-test, *z*-test for proportions, umm, a two-sample *t*-test" (Video Transcript, 4/14/15). After that, she described when it is appropriate to use each of the tests. "You use a one-proportion *t*-test for proportions of one-sample, use a two-sample *t*-test for a two-sample mean . . . you do two-proportion *z*-test proportions of two samples" (Video Transcript, 4/14/15).

Dr. Anderson once again began a class discussion, asking the class what is the purpose of a one-proportion *z*-test. Chelsea immediately answered with, "It was one-sample mean" (Video Transcript, 4/14/15). After a few clarifying comments from the instructor, Chelsea realized she had misunderstood Dr. Anderson's question and said, "Oh it's *z*-test. I am sorry. So it's one-sample proportions. I looked at the wrong thing. My fault" (Video Transcript, 4/14/15). The instructor continued with the class discussion, and Chelsea again responded to the instructor's question, "So each one of the tests has a corresponding what? A partner" with, "confidence interval," which the instructor said was correct (Video Transcript, 4/14/15).

When it was time for group discussion, Chelsea once again initiated the conversation by presenting what she believed to be the claim. She stated, "So the claim is people can correctly identify a drink as either a Coke or a Pepsi. And then you have to decide your hypothesis and your null?" (Video Transcript, 4/14/15). She continued to hold a role in the conversation with her continual inputs and clarifying statements for her other group members. For example, after a group member was confused about the research question, Chelsea stepped in and clarified by saying, "If they can tell the difference between them. Like if you don't have a label, can you tell if this is Coke or Pepsi?" (Video Transcript, 4/14/15). As group discussion transitioned into designing the study, she continued to take an active role in the discussion by providing the suggestion to, "make them try both and label this one is Coke or this one is Pepsi" (Video Transcript, 4/14/15).

In her attempts to clarify other group members' thinking, Chelsea questioned her group members to explain what they were trying to say. For example, when a group member questioned the wording for the hypothesis for having a difference, Chelsea asked, "How would you put it then?" (Video Transcript, 4/14/15). When Dr. Anderson came by to ask the group a question, Chelsea was the one to answer, "The proportion of people that can identify or the proportion of people" (Video Transcript, 4/14/15). She was often the one to transition her group to the next phase, such as when she moved her group from the design of collecting data to what they were choosing as their significance level for their study when she said, "So now we need to pick an alpha level" (Video Transcript, 4/14/15). As part of her leadership role within her group, Chelsea also made many decisions for her group. For example, one group member suggested using a 90% significance level for the study, but Chelsea said, "I say just go with . . . 0.05" (Video Transcript, 4/14/15), and there was no further discussion on the topic.

In addition to what appeared to be her confidence in answering questions in class discussions, Chelsea also seemed to be comfortable asking questions in class discussion by asking the question, "Would you, like, in a test like this, would you want to I guess ask, sample people individually so that you don't have the influence of the person. Would that make a difference in your data?" (Video Transcript, 4/14/15) and engaging in a discussion with the instructor related to her question. Chelsea's leadership and confidence in her knowledge continued to be exhibited throughout the group work. At one point, Megan copied Chelsea's paper in order to get the wording correct (Field Notes, 4/14/15).

Use of Argumentation During the Coke/Pepsi Lesson

Throughout the Coke/Pepsi lesson, Chelsea's participation in argumentation consisted mostly of claims. For example, during a discussion of what is important to do when solving a statistical problem, Chelsea made claims such as "collect a reasonable amount of data," "the significance level," and "make sure you don't have Type I or Type II error" (Video Transcript, 4/14/15). She did not use any of the other components of

Toulmin's Model of Argument, such as data, to support her claims, during this discussion.

In another conversation during the Coke/Pepsi lesson, the instructor asked groups to review statistical tests learned throughout the semester. Chelsea made claims, such as when she stated, "Mean tests, proportions or, ah, *t*-test, *z*-test for proportions, umm, a two-sample *t*-test" (Video Transcript, 4/14/15). However, this time she referred to her notes as a data source for her claims. She continued to refer to her notes as data as she further made the claims of "*t*-test for mean" and "use a one-proportion *t*-test for proportions of one-sample [and] a two-sample *t*-test for two-sample means" (Video Transcript, 4/14/15). Unlike with the prior claims Chelsea made, Chelsea used her notes as a data source when making the claims just mentioned. However, Chelsea did not reference her notes as her data source when making her claims. She also did not use a warrant to link her notes as her data source to her claims.

In reference to the Coke/Pepsi study, Chelsea was the group member who stated the claim for the overall study for the group. She said, "So the claim is people can correctly identify a drink as either a Coke or a Pepsi" (Video Transcript, 4/14/15). After sharing ideas among each other for the design of the study, Chelsea once again stated, "So the claim is that people can tell a difference" (Video Transcript, 4/14/15). After another group member provided a rebuttal to this claim by saying you cannot tell if people can tell a difference because you would not know until you perform the study, Chelsea then made a rebuttal by saying, "But you are researching, so [it's] like your research hypothesis" (Video Transcript, 4/14/15). She then continued the conversation when she provided data and warrants to support her claim and rebuttal. She stated, "So if you are saying people can tell a difference, your null [hypothesis] is equal to 0.5 and your research is greater than 0.5 . . . if you wanted to word it where people couldn't tell a difference, your null [hypothesis] would still be equal to and your research would be less than 0.5" (Video Transcript, 4/14/15). In this case, the data were her notes and knowledge about the null and alternative hypothesis and the warrants were the interpretation of the null and alternatives in relation to the claim. However, she once again did not acknowledge the data or a warrant for the claims she had made. For the remainder of this portion of the conversation, Chelsea continued to make claims. For example, when Bailey asked what the null would be, Chelsea responded with the claim that "*p* equals 0.05" (Video Transcript, 4/14/15). She made this statement without providing data or a warrant to justify her response as being correct. Another claim made by Chelsea occurred after the instructor walked away when she stated to her group members, "The proportion of people that correctly identify would be greater than 0.5, yea, because that means people can definitely tell a difference if it's significantly greater than 0.5" (Video Transcript, 4/14/15). She again did not attempt to provide data supporting her claim.

In a later conversation in reference to which test should be used to analyze the data collected by the class, Chelsea provided another rebuttal to a claim made by a group member which implied the significance level was not important for the statistical test. She replied, "Yes, it is because it is lower. If your *p*-value is lower than the 0.05, then you reject the null [hypothesis] and take the alternative [hypothesis]" (Video Transcript, 4/14/15). After she made her rebuttal, she then provided the data from her notes, which she read, "Yea, because it's lower than alpha, you reject null [hypothesis] and take alternative [hypothesis]" (Video Transcript, 4/14/15). She continued to use her notes as data throughout this conversation to support her claims but did not verbally acknowledge her notes as her data source. For example, Chelsea stated, "I looked at these two examples. So because the *p*-value was smaller, we reject the null [hypothesis] and take the alternative [hypothesis]" (Video Transcript, 4/14/15). In this statement, her claim was that the null should be rejected in favor of the alternative. Chelsea's data were once again her notes, and the warrants were the examples she used in her notes to make the connections to the claim. Chelsea later made the claim for the interpretation of the results from the study when she stated, "At the 5% level, the true population proportion of [University] students who can correctly identify a drink as Coke or Pepsi is significantly greater than 50%" (Video Transcript, 4/14/15).

In a conversation in which the group members discussed the most important components of designing or conducting a study, Chelsea made the claim that choosing your correct test was most important (Video Transcript, 4/14/15). She then also made the claim "it is all pretty much important" (Video Transcript, 4/14/15).

Overall, Chelsea made several unsupported claims throughout the Coke/Pepsi lesson. When she did support her claims with data, she used her notes as her data source, but she did not verbally acknowledge her notes as her data source. Chelsea did not explicitly provide warrants between her data and her claims. However, she did provide a rebuttal with supported data when she felt another group member's claims were incorrect. Her use of argumentation occurred mostly in the first two stages of the statistical problem-solving process. Therefore, Chelsea's level of understanding of argumentation throughout the Coke/Pepsi lesson appeared to align best with a Level 1. Although she did use the word claim at the beginning of the lesson in reference to the overall claim for the research study, Chelsea did not use any other argumentation terminology throughout the Coke/Pepsi lesson. As mentioned above, Chelsea used data, warrants, and a rebuttal during the lesson, which indicates her movement towards seeing argumentation in terms of Toulmin's Model of Argument. However, her lack of identifying data and warrants when making claims and lack of using the proper terminology of argumentation resulted in the assignment of Level 1 in terms of her understanding of argumentation at this stage of the study.

Post Coke/Pepsi Lesson

In this section, Chelsea's data from writing prompt 3 and her interview are provided.

Writing Prompt 3. At the conclusion of the Coke/Pepsi lesson, Chelsea responded to a final writing prompt that asked, "Do you think you used argumentation in the Coke/Pepsi Taste Testing lesson? If so, how? If not, why did you not use it?" (Writing Prompt 3, 4/14/15). She responded, "Yes. We made a claim, collected data to support the claim and created a warrant to link the data to the claim, which is a justification or test for the claim" (Writing Prompt 3, 4/14/15). In her response to this writing prompt, Chelsea incorporated the argumentation vocabulary term, warrant, which was not used in the previous writing prompts (Writing Prompt 3, 4/14/15). She also described the structure of a simple argument and used the proper argumentation terminology. Through her description, Chelsea appeared to be describing argumentation in terms of the statistical problem-solving process. It is uncertain whether Chelsea believed she used argumentation within her group in different ways other than engaging in the statistical problem-solving process. Therefore, Chelsea's level of understanding of argumentation was coded as a Level 2.

Interview. Two days after she participated in the Coke/Pepsi lesson, Chelsea completed an interview with the researcher. In this interview, she did not acknowledge her role as a leader and main decision-maker within her group, but instead used *we* when answering questions about the group work. For example, when asked how her group decided which plan to present to the class, Chelsea answered, "We looked over our notes to make sure that would make sense, and then everyone kind of made their input. And then we decided what was best to give." When asked if she believed she used argumentation to help make decisions within her group, Chelsea answered, "I mean you had to prove why that was the best decision. So, yeah, in a way, but not like hard core" (Chelsea Interview Transcript, 4/16/15).

Chelsea continued with her description of using argumentation as a way to prove the best decision as to how she used argumentation throughout the Coke/Pepsi lesson which was exemplified in her answer.

> To try to prove, like, my point or my idea or whatever, and the thing to try to prove that was right or try to prove, like, the reasoning behind what I was trying to say to make them believe or know that was the correct way to go. Like, even if it wasn't, like, they would then argue back and tell me, hey, no, I think it's this. (Chelsea Interview Transcript, 4/16/15)

When asked how her group approached the questions posed to the group and the class as a whole by Dr. Anderson, Chelsea responded, "Just reviewing your notes and talking to your group members to make sure they agree before you yell out something that's not going to be correct" (Chelsea Interview Transcript, 4/16/15). However, she responded "No, not really" (Chelsea Interview Transcript, 4/16/15) when asked if she believed argumentation was used to answer the discussion questions asked by Dr. Anderson.

In addition to answering questions during the interview, Chelsea was shown three video clips from the Coke/Pepsi lesson and asked to identify any argumentation that occurred during each clip. Of the three clips, Chelsea believed she saw argumentation in two clips and identified what that argumentation was. In the first clip, she believed all three group members used argumentation.

> We were trying to come up with a test if someone had an idea then they would try and tell you why this idea would work and everything. And then other people would disagree, and then that person would then tell, ok, well then what if we did this. So I think we were all kind of using [argumentation] even if you couldn't hear from what I remember. Yea, I think we were all using some sort or form of argumentation. (Chelsea Interview Transcript, 4/16/15)

When asked to identify the data in the argumentation from the clip, she answered, "Previous examples or things we did in class, or like previous lessons I guess, to how to set up a test. And then we used our notes and everything" (Chelsea Interview Transcript, 4/16/15).

The third clip was the remaining clip in which Chelsea believed argumentation occurred. She described the argumentation as "whenever she was asking about, I think it was *p*-value or something like that, and I had to explain why I came up with this by looking back at my notes or looking back at what Dr. Anderson said. So, yea, I think so" (Chelsea Interview Transcript, 4/16/15).

In both of these descriptions of what she believed to be argumentation in the clips, Chelsea did not use any argumentation terminology. She identified her data source when asked to by the researcher in the first clip and provided an implied data source in the third clip. However, in neither clip description did she provide a warrant or mention providing a claim explicitly. Once again, Chelsea was describing some components of argumentation without using the proper terminology of argumentation. Therefore, her level of understanding of argumentation during the interview was coded as a Level 1.

Summary

In summary, prior to being introduced to argumentation, Chelsea described some components of argumentation, but she did not use the correct terminology for any of the components of Toulmin's Model of Argument when asked to describe argumentation. Therefore, her understanding of argumentation was coded as a Level 1. After engaging in argumentation training, Chelsea described all three components of the simple structure of Toulmin's Model of Argument and used the correct terminology for two of the components in her description. Her level of understanding of argumentation at this point was coded as a Level 1.5, which indicated some gains in terms of understanding. While participating in the Coke/Pepsi lesson, Chelsea's use of argumentation consisted mostly of making claims that were unsupported by data, and, therefore, not connected to that data by a warrant. Based on her use of argumentation during the Coke/Pepsi lesson, her understanding of argumentation during the Coke/Pepsi lesson was coded as a Level 1. After participating in the Coke/Pepsi lesson, Chelsea stated that she believed her group used argumentation during the Coke/Pepsi lesson and mentioned all three components in the simple structure of Toulmin's Model of Argument in her description of argumentation. Therefore, her level of understanding of argumentation was coded as Level 2. Also after participating in the Coke/Pepsi lesson, Chelsea indicated in an interview with the researcher that she believed there was argumentation in both the first and third video clips shown. However, in her description of the argumentation in these clips, Chelsea described some of the components of argumentation, but did not explicitly identify components of the argument, such as warrants and claims. She also did not use any of the proper terminology of argumentation. Therefore, Chelsea's understanding of argumentation during the interviews was coded as Level 1. Figure 4 below provides a graph of Chelsea's progression of understanding argumentation throughout the study. Her highest levels of understanding of argumentation occurred when writing about argumentation while her lowest levels of understanding of argumentation occurred prior to argumentation training, when using argumentation, and when asked to identify or discuss argumentation.

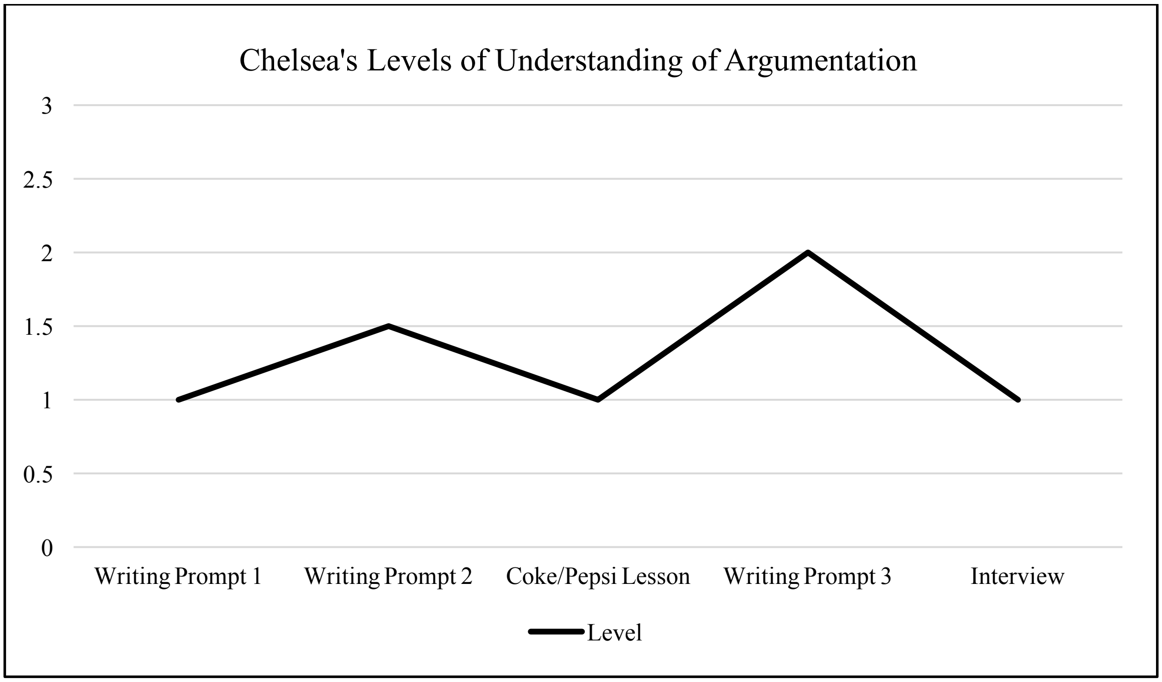

Figure 4. Progression of Chelsea's levels of understanding of argumentation throughout the study.

Participant Two – Bailey

In this section, a description of Bailey's data from prior to argumentation training, during argumentation training, post argumentation training, Coke/Pepsi lesson, and post Coke/Pepsi lesson are provided.

Prior to Argumentation Training

During an observation of students working in small groups prior to argumentation training, the researcher identified Bailey as a moderately active member of group work and discussions. Bailey took part in several group discussions that occurred within the group (Field Notes, 3/26/15). Her participation mostly consisted of answering questions that were posed by her group members and engaging in discussions with Chelsea about possible solutions (Field Notes, 3/26/15).

After being identified as a moderately active participant in group discussions by the researcher, Bailey was given a writing prompt that asked, "In your opinion, what is argumentation?" (Writing Prompt 1, 4/7/15). Her response to this writing prompt stated,

> Argumentation is an attempt at persuasion in which an individual would like to persuade another by learning about two sides and presenting the information. When done properly, it can be civil and persuasive. When done incorrectly, it can lead to anger and further misinformation. When arguing, it is important to avoid logical fallacies. (Writing Prompt 1, 4/7/15)

In this statement, Bailey described argumentation as a way to persuade someone to believe your statements. Although she did not mention any of the components of Toulmin's Model of Argument by name, she did hint at the use of data and claims as part of argumentation. Her description of argumentation was beyond that of simple disagreement between people. Therefore, Bailey's response was coded as a Level 1 on the levels of understanding of argumentation.

Argumentation Training

After the first writing prompt was given, Bailey participated in argumentation training with the rest of her classmates. During this training, she was active in many of her group discussions that occurred by engaging in conversation with Chelsea (Field Notes, 4/9/15). However, Bailey did not attempt to engage in the class discussion by answering or asking a question throughout the entire training (Field Notes, 4/9/15).

Post Argumentation Training

At the conclusion of the argumentation training, Bailey was given a second writing prompt that asked, “What do you think argumentation is now?” (Writing Prompt 2, 4/9/15). Her response to this writing prompt stated, “Argumentation is making a claim and using data to support it. Also, the claim must be connected to the data through a warrant. Argumentation will be met with rebuttals causing there to be a qualifier added to the claim” (Writing Prompt 2, 4/9/15). In her response to this writing prompt, Bailey identified all three components (i.e., claim, data, and warrant) of the simple model of Toulmin’s Model of Argument, as well as two components (i.e., rebuttal and qualifier) of the complex model of Toulmin’s Model of Argument to define argumentation (Writing Prompt 2, 4/9/15). Therefore, Bailey’s response was coded as a Level 2 on the levels of understanding of argumentation. In her previous response prior to this training, Bailey had described a similar concept, but without using the terminology of argumentation and without the inclusion of a rebuttal and qualifier in her description.

Engagement in the Coke/Pepsi Lesson

During the next class meeting following the argumentation training, Bailey engaged in the Coke/Pepsi lesson. After being asked by Dr. Anderson to discuss within each group what is important to do when solving a statistical problem, Bailey engaged in a conversation with Chelsea. After Chelsea initiated the conversation by providing the first suggestion, Bailey followed with her suggestion, “Yea. And then you have to figure out the right test. That’s the hard part” (Video Transcript, 4/14/15). Following input from Chelsea, Bailey added, “And what kind of test you are going to take. Remember she [the instructor] made that chart,” as well as, “And then your interpretation is important too" (Video Transcript, 4/14/15). When the instructor approached the group, Bailey took a more active role in the conversation by adding more suggestions and bringing the focus of the group conversation to the lesson for the day by saying, "What are we going to be doing today with the Coke and Pepsi stuff because I guess the most important stuff in doing a statistical study is having a clear question" (Video Transcript, 4/14/15). Throughout the conversation that occurred among the group members about what they believed was important when solving a statistical problem, Bailey provided suggestions both when the instructor was present and when she was not. She did not search through her notes, and she quickly responded to Chelsea each time. However, Bailey spoke extremely quietly and was, at times, unable to be heard by the researcher or the audio recorder.

After leaving the group, Dr. Anderson brought the class together for a whole-class discussion. When Dr. Anderson called on Bailey's group to answer her question about what is important to do when solving a statistical problem, both Chelsea and Megan turned to Bailey to respond (Field Notes, 4/14/15). Bailey leaned back in her chair, and then responded by saying, "Interpreting our results" (Field Notes, 4/14/15; Video Transcript, 4/14/15). The instructor responded, "Well, of course. Interpreting the results. That is very good. I didn't hear you say that one but that is very good because that would be the perfect next step" (Video Transcript, 4/14/15). Bailey then added, "Oh and making sure you didn't make an error. A Type I or II" (Video Transcript, 4/14/15).

After finishing the whole-class discussions, Dr. Anderson asked the students to discuss with their group members different tests they had learned about during the semester and when to use them. Bailey began naming tests from memory while her group members both took out their notes (Field Notes, 4/14/15). She named "one-sample, two-sample" (Video Transcript, 4/14/15). Chelsea named some tests and when to use them from her notes, and Bailey added, "And then more than one-sample you use chi-squared for proportions" (Video Transcript, 4/14/15). Bailey then looked through her calculator.

Dr. Anderson once again began a whole-class discussion. She asked the class what the purpose is of a one-proportion *z*-test, and Chelsea quickly answered incorrectly after misunderstanding the question. Bailey whispered "one-sample proportion" to Chelsea, and Chelsea realized her misunderstanding and changed her response (Video Transcript, 4/14/15). After a few moments, the instructor called on Megan to answer a question during the class discussion about the null hypothesis for the means. After Megan responded, "I didn't say anything," Bailey quickly answered with, "It could be equal to" (Video Transcript, 4/14/15). Dr. Anderson continued the class discussion and wrote key points on the board. During this time, Bailey was on her cell phone (Field Notes, 4/14/15).

When it was time for group discussion, the group members discussed how they would design their study. Once the conversation began, Bailey was the last in her group to speak; however, once she began providing suggestions for the design of the study, she held an active and knowledgeable role in the conversation. This can be seen in the following passage.

> Megan: You have to have people. Your sample.
>
> Bailey: Yea, and I'm assuming everybody in this class. Like one. So our A would be [inaudible]. So we would need two samples. (She counts

> students in classroom.) Sixteen. So 16. Then our sample is 16. That's, like, very small.
>
> Chelsea: Yea.
>
> Bailey: Given our circumstances, 16 is probably going to be our sample.
>
> Chelsea: Mm hmm. (Video Transcript, 4/14/15)

Bailey continued her role of providing ideas and information for the group throughout the group discussion. For example, when the instructor came over to ask the group a question, Chelsea began to answer. However, she paused during her response and looked at Bailey for help in which Bailey responded "correctly identify" and continued with, "So our claim was people can tell a difference between Coke and Pepsi. So our null hypothesis would be that the proportion of people that can tell a difference would be greater than 0.5" (Video Transcript, 4/14/15). When Megan asked a question a few moments later about why the alternative hypothesis was greater than 0.5, Bailey quickly explained by saying,

> Like, 0.5 is no preference. So if they just randomly choose one, and it could have been either one, then they didn't know. But if you can tell a difference, then that means there is something guiding your response. So it's not like you just chose something like heads or tails. (Video Transcript, 4/14/15)

In addition to answering questions and providing suggestions within her group discussions, Bailey also answered questions in whole-class discussions. She was often the spokesperson for her group when called upon. For example, when Dr. Anderson called on Bailey's group to describe the study they had designed, Bailey answered,

> We said we were going to set up a booth somewhere on campus, and we are going to hand each person two cups. One that had number one and one that has number two. And we are going to ask them which one was Coke and which one was Pepsi. (Video Transcript, 4/14/15)

Bailey's role of providing knowledge and information to the group continued throughout the lesson. When writing on their Coke/Pepsi lesson worksheet, both Megan and Chelsea looked at Bailey's paper to read her answer (Field Notes, 4/14/15).

Use of Argumentation During the Coke/Pepsi Lesson

Throughout the Coke/Pepsi lesson, Bailey's participation in argumentation consisted mostly of claims. For example, during a discussion of what is important to do when solving a statistical problem, Bailey made claims such as "figure out the right test," "interpretation," and "having a clear question" (Video Transcript, 4/14/15). She did not use any of the other components of Toulmin's Model of Argument, such as data, to support her claims, during this discussion.

In another conversation during the Coke/Pepsi lesson, the instructor asked groups to review statistical tests learned throughout the semester. Bailey again made a claim of "more than one sample you use chi-square for proportions," without using data to support her claim (Video Transcript, 4/14/15). After this, Bailey took out her calculator and used it as a data source for her next claim of "confidence intervals" (Video Transcript, 4/14/15). This was the first time Bailey used a data source when making a claim. However, Bailey did not mention her calculator as her data source when making her claim, nor did she use a warrant to link the calculator as her data to her claim.

During the group discussion in which the group members decided on the design of their study, Bailey made the claim, “Our sample is 16,” after counting the number of students in the classroom (Video Transcript, 4/14/15). In this case, Bailey used the number of students in the classroom as her data. In this same conversation, Bailey made several other claims that were not supported by data. Some of these claims included, “A 50-50 chance that they would get it right or they would get it wrong if they can’t tell a difference [between Coke and Pepsi],” “Our null hypothesis would be that the proportion of people that can tell a difference would be greater than 0.5,” and “our research would be greater than 0.5 for people who got it” (Video Transcript, 4/14/15). Bailey also made a rebuttal to a claim made by Megan during this conversation when she said, “We already know they [Coke and Pepsi] are not the same product, so we already know there is a difference” (Video Transcript, 4/14/15).

In a later conversation in reference to which test should be used to analyze the data collected by the class, Bailey made the claims to use a “one-prop *z*-test” to analyze the class’ data and “It’s easy to interpret the one-prop *z* because the interval doesn’t even matter for the *p*-test” (Video Transcript, 4/14/15). After a rebuttal by Chelsea clarifying the importance of the significance level and *p*-value, Bailey made a qualifier/new claim when she said, “So this *p*-value it [the calculator] gave us is way lower than our significance level. So we are going to reject the null [hypothesis]” (Video Transcript, 4/14/15). Bailey later made a claim for her interpretation of the results from the study when she stated, “At the 5% confidence level, we have strong evidence to suggest that the population proportion of [University] students who can correctly identify a drink as Coke or Pepsi is higher than 50%” (Video Transcript, 4/14/15).

Overall, Bailey made several unsupported claims throughout the Coke/Pepsi lesson. When she did use a data source when making a claim, she used her calculator. Bailey did not explicitly identify any data source, nor identify warrants to link her data to her claims. However, she did provide a rebuttal and qualifier in response to her group members. Her use of argumentation occurred within all stages of the statistical problem-solving process.

Although Bailey provided many claims for her group throughout the lesson, she did not identify data and warrants for those claims. She also did not use any of the terminology for argumentation throughout the entire Coke/Pepsi lesson. Although Bailey appeared to view argumentation in terms of a model instead of a disagreement between people, she did not use the proper terminology of this model and identify components of the model as she used them. Therefore, Bailey's level of understanding of argumentation was coded as a Level 1.

Post Coke/Pepsi Lesson

In this section, Bailey's data from writing prompt 3 and her interview are provided.

Writing Prompt 3. At the conclusion of the Coke/Pepsi lesson, Bailey responded to a final writing prompt that asked, "Do you think you used argumentation in the Coke/Pepsi Taste Testing lesson? If so, how? If not, why did you not use it?" (Writing Prompt 3, 4/14/15). She responded,

> I definitely did. I had to establish my opinion and why I thought it was correct and used a warrant to link my data to my claim. The test we conducted help[ed] to provide good evidence to support my claim, and ultimately I was able to

> interpret my results and offer it to possible opposing views. (Writing Prompt 3, 4/14/15)

In her response to this writing prompt, Bailey directly connected her use of argumentation during the Coke/Pepsi lesson to the statistical problem-solving process. She related the argumentation terms to how she believed she used argumentation in the lesson through describing how each term was related to some part of the lesson. Because she related her use of argumentation to only the experiment that occurred within the lesson, it is unclear as to whether Bailey recognized her use of argumentation outside of the statistical problem-solving process. Also, Bailey once again incorporated all of the argumentation vocabulary used in the simple model of Toulmin's Model of Argument and described the relationship of the model. Because of this, Bailey's level of understanding of argumentation was coded as a Level 2.

Interview. Two days after she participated in the Coke/Pepsi lesson, Bailey completed an interview with the researcher. In this interview, she did not acknowledge her role as a significant contributor of knowledge for the group. For example, after saying her idea was to have a booth, Bailey responded to being asked if her group members agreed with her plan or went with someone else's plan to present to the class by saying,

> Somebody else's plan was selected, but we ended up choosing to have a booth somewhere on campus actually, and each person would have Coke and then Pepsi, but not in that specific order. And then we would see if they figured it out by comparing the two. (Bailey Interview Transcript, 4/16/15)

In reference to her or her group's use of argumentation during the Coke/Pepsi lesson, Bailey appeared to have a difficult time identifying argumentation. For example, when asked if she believed her group used argumentation to decide which study design to present to the class, Bailey responded,

> Yea, we did. We used the model that we went over in class. So we just made sure that, like, when we went over certain things, we would like say something or, like, oh no let's fix that. Let's fix our claim or whatever. But our claim stayed the same I think. Yea. (Bailey Interview Transcript, 4/16/15)

In this response, Bailey referred to Toulmin's Model of Argument and indicated that was how argumentation was used when deciding which study design to present to the class. However, she did not then explain how the model was used in this manner. Instead, she indicated her group would fix something if not correct, and that the claim did not change for the study. This does not provide an explanation for how Toulmin's Model of Argument was used to decide which study design was presented to the class like Bailey initially answered.

Similarly, when asked if her group used argumentation to come to a conclusion about whether or not a person could tell a difference between Coke and Pepsi, Bailey answered,

> Umm, I don't know specifically about my group, but as the class as a whole when we were discussing it. Because I think we were lost at that point, but then when Dr. Anderson was going over it, we were like, oh yea, that makes sense. (Bailey Interview Transcript, 4/16/15)

In this response, Bailey indicated the class as a whole used argumentation when coming to the conclusion about whether a person could tell the difference between Coke and Pepsi. However, she again did not provide further explanation to support how argumentation was used to do this. Instead, she explained that Dr. Anderson was what helped the class come to a conclusion in the Coke/Pepsi lesson.

In reference to how her group used argumentation to answer the discussion questions posed by Dr. Anderson, as well as how she personally engaged in argumentation, Bailey described argumentation differently. Her response to how her group used argumentation to answer the discussion questions posed by Dr. Anderson was,

> Yes, but not - I don't know if it's the same argumentation we were using in class, but it's sort of like we would tell her something and she would be like why, so we would have to argue our point even when we were answering questions about argumentation. (Bailey Interview Transcript, 4/16/15)

In reference to how she personally engaged in argumentation, Bailey responded,

> Yea, when we were picking which study to design, we were sort of like arguing our points. Like I think this one is better because it has randomization. That kind of thing. And then also when we were answering some of the individual questions, our group would come up with maybe two people had different answers, and then we would be like I think mine is correct because and I think yours is correct because. (Bailey Interview Transcript, 4/16/15)

In both of her responses, Bailey described argumentation as a way of providing her opinion to her group members or to the instructor and having a discussion about who was correct. Although she described some of the components of argumentation in her responses, such as claims and data, she never used any of the terminology of Toulmin's Model of Argument nor did she discuss using a source of data to support her point to her group members.

In addition to answering questions during the interview, Bailey was shown three clips from the Coke/Pepsi lesson and asked to identify any argumentation that occurred during each clip. Of the three clips, Bailey believed she saw argumentation in one clip and identified what she believed was that argumentation. In the first clip, she described the group's use of argumentation as follows.

> So yea in that clip I think definitely we were using argumentation, and it was, I don't know. I think that we did a really good job with it there because we would always be asking a question. Well, like somebody would bring their idea forth, and then I would be like why do you do that, and she would explain why and then I would be like, well, what about this, and we would combine our two ideas. (Bailey Interview Transcript, 4/16/15)

When asked to identify the data in the argumentation from the clip, she answered,

> I think that making sure the things were random was a big factor for me, and also making sure there wasn't too many confounding variables inside our own study. So like if we are going to do the same people twice, we wouldn't want them to be like well he did it, and then we give them another one immediately afterwards, because that might confound the experiment. So we would have like everybody would get one, and then everybody would go back in line and get another one. (Bailey Interview Transcript, 4/16/15)

In this clip, Bailey again described some components of argumentation, such as claims, rebuttals, and qualifiers. However, when asked to identify the data used in the argumentation that she believed occurred in the video clip, Bailey instead described how data were collected for the Coke/Pepsi lesson experiment. She did not describe the data used to support her statements from the video clip. She also did not use any of the argumentation terminology during her description of argumentation. Therefore, Bailey's level of understanding of argumentation during the interview was coded as a Level 1.

Summary

In summary, prior to being introduced to argumentation, Bailey described some of the components of argumentation, but she did not use the correct terminology for any of the components of Toulmin's Model of Argument when asked to describe argumentation. Therefore, her level of understanding of argumentation was coded as a Level 1. After engaging in argumentation training, Bailey described all three components of the simple structure of Toulmin's Model of Argument and two of the three components of the complex structure of Toulmin's Model of Argument. This time she also included the correct argumentation terminology when describing the components. Therefore, Bailey's understanding of argumentation was coded as a Level 2. While participating in the Coke/Pepsi lesson, Bailey's use of argumentation consisted mostly of making claims that were unsupported by data, and, therefore, not connected to that data by a warrant. Based on her use of argumentation during the Coke/Pepsi lesson, Bailey's understanding of argumentation during the Coke/Pepsi lesson was coded as a Level 1. After participating in the Coke/Pepsi lesson, Bailey stated that she believed her group used argumentation during the Coke/Pepsi lesson and described and named all three components in the simple structure of Toulmin's Model of Argument in her description of argumentation. Therefore, her level of understanding of argumentation was coded as a Level 2. Also after participating in the Coke/Pepsi lesson, Bailey indicated in an interview with the researcher that she believed there was argumentation in the first video clip shown. However, in her description of the argumentation in this clip, Bailey described some of the components of argumentation, but did not include the proper argumentation terminology in her description. When asked to identify the data used in the argumentation from the clip, she described how data was collected for the experiment. Therefore, Bailey's level of understanding of argumentation during the interview was coded as a Level 1. Figure 5 below provides a graph of Bailey's progression of understanding argumentation throughout the study. Her highest levels of understanding of argumentation occur when writing about argumentation while her lowest levels of understanding of argumentation occur prior to argumentation training, when using argumentation, and when asked to identify or discuss argumentation.

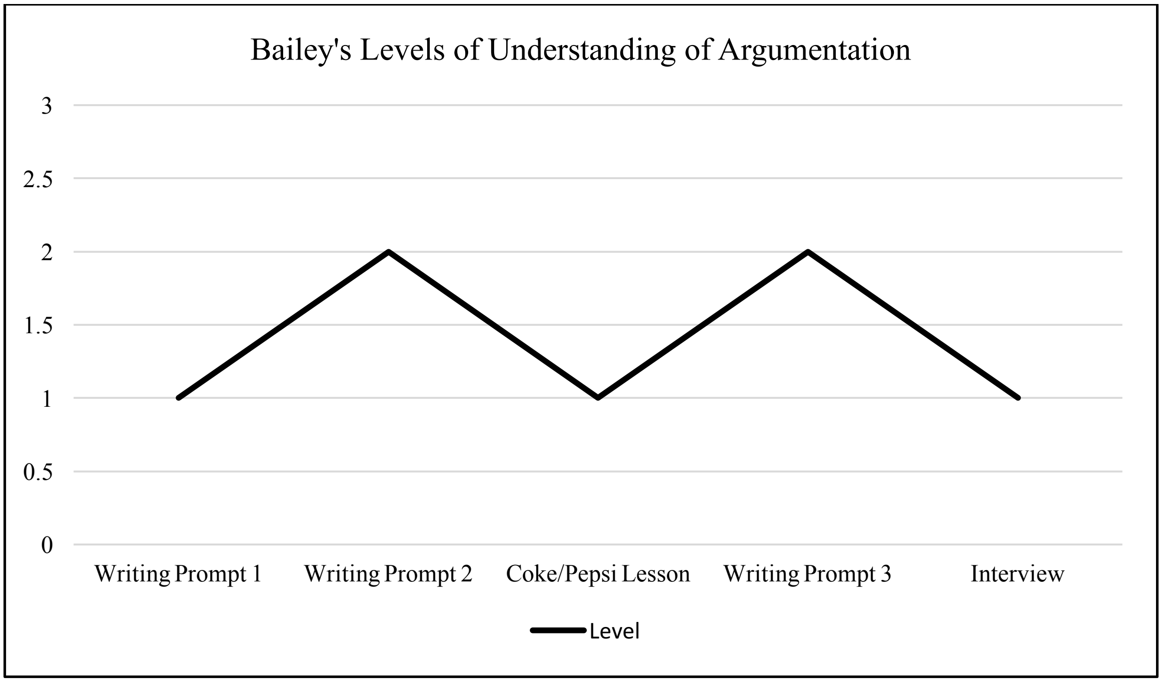

Figure 5. Progression of Bailey's levels of understanding of argumentation throughout the study.

Participant Three – Megan

In this section a description of Megan's data from prior to argumentation training, during argumentation training, post argumentation training, Coke/Pepsi lesson, and post Coke/Pepsi lesson is provided.

Prior to Argumentation Training

During an observation of students working in small groups prior to argumentation training, the researcher identified Megan as a less active member of group work and discussions. Megan participated in some group discussions by asking a question when she was confused (Field Notes, 3/26/15). However, she mostly looked through her notes and only occasionally provided input in the group discussions (Field Notes, 3/26/15).

After being identified as a less active participant in group discussions by the researcher, Megan was given a writing prompt that asked, “In your opinion, what is argumentation?” (Writing Prompt 1, 4/7/15). Her response to this writing prompt stated, “Argumentation is when two people have different opinions on a certain topic. There are different ways to argue with a person. You can be respectful or disrespectful in an argument. People argue every day about things. Arguments can strain or ruin relationships you have with people, but some people are able to respect others’ opinions and not let the arguments escalate too much” (Writing Prompt 1, 4/7/15). In her response, Megan described argumentation as an argument or disagreement between two people. Therefore, Megan’s level of understanding of argumentation was coded as a Level 0. Megan did not describe or mention any of the components of Toulmin’s Model of Argument.

Argumentation Training

After the first writing prompt was given, Megan participated in argumentation training with the rest of her classmates. During this training, her role in the group was more as an observer than that of an active participant (Field Notes, 4/9/15). Megan did not participate in any of the whole-class discussions throughout the entire training (Field Notes, 4/9/15). She was mainly quiet during the entire training and occasionally appeared to take notes (Field Notes, 4/9/15).

Post Argumentation Training

At the conclusion of the argumentation training, Megan was given a second writing prompt that asked, “What do you think argumentation is now?” (Writing Prompt 2, 4/9/15). Her response to this writing prompt stated, “Making a claim about something. Then having evidence to support that claim. Then making a warrant, which connects data with the claim. This is performed every day in several situations" (Writing Prompt 2, 4/9/15). In her response to this writing prompt, Megan described the simple model of Toulmin's Model of Argument and mentioned all of the components of the model, none of which were mentioned in Writing Prompt 1 when asked to define argumentation (Writing Prompt 2, 4/9/15). Therefore, Megan's response was coded as a Level 2 on the levels of understanding of argumentation.

Engagement in the Coke/Pepsi Lesson

During the next class meeting following the argumentation training, Megan engaged in the Coke/Pepsi lesson. During the first group discussion in which the instructor asked the groups to discuss within each group what is important to do when solving a statistical problem, Megan did not participate in the discussions (Video Transcript, 4/14/15). She sat back in her chair while Chelsea and Bailey engaged in conversation (Field Notes, 4/14/15). When the instructor asked the group a question during the whole-class discussion, Megan looked at Bailey (Field Notes, 4/14/15). After the instructor asked students to discuss with their small groups different statistical tests they have learned, Megan took out her notes and turned pages until she stopped and stared down at a page while Chelsea and Bailey named different statistical tests (Field Notes, 4/4/15). During their conversation, the instructor approached and asked the group, "What corresponds to the test?" (Video Transcript, 4/14/15). Bailey answered, "Umm, confidence intervals" (Video Transcript, 4/14/15). A few moments later Megan said to Bailey, "You were right" (Video Transcript, 4/14/15). This was the first time Megan spoke during the Coke/Pepsi lesson. After this, Megan began to have an off-topic conversation with Chelsea and Bailey (Field Notes, 4/14/15).

When the instructor began the whole-class discussion about the statistical tests, Megan wrote in her notebook and looked at the board (Field Notes, 4/14/15). She appeared to be taking notes on what the instructor was saying and writing on the board (Field Notes, 4/14/15). During the whole-class discussion, Dr. Anderson called on Megan to answer a question by saying, “So then what’s the null [hypothesis] for the means? Megan, what was it?” (Video Transcript, 4/14/15). Megan responded, “I didn’t say anything,” and Bailey followed with an answer (Video Transcript, 4/14/15). The instructor then asked Megan a second question by saying, “Now, Megan, how many choices do I have for my alternative [hypothesis]?” (Video Transcript, 4/14/15). Megan responded, “Uh, three,” and Dr. Anderson replied, “Yes. And so it would be greater than, less than, and not equal to” (Video Transcript, 4/14/15).

The instructor ended the whole-class discussion by asking the groups to design a study to determine if people could correctly identify a drink as either Coke or Pepsi. After Chelsea initiated the discussion within the group, Megan responded, “Yea, so right now what we are doing is deciding how we would go about designing a study” (Video Transcript, 4/14/15). Chelsea replied, “Oh ok,” and Megan continued by saying, “You have to have people. Your sample,” (Video Transcript, 4/14/15). Bailey and Chelsea then began to talk back and forth discussing the design of their study while Megan appeared to watch and listen (Field Notes, 4/14/15). She interrupted the conversation in which Bailey and Chelsea were discussing the null hypothesis and significance level for their study to ask the question, “But aren't we supposed to be deciding how we would do a test though? We aren't supposed to be finding all this” (Video Transcript, 4/14/15). Chelsea responded, “But we are just kind of saying to design the study we would have”

(Video Transcript, 4/14/15). After this, the instructor walked over and Megan asked, "I thought we were going to conduct a test. Is that what you were trying to say?" (Video Transcript, 4/14/15). Dr. Anderson agreed and walked away.

The group continued their conversation about the design of their study. Megan took a more active role in the conversation than previously during the lesson by contributing ideas. For example, she suggested, "We go on campus and have like a booth for people to come up and try it" (Video Transcript, 4/14/15). She also suggested to have participants try both Coke and Pepsi and to "put numbers like one or two [on the samples]" (Video Transcript, 4/14/15). When the instructor returned to the group, Megan spoke on behalf of the group and said, "We said that we would set up a booth on campus and then assign a number to a product and let them taste it and tell us which is which" (Video Transcript, 4/14/15).

After the instructor walked away, the group members began writing. Megan looked at Bailey's paper while she was writing (Field Notes, 4/14/15). She then asked, "Wouldn't the claim be that is there a difference?" (Video Transcript, 4/14/15). After being asked how she would word her claim, Megan responded, "I would say, like, is there a difference between Coke and Pepsi instead of people can tell a difference" (Video Transcript, 4/14/15). Bailey clarified when she responded, "Because they are not the same product so we already know there is a difference" (Video Transcript, 4/14/15). Once the group members continued writing, Megan again looked at Bailey's paper and asked a question referring to why the null hypothesis was greater than 0.5. Bailey answered her question once again. While Bailey and Chelsea were talking, Megan was looking at Chelsea's paper and writing. She then looked at Bailey's paper again and asked, “Did we pick 95 because that’s what we always do?” (Video Transcript, 4/14/15). Chelsea and Bailey both responded, “Yea,” but Bailey continued, “We could pick 90 because it’s not a life or death situation” (Video Transcript, 4/14/15). Megan continued to look at Bailey and Chelsea’s paper while they were writing when the group members were completing the Coke/Pepsi lesson worksheet.

When the group members discussed their interpretation from their statistical test, Megan asked, “So if it would have been - if the *p*-value would have been greater than, then we would have put it significantly. We would say, well, how would you word it?” (Video Transcript, 4/14/15). Chelsea responded by reading an example from her notes. Megan then stated to her group members, “I get this better after today. Like, I didn’t understand it at all before” (Video Transcript, 4/14/15).

Use of Argumentation During the Coke/Pepsi Lesson

Similar to Megan’s less active role in discussion during the Coke/Pepsi lesson, her use of argumentation was also limited throughout the Coke/Pepsi lesson. When Megan did participate in argumentation, she mostly made claims. Her first use of a component of argumentation occurred during a group discussion where group members decided on the design of their study. In this discussion, Megan made a claim, “You have to have people, like your sample” (Video Transcript, 4/14/15). During this discussion, Megan and Chelsea both interpreted the instructor’s directions differently. Megan used the instructor as a data source when she asked her, “I thought we were going to conduct a test. Is that what you were trying to say?” (Video Transcript, 4/14/15). However, she did not mention her use of the instructor as her data source to support her claim. Megan then provided a rebuttal to Chelsea’s claim about the research hypothesis when she said, “I would word it like you're saying if people can tell a difference, but you wouldn't know that unless you have already done it" (Video Transcript, 4/14/15). Megan provided another claim during this same conversation when she stated, "I would say is there a difference between Coke and Pepsi instead of people can tell a difference" (Video Transcript, 4/14/15). She also made the claim to use a "one-prop *z*-test," but this time she used her notes as a data source. However, Megan did not mention her notes as her data source when making her claim, nor did she use a warrant to link her data to her claim.

Overall, Megan rarely used components of argumentation during the Coke/Pepsi lesson. When she did use argumentation, she mostly made unsupported claims. However, Megan did use both her notes and the instructor as a source of data, though they were not identified by her in that way, and she did provide a rebuttal to another participant's claim. Therefore, Megan's level of understanding of argumentation was coded as a Level 1.

Post Coke/Pepsi Lesson

In this section, Megan's data from writing prompt 3 and her interview are provided.

Writing Prompt 3. At the conclusion of the Coke/Pepsi lesson, Megan responded to a final writing prompt that asked, "Do you think you used argumentation in the Coke/Pepsi Taste Testing lesson? If so, how? If not, why did you not use it?" (Writing Prompt 3, 4/14/15). She responded, "Yes, we made a claim, picked a warrant, collected data, and came to a conclusion about Coke versus Pepsi. We used argumentation and statistics to come to a conclusion about research" (Writing Prompt 3, 4/14/15). In her response to this writing prompt, Megan directly related argumentation to the statistical problem-solving process. Since she only related argumentation to the statistical problem-solving process, it is unclear whether or not Megan understood the role or use of argumentation outside of this process. However, Megan again mentioned all of the components by name of the simple structure of Toulmin's Model of Argument. Therefore, Megan's understanding of argumentation was coded as a Level 2.

Interview. Two weeks after she participated in the Coke/Pepsi lesson, Megan completed an interview with the researcher. When asked by the researcher if all of the group members agreed on a research plan to present to the class, Megan responded, "We kind of all had different ideas. We kind of had to communicate to explain what we both thought" (Megan Interview Transcript, 4/28/15). Megan then acknowledged the leadership role of Chelsea and Bailey when she said, "Both the people in my group are kind of like leaders in the class, so we kind of had to agree on different things" (Megan Interview Transcript, 4/28/15). Megan was then asked if she believed her group used argumentation to decide which study design to present to the class, and was reminded by the researcher that argumentation consisted of claims, data, and warrants. She responded.

> Yea, we did all of that. I mean we listened to other examples from other groups, and we came up with it on our own. But then we decided it based on what other groups were also saying because the whole class was kind of communicating. So we did that to decide the claim and all that. (Megan Interview Transcript, 4/28/15)

Megan appeared to have a difficult time understanding the question because she answered in reference to the whole-class discussion instead of her group discussion. She also indicated argumentation was used, but did not identify how or when it was used.

In reference to her or her group's use of argumentation during the Coke/Pepsi lesson to decide if a person could tell the difference between Coke and Pepsi, Megan answered, "Umm, because you had to compare, was it the sigma?" (Megan Interview Transcript, 4/28/15). The researcher responded, "The *p*-value to the alpha?" Megan continued, "Alpha, yes, sorry. *p*-value to the alpha, and if one was higher than the other, so yea we had to do that" (Megan Interview Transcript, 4/28/15). In her response, Megan again related argumentation directly to a component within the statistical problem-solving process. Similarly, when asked if her group used argumentation to answer the discussion questions posed by Dr. Anderson, Megan responded,

> Yes, because we had to find the claim and all the things within. And then all the groups had to come up with what they thought the claim was, and then we all had to combine it and see what the best outcome was. (Megan Interview Transcript, 4/28/15)

In both of these responses, Megan again referred to the overall research experiment when asked about argumentation instead of her individual use of argumentation during the Coke/Pepsi lesson. She directly related argumentation to the statistical problem-solving process.

Megan was also asked if she thought she personally engaged in argumentation during the Coke/Pepsi lesson to which she answered,

> Yes. In the very beginning we were coming up with the null [hypothesis] and we kind of argued on what the null [hypothesis] would be because I thought that, I am trying to think of the specific of what I thought and what they thought. I thought that it was more. I am trying to give you details. I think that one of us

> thought that we were actually comparing the two between each other and the other thought that we were just, I can't remember exactly what the other argument was, but we just had different ideas of what the null [hypothesis] would be and we had to conclude it. And then Dr. Anderson came and talked to us and told us what the best thing was. (Megan Interview Transcript, 4/28/15)

In her response, Megan described argumentation as a disagreement between herself and another student. Although she mentioned the instructor as clarifying the disagreement, she did not describe the instructor as the data source to support her claim. Similarly, she did not mention that either her or the other participant provided any type of support for their opinion prior to Dr. Anderson clarifying who was correct. In all of her responses, Megan appeared to have a difficult time explaining how argumentation was used.

In addition to answering questions during the interview, Megan was shown three video clips from the Coke/Pepsi lesson and asked to identify any argumentation that occurred during each clip. Of the three clips, Megan believed she saw argumentation in all three video clips. In the first video clip, Megan described the group's use of argumentation as follows:

> Umm, it was hard to hear exactly, but it sounded like it was kind of what I was talking about. I don't know if it were with the null [hypothesis] or what, but it was something about deciding what the study was about. And we were arguing about that, and then Dr. Anderson came over and told us which one was right. And then it also sounded like there was argumentation at the beginning. I might have heard it wrong, but it sounded like they were trying to decide if it were

> greater than, less than, or equal to on the alternative. (Megan Interview Transcript, 4/28/15)

In her response, Megan described argumentation as a form of decision making. However, it is unclear how she used the term arguing in her response because she did not describe what occurred as part of the arguing.

In the second clip, Megan identified what she believed to be the group's argumentation when she said, "Yea, between the two of them, like both of their different opinions about it" (Megan Interview Transcript, 4/28/15). When asked to identify the data in the argumentation from the clip, she responded, "Oh would it be the, like, because our thing was to do a booth that they had on campus. And they would test that" (Megan Interview Transcript, 4/28/15). In her response to this clip, Megan described the argumentation as two different opinions. However, when asked to identify the data of the argument, she once again related back to the statistical problem-solving process and described the data for the experiment in the Coke/Pepsi lesson instead of the data from the argument she identified in the clip.

In the third video clip, Megan identified what she believed to be the group's argumentation when she responded, "We were deciding if we should reject or accept the null [hypothesis]. And we were talking about why" (Megan Interview Transcript, 4/28/15). When asked to identify any data or warrants involved in making those decisions, she responded, "Yes, because that all comes to the final [decision]. You have to have all of that to get to your final [decision] to know if you can accept it or reject it [the null hypothesis] and all that stuff" (Megan Interview Transcript, 4/28/15). In her response to this clip, Megan again referred to argumentation as a form of decision making. When asked if she could identify data or warrants that occurred during the argumentation she identified, Megan once again related argumentation back to the statistical problem-solving process during the Coke/Pepsi lesson.

In all three video clips, Megan related argumentation to components of the overall research experiment for the Coke/Pepsi lesson. Additionally, in the first video clip, she also related argumentation to deciding who was right when "arguing" (Megan Interview Transcript, 4/28/15). When asked by the researcher to identify data or warrants, Megan provided answers that were related to the research experiment. Because of this, Megan's level of understanding of argumentation appeared to be between a Level 0 and Level 1. Although she implied that argumentation is a disagreement between two people, she also related argumentation to the statistical problem-solving process. Megan's understanding was somewhere between viewing argumentation as a disagreement and viewing argumentation as a model. Therefore, her level of understanding of argumentation during the interview was coded as a Level 0.5.

Summary

In summary, prior to being introduced to argumentation, Megan described argumentation as a disagreement between two people and did not mention any of the components of Toulmin's Model of Argument when asked to describe argumentation. Therefore, her understanding of argumentation was coded as a Level 0. After engaging in argumentation training, Megan mentioned and described all three components of the simple structure of Toulmin's Model of Argument when asked to define argumentation. Therefore, her level of understanding of argumentation was coded as a Level 2. While participating in the Coke/Pepsi lesson, Megan's use of argumentation was limited. When she did participate, her use of argumentation consisted mostly of making claims that were unsupported by data, and, therefore, not connected to that data by a warrant. Based on her use of argumentation during the Coke/Pepsi lesson, Megan's understanding of argumentation during the Coke/Pepsi lesson was coded as a Level 1. After participating in the Coke/Pepsi lesson, Megan stated that she believed her group used argumentation during the Coke/Pepsi lesson and mentioned all three components in the simple structure of Toulmin's Model of Argument in her description of argumentation. Therefore, her level of understanding of argumentation was coded as a Level 2. Also after participating in the Coke/Pepsi lesson, Megan indicated in an interview with the researcher that she believed there was argumentation in all three video clips shown. However, in her description of the argumentation in the video clips, Megan described argumentation more as a disagreement between two people. She then related the argumentation she identified to the statistical problem-solving process that occurred in the Coke/Pepsi lesson. She did not use any proper argumentation terminology in her descriptions. Therefore, Megan's level of understanding of argumentation during the interview was coded as a Level 0.5. Figure 6 below provides a graph of Megan's progression of understanding argumentation throughout the study. Megan's lowest level of understanding of argumentation occurred prior to the argumentation training. Her highest levels of understanding of argumentation occurred when asked to write about argumentation.

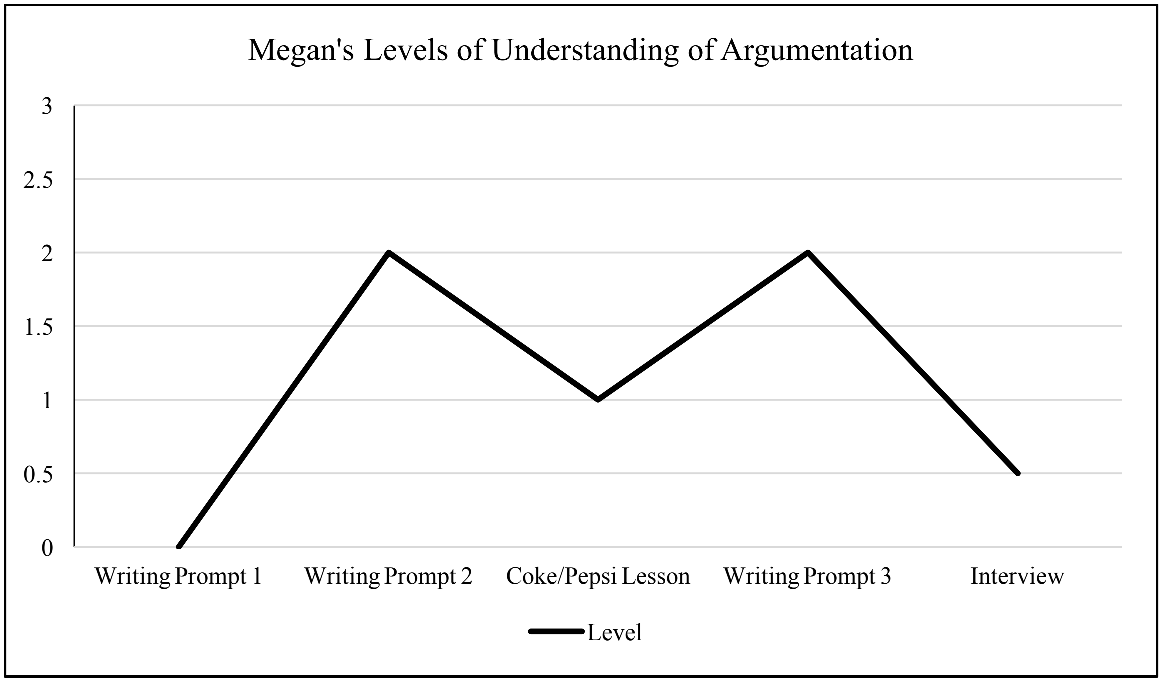

Figure 6. Progression of Megan's levels of understanding of argumentation throughout the study.

Cross-Case Analysis

The three cases were analyzed for similarities and differences that occurred among the three participants' knowledge of argumentation. These components were looked at from the view of participants' baseline knowledge of argumentation, knowledge of argumentation post argumentation training, use of argumentation during the Coke/Pepsi lesson, and knowledge of argumentation post Coke/Pepsi lesson. In what follows, these findings are discussed.

Baseline Knowledge of Argumentation

Prior to being introduced to argumentation, none of the three participants used the proper terminology for any of the components of argumentation when asked to describe argumentation. Both Chelsea and Bailey did describe some of the components of argumentation in their responses. Both Chelsea and Bailey described a type of situation that involved having a different opinion than someone else and supporting your opinion. However, Megan's response described more of a disagreement between two people. In all three cases, the participants appeared to be unaware of a formal structure of argumentation. Chelsea and Bailey's understanding of argumentation was coded as a Level 1 while Megan's understanding of argumentation was coded as a Level 0.

Knowledge of Argumentation Post Training

After engaging in argumentation training with their classmates, the participants were once again asked to describe argumentation. This time, all three participants mentioned and described at least two of the three components of the simple structure of Toulmin's Model of Argument. Bailey provided the most advanced description of argumentation by not only giving a complete description of the simple structure of Toulmin's Model of Argument, but also by providing a description that involved two of the three components of the complex structure of Toulmin's Model of Argument. Neither Chelsea nor Megan mentioned or described any of the components of the complex structure of Toulmin's Model of Argument. Both Bailey and Megan's understanding of argumentation was coded as a Level 2 while Chelsea's understanding of argumentation was coded as a Level 1.5.

Use of Argumentation During the Coke/Pepsi Lesson

While participating in the Coke/Pepsi lesson, none of the three participants made a complete argument consisting of a claim, data, and warrant. Their use of argumentation was extremely limited. When the participants did use components of argumentation, all three participants mostly made claims that were unsupported by data, and, therefore, not connected to that data by a warrant. Bailey and Chelsea provided the majority of the claims that were made. When data were used, the data source was not explicitly identified by the participant. For example, Chelsea made the claim, “You use the *t*-test for mean” (Video Transcript, 4/14/15). When she made this claim, she was reading from her notes. However, she never identified her notes as the data source to support her claim. Throughout the lesson, both Chelsea and Megan used their notes in this fashion while Bailey used her calculator as her data source. Since none of the participants identified their data, warrants were likewise never identified by the participants. However, all three participants provided a rebuttal during the lesson to another participant’s claim. Since all three participants used components of argumentation during the lesson, but none of the participants identified these components that were used, all three participants level of understanding of argumentation was coded as a Level 1 for the Coke/Pepsi lesson.

Knowledge of Argumentation Post Coke/Pepsi Lesson

After participating in the Coke/Pepsi lesson, the participants were asked about their use of argumentation during the lesson. All three participants believed the group used argumentation during the lesson. Similarly, all three participants mentioned and described all of the components of the simple structure of Toulmin’s Model of Argument in their description. Therefore, all three participants understanding of argumentation was coded as a Level 2.

In the days following the Coke/Pepsi lesson, the participants completed an interview with the researcher. In this interview, all three participants mentioned argumentation was used at some point. However, none of the participants identified a complete argument with the components of the simple structure of Toulmin's Model of Argument. For example, all three participants believed argumentation was used in the first video clip shown to them. However, Bailey described the argumentation as, "Somebody would bring their idea forth, and then I would be like why do you do that. And she would explain why. And then I would be like, well, what about this. And we would combine our two ideas" (Bailey Interview Transcript, 4/16/15). Chelsea described the argumentation as, "We were trying to come up with a test, and if someone had an idea then they would try and tell you why this idea would work and everything. And then other people would disagree. And then that person would then tell them, ok, well, what if we did this then" (Chelsea Interview Transcript, 4/16/15). Megan described the argumentation as, "It was something about deciding what the study was about, and we were arguing about that. And then Dr. Anderson came over and told us which one was right" (Megan Interview Transcript, 4/28/15). In all three cases, the participant believed argumentation was used by the group during the video clip. However, none of the three participants used the proper terminology for any of the components of the Toulmin's Model of Argument to describe the argumentation they believed was occurring. Both Chelsea and Bailey described some of the components of Toulmin's Model of Argument without using the terminology of the model.

In the second video clip, Megan was the only participant who indicated she believed there was argumentation. In the third video clip, both Chelsea and Megan indicated they believed there was argumentation. This time, Chelsea described her use of argumentation as explaining her idea by looking at her notes (Chelsea Interview

Transcript, 4/16/15). Megan described her use of argumentation as, "Deciding if we should reject or accept the null [hypothesis]" (Megan Interview Transcript, 4/28/15). Again, none of the proper terminology for the components of Toulmin's Model of Argument were mentioned in their description of the argumentation in the video clip. Chelsea described some of the components without using the terminology while Megan described argumentation as making a decision.

Throughout the interview, both Chelsea and Bailey often described components of argumentation without using the terminology of Toulmin's Model of Argument. Therefore, they were both coded as a Level 1 of understanding argumentation. Megan, however, described argumentation at times as a disagreement between two people. At other times, she related argumentation to the statistical problem-solving process. Therefore, Megan was coded as a Level 0.5 for her understanding of argumentation during the interview.

Summary

Through their engagement in both argumentation training and the Coke/Pepsi lesson, all three participants made gains in their knowledge of argumentation by being able to describe the simple structure of Toulmin's Model of Argument. However, none of the participants were able to engage in an argument during the Coke/Pepsi lesson. Similarly, the participants had difficulty identifying and describing argumentation when watching video clips from the Coke/Pepsi lesson. This difficulty can further be seen in the different levels of understanding of argumentation that each participant possessed as they progressed throughout the study, as seen in Figure 7.

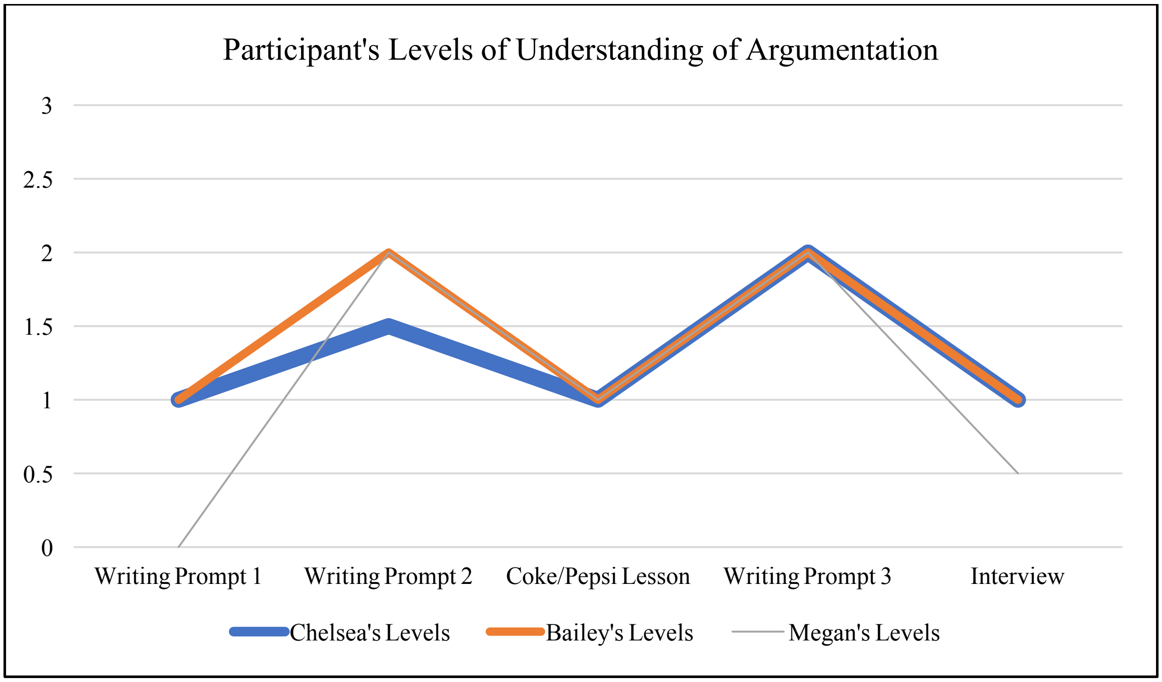

Figure 7. Progression of all three participants' levels of understanding of argumentation throughout the study.

Chapter Summary

In this chapter, a detailed description of the Coke/Pepsi Taste Test lesson was provided. Following this, each of the three cases were presented. Finally, a cross case analysis of the three cases was provided.

Prior to being introduced to argumentation, both Chelsea and Bailey described some components of argumentation in their description of argumentation, and their understanding of argumentation was coded as a Level 1. Megan, however, described argumentation as more of a disagreement between two people, and her understanding of argumentation was coded as a Level 0.

After engaging in argumentation training, both Bailey and Megan described all three components in the simple structure of Toulmin's Model of Argument and used proper terminology for the components in their description. Their understanding of argumentation was coded as a Level 2. Chelsea also described the three components of the simple structure, but only used the proper terminology for two of the components. Her understanding of argumentation was coded as a Level 1.5.

While participating in the Coke/Pepsi Taste Test lesson, all three participants made mostly claims that were unsupported by data when using argumentation. Since all three participants used components of argumentation during the lesson, but none of the participants identified these components that were used, all three participants' level of understanding of argumentation was coded as a Level 1 for the Coke/Pepsi lesson.

After participating in the Coke/Pepsi lesson, all three participants indicated they believed they used argumentation during the lesson and described the simple structure of Toulmin's Model of Argument using the proper terminology for the components. Therefore, the three participants' understanding of argumentation were all coded as a Level 2.

In the days following the Coke/Pepsi lesson, the participants completed an interview with the researcher. During this interview, all three participants indicated they believed argumentation was used by their group during the Coke/Pepsi lesson, but none of the three identified an argument. Both Chelsea and Bailey's level of understanding of argumentation was coded as a Level 1 and Megan's was coded as a Level 0.5.

The descriptions in the three cases revealed that through their engagement in both argumentation training and the Coke/Pepsi lesson, all three participants made gains in their knowledge of argumentation by being able to describe the simple structure of Toulmin's Model of Argument. However, none of the participants engaged in an argument during the Coke/Pepsi lesson, and all three participants had difficulty identifying arguments during the interview with the researcher. In the next chapter, a discussion of these results will be provided.

CHAPTER FIVE: DISCUSSION

Introduction

As part of the recent history of the mathematics education curriculum, reasoning and argument have been emphasized throughout mathematics curriculum standards (c.f. CCSSI, 2010; NCTM, 1989, 2000). Due to its emphasis on reasoning and use of argument, argumentation is one possible way to improve students' reasoning and argument skills in the mathematics classroom. However, when used in mathematics, argumentation has been commonly limited to proofs (e.g., Boero, 1999; Conner, 2007; Smith, 2010). Therefore, it is important to introduce argumentation through venues other than proof in order to help build students' understanding of argumentation. One possible place to develop students' understanding of argumentation for use in mathematics is through engaging students in the statistical problem-solving process.

This chapter will begin with a review of the purpose and methodology of the study. Then, a review of the results of the study, which will include a review of each of the three individual cases and the cross-case analysis, will be provided. Following this, a discussion of the results of the study and implications for mathematics education will be provided. Last, suggestions for future research will be discussed.

Research Purpose

The purpose of this study was to investigate students' engagement in the statistical problem-solving process as a way to develop students' understanding of argumentation. Specifically, this study sought to examine students' understanding of argumentation while being engaged in the four components of the statistical problem-solving process. In order to do this, an exploratory case study was used. The specific research question for this study was: How does students' engagement in the statistical problem-solving process support students' understanding of argumentation, if at all?

Review of Methodology

In order to explore and understand how engaging students in the statistical problem-solving process could enhance students' understanding of argumentation in a statistics classroom, this study utilized an exploratory qualitative case study, as defined by Yin (2014). Due to the desire to examine different types of students' understanding of argumentation while engaging in the statistical problem-solving process, a multiple case study approach was used. This holistic multiple case study included three cases, one case for each of the three participants in the introductory statistics class. The participants were selected by the researcher for the study based on prior observations in order to gain a sample of three participants as follows: one participant who was very verbal and active in group discussions; one participant who was moderately active in group discussions; and one participant who was rarely active in group discussions.

Data for the study was collected from five sources: lesson documents, writing prompts, interviews, observations, and the researcher. The researcher introduced the instructor of the introductory statistics course, Dr. Anderson, to argumentation through an informal training in which the researcher and instructor discussed Toulmin's Model of Argument and examined examples of argumentation. The instructor then engaged the participants in the argumentation training during a class period. During this training, the participants were introduced to Toulmin's Model of Argument and worked with one another to create and participate in mock arguments with one another. The participants completed their first of three writing prompts in a class period prior to this argumentation training. Immediately following the argumentation training, the participants completed the second writing prompt. The participants then completed the Coke/Pepsi Taste Test lesson during a single 85-minute class period. While the participants participated in the Coke/Pepsi Taste Test lesson, the researcher both video recorded the participants as well as took observational field notes. At the end of this lesson, the participants completed the third writing prompt. Following the Coke/Pepsi lesson, the three participants completed an interview with the researcher during which they were asked questions about their engagement in argumentation during the lesson. They were also shown three video segments from the lesson and asked to identify any argumentation they saw in the segments.

The data collected were analyzed following the five general steps of analyzing qualitative data as described by Creswell (2013): organizing the data; reading and memoing; describing, classifying, and interpreting data into codes and themes; interpreting the data; and representing and visualizing the data. Toulmin's Model of Argument was used to analyze the discussion among participants during the Coke/Pepsi lesson. An analysis of when argumentation was used in the lesson by each participant was then compared to the participant answers from the interview questions and their video analysis for consistency with the results. The interview was also examined to determine each participant's view of how argumentation was used during the lesson and to help determine each participant's understanding of argumentation. From the detailed description of each participant that was formed from these analyses, codes emerged in relation to the level of understanding that occurred for each participant. The cases were then reviewed once again, and each participant's understanding of argumentation was coded based upon these levels of understanding argumentation.

Review of Results

The results for this study were presented through the three individual case reports as well as the cross-case report presented in Chapter Four. In what follows, a review of the three individual cases and the cross-case analysis will be provided.

Chelsea's Results

Chelsea was a 20-year-old female majoring in pre-nursing. She was selected by the researcher because she was very active in group discussions as observed by the researcher. Prior to being introduced to argumentation, Chelsea described some of the components of argumentation but did not use any of the proper terminology of Toulmin's Model of Argument in her description. Therefore, her level of understanding of argumentation was coded as a Level 1. After engaging in argumentation training, Chelsea completed the second writing prompt where she again described components of Toulmin's Model of Argument. This time, Chelsea included the proper terminology for some of these terms, so her level of understanding of argumentation was coded as a Level 1.5.

During the Coke/Pepsi lesson, Chelsea's use of argumentation consisted mostly of making claims that were unsupported by data and, hence, not connected to that data by a warrant. Therefore, Chelsea's level of understanding of argumentation during the Coke/Pepsi lesson was coded as a Level 1. Following the Coke/Pepsi lesson, Chelsea completed a third writing prompt where she described and mentioned by name all three components of Toulmin's Model of Argument in her description of argumentation. Therefore, Chelsea's level of understanding of argumentation was coded as a Level 2.

In an interview with the researcher, Chelsea indicated that she believed there was argumentation in both the first and third video segments shown. In her description of the argumentation she saw in these segments, Chelsea described some of the components of Toulmin's Model of Argument, but she did not use the proper terminology in reference to these components. After being asked by the researcher to identify the data, Chelsea identified the data in the video segment; however, she did not explicitly identify any claims or warrants. Therefore, Chelsea's level of understanding of argumentation during the interview was coded as a Level 1.

Overall, Chelsea's highest levels of understanding of argumentation occurred when writing about argumentation. Her lowest levels of understanding of argumentation occurred prior to argumentation training, when using argumentation, and when asked to identify or discuss argumentation.

Bailey's Results

Bailey was an 18-year-old female majoring in biology with a concentration in physiology. She was selected by the researcher because she was moderately active in group discussions as observed by the researcher. Prior to being introduced to argumentation, Bailey described some of the components of argumentation, but did not use any of the proper terminology of Toulmin's Model of Argument in her description. Therefore, her level of understanding of argumentation was coded as a Level 1. After engaging in argumentation training, Bailey completed the second writing prompt where she again described components of Toulmin's Model of Argument, including all three components of the simple structure of the model and two components of the complex structure. This time, Bailey also included the proper terminology for these terms, so her level of understanding of argumentation was coded as a Level 2.

During the Coke/Pepsi lesson, Bailey's use of argumentation consisted mostly of making claims that were unsupported by data, and, hence, not connected to that data by a warrant. Therefore, Bailey's level of understanding of argumentation during the Coke/Pepsi lesson was coded as a Level 1. Following the Coke/Pepsi lesson, Bailey completed a third writing prompt where she described and mentioned by name all three components of the simple structure of Toulmin's Model of Argument in her description of argumentation. Therefore, Bailey's level of understanding of argumentation was coded as a Level 2.

In an interview with the researcher, Bailey indicated that she believed there was argumentation in the first video segment shown. In her description of the argumentation she saw in the segment, Bailey described some of the components of Toulmin's Model of Argument, but she did not use the proper terminology in reference to these components. She did not explicitly identify components of the argument, such as claim, data, or warrant. After being asked by the researcher to identify the data, Bailey described how data were collected for the experiment in the Coke/Pepsi lesson instead. Therefore, Bailey's level of understanding of argumentation during the interview was coded as a Level 1.

Overall, Bailey's highest levels of understanding of argumentation occurred when writing about argumentation. Her lowest levels of understanding of argumentation occurred prior to argumentation training, when using argumentation, and when asked to identify or discuss argumentation.

Megan's Results

Megan was a female and appeared to be a traditional college student in terms of age. She was selected by the researcher because she rarely participated in group discussions as observed by the researcher. Prior to being introduced to argumentation, Megan described argumentation as a disagreement between two people and did not mention or describe any of the components of Toulmin's Model of Argument. Therefore, her level of understanding of argumentation was coded as a Level 0. After engaging in argumentation training, Megan completed the second writing prompt where she described all three components of the simple structure of Toulmin's Model of Argument. Megan also included the proper terminology for some of these terms, so her level of understanding of argumentation was coded as a Level 2.

During the Coke/Pepsi lesson, Megan's participation and use of argumentation was limited. Her use of argumentation consisted mostly of making claims that were unsupported by data, and, hence, not connected to that data by a warrant. Therefore, Megan's level of understanding of argumentation during the Coke/Pepsi lesson was coded as a Level 1. Following the Coke/Pepsi lesson, Megan completed a third writing prompt where she again described and mentioned by name all three components of the simple structure of Toulmin's Model of Argument in her description of argumentation. Therefore, Megan's level of understanding of argumentation was coded as a Level 2.

In an interview with the researcher, Megan indicated that she believed there was argumentation in all three video segments shown. In her description of the argumentation she saw in these segments, Megan described argumentation as a disagreement between two people. When asked by the researcher to identify the data for the argumentation shown in the segment, Megan then related argumentation to the statistical problem-solving process that occurred in the Coke/Pepsi lesson. Therefore, Megan's level of understanding of argumentation during the interview was coded as a Level 0.5.

Overall, Megan's highest levels of understanding of argumentation occurred when writing about argumentation. Her lowest level of understanding of argumentation occurred prior to the argumentation training.

Cross-Case Results

Prior to being introduced to argumentation, both Chelsea and Bailey described some of the components of argumentation. They both described a type of situation that involved having a different opinion than someone else and supporting your opinion during the discussion. Megan, however, described argumentation as a disagreement between two people. None of the three participants used proper argumentation terminology in their description. After engaging in argumentation training, all three participants described the three components of the simple structure of Toulmin's Model of Argument. In addition, Bailey also described two of the components of the complex structure of the model as well. This time, all three participants used at least some of the proper terminology of argumentation in their descriptions.

During the Coke/Pepsi Taste Test lesson, none of the participants made a complete argument consisting of a claim, data, and warrant. Instead, the participants mostly made claims that were unsupported by data and not connected to the data by a warrant. The participants used data at times, such as their notes or calculator, but the data was not identified by the participant. All three participants provided a rebuttal during the lesson. In the writing prompt following the Coke/Pepsi lesson, all three participants indicated they believed argumentation was used in the lesson. In their description of argumentation, all three participants mentioned and described all of the components of the simple structure of Toulmin's Model of Argument.

In the days following the Coke/Pepsi lesson, the participants completed an interview with the researcher during which they all three indicated they believed argumentation was used at some point by their group in the Coke/Pepsi lesson. However, none of the participants identified a complete argument with the components of the simple structure of Toulmin's Model of Argument. Also, none of the participants used any of the terminology when discussing the argumentation in the video segments.

Overall, through their engagement in both argumentation training and the Coke/Pepsi lesson, all three participants made gains in their knowledge of argumentation by being able to describe the simple structure of Toulmin's Model of Argument. However, none of the participants engaged in an argument during the Coke/Pepsi lesson, and the participants had difficulty identifying and describing argumentation when watching video clips from the lesson.

Discussion

The research question for this study was: How does students' engagement in the statistical problem-solving process support students' understanding of argumentation, if at all? From reviewing the results presented in Chapter Four, three main discussion points emerged. These three points were related to engagement in argumentation training and the statistical problem-solving process as a means to support an increased understanding of argumentation in all participants; the use of argumentation during the Coke/Pepsi Taste Test lesson; and the use of the levels of understanding of argumentation for coding throughout the study. In what follows, a discussion related to each of these three points will be provided.

In reference to the research question, after engaging in argumentation training and the statistical problem-solving process through the Coke/Pepsi Taste Test lesson, all three participants displayed an increased understanding of argumentation. Evidence for this increased understanding was found in their responses to the three writing prompts. However, engagement in argumentation training and the statistical problem-solving process was not sufficient for the participants to develop a deep understanding of argumentation. Evidence of this can be found in participants' limited recognition of Toulmin's Model of Argument to occurrences within the statistical problem-solving process. Their recognition of the model and of components of argumentation did not extend outside of the statistical problem-solving process. Therefore, engagement in argumentation training and the statistical problem-solving process increased participants' understanding of argumentation; however, it was not sufficient to support them in developing a deep understanding of argumentation.

The second discussion point that emerged from the results was the participants' limited use of argumentation during the Coke/Pepsi Taste Test lesson. The participants' use of argumentation consisted mostly of claims with the occasional use of data to support the claim. This use of data came in the form of a classroom source (e.g., notes, calculator, instructor). These results are consistent with the results from Jimenez-Aleixandre et al.'s (2000) study that examined how students use the components of argumentation in a high school genetics class in Spain. The researchers found that students mainly made claims, but rarely supported these claims with data or warrants.

One possible reason the participants used claims without supporting those claims with data could be related to the role of authority in the classroom. When data were used, it was in the form of notes, calculator, and the instructor. Therefore, it appeared as if the instructor was viewed as a form of authority by the participants. If information was given to the participants by the instructor, either directly or through notes taken in a previous class, it was not questioned by the participants. Claims made by the participants related to these were, therefore, also not questioned by other group members. Similarly, because the participants viewed the instructor as an authority, when making claims based upon information from notes or other instances that involved the instructor as being the authority, the participants likely did not feel they needed to provide support for their claim.

Just as the role of the instructor's authority could have played a role in the participants' use of argumentation, the role of authority within the small group could have also played a role in the participants' use of argumentation. In a study that examined interactions and roles within a small group, Larriva (1998) found that in reference to who was in charge within a small group a student's perception of herself and of her other group members tended to influence the roles within the group. Of the two participants in the study, one of which had a low average and the other a high average, Larriva (1998) found that the student with the lower average was less assertive in her communication within the group and had a difficult time gaining the acceptance of her contributions from her group members. Although the participants in the current research study were selected based on how active they were in small group discussions and not on their grade average as with Larriva's (1998) study, the participant who was selected based on being rarely active in group discussions, Megan, likely viewed the other two participants as having the authority within the group, just as the student with the lower grade average did in Larriva's (1998) study. This is exemplified when Megan stated, "Both the people in my group are kind of like leaders in the class" (Megan Interview Transcript, 4/28/15). This role was further seen through the description of the Coke/Pepsi lesson. In many instances, Chelsea and Bailey stated a claim and were not questioned despite not providing data to support their claim. However, Megan was questioned or asked to explain her thinking by Chelsea and Bailey many times throughout the lesson, just as seen in Larriva's (1998) study. Therefore, through both the example of the role of authority of the instructor and the example of the role of authority within the group, it is believed that authority may influence engagement in argumentation. This influence of authority on engagement in argumentation, in turn, influences the development of students' understanding of argumentation. It appeared as if the participants did not view argumentation as a useful tool within their group discussions despite engaging in argumentation training where they created arguments as a way to support a claim. Instead, the participants were only able to see argumentation through the statistical problem-solving process.

The third discussion point that emerged from the results was the use of the levels of understanding of argumentation as a way to code the participants' understanding of argumentation throughout the study. The levels were defined based on how the participants described argumentation and recognized and used the proper terminology for argumentation within these descriptions. Due to the descriptive nature of the writing prompts, the levels of understanding seemed appropriate for coding the responses to these writing prompts. Additionally, because the participants were both describing and identifying argumentation in the interviews, the levels of understanding of argumentation also seemed appropriate for coding the interview responses as well. However, the levels of understanding of argumentation did not seem appropriate for coding the participants' use of argumentation during the Coke/Pepsi Taste Test lesson as a means for defining their levels of understanding of argumentation. Reflecting on the participants' use of argumentation during the Coke/Pepsi lesson, it seems unreasonable to think a participant would explicitly say as part of a conversation, "I am making a claim and here is my data to support my claim and warrant to link my data to my claim." Similarly, if a participant disagrees with a statement made by another participant, it is unlikely they would say, "I have a rebuttal to your claim." However, based on the definitions of the levels of understanding of argumentation, if a participant did not use proper argumentation terminology, their level of understanding of argumentation could not be coded above a Level 1. A lack of identifying components of Toulmin's Model of Argument while engaging in a discussion and using argumentation in that discussion does not necessarily mean a participant does not have an understanding of argumentation. Therefore, the levels of understanding of argumentation do not accurately capture the understanding that is evidenced while in the act of engaging in argumentation.

Implications

This multiple case study explored the question of how students' engagement in the statistical problem-solving process supports students' understanding of argumentation. The importance of students developing an understanding of argumentation can be seen through the recent call for students to “construct viable arguments and critique the reasoning of others” (CCSSI, 2010, p. 6) as part of the Standards for Mathematics Practice. However, when used in mathematics, argumentation is commonly limited to proofs (e.g., Boero, 1999; Conner; 2007; Smith, 2010). Therefore, it is important to examine different avenues in mathematics as a means for developing and understanding argumentation and using argumentation in mathematics.

This study examined engagement in the statistical problem-solving process as a means for developing students’ understanding of argumentation. Although a multiple case study cannot provide a basis for determining if engagement in the statistical problem-solving process does increase all students’ understanding of argumentation, this study found that for the three participants involved in the study, this was the case. However, developing an increased understanding of argumentation should not be confused with developing a deep understanding of argumentation. That was not the case for this study. The participants only exemplified an increased understanding of argumentation and not a deep understanding of argumentation as participants’ recognition of argumentation was limited to occurrences within the statistical problem-solving process. Therefore, a practical implication of this research is to enhance the training on argumentation so that it includes argumentation in the context of the classroom setting and explicit reflection, both individually and within small groups, that includes the identification of arguments and description of how argumentation was used during the training.

A theoretical implication for mathematics education from this research is the use of the levels of understanding of argumentation for coding a student's understanding of argumentation. Although the use of these codes in a multiple case study does not provide a basis for the validity of these codes for use more broadly, the research indicates these codes could be useful for describing students' understanding of argumentation related to written and verbal descriptions of argumentation.

The results of this study provide insight to instructors and researchers seeking to further develop students' understanding of argumentation through engagement in the statistical problem-solving process. In the next section, ideas regarding future research are provided.

Future Research

To contribute to the use of engaging students in the statistical problem-solving process as a way to develop their understanding of argumentation, further research is needed. For future research, it would be beneficial to examine the development of students' understanding of argumentation after participating in multiple argumentation trainings and engaging in the statistical problem-solving process over a longer period of time. One class period of argumentation training and engagement in the statistical problem-solving process was not enough to support students in developing a deep understanding of argumentation. Therefore, additional sessions of both should be examined to determine their effect on students' understanding of argumentation.

Additionally, future research should be conducted to develop a tool for measuring students' understanding of argumentation while using argumentation. The levels of understanding of argumentation proved to not be a successful way to measure students' understanding of argumentation during the Coke/Pepsi lesson. Therefore, a tool should be developed that can measure this understanding.

A final recommendation for future research is related to the authority within the classroom and its effects on argumentation. The role of authority appeared to impact the engagement in argumentation during the Coke/Pepsi lesson by limiting the need for data sources and warrants to support claims that stemmed from the authoritative sources. Therefore, this relationship should be researched for shaping the future use of argumentation in mathematics classrooms.

Each of these suggestions for future research would further inform the field of mathematics education in relation to the results of this multiple case study. These suggestions would help further investigate engaging students in the statistical problem-solving process as a way to develop students' understanding of argumentation.

Chapter Summary

In this chapter, a review of the methodology and procedures of study was provided. Next, results from all three cases and the cross-case analysis were presented. Following this, a discussion of these results was provided. Finally, implications for mathematics education and future research in mathematics education that emerged from this study were discussed.

This study was a multiple case study with three cases. Data for the study was collected from five different sources: lesson documents, writing prompts, interviews, observations, and the researcher. This data were analyzed based on the Creswell's (2013) five general steps of analyzing qualitative data, Toulmin's Model of Argument, and codes for levels of understanding of argumentation, which emerged from the data. Results from the study indicated that through their engagement in both argumentation training and the Coke/Pepsi lesson, all three participants made gains in their knowledge of argumentation by being able to describe the simple structure of Toulmin's Model of Argument. However, none of the participants engaged in a complete argument during the Coke/Pepsi lesson, and the participants had difficulty identifying and describing argumentation when watching video clips from the lesson.

Overall, although all three participants made gains in their understanding of argumentation after participating in argumentation training and engaging in the statistical problem-solving process, none of the participants provided evidence of developing a deep understanding of argumentation through this process. Some implications for mathematics education are present from this research. Additionally, this research study presents several areas for future research to be conducted.

REFERENCES

American Statistical Association. (2005). *Guidelines for assessment and instruction in statistics education (GAISE): College report*. Alexandria, VA: Author.

Aydeniz, M., Pabuccu, A., Cetin, P. S., & Kaya, E. (2012). Argumentation and students' conceptual understanding of properties and behaviors of gases. *International Journal of Science and Mathematics Education, 10*, 1-24. DOI: 10.1007/s10763-012-9336-1.

Banegas, J. A. (2013). Argumentation in mathematics. In A. Aberdein & I. Dove (Eds.), *The argument of mathematics* (pp. 47-60). Netherlands: Springer.

Bell, P., & Linn, M. C. (2000). Scientific arguments as learning artifacts: Designing for learning from the web with KIE. *International Journal of Science Education, 22*, 797–817.

Ben-Zvi, D. (2006). Scaffolding students’ informal inference and argumentation. In A. Rossman & B. Chance (Eds.), *Working cooperatively in statistics education: Proceedings of the Seventh International Conference on Teaching Statistics* (pp. 1-22). Salvador, Brazil and Voorburg, The Netherlands: International Statistical Institute.

Billig, M. (1987). *Arguing and thinking: A rhetorical approach to social psychology*. Cambridge: Cambridge University Press.

Boero, P. (1999, July/August). Argumentation and mathematical proof: A complex, productive, unavoidable relationship in mathematics and mathematics education. *International Newsletter on the Teaching and Learning of Mathematical Proof*. Retrieved November 12, 2014, from http://www.lettredelapreuve.org/OldPreuve/Newsletter/990708Theme/990708ThemeUK.html

Bradford, S. (2007). *The use of mathematics dialogues to support student learning in high school prealgebra classes* (Doctoral dissertation). The University of Montana, Montana, United States.

Bricker, L., & Bell, P. (2008). Conceptualizations of argumentation from science studies and the learning sciences and their implications for the practices of science education. *Science Education, 92*, 473–498.

Brownell, W. A. (1947). The place of meaning in the teaching of arithmetic. *Elementary School Journal, 47,* 256–265.

Common Core State Standards Initiative. (2010). *Common core state standards for mathematics*. Washington, DC: National Governors Association Center for Best Practices and the Council of Chief State School Officers. http://www.corestandards.org/Math.

Common Core State Standards Initiative. (2015). *Standards in your state*. Retrieved March 5, 2015, from http://www.corestandards.org/standards-in-your-state

Conner, A. (2007). *Student teachers' conceptions of proof and facilitation of argumentation in secondary mathematics classrooms* (Doctoral dissertation). The Pennsylvania State University, Pennsylvania, United States.

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

Cross, D., Taasoobshirazi, G., Hendricks, S., & Hickey, D. T. (2008). Argumentation: A strategy for improving achievement and revealing scientific identities. *International Journal of Science Education*, *30*, 837-861.

Driver, R., Newton, P., & Osborne, J. (2000). Establishing the norms of scientific argumentation in classrooms. *Science Education, 84*, 287–312.

Eisenhardt, K. M. (1989). Building theories from case study research. *Academy of Management Review*, *14*, 532-550.

Erduran, S. & Jimenez-Aleixandre, M. P. (Eds.). (2008). *Argumentation in science education: Perspectives from classroom-based research*. Dordrecht, the Netherlands: Springer.

Erduran, S., Simon, S., & Osborne, J. (2004). TAPping into argumentation: Developments in the application of Toulmin's argument pattern for studying science discourse. *Science Education, 88*, 915-933.

Forman, E. A. (2003). A sociocultural approach to mathematics reform: Speaking, inscribing, and doing mathematics within communities of practice. In J. Kilpatrick, W. G. Martin, & D. Schifter (Eds.), *A research companion to* Principles and Standards for School Mathematics (pp. 333–352). Reston, VA: National Council of Teachers of Mathematics.

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

Garfield, J. B., & Ben-Zvi, D. (2007). How students learn statistics, revisited. *International Statistical Review, 75,* 372-396.

Garfield, J., DelMas, R., & Zieffler, A. (2007). Coke vs. Pepsi Taste Test: Experiments and Inference about Cause. Retrieved November 12, 2014, from http://serc.carleton.edu/sp/cause/datasim/examples/cokepepsi.html

Jimenez-Aleixandre, M. P., Bugallo-Rodriguez, A., & Duschl, R. A. (2000). 'Doing the lesson' or 'doing science': Argument in high school genetics. *Science Education, 84*, 757–792.

Klein, D. (2003). A brief history of American k-12 mathematics education in the 20th century. In James M. Royer (Ed.), *Mathematical cognition* (pp. 175-225). Charlotte, NC: Information Age Publishing.

Kuhn, D. (2010). Teaching and learning science as argument. *Science Education, 94*, 810–824.

Kysh, J. M. (1999). *Discourse in small groups in an algebra 1 class* (Doctoral dissertation). University of California, Davis, California, United States.

Lambdin, D. V., & Walcott, C. (2007). Changes through the years: Connections between psychological learning theories and the school mathematics curriculum. In W. G. Martin & M. E. Strutchens (Eds.), *The learning of mathematics: Sixty-ninth yearbook* (pp. 3–26). Reston, VA: National Council of Teachers of Mathematics.

Larriva, C. M. (1998). *A situated view of participation in a high school mathematics classroom* (Doctoral dissertation). Stanford University, California, United States.

Lefrançois, G. R. (2006). *Theories of human learning: What the old woman said* (5th ed.). Belmont, CA: Wadsworth.

Mason, L. (1996). An analysis of children's construction of new knowledge through their use of reasoning and arguing in classroom discussions. *Qualitative Studies in Education*, *9,* 411-433.

Merriam-Webster. (2017). Definition of Argumentation. Retrieved January 21, 2017, from https://merriam-webster.com/dictionary/argumentation

Moore, J. B. (2000). *The role of student discourse in the mathematics achievement of African American male high school students* (Doctoral dissertation). University of Cincinnati, Ohio, United States.

National Council of Teachers of Mathematics. (1980). *Agenda for Action*. Reston, VA: Author.

National Council of Teachers of Mathematics. (1989). *Curriculum and evaluation standards for school mathematics*, Reston, VA: Author.

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics.* Reston, VA: Author.

National Council of Teachers of Mathematics. (2014). *Principles to actions: Ensuring mathematical success for all*. Reston, VA: Author.

National Research Council. (2001). *Adding it up: Helping children learn mathematics.* J. Kilpatrick, J. Swafford, and B. Findell (Eds.). Mathematics Learning Study Committee, Center for Education, Division of Behavioral and Social Sciences and Education. Washington, DC: National Academy Press.

Palincsar, A. S. (1998). Social constructivist perspectives on teaching and learning. *Annual Reviews Psychology*, *49*, 345-375.

Patton, M. Q. (2002). *Qualitative research and evaluation methods* (3rd ed.). Thousand Oaks, CA: Sage.

Skemp, R. R. (1976). Relational understanding and instrumental understanding. *Mathematics Teaching, 77,* 20–26.

Smagorinsky, P. (1995). The social construction of data: Methodological problems of investigating learning in the zone of proximal development. *Review of Educational Research*, *65*, 191-212.

Smith, R. C. (2010). *A comparison of middle school students' mathematical arguments in technological and non-technological environments* (Doctoral dissertation). North Carolina State University, North Carolina, United States.

Toulmin, S. (1958). *The uses of argument*. Cambridge, NY: Cambridge University Press.

Usiskin, Z. (2012, April). *Unpacking mathematical understanding in the Common Core*. Paper presented at the annual meeting of the National Council of Supervisors of Mathematics, Philadelphia, PA.

Venville, G. J., & Dawson, V. M. (2010). The impact of a classroom intervention on grade 10 students' argumentation skills, informal reasoning, and conceptual understanding of science. *Journal of Research in Science Teaching, 47*, 952–977. doi:10.1002/tea.20358.

Vygotsky, L. S. (1978). *Mind in society: The development of higher mental processes*. Cambridge, MA: Harvard University Press.

Webb, N. (1991). Task-related verbal interaction and mathematics learning in small groups. *Journal for Research in Mathematics Education, 22*, 366-389.

Wertsch, J. (1991). *Voices of the mind: A sociocultural approach to mediated action*. Cambridge, MA: Harvard University Press.

Yackel, E. (1997, April). *Explanation as an interactive accomplishment: A case study of one second-grade mathematics classroom*. Paper presented at the Annual Meeting of the American Educational Research Association, Chicago, Illinois.

Yackel, E. (2001). Explanation, justification and argumentation in mathematics classrooms. In M. Van den Heuvel-Panhuizen (Ed.), *Proceedings of the 25th conference of the International Group for the Psychology of Mathematics Education Vol. 1* (pp. 9 - 24). Utrecht, The Netherlands: PME.

Yin, R. K. (2014). *Case study research: Design and methods* (5th ed.). Los Angeles, CA: Sage.

APPENDICES

APPENDIX A: COKE/PEPSI TASTE TEST LESSON

Coke/Pepsi Taste Test

Part I: Designing the Study

1. How could you design a study that would determine if someone could actually tell the difference between Coke and Pepsi? Explain.

NULL MODEL:
ALTERNATIVE MODEL:

Conduct an experiment given the materials you have been provided.

How many sodas did your taster correctly identify? __________

2. What is your conclusion based upon the results of your study? Explain.

Part II: Critiquing the Study

3. Was this taste test an experiment or an observational study? Explain.

4. Critique the study on each of the four elements of a good experiment (random sample, randomization, control, and replication). Which were met in our study? Which were not? Explain.

5. What are three ways in which we could improve our study for next time? How would each of these suggestions improve the study?

Part III: The Guessing Model

6. What if your taster really could not tell the difference between Coke and Pepsi, how many would you expect him/her to correctly identify? Explain.

7. If we replicated the experiment many times, and each time our taster tried the soda he/she was guessing, what would the distribution of correct identifications look like? Sketch a graph.

8. Could a student correctly identify all ten sodas just by guessing? Explain.

9. How many sodas are likely to be correctly identified if our taster was guessing each time? What are likely values?

10. What are unlikely values (the number of correct identifications that don't occur as often if they were guessing)?

11. Is it likely or unlikely that your taster would have gotten the result he/she did if he/she was just guessing? Explain.

APPENDIX B: WRITING PROMPTS

1. In your opinion, what is argumentation?

2. What do you think argumentation is now?

3. Do you think you used argumentation in the Coke/Pepsi Taste Testing lesson? If so, how? If not, why did you not use it?

APPENDIX C: INTERVIEW PROTOCOL

1. What study did you yourself think of to determine if someone could taste the difference between Coke and Pepsi?
2. Did your group members all agree with your plan or was someone else's plan selected?
3. How did you decide as a group whose study design you would use?
4. Do you think your group used any form of argumentation to decide which study design to use? If so, can you explain how?
5. Once you collected your data, how did you decide on whether the taster could tell the difference between Coke and Pepsi?
6. Do you think your group used any form of argumentation to come to this conclusion? If so, can you explain how?
7. How did your group approach the discussion questions when trying to answer them?
8. Do you think your group used any form of argumentation to answer these questions? If so, how?
9. Do you think you personally engaged in any form of argumentation during any part of the taste testing activity?
10. If so, can you tell me when?
11. Do you think argumentation is a useful tool for discussing different aspects of statistics? If so, why? If not, why not?

APPENDIX D: INSTITUTIONAL REVIEW BOARD APPROVAL

3/30/2015

Investigator(s): Brittany Baum and Angela Barlow
Department: Mathematics Science Education
Investigator(s) Email: bds2x@mtmail.mtsu.edu and Angela.Barlow@mtsu.edu

Protocol Title: "Argumentation in an introductory statistics classroom "

Protocol Number: 15-249

Dear Investigator(s),

The MTSU Institutional Review Board, or a representative of the IRB, has reviewed the research proposal identified above. The MTSU IRB or its representative has determined that the study poses minimal risk to participants and qualifies for an expedited review under 45 CFR 46.110 and 21 CFR 56.110, and you have satisfactorily addressed all of the points brought up during the review.

Approval is granted for one (1) year from the date of this letter for 5 (FIVE) participants.

Please note that any unanticipated harms to participants or adverse events must be reported to the Office of Compliance at (615) 494-8918. Any change to the protocol must be submitted to the IRB before implementing this change.

You will need to submit an end-of-project form to the Office of Compliance upon completion of your research located on the IRB website. Complete research means that you have finished collecting and analyzing data. **Should you not finish your research within the one (1) year period, you must submit a Progress Report and request a continuation prior to the expiration date.** Please allow time for review and requested revisions. Failure to submit a Progress Report and request for continuation will automatically result in cancellation of your research study. Therefore, you will not be able to use any data and/or collect any data. Your study expires **3/31/2016**.

According to MTSU Policy, a researcher is defined as anyone who works with data or has contact with participants. Anyone meeting this definition needs to be listed on the protocol and needs to complete the required training. **If you add researchers to an approved project, please forward an updated list of researchers to the Office of Compliance before they begin to work on the project.**

All research materials must be retained by the PI or faculty advisor (if the PI is a student) for at least three (3) years after study completion and then destroyed in a manner that maintains confidentiality and anonymity.

Sincerely,

Institutional Review Board
Middle Tennessee State University

APPENDIX E: ARGUMENTATION TRAINING LESSON PLAN

Handouts:	**Materials:** • White board and markers

Lesson Goal: The goal of this lesson is to introduce students to argumentation in a meaningful way.

Warm-up (10 minutes)

Think-pair-share: How could you convince a group of high school students that [our university] is the best college in the state of [redacted]? (Claim: [Our university] is the best college in the state of [redacted].)

After approximately 5 minutes, ask the students to provide responses to the claim and write the responses on the board.

Argumentation Training (20 minutes)

Think-pair-share: Ask the students what they believe to be the most convincing response to the claim in the warm-up exercise.

Use student responses to explain the relevance of using data to support a claim.

Think-pair-share: Ask students How or why is the given data relevant to the claim of [our university] is the best college in the state of [redacted]? (In other words, what links or justifies the data and the claim?)

After approximately 5 minutes, ask students to provide responses.

Explain to the students that these responses are considered to be warrants.

Explain to students that through the claim of [our university] is the best college in the state of [redacted], the data used to support the claim, and the warrants they have just provided to link the clam to the data, they have just developed a simple argument.

Present Toulmin's Model of Argument (simple argument) and define claim, data, and warrant.

- Claim – statements that are being argued
- Data – facts or evidence used to prove the argument
- Warrant – justifications one makes to connect the data and the claim

Think-pair-share: Ask students to now consider how they would respond if someone presented data that shows [a different university] has the best agricultural program in the state. (How does this affect their original claim?)

After approximately 5 minutes, ask students to present their responses.

Explain to students the statement about [a different university] having the best agricultural program is a rebuttal to the original claim, which sometimes affect the claim and call for the need of a qualifier to be added to the original claim (link this to student responses of qualifiers that were given).

Present the full model of Toulmin's Model of Argument and define the complex components of qualifier, backing, and rebuttal.

- Qualifier – a statement or condition that limits when the claim is true
- Backing – statements that are used to support the warrant
- Rebuttal – a counter-argument or claim that negates any part of the argument

Debrief (20 minutes)

Think-pair-share: Ask students to develop their own argument (simple argument).

After approximately 10 minutes, select a group to present their argument. Have non-presenting students identify the claim, data, and warrant.

Ask students if they can think of any possible rebuttal to this argument.

Think-pair-share: Select one rebuttal from student responses and ask students to consider if the rebuttal can be clarified with backing for the warrant or new data and/or warrants, or if the rebuttal requires a qualifier for the claim.

Explain to students that they have just been engaged in argumentation.

Questions to pose to students:

- What role does argumentation have in the real world?
- Why is it necessary to use data to support a claim?
- What role does argumentation have in statistics?

Summary (5 minutes)

Distribute the writing prompt. Ask students to respond individually in writing. Collect as students exit the class.