AN INVESTIGATION INTO FOUNDATIONAL CONCEPTS RELATED TO SLOPE:

AN APPLICATION OF THE ATTRIBUTE HIERARCHY METHOD

By

Angela E. Broaddus

Submitted to the graduate degree program in the Department of Curriculum and Teaching and the Graduate Faculty of the University of Kansas in partial fulfillment of the requirements for the degree of Doctor of Philosophy

Dissertation Committee:

____

Chairperson Susan Gay

____

Phil McKnight

____

Ingrid Peterson

____

William P. Skorupski

____

Jack Porter

Date Defended:____________________________

The Dissertation Committee for Angela E. Broaddus

certifies that this is the approved version of the following dissertation:

AN INVESTIGATION INTO FOUNDATIONAL CONCEPTS RELATED TO SLOPE:

AN APPLICATION OF THE ATTRIBUTE HIERARCHY METHOD

____

Chairperson Susan Gay

Date approved: ________________________

ABSTRACT

The mathematics education community has been working for more than two decades to reform K – 12 school mathematics programs. The literature consistently emphasizes the importance of students understanding and making sense of mathematics, which is characterized by complex networks of knowledge reflecting the conscious organization of related facts and processes. Understanding of slope is an essential milestone in a person's mathematical development, yet assessments often measure student knowledge of slope in terms of their procedural fluency rather than their conceptual understanding.

The purpose of this study was to explore the concepts students should possess to demonstrate understanding of selected foundational concepts related to slope, to determine a cognitive model hypothesizing how this understanding develops, and to design an instrument to assess understanding of selected foundational concepts related to understanding slope as described in the model. The instrument was administered to a sample of Kansas students in middle and high school mathematics courses.

This study provided an example of one way to implement components of Evidence-Centered Design. The study was conducted in two phases. The first phase included a domain analysis and yielded a theoretical cognitive model of how selected foundational concepts related to slope are acquired. The second phase included a task analysis and the development of an assessment containing items that targeted the knowledge described in the cognitive model. Test responses were analyzed using Item Response Theory, and students were classified into knowledge states based on their test response data using the Attribute Hierarchy Method (AHM).

Students demonstrated varying levels of knowledge with regard to the selected foundational concepts of slope. The AHM revealed that students who participated in this study demonstrated three main levels of understanding of the selected foundational concepts of slope. First, students demonstrate the ability to identify quantities that are related as covariates. Second, students demonstrate the ability to identify the direction of covariation in a problem setting. Third, students demonstrate the ability to interpret a slope ratio in terms of a problem's context variables.

ACKNOWLEDGEMENTS

I am grateful to my committee, whose members have guided me with their wisdom and expertise. Dr. Porter's patience and insight caused me to reflect on my own mathematical understanding and the goals of this study. Dr. McKnight's propensity to think broadly and recommend seminal literature allowed me to develop sensitivity to the history of curriculum and instruction in our country. Dr. Peterson's interest in and awareness of how students learn reminded me to maintain my focus on student understanding of mathematics. Dr. Skorupski's insistence on good measurement practice combined with his innovative thinking guided my design and solution to the research problem. I am indebted to Dr. Susan Gay, whose tireless efforts insured the research integrity and technical quality of this dissertation. While Dr. Gay has advised me as a mathematics educator and researcher, I have been inspired by her tremendous passion for the teaching profession and the preparation of mathematics teachers.

I am thankful to the many people who throughout my life have provided me with the motivation and skills to complete this dissertation and degree. I am indebted to my parents, Alice and Bernard G. Castle. My father taught me by example the value of service to others through his work as a caring physician, while my mother instilled in me the practice of completing every task to the very best of my ability. I am thankful to my teacher colleagues, who throughout my teaching career showed me the meaning of teaching for understanding and caring for students. I am grateful to school administrators Frank Till and James Williams for their superb leadership and generous support.

I am beholden to my colleagues and friends, who have witnessed and supported me during the past several years of my studies. My colleagues at the Center for Educational Testing and Evaluation not only have helped me with many aspects of this study but also have supported me with their kindness and patience. I am grateful to my friends in our community and church, who have continuously encouraged me with smiles and kindness.

I express my deepest thanks to my children, Elise and Daniel Broaddus, who have patiently supported me throughout my studies. I hope they have learned about patience, persistence, and faith as they have observed my work during these years.

There are no words to suitably express my gratitude to my husband, Matthew Broaddus - my dearest friend and loyal companion. His unwavering support, encouragement, and love have sustained me, for which I am eternally grateful.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

THE RESEARCH PROBLEM

Introduction

The mathematics education community has been working for more than two decades to reform K – 12 school mathematics programs. The National Council of Teachers of Mathematics (NCTM, 1989) voiced its recommendations for how to improve mathematics education when it published the *Curriculum and Evaluation Standards for School Mathematics*. This document promoted the idea that students should learn and understand important mathematics. Since 1989, mathematics educators have been working to realize the ambitious goals articulated in this document. Through extensive collaboration and commitment, the NCTM continued the dialogue regarding how best to mathematically educate children by publishing coordinated materials and updates to the original *Standards.* The *Professional Standards for Teaching Mathematics* (1991) and the *Assessment Standards for School Mathematics* (1995) offered information for teachers and school officials to assist them in designing and pursuing educational programs aligned with the *Standards*. More recently the NCTM released updates to its original documents; these included *Principles and Standards for School Mathematics* (PSSM) (2000), *Mathematics Teaching Today* (2007), and *Focus in High School Mathematics: Reasoning and Sense Making* (2009). In 2010, the Council of Chief State School Officers (CCSSO) and the National Governors Association (NGA) jointly published the *Common Core State Standards for Mathematics* (CCSSM) (CCSSO/NGA, 2010). This document is the most recent effort to "define what students should understand and be able to do in their study of mathematics" (CCSSO/NGA, 2010, p. 4), by placing an equal amount of emphasis on conceptual understanding and procedural fluency.

The similarities in the recommendations contained in all of these volumes are many, and they consistently emphasize understanding and making sense of mathematics. In all of the documents, teachers are encouraged to create classroom environments that cultivate inquiry, discourse, and collaboration in support of children understanding important mathematics. Understanding mathematical ideas is characterized by complex networks of knowledge reflecting the conscious organization of related facts and processes (Hiebert & Carpenter, 1992; Hiebert & Lefevre, 1986; Klausmeier, Harris, Davis, Schwenn, & Frayer, 1968; Marshall, 1990; Messick, 1984; Skemp, 2006; Webb & Romberg, 1992).

Over the course of the last 20 years, federal initiatives have prompted state agencies to create frameworks similar to the national documents to guide the development of curricula and assessments at state and local levels (Marshall, Sears, Allen, Roberts, & Schubert, 2006). In many cases, including Kansas, state assessments have been implemented to measure student achievement of certain skills delineated in the state standards documents (Neill, 2003). However, the skills identified to be tested often consist of specific routines or procedures rather than queries into what students understand (Neill, 2003). Unlike procedural knowledge, assessing conceptual understanding does not lend itself easily to classically reliable measurement practices and may require development of alternative measurement tools (Webb, 1992; Webb & Romberg, 1992).

Researchers agree that assessments influence instruction (Battista, 1999; Bond, Moss, & Carr, 1995; Harlen, 2007; Jaeger, Mullis, Bourque, & Shakrani, 1995; Moss, 1994; Suurtamm, Lawson, & Koch, 2008; Webb, 1992). In particular, when teachers are focused on students' mastery of assessed material, the classroom instruction often concentrates on performance of skills and routines rather than understanding the underlying concepts (Harlen, 2007). Teachers are compelled to insure that their students can carry out these routines, whether or not their students develop understanding of why the routines work and how they arise from related mathematical concepts (Battista, 1999). One consequence of such instruction is the reduced likelihood that students will be able to apply what they have learned in class to novel situations (Harlen & James, 1997). The attention devoted to preparing students for state assessments may undermine classroom teachers' success in promoting rich experiences for students, thereby placing at risk the potential for students to construct deep understanding of the mathematics they study (Battista, 1999; Haertel, 1985; Hancock & Kilpatrick, 1993). Restricted instruction produces restricted learning potential for students (Harlen & James, 1997).

Educators who aim to evaluate what people know or can do are obliged to scrutinize whether their measurement methods in fact produce reliable information that can be used to make valid decisions. The nature of validation studies has developed during the last half century to be characterized by arguments that describe each step in the evaluation process (Kane, 2001). That is, validation is supported by a trail of evidence that includes descriptions of the knowledge or abilities to be measured, the means by which they will be measured, i.e., assessment or test development, the interpretations of the scores or judgments made based on the assessments, and the consequences or decisions of the interpretations (Kane, 1992, 2001).

One framework that is used to structure validity arguments is evidence-centered design (ECD) (Mislevy, Almond, & Lukas, 2003; Mislevy & Haertel, 2006; Mislevy, Steinberg, & Almond, 2003; Pellegrino & Huff, 2010). This framework encourages educators to prioritize the nature of understanding in the knowledge domain of interest in all stages of instructional planning and assessment. The different levels of the ECD process result in descriptions of the knowledge to be assessed and how this knowledge can be demonstrated in addition to descriptions of how reviewers should interpret different demonstrations of knowledge to make fair and appropriate decisions. By leading educators and measurement specialists through the process of building evidentiary arguments throughout the process of building and administering assessments, ECD fosters a focus on scientific reasoning in educational evaluation practices.

Different cognitive models are available to depict how understanding of a particular body of knowledge develops over time (Gierl, Wang, & Zhou, 2008). Cognitive models are used to show the components of a particular body of knowledge and the connections among the components that lead to complete understanding of that body of knowledge, such as the knowledge held by an expert. Cognitive models contain the names of the concepts and skills and descriptions of how these should be interconnected to support the development of a person's knowledge. The concepts and skills that make up a person's knowledge are also referred to as the attributes comprising that person's knowledge. Assessments developed to measure knowledge in terms of cognitive models are used to provide diagnostic feedback to examinees in terms of the attributes they possess and those they need to acquire.

Understanding of algebra is a gateway to college and career success (Benbow, 2008; National Mathematics Advisory Panel, 2008). A key component of algebraic reasoning is a working understanding of linear functions (National Mathematics Advisory Panel, 2008; NCTM, 2009). Understanding linear function behavior is a prerequisite for learning more advanced mathematics topics such as calculus and statistics (Wilhelm & Confrey, 2003). Linear functions are characterized by constant rates of change, or slopes. Therefore, understanding the concept of slope is deemed to be "one of the most important mathematical concepts students encounter" (Joram & Oleson, 2007). As such, a comprehensive understanding of slope is recommended in prominent curricular standards, namely the CCSSM and the PSSM.

Understanding of slope depends on a student's ability to reason with ratios (Lobato & Thanheiser, 2002). Two types of reasoning with ratios are covariational reasoning and proportional reasoning (Hoffer, 1988). Covariational reasoning concerns a person's ability to perceive and interpret relationships among quantities that vary in correspondence to one another (Carlson, Jacobs, Coe, Larsen, & Hsu, 2002). Proportional reasoning concerns a person's ability to compare ratios (Hoffer, 1988; Lesh, Post, & Behr, 1988), which is a significant milestone in a person's mathematical development (Lamon, 1993). The simplest instances of slope are present in situations of direct variation, where the slope value equals the constant of proportionality (Lamon, 2005). Students should develop the ability to determine and interpret slope, or the constant of proportionality, for problems presented in any appropriate mathematical representation (Lamon, 2005).

Statement of the Problem

The goal of this study was to explore the concepts students should possess to demonstrate the ability to consider a problem or a graph depicting a direct variation relationship and to interpret the rate or slope that governs the variation relationship in terms of the problem's context variables. The primary objectives of the study were to determine a model to represent selected foundational concepts related to understanding slopes depicted in graphs or described in problems and to develop a means by which such understanding may be assessed. The study included the development of an instrument to assess selected foundational concepts related to understanding slope.

Research Questions

This investigation answered these research questions.

1. What insight is gained about the validity of the proposed cognitive model from an analysis of student data generated from an assessment informed by the model?
2. To what extent did student participants exhibit common misconceptions regarding slope?

The study consisted of two phases of research. During the first phase, the slope concept was analyzed by the researcher in order to develop a theory about the foundational concepts related to understanding slopes depicted in graphs or described in problems concerning direct variation. This analysis led to the development of a cognitive model representing five foundational concepts as attributes that support a person's ability to recognize and consider two quantities that vary, to detect the direction of the covariation relationship, and to relate the slope ratio to the context of the problem. During the second phase, the researcher developed an assessment containing test items that targeted the concepts and connections students have acquired to support their understanding of the five foundational concepts related to slope, using the cognitive model developed in the first phase as a guide. Subject matter experts were consulted to confirm the cognitive model developed in the first stage and the quality of the assessment developed in the second stage. The assessment was administered to a sample of middle and high school students in Kansas. Test responses were analyzed using item response theory and the attribute hierarchy method (Leighton, Gierl, & Hunka, 2004) to classify students according to their knowledge and to validate the hypothesized cognitive model. Test responses were also analyzed in order to describe the knowledge and misconceptions held by the students in the sample with regard to the slope concept.

Rationale

Traditional assessments fail to assess understanding of mathematics concepts (de Lange & Romberg, 2005; Mathematical Sciences Education Board, 1993; Romberg, Carpenter, & Kwako, 2005; Suurtamm et al., 2008). In particular, Suurtamm, Lawson, and Koch (2008) questioned whether current assessments address reform ideals such as conceptual understanding. Assessments that are not aligned with the ambitious goal of learning with understanding fail to measure the progress students are making toward that end (Mathematical Sciences Education Board, 1993). Few studies have examined how different measurement methods make explicit what students understand about mathematical constructs (Haertel, 1985). This study provided an example of how to develop and analyze an assessment focused on understanding.

Marshall (1990) found that most assessments addressed only the components of knowledge rather than how these components of knowledge were cognitively connected by the student and recommended the use of assessment tasks targeting the relationships among knowledge components. Assessments focused on connections are needed to contribute meaningfully to the learning process and to provide information about the connections students have made among concepts and skills being studied (de Lange & Romberg, 2005; Mathematical Sciences Education Board, 1993). This study yielded an instrument deliberately designed to assess the connections students have made among selected foundational concepts related to slope.

A significant milestone in a person's mathematical achievement is that person's understanding of constant rate or slope (Hauger, 1999; Thompson, 1994; Vergnaud, 1994). Constant rate is among the most important concepts a student learns and is related to multiplication, division, ratio, and linear functions, all of which are members of the multiplicative conceptual field (Vergnaud, 1994). Several studies have found that students can often compute slope and write equations of lines without having developed an understanding of the slope concept (Crawford & Scott, 2000; Orton, 1984; Stump, 2001). Other studies have found that students experience confusion when learning about slope (Clement, 1989; Moschkovich, 1996; Moschkovich, Schoenfeld, & Arcavi, 1993). This study described some components of the slope concept and some things a student should know and be able to demonstrate as evidence of understanding slope. In addition, this study provided insight into how educators can measure understanding through questioning aimed at the critical components of knowledge that support understanding of the concept of slope.

The ability to form connections among mathematics concepts and skills supports understanding and consequently is a valued mathematics practice (Mathematical Sciences Education Board, 1993). The early work in this study included an analysis of what a person needs to know in order to understand foundational concepts related to slope. The study yielded a cognitive model containing five attributes and the cognitive connections students need in order to to understand selected foundational concepts related to slope. The cognitive model guided the design of an assessment that was developed to gauge what students did and did not understand about the selected foundational concepts related to slope. These outcomes, i.e., the cognitive model and the assessment, serve as examples of how to study the learning and understanding of mathematics concepts and how to develop assessments that target students' developing understanding of important mathematics.

Definitions

Terms Related to Mathematics Content

Additive inverses (opposites) are two numbers whose sum is zero (Carson & Jordan, 2011).

Additive reasoning describes a problem solving strategy in which a person applies addition or subtraction in order to determine the relationship between a pair of numbers (Baroody & Coslick, 1998). For example, instead of considering how many times larger six is than two, an additive reasoner would determine that six is four more than two.

A **concept** is a cognitive representation of a real object or experience (Ausubel, 1968).

Covariation is the correspondence of variation, or equivalently, the correspondence of two variables (Moritz, 2005).

Covariational reasoning occurs when a student is able to understand and explain how one quantity changes as another quantity changes (Adamson, 2005).

Multiplicative reasoning describes a problem solving strategy in which a person applies multiplication or division in order to determine the relationship between a pair of numbers (Baroody & Coslick, 1998). For example, six is the same as three times two.

A mathematical **procedure** specifies how to proceed with given inputs to produce desired outputs (National Research Council, 2001).

A **proportion** consists of two equal ratios (Lovell & Butterworth, 1966).

A **rate** is "a reflectively abstracted constant ratio" (Thompson, 1994, p. 190), that is, a multiplicative comparison between two quantities that yields a single, constant quantity (Thompson, 1994).

A **ratio** is the result of a multiplicative comparison between two quantities (Thompson, 1994).

Reciprocal quantities (multiplicative inverses) are "those quantities which when multiplied together produce unity" (McKechnie, 1979, p. 1505).

A mathematical **representation** is a way to view a mathematics concept (NCTM, 2000).

Slope is "the measure of the steepness of a nonvertical line…The slope of a line is the ratio of the change in y, Δy, to the change in x, Δx. The slope is therefore the rate of change of y with respect to x" (Stewart, 2003, p. A12).

Terms Related to Measurement Methods

An **attribute** is a specific cognitive competency, e.g. concept or skill, that contributes to a person's knowledge in a particular content domain (Gierl et al., 2008).

The **attribute hierarchy method** is an educational measurement method based on the theory that knowledge is structured hierarchically with dependent relationships among concepts and skills (Gierl et al., 2008).

Cognitive skills consist of both facts and procedures (Royer, Cisero, & Carlo, 1993). Cognitive skills are thereby made up of both declarative and procedural knowledge, and they are acquired through instruction and practice. They develop from slow, deliberate cognitive actions into efficient, automated sets of cognitive actions.

A **mental model** is a hypothetical representation of some body of knowledge that is known along with how it is organized in a person's mind (Pellegrino, 1988). It specifies a person's representation of a body of knowledge in terms of the required knowledge, skills, and abilities (Gorin, 2006).

Terms Specific to the Present Study

Knowledge state refers to the level of knowledge a person has with regard to a specific content domain.

Foundational Concepts of Slope Attribute Hierarchy (FCSAH) refers to the hierarchical arrangement of five attributes representing selected foundational concepts related to understanding slope.

Foundational Concepts of Slope Assessment (FCSA) is the instrument that was developed during this study to assess student understanding of selected foundational concepts of slope in a manner that was consistent with the FCSAH.

Assumptions

It was assumed the sample of students who participated in this study was representative of Kansas students studying Pre-algebra, Algebra 1, Geometry, Algebra 2, and Pre-calculus. It was assumed that the different settings in which students took the FCSA did not impact their responses and that students demonstrated their best effort and work when taking the FCSA, especially given that this assessment was administered during the last weeks of the school year. It was assumed that student participants understood the questions presented on the FCSA, that is, they were able to read the sentences and graphs and were able to understand the situations presented in the problems. It was assumed that students who participated possessed some knowledge of the slope concept and that student participants had a range of ability and knowledge of the attributes identified in the FCSAH. It was assumed that student responses indicated their understanding of the attributes that were measured, that is, student responses were based on student knowledge of the foundational concepts related to slope and not on guessing.

The development of the FCSAH incorporated previous research findings regarding student understanding of concepts related to slope. It was assumed that the researcher in the present study accurately interpreted previous studies' results and represented these findings appropriately in the FCSAH and the FCSA. Furthermore, it was assumed that the subject matter experts (SMEs) consulted during this study genuinely endeavored to use their professional expertise and insight to guide the review and refinement of the FCSAH and the FCSA.

Limitations

This study was conducted with students who attended middle schools and high schools in Kansas. The assessed curriculum in Kansas (Kansas State Board of Education, 2003) includes the slope concept at Grade 8 and in high school. The results of the study may not be able to be generalized beyond the types of settings and the curricular expectations of the classrooms of the teachers who participated. Students were not randomly assigned; rather, students assigned to the teachers who agreed to participate were therefore participants in the study. Many of the questions on the FCSA required students to interpret problems presented in sentences. Therefore, students' reading levels may have influenced their performance on the FCSA. Additionally, students who were more familiar with mathematics problems that mainly required use of procedural knowledge may have lacked familiarity with the conceptual questions posed on the FCSA.

All of the items on the FCSA were multiple-choice items. This format restricted the range of item responses that were collected by forcing students to select one of the answer choices rather than allowing them to respond to the items more freely. The items on the FCSA used a limited set of item types to assess the five attributes listed in the FCSAH. The use of different item types may have produced different results. The results of this study cannot be generalized beyond the scope of the five attributes listed in the FCSAH and the scope of the items on the FCSA. Specifically, all but one of the items on the FCSA concerned positive slope values. Therefore, success on the FCSA does not indicate that a student necessarily would have success with the entire slope concept. While the contexts used in the items were similar to the contexts that other researchers reportedly used and found to be common in many classrooms, it is possible that one or more of the problem contexts on the FCSA were unfamiliar to some students. There were deliberate choices made during the study with regard to the statistical method used to classify students into different knowledge states. Other methods or choices might produce a different, possibly more robust, result with regard to classifying students into different knowledge states.

Overview

This dissertation is divided into five chapters. This first chapter introduces the problem that was considered. The second chapter presents a review of related literature addressing concepts, connected knowledge, connections in mathematics, slope, validity in educational measurement, and models for developing assessments that target conceptual understanding. The third chapter presents the methodology and describes the two phases of the study. The fourth chapter presents the results. The fifth chapter presents a summary of the study with conclusions and recommendations for further investigations.

CHAPTER 2

REVIEW OF LITERATURE

Introduction

This chapter will first discuss theories related to concepts, connections that support learning concepts with understanding, and connections that enhance mathematical understanding. The next section will discuss issues related to validity in educational measurement, evaluating what people know and understand, and the use of cognitive models to inform the development of educational assessments. A survey of prominent curriculum documents and where each of these places the concept of slope will follow. A review of the literature related to student understanding of covariation, proportions, and slope will then be presented. The chapter will conclude with a summary.

Concepts

A concept is a cognitive representation of a real object or experience (Ausubel, 1968; Bruner, Goodnow, & Austin, 1956; Martorella, 1972). As such, a concept belongs exclusively to its knower, although one person's conception of a certain object or idea may closely resemble another person's conception of the same thing (Martorella, 1972). According to Piaget, effective concept learning is based on concrete experiences (Flavell, 1963). A person's conceptions, although rooted in concrete models, change and mature as that person is faced with experiences or perceptions that are inconsistent with previous conceptions (Martorella, 1972). Such experiences cause conflicts that are resolved through adaptation, whereby learners first accommodate their current schemes to situate new information (Adler, 1971). Then, through assimilation, learners internalize their new frameworks and use these more mature conceptions to evaluate future experiences (Martorella, 1972).

Klausmeier (1992) characterized a concept as a mental construct in which a person has consciously associated a given term with the accepted meaning of that term. This interpretation varies from Piaget's by asserting that a person does not know a concept until that person has grasped the concept's meaning and associated it with an accepted name. According to Klausmeier, students pass through four stages as they acquire concepts, and the process of attaining a concept may span multiple years. At the concrete stage the student is able to recognize an example; at the identity stage a student is able to recognize an example from a variety of views; at the classification stage, a student is able to distinguish an example from a non-example; and at the formal stage a student is able to name the concept, list its defining attributes, and describe distinctions from similar concepts. Klausmeier found that students who reached the formal stage demonstrated a six-fold increase in their ability to relate multiple concepts to one another and to use their knowledge of concepts to solve problems.

Stages of Concept Learning

Different authors asserted that learning concepts requires passage through sequential stages. While the stages' descriptions overlap among investigators, subtle nuances exist between the terms used to describe the steps in learning concepts. The terms used frequently in the literature are concept formation, concept attainment, and concept assimilation. The next section contains descriptions of how these terms are used by different authors in addition to how these stages should be sequenced for effective learning.

Concept formation. Concept formation occurs when an individual internalizes the meaning of an object or idea from concrete, empirical experiences (Ausubel, 1968; Bruner et al., 1956). Conceptual meanings, the basis of all knowledge, are constructed from experiences during which a person derives a sense of the regularities among objects or events grouped under a concept label or name (Novak, 1998). While Piaget argued that such sense making experiences must precede naming a concept, Vygotsky asserted that concept names aid in meaningful acquisition of concepts (Novak, 1998). They agreed, however, that naming a concept and conceiving of its meaning are not the same things. Concept formation based on concrete experiences predominantly takes place during early childhood, when children learn exclusively through concrete, real world interactions with the objects and events that occur in their lives (Adler, 1971; Ausubel, 1968). After the first three years, individuals apply their existing knowledge when presented with opportunities to learn new concepts and construct new meanings through adaptation (Adler, 1971) or assimilation (Ausubel, 1968).

Concept attainment. Children first experience concepts through perceptual connections, but a meaningful grasp of any concept requires more than perceptual cues (Macnamara, 1982). Concept attainment involves analysis and evaluation of the criterial attributes that qualify an object or idea for membership into a category with the concept's name (Bruner et al., 1956; Martorella, 1972; Novak, 1998). This process requires the person to be able to distinguish exemplars from non-exemplars by identifying attribute qualities that are common among the exemplars and distinct from some of the attributes present in non-exemplars (Bruner et al., 1956). Concepts are attained when a person can classify an object as a member of a conceptual grouping or not. Meaningful conceptions have as their basis some reference to the conditions required for membership into a specific category, which children perceive by reviewing sets of examples (Ausubel, 1968). Concept attainment is enhanced when instructional examples contain several positive instances of the concept (Bruner et al., 1956). Although children frequently learn the names of things as they perceive concepts, meaning cannot be inferred by simply using the correct name for a concept (Martorella, 1972; Novak, 1998). Rather one must demonstrate knowledge of the necessary and sufficient conditions for class membership by correctly distinguishing exemplars from non-exemplars (Bruner et al., 1956; Macnamara, 1982).

Human beings are naturally inclined at a very basic level to cognitively categorize what they experience in an effort to store knowledge for future use (Bruner et al., 1956). Categorizing is a way of simplifying what one knows for efficient retrieval. Over time and experience a person groups together concurrent attributes of a particular concept so that when a subset of the attributes is present, the person immediately looks for the other attributes in the group. For example, when learning the concept of bird, children see feathers, wings, a bill, and legs. After repeated experiences with birds, a child might see wings and feathers and then predict that the object has a bill and legs. Eventually the child knows that all the mentioned attributes are present for birds. The fusion of the attributes into a unified concept allows more efficient use of the mature concept. Concepts are eventually freed from the specific contexts in which they were first experienced, as they are stored and categorized in memory in more abstract forms (Ausubel, 1968; Bruner et al., 1956).

Concept assimilation. Meaningful learning occurs when a person relates a new abstract idea substantively and non-arbitrarily to what the person already knows (Ausubel, 1968). During this process there exists a constructive interaction between the new content and the existing cognitive structure, as the person fits the new information in with existing knowledge. The opportunity for meaningful learning is dependent on what is to be learned and what the learner already knows that is relevant to the learning goal. Prior relevant and structured knowledge is the most significant variable influencing the possibility of meaningful learning (Ausubel, 1968; Bruner et al., 1956).

From birth to about the age of three, children form concepts directly from their concrete experiences (Novak, 1998). However, after this period most conceptual learning consists of concept assimilation, whereby a person relates new ideas to existing, structured knowledge. Factors influencing the success of knowledge assimilation include the learner's deliberate effort to relate new information to existing knowledge and the quality of the structure of this existing knowledge. As knowledge within a topic is accumulated and organized into a logical framework, it becomes easier to acquire new meanings. A proven strategy is to introduce more specific concepts following more general concepts, allowing prior knowledge to subsume new knowledge (Ausubel, 1968; Bruner et al., 1956; Novak, 1998).

As a person's conceptions improve and mature, that person's cognitive structure undergoes concept assimilation in the form of "integrative reconciliations" (Novak, 1998, p. 66), whereby new information is logically and structurally incorporated with existing knowledge. This process typically involves the learner acknowledging and using similarities and differences to organize concepts in increasingly sophisticated webs indicating appropriate hierarchical relationships. There are three mental actions used to form logical relationships among concepts. Subsumption describes the alignment of newly acquired details under more comprehensive concepts; progressive differentiation describes the distinction between similar though different concepts; and integrative reconciliation describes the construction of a coherent conceptual framework.

Misconceptions

According to the literature, misconceptions occur when flawed information or erroneous connections are associated with a concept. Misconceptions can be described in terms of what information a person has cognitively stored and how it is interconnected, a person's success in applying a concept's definition to classify, or the maturity of a conception.

Glaser (1986) asserted that misconceptions arise when students have incomplete or incorrect knowledge of the components that make up a concept or have constructed erroneous connections among them. In these cases, students' flawed conceptions may cause them to apply the concept incorrectly when analyzing or solving a problem. For example, students demonstrate error patterns when learning arithmetic operations such as multi-digit subtraction. One particularly common error is for students to subtract the smaller digit from the larger digit in a column instead of applying the rule to subtract the top digit minus the bottom digit.

A second way to view a misconception is that a person has a flawed sense of the requirements for set membership (Henderson, 1970). In such a case a person might classify exemplars of a concept into the wrong classes because that person fails to observe one of the required criteria. In this case, the person might place the exemplar in a subordinate category rather than the concept class in which the exemplar belongs. In another case a person might classify exemplars of a concept into the wrong class because that person learned the concept by reviewing a restricted set of exemplars (Klausmeier, 1992). For example, a student might incorrectly classify an equilateral triangle if the triangle's image is not oriented to have a horizontal base and centered vertex at the top of the image. This student would correctly classify Triangle A as equilateral but incorrectly classify Triangle B as not equilateral.

Triangle A Triangle B

A third way to view misconceptions is to view them as concepts in progress (Klausmeier, 1992; Wilson, 2009). Klausmeier (1992) asserted that students pass through four stages as they acquire concepts. In the first stage, a student correctly classifies exemplars that are identical in appearance and orientation to the exemplars from that student's first encounter with the concept. In the second stage, a student correctly classifies exemplars that are identical in appearance but not in the same orientation as the exemplars from the student's first encounter with the concept. In the third stage, a student correctly classifies exemplars that are neither identical in appearance nor in the same orientation as the exemplars from the student's first encounter with the concept, but the student may not be able to describe the reasons for each classification. In the fourth stage, a student correctly classifies exemplars in any appearance or orientation and is able to explain each classification. Each successive stage is associated with more sophisticated conceptions, and each stage subsumes prior stages of the conception. Thus gauging what a person understands about a concept is a matter of identifying the stage of that person's particular conception.

Models that Depict Concept Knowledge

A variety of models are used to display how concept knowledge might be organized in a person's cognitive structure. Concept maps (Baroody & Bartels, 2001), learning hierarchies (Gagné, 1968), construct maps (Wilson, 1992, 2009), and learning progressions (Wilson, 1992, 2009) are models containing descriptions of the concepts and skills that make up larger learning targets (Popham, 2008). They define what students must learn to understand concepts and map out optimal learning sequences (Wilson, 2009).

Concept maps often look like webs and contain links among the concepts and skills that contribute to the body of knowledge modeled by the map (Baroody & Bartels, 2001). They are very useful to display how knowledge in a particular content domain is organized in a person's mind.

Gagné (1968) believed that students were most able to learn material if they had mastered all prerequisite knowledge. His learning hierarchies were used to arrange the specific intellectual abilities a person had to possess in order to be able to learn a particular intellectual skill. He urged educators to complete task analyses of the learning objectives they planned to teach in order to describe the sequences in which subordinate skills should be acquired to facilitate the learning of the terminal objectives of their instruction. These models could be used to plan optimal sequences of learning activities. He cautioned, however, that learning hierarchies were most appropriate to model the learning of intellectual skills or cognitive strategies, which can be described in terms of how people should be able to demonstrate their knowledge or skills.

Construct maps, learning hierarchies, and learning progressions often consist of linear sequences of the pieces of knowledge pertaining to particular concepts or skills (Popham, 2008; Wilson, 2009). In particular, each level of a construct map is associated with a more sophisticated conception and subsumes the levels below it in the construct map. Learning progressions that model complicated concepts or ideas, however, may consist of several, interrelated construct maps (Wilson, 2009), which can produce displays that appear more like webs than linear sequences.

Connected Knowledge Supports Understanding

Educational policy, curriculum documents, and other literature from the past century revealed several attempts to describe what it means to understand and how people learn things with understanding. Different models were developed to represent how learning with understanding may optimally occur, and in some cases these models were used to describe instruction that could foster learning with understanding. A common theme emphasized by many authors was that conceptual understanding depends on connected knowledge. Specifically, connections between existing knowledge and new information have been and continue to be a prominent feature in the recommendations for how to develop and foster conceptual understanding.

Progressive Education

Dewey (1938) described authentic learning as a cyclic process in which a person uses what that person already knows and has experienced to perceive and organize the facts and ideas in his/her present experience. A person's past and present experiences similarly influence what that person perceives and is able to learn in the future. Understanding emerges as a person develops an increasingly sophisticated organization of more facts and ideas related to that person's current and previous experiences. In order to foster learning with understanding Dewey recommended infusing the curriculum with deliberately designed, meaningful experiences for students (Willis, Schubert, Bullough, Kridel, & Holton, 1994). After many such experiences, the child would develop an "arranged, orderly view of previous experiences" (Willis et al., 1994, p. 126) to serve as a guide to future experiences. While this discussion came from *The Child and the Curriculum* (Dewey, 1902, reprinted in Willis et al., 1994), Dewey later revisited the notion of how students fit newly acquired ideas into existing knowledge when he stated that "the achievements from the past provide the only means at command for understanding the present" (Dewey, 1938, p. 77)… "and for production of new ideas" (p. 79).

Pratt described understanding of material to be that which remained with a child as a result of learning experiences (Willis et al., 1994). She insisted that all knowledge understood and valued by a person related to that person's experiences, had been tested against his/her experience and prior understandings, and was organized according to these relationships and tests. Therefore, Pratt advocated that students were best able to form meaningful understandings when newly introduced material related to their previous experiences. In her school, learning activities were coordinated among sequential grades so that one grade's activities built upon the activities explored in previous grades. Teachers considered both past and future learning experiences of their students when planning classroom activities. An example of such coordination concerned geography and orientation. In Pratt's school, the four-year old children learned about and represented in diagrams where the different parts of the school were in relation to their classroom. Then as five-year old students, they added on the names and locations of the streets outside the school and where their homes were in relation to the school. Later, when they studied history and geography, they were able to relate how to use large-scale maps to these early experiences with orientation and pictorial representations of their surroundings. Their prior experiences infused meaning into the nature of graphical representations of space in general.

Piaget

Piaget conducted a vast amount of research aimed at describing and explaining behavior and cognition (Flavell, 1963). His theory is built on the notion that intellectual development is inherently linked to a person's experience. Piaget's work provided evidence to suggest that cognition develops as a person's intellectual structures pass iteratively through successively more advanced cycles of sophistication, each of which derives coherently from all previous cycles. A cycle begins and ends in states of equilibrium, between which a person's cognitive structure adapts previous knowledge to accommodate information gained through new experiences. Equilibrium is a state reached through adaptation to something new, when the closely associated cognitive actions of assimilation and accommodation are balanced.

Piaget asserted that "intelligent activity is always an active, organized process of assimilating the new to the old and of accommodating the old to the new" (Flavell, 1963, p. 17). He claimed that the nature of cognitive development mirrors that of biological organisms in general. As organisms adapt to their environments, so too a person's intellectual structure adapts to that person's experiences. Adaptation describes how a person cognitively changes what is previously intellectually organized to reflect perceived experiences, to include new information, and to incorporate that new information with the old in such a way as to address new experiences using a revised cognitive organization. Structures from early stages of cognitive development become integrated at higher stages of cognitive development, and successive structures necessarily are built on and continuously linked to previous structures. The culmination of successive adaptations is a sophisticated, networked, highly organized intellectual structure.

Constructivism

Building on the work of Piaget, constructivists insist that human beings actively construct new knowledge in response to their personal concrete, perceptual, or abstract experiences (Confrey, 1990; Noddings, 1990; von Glasersfeld, 1990). There exists an essential connection between how a person experiences and acts on new information and the development of that person's cognition (Noddings, 1990), making the learning process not only active but also an individual, constructive experience. Learning with understanding must be viewed as a continuous process of reconstructing a person's record of experiences (Dewey, 1938).

Students benefit from learning experiences that offer them opportunities to interact socially in order to construct new understandings (Cobb, Wood, & Yackel, 1990; Noddings, 1990; Steffe, 1990; Vygotsky, 1978). When faced with a new experience, the student must consciously engage in a conversation or an activity in order to connect something new to what is already understood. Neither the motivation to learn nor the connections between new and prior knowledge are able to take place without social interactions between the student and expert or teacher (Vygotsky, 1978). The discussions that take place during a learning activity help the student describe and clarify his/her conception so that the student can make sense of the newly acquired knowledge and fit it into his/her cognitive structure (Vygotsky, 1978).

In the course of acquiring new understandings, students may form intermediate, personal representations or informal conceptions (Battista, 1994; Carpenter & Lehrer, 1999). Although these may be different from those possessed by a teacher or other students, they are valid for the students who hold them and should form the basis for learning more about the idea being introduced (Confrey, 1990). Communicating about these conceptions and representations gives students the chance to check how their understanding compares with that of their classmates and of their teacher (von Glasersfeld, 1990), and listening to these communications offers the teacher a glimpse into the students' understandings (Confrey, 1990). Students' communications may transition from speaking aloud to themselves, to conversing with another student, and eventually to silent, abstract reflection (Steffe, 1990). It is through communication and social interactions that students adjust their conceptions and relate them to other things they know (von Glasersfeld, 1990). When students understand things in the same ways as other students and their teacher, there is consensus within the classroom culture (Cobb, Yackel, & Wood, 1992).

Conceptual knowledge that is deliberately constructed through constructivist learning experiences is inherently connected and integrated with other concepts (Confrey, 1990). This connected conceptual knowledge relates and embodies all the contexts and representations in which the concept has been experienced by the person, and it reflects the history of how that person constructed his/her understanding. Future learning is enhanced by the presence of networked knowledge, supporting the cognitive actions that form links to more sophisticated understandings.

Conceptual Connections

Piaget claimed that assimilation is responsible for connecting an image or symbol to the object or experience that it represents (Flavell, 1963). As symbols are introduced, both assimilation and accommodation must take into account prior cognitive organizations, current perceptions along with their images or representations, and all associations among these. Piaget distinguished between language used to convey meaning among people, e.g. words and mathematical notations, and the representations used by children to intellectually record their experiences. He insisted that children's representations must be in place before they can associate commonly used words or notations with their own intellectual representations. As part of this process, over time children assimilate words and mathematical notations into their personal symbolic representations.

Tall and Vinner (1981) asserted that concept learning requires both attention to the concept image and the concept definition. The concept image exists in a person's cognitive structure and includes everything the person has experienced with the concept. At early stages of knowing a concept, there may not be a name or precise definition associated with the concept image, but the image may be useable even if it is incomplete. Once a concept image is formed, it may be expanded to include a name, description, and definition. People can use concept images lacking names and definitions in a productive way. An example of such an incomplete concept image is when a student's preliminary experience with subtraction of whole numbers causes that student to believe that taking one value from another value always results in a smaller number.

This act demonstrates knowledge of the concept of subtraction without requiring that the concept be associated with a name or its complete definition.

Concept definitions are very different from concept images because they are specific, verbal representations of concepts (Tall & Vinner, 1981). People may use personal or formal definitions of concepts, the latter of which in mathematics would be one that is accepted by the mathematical community. Problems arise when a person's operating definition and concept image lack agreement, as is the case when students learn abstract concepts formulaically rather than building on more concrete previous experiences. Tall and Vinner recommended that instructional sequences proceed inductively from concrete exemplars, to constructed definitions that are sensible based on a breadth of experience with examples, and finally to formal definitions.

Gagné

Gagné (1971) advocated the idea that learning new information depended on the connections a person could build to relate the new information to what that person already knew. He developed a model of concept learning based on the notion that in order to acquire a concept, one must have learned all pre-requisite concepts; otherwise it would be impossible for such learning to occur. Results from Gagné's research showed that in more than 97% of cases, gaps in pre-requisite knowledge accounted for failure to demonstrate mastery of a given learning objective.

Gagné's (1968) theory of learning hierarchies required the identification of a hierarchy of subordinate knowledge components for each terminal learning outcome. Such a hierarchy often resembles a tree with branches rather than a linear sequence. Gagné described three essential characteristics of a learning hierarchy. First, the components of the hierarchy are identified by answering the question, “What would an individual have to know how to do…” (p. 66). Second, the components of the hierarchy must be organized in a sequence that would yield successful learning. That is, if someone learned the components in the order they appeared in the hierarchy, then that person would successfully learn the intended learning objective and all its prerequisites. Third, hierarchies represent how many people might effectively learn, but a hierarchy does not necessarily represent the most efficient route of learning for any one individual.

Gagné (1968, 1971) suggested that by evaluating students’ knowledge of the components of a hierarchy, teachers could accurately determine what subordinate knowledge is known and identify what needed to be learned. He recommended two types of inquiries concerning hierarchies. One type of inquiry would investigate two levels of a hierarchy at a time, thereby studying the connections among the knowledge described in two neighboring components in a hierarchy, or those delineated as having a direct connection. A second type of inquiry would investigate longer sequences of learning targets arranged as the components of a hierarchy, thereby studying the multiple connections among the components and their sequencing in the hierarchy.

Connections that Enhance Mathematical Understanding

Understanding mathematics requires much more than acquiring a host of unrelated facts and procedures; it depends upon richly connected knowledge (Carpenter & Fennema, 1991; Marshall, 1990; NCTM, 1989, 2000). Understanding mathematics involves knowing concepts and using procedures as well as deliberately building progressively sophisticated cognitive networks or schema relating concepts and procedures (Beyer, 1993; Hiebert & Carpenter, 1992; Hiebert & Lefevre, 1986; Marshall, 1990; Messick, 1984; National Research Council, 2001; Vergnaud, 1997; Webb & Romberg, 1992).

Knowledge of facts and procedures is instrumental for working mathematically, but symbol manipulations are often required as the chief demonstration of students' mathematical proficiency (Battista, 1999), which overlooks the importance of the conceptual underpinnings of mathematical processes. The consequence of this oversight is the dominant view that proficiency in mathematics can be explained by procedural fluency, which conveys little about a student's mathematical understanding (Boaler, 2002). Rote manipulation of numbers and symbols does not constitute demonstration of mathematical understanding and should be replaced by efforts and materials that draw on connections among concepts, representations, and procedures to provide evidence of understanding (Battista, 1999).

Ideas about the Role of Connections in How Mathematics Concepts are Learned

When presented with a new mathematics concept, the learner must be able to situate new knowledge in a scheme, which essentially becomes the home of a newly introduced concept (Vergnaud, 1997). A scheme consists of the items that are related to one another in a person's cognitive structure. Each concept or individual idea that is part of the scheme is called a node. The depth of a person's understanding of a concept depends upon the number of nodes that person has arranged for the concept, the number and strength of these nodes' interconnections, and the person's acknowledgement of these nodes and their interconnections. Over repeated use, concepts emerge as flexible resources containing all the attributes acquired through prior trials and experiences. Essentially when a concept is activated, its entire scheme or network is awakened and made available for learning. Students deepen their understanding of mathematical concepts as they relate newly introduced ideas within their existing schemes (Arzarello, Robutti, & Bazzini, 2005; Marshall, 1990).

Dienes (1971) claimed that the evolution of each person's knowledge of a mathematics concept mirrors the human maturation process suggested by Piaget. Dienes asserted that conceptions grow in sophistication from concrete experiences that can be based on actions or real objects to ideas that can be expressed by symbolic notation and eventually to cognitive abstractions that can be used to interpret new information or can be applied to problems. There exists a concern, however, that when a mathematics concept is introduced to a student who possesses no relevant cognitive structure or prior experience on which to draw, that such a student acquires a set of notations or symbols void of authentic meaning. The student may be able to name such a concept or even carry out a procedure, but without a connection to something the student already understands, this new knowledge lacks meaning.

Different Types of Connections in Mathematics

Connections that promote understanding of mathematical concepts include (1) ways in which different representations are used to express the meaning of a concept, (2) ways in which concepts are associated with procedures, and (3) ways in which prior knowledge is associated with new information. Connections support fluency with concepts, procedures and representations, which is an important indicator of mathematical understanding. Connections are critical to building mathematical understanding, but they "are frequently overlooked in curriculum planning and assessment" (Boaler, 2002, p. 14).

Representations. Representations offer people different ways to view mathematics concepts (NCTM, 2000). Different types of representations include physical manifestations, actions, verbal descriptions, numerical values, symbols, pictures, charts, and graphs (National Research Council, 2001). Many concepts can be conveyed by more than one representation, each of which can clarify or improve a student's existing conception. In fact, "mathematical ideas are enhanced through multiple representations" (National Research Council, 2001, p. 95). The proficient student must be able to work with a concept in any of its representations and translate among different representations (National Research Council, 2001; Niemi, 1996).

Effective instruction illuminates the relationships among various representations of mathematical concepts (Hiebert & Carpenter, 1992; National Research Council, 2001; NCTM, 2000). Student-created representations may aid in elementary understanding and should not be overlooked by teachers trying to glean information about students' pre-existing conceptions about a topic (NCTM, 2000). Students' personal representations can help them organize their understanding and aid their transition to more conventional representations (NCTM, 2000). Instructional and assessment tasks that specifically target how students have connected their knowledge should be used to assess whether students have correctly matched different representations of mathematics concepts (Niemi, 1996).

Concepts and procedures. Understanding and doing mathematics require students to actively construct and acknowledge connections among concepts and procedures (Boaler, 2002; National Research Council, 2001). Knowledge of procedures includes the rules, algorithms, and sequences of steps used in mathematics (Hiebert & Lefevre, 1986). Hiebert and Lefevre (1986) suggested that procedures that are meaningful are linked to conceptual knowledge.

The connections among procedures and concepts and their different representations support a person's mathematical understanding (National Research Council, 2001). The proficient student must be able to choose an appropriate procedure for use with a concept in a particular representation (National Research Council, 2001). Effective instruction illuminates how different mathematical procedures relate to each other and to mathematical concepts in their various representations (Hiebert & Carpenter, 1992). The Mathematical Sciences Education Board

(1993) echoed the call for a well-connected curriculum so that students might grow to understand mathematics as an integrated subject in which concepts and procedures reinforce each other and meld together into ways to explain the world.

Prior knowledge. Understanding mathematics is supported by connections between new and existing knowledge (NCTM, 2000). Consequently, the NCTM (1989, 2000) emphasized the importance of relating new information to prior knowledge to support learning mathematics with understanding by advocating a curriculum that is coherent, focused, and articulated across the grades. Concepts are linked both vertically across grade levels and horizontally among the overarching themes of the curriculum, and interconnections among topics should be targeted by instruction and curricular materials (NCTM, 2000).

In order to help students build understanding, teachers must design lessons that draw out students' prior knowledge (NCTM, 2000). Teachers who purposefully identify their students' existing knowledge are better equipped to make instructional choices that help students develop mathematical understanding by effectively building on what they already know (Hiebert & Lefevre, 1986; Vergnaud, 1997). Instructional activities designed to provide effective scaffolds allow students to relate new experiences to their existing cognitive schemes (Arzarello et al., 2005).

Measuring and Describing What People Know

Educational measurement theory and practice aim to develop means by which educators can describe and interpret what students know and can do. The process of measuring what people know and can do is complex. In order for the measurement process to inform valid decisions and appropriate consequences, judgments made at every stage in the measurement process must be valid.

Validity

Validity is a holistic judgment of how and why decisions are made (Messick, 1989, 1994a). The term validity refers to the level to which a test or evaluation does the task for which it is intended. Validity studies concern whether judgments made in the evaluation process are based upon good measurement procedures and how well the measurement actually relates to whatever is being judged. The most convincing validity rationales contain evidence from all possible means along with descriptions of the connections between the constructs assessed, the tasks used to measure them, the empirical data collected and scored, the interpretations of the data, the decisions based on these interpretations, and the consequences of these decisions.

Valid educational measurement should be construct centered, that is, it should originate with descriptions of the knowledge, skills, and abilities students are expected to learn (Haertel, 1985; Kane, 1992, 2001; Messick, 1984, 1994a, 1994b, 1995b; Mislevy, Almond, et al., 2003; Mislevy & Haertel, 2006; Mislevy, Steinberg, & Almond, 1999; Mislevy, Steinberg, et al., 2003; Mislevy, Steinberg, Almond, Haertel, & Penuel, 2000). These descriptions should include not only aspects of the content domain, but also how typical students build increasingly sophisticated cognitive networks relating the essential concepts and skills in the content domain (Messick, 1984, 1994b, 1995a). The measurement process should also take into account the nature of how people come to know things, how they demonstrate their knowledge, and how observations are interpreted in terms of decisions and consequences (Mislevy, Almond, et al., 2003; Mislevy, Steinberg, et al., 2003; Mislevy et al., 2000). A construct-centered approach to assessment promotes a comprehensive conception of validity that can be used as a framework for appraising the validity of every step in the measurement process and of the process as a whole (Kane, 2001;

Messick, 1984, 1995a). Messick (1994a) published a concise and complete description of a construct-centered approach to validity.

> A construct-centered approach would begin by asking what complex of knowledge, skills, or other attributes should be assessed, presumably because they are tied to explicit or implicit objectives of instruction or are otherwise valued by society. Next, what behaviors or performances should reveal those constructs, and what tasks or situations should elicit those behaviors? Thus, the nature of the construct guides the selection or construction of relevant tasks as well as the rational development of construct-based scoring criteria and rubrics. (p. 16)

Messick (1989, 1994b) described construct validity as a unitary concept comprising the classically accepted components of validity in addition to the interpretation and use of scores. In doing so, he acknowledged six aspects of construct validity that could be used to form a framework for validating evaluation processes. The six aspects of construct validity he identified are content, substantive, structural, generalizable, external, and consequential. Messick (1995a) described each of these explicitly in terms of the tasks used in testing, test scores, and the interpretations of test scores.

- The content aspect concerns evidence of the tasks' content relevance, representativeness, and technical quality.
- The substantive aspect references how effectively the tasks cause students to use the mental processes the tasks intend to measure and whether examinee responses are consistent with their knowledge.
- The structural aspect judges the consistency between a task, how it is scored, and the structure of the construct being measured.

- The generalizability aspect inspects whether score properties and interpretations of scores can be generalized reliably to different populations, settings, and tasks.
- The external aspect accounts for convergent and discriminant evidence to support score interpretations consistent with test design and to discount plausible rival interpretations.
- The consequential aspect judges the implications of how scores are used to make decisions and determine actions, particularly in regard to bias and fairness.

Messick (1995a) encouraged educators to use the six aspects of construct validity to frame empirical studies of the meaning and utility of test scores and to build rational arguments to justify score interpretations and uses. Kane (1992) used the term *interpretive argument* to describe an argument leading from test scores to score-based judgments or decisions. The type of reasoning used in building interpretive arguments has been associated with a variety of terms over the course of the last 40 years, including *practical reasoning, informal logic,* and *rhetoric* (Kane, 1992). The validity of a test-score interpretation depends upon the plausibility of the interpretive argument (Kane, 1992). Strong interpretive arguments contain multiple sources of evidence supporting score interpretations, and the claims and assumptions made in such arguments are stated clearly (Kane, 1992).

Evidence-Centered Design

Evidence-centered design (ECD) is an assessment design framework reflecting a construct-centered approach (Mislevy, Almond, et al., 2003; Mislevy & Haertel, 2006; Mislevy, Steinberg, et al., 2003; Pellegrino & Huff, 2010). This framework for building assessment systems requires test developers to build assessments by first analyzing the domain of knowledge that is to be measured and having this analysis drive test design, test delivery, scoring, reporting, and interpretation of test scores. The ECD framework enables test designers to build evidentiary arguments for score interpretations while engaging in the processes associated with test development, test delivery, test scoring, and score reporting, thereby binding each stage of the process to a case for valid score interpretations.

Evidence-centered assessment design uses five layers to frame evidentiary arguments and cases for validity (Mislevy & Haertel, 2006). The layers needed to establish the validity argument are domain analysis, domain modeling, conceptual assessment framework, assessment implementation, and assessment delivery. Although the layers often proceed in sequence, the processes within and among the layers are iterative in nature. As new test items or tasks are added to an assessment, each layer is reexamined to accommodate the new item or task, and the relationship among the layers is reevaluated to insure consistency with the purpose of the assessment and proposed interpretations. Because this process requires test developers to repeatedly evaluate the relationship between the assessment's intended purpose and its proposed interpretations, assessments built using ECD develop in a manner consistent with an underlying validity claim.

Domain analysis. Domain analysis consists of gathering information about the domain of knowledge, skills, and abilities to be assessed (Mislevy, Almond, et al., 2003; Mislevy & Haertel, 2006; Mislevy, Steinberg, et al., 2003; Pellegrino & Huff, 2010). This analysis leads test developers to understand the knowledge needed to be proficient in the domain, its various representations, and the types of situations or activities that evoke the use of relevant conceptual knowledge, procedures, and strategies. Subject matter experts assist test developers during domain analysis by raising their awareness of certain student characteristics such as probable existing knowledge, common misconceptions, instructional patterns, and equity issues (Pellegrino & Huff, 2010).

Domain modeling. Domain modeling involves organizing the information gathered during domain analysis and relating the knowledge, skills, and abilities to be assessed to item types and tasks that will cause students to demonstrate their relevant knowledge, procedures, and strategies (Mislevy & Haertel, 2006). Domain modeling contains three critical components, the first of which is to identify the focal knowledge, skills, and abilities (KSAs) the assessment aims to measure. Secondly, potential test items and tasks are described and related to the focal KSAs. Thirdly, characteristic and variable features of tasks that will cause respondents to apply their KSAs are identified and described. One type of outcome of domain modeling is called a *proficiency paradigm* (Mislevy, Steinberg, et al., 2003), which structures the claims about students' proficiency with regard to the KSAs targeted by an assessment. A proficiency paradigm associates test items and tasks to likely student responses and to a corresponding range of proficiency levels.

Conceptual assessment framework. The conceptual assessment framework (CAF) transforms the work done at the domain modeling layer and outlines the technical specifications for an assessment (Mislevy & Haertel, 2006). Three types of internal models give structure to the CAF: student models, task models, and evidence models.

Student models. The student model variables describe the kinds of proficiencies the assessment aims to address (Mislevy, Almond, et al., 2003; Mislevy & Haertel, 2006). These are described in terms of what educators want to know about what students know but cannot measure directly (Mislevy et al., 2000). Student models relate directly to how student knowledge in a domain is acquired and structured over the course of a student's development and education.

These models are informed by cognitive learning theory as well as educational opportunity to produce descriptions of the knowledge structures, cognitive processes, procedural skills, and their integration as they relate to expertise in the targeted knowledge domain.

Task models. The task models describe the testing environment and how students will respond in an assessment situation to indicate their knowledge in the knowledge domain being tested (Mislevy & Haertel, 2006). These consist of descriptions of the tasks and situations that will be used to evoke the behaviors or KSAs targeted by an assessment (Mislevy et al., 2000). Reflective of the domain modeling layer, test items and tasks are designed to systematically target the declarative, procedural, and strategic knowledge about which test users intend to make judgments. The task model variables offer principled links between the observable student behaviors targeted by an assessment and the features of the items and tasks used to measure them.

Evidence models. Evidence models bridge the student and task models (Mislevy & Haertel, 2006). They define arguments for why and how observations collected from tasks offer evidence about what students know or can do as described by the student model variables (Mislevy et al., 2000). They instruct test users about how to update their information about student model variables based on student responses to given sets of tasks (Mislevy, Almond, et al., 2003). Cognitive task analysis provides descriptions of how the test items and tasks cause students to engage the focal KSAs identified in student models (Mislevy, Steinberg, et al., 2003). The evaluation component of an evidence model describes how student work is qualitatively judged and scored in terms of its observable attributes (Mislevy & Haertel, 2006). Once tasks and items have been evaluated and scored, the data is aggregated into a measurement model, which produces a test score.

Assessment implementation. The assessment implementation layer concerns the preparation of all the operational elements of the assessment that are outlined in the CAF (Mislevy & Haertel, 2006). Many of these tasks such as item development, scoring rules, measurement models, and standard setting are common to large-scale assessments. However, ECD requires the rationale for each process in the assessment implementation phase to be directly related to the validity argument.

Assessment delivery. The assessment delivery layer consists of students taking tests, their responses being scored, and reports being produced and interpreted by test users (Mislevy & Haertel, 2006). Although these operations are common to all large-scale assessments, ECD requires stakeholders to reflect on test score interpretations and consequential decisions by considering their consistency with the CAF variables, domain models, and domain analysis. Such reflection should assess the fit of the assessment to its purpose, while tracing the evidentiary argument back to its origin.

Models in Educational Measurement

Models can be used to represent what knowledge people possess, how knowledge is typically acquired, and how knowledge is measured. Learning progressions are useful in describing the knowledge people possess, how it is typically learned, and common misconceptions. Construct maps and attribute hierarchies are useful for developing assessment items and tasks that measure what people know in a manner that is consistent with how knowledge is acquired.

Theoretical mental models of the knowledge and cognitive processes people use when they work through problems can inform test item development because these mental models delineate the components of the knowledge and processes being targeted by a particular test

(Pellegrino, 1988). Armed with such detailed information, test developers are able to produce carefully designed problem sets that elicit the targeted knowledge, provide useful data, and substantiate valid interpretations. This approach yields a science of test item development that can be the foundation for diagnostic tests that not only measure what people know or can do, but also identify knowledge gaps or misconceptions. The diagnostic capacity of such a test is directly related to the fact that the items developed to correspond to a mental model permit test administrators to review the meaning of test item responses in terms of the mental model used to design the test. That is, the mental model serves as a guide for assessment design and analysis, which in turn can influence instruction by informing educators more specifically about what students know and can do (Pellegrino, Baxter, & Glaser, 1999).

Students develop expertise in a particular knowledge domain over the course of their development and educational experience (Wilson, 2009). As students learn and their understanding grows, they replace less sophisticated models of their understanding with more sophisticated representations (Wilson, 1992). When students are first introduced to a body of knowledge or a particular skill, their cognitive representation is often quite unstructured (Royer et al., 1993). As this knowledge becomes more sophisticated, a student’s cognitive representation becomes increasingly networked and integrated (Royer et al., 1993).

The development of knowledge in a domain can be represented by a learning progression, which includes detailed descriptions of the knowledge that should be learned and suggests a sequence in which the knowledge may be optimally acquired (Wilson, 1992, 2009). Experts and novices differ in how much they know and in how many connections exist among the pieces of knowledge they possess (Wilson, 1992). Where naive conceptions are often judged as misconceptions, or wrong (Wilson, 1992), the use of learning progressions offers a means to view students with such conceptions as having partial knowledge. In order to plan future instruction and learning opportunities aimed at increasing student knowledge, it is important for educators to identify what is present and what is missing. Educators should view understanding on a continuum rather than by benchmarks of success or failure (Wilson, 1992).

Assessments based on cognitive models are needed in order to distinguish the varying levels of knowledge between that of a novice and that of an expert (Pellegrino et al., 1999; Royer et al., 1993). Such assessments must be designed to carefully probe student understandings, and student responses must be analyzed in order to inform instructional decisions (Pellegrino et al., 1999).

When planning an assessment, it is useful to design ordered categories of understanding of the knowledge being assessed (Wilson, 1992). Where learning progressions describe relatively large bodies of knowledge, construct maps depict smaller pieces of knowledge in fine detail (Wilson, 2009). A construct map is focused on a particular skill or concept and describes what students with different levels of understanding should know and be able to do (Wilson, 2009). In this way a construct map relates how learning likely occurs, as described by a learning progression, to how that learning is measured and evaluated (Wilson, 2009). “A construct map is a well thought out and researched ordering of qualitatively different levels of performance focusing on one characteristic” (Wilson, 2009, p. 718). Construct maps are effective for designing assessments. They also can influence instruction while remaining independent of any particular curriculum. The use of construct maps can improve the validity claims of an assessment system, as they tie curricular expectations to how students learn and demonstrate their expertise.

Construct maps incorporate common student misconceptions by identifying them at lower achievement levels as likely errors that are resolved at higher achievement levels (Wilson, 2009). As assessment items and tasks are developed, typical student responses are aligned to appropriate achievement levels in the construct map. Effective distractors are written to represent different levels of achievement so that test items accurately distinguish students performing at different achievement levels.

Cognitive attribute hierarchies can be used to model the cognitive processes or skills students need to possess in order to be prepared to successfully complete test items and tasks (Gierl, 2007; Gierl, Leighton, & Hunka, 2000; Gierl et al., 2008; Leighton & Gierl, 2007; Leighton et al., 2004). A cognitive attribute describes an essential piece of knowledge, skill, or ability needed to perform in a particular content domain. Cognitive attribute hierarchies can be built to represent how knowledge is acquired in a particular domain. It is important to carefully arrange the attributes so that the hierarchy explicates the dependencies among the attributes in terms of how and in what sequence people likely acquire domain knowledge.

The attribute hierarchy method (AHM) is an educational measurement model based on the theory that knowledge is structured hierarchically with dependent relationships among concepts and skills (Gierl, 2007; Gierl et al., 2000; Gierl et al., 2008; Leighton & Gierl, 2007; Leighton et al., 2004). AHM requires that test items and tasks are sensitive to the theoretical hierarchy of attributes a test is designed to measure, that is, the items should measure the subsets of attributes in the hierarchy that theoretically depend upon one another according to the hierarchy. This is most likely to occur if the attribute hierarchy is built prior to test construction and used as a guide during test item and task development.

Studies using the AHM assume that the attribute hierarchy developed to model cognitive organization of knowledge in a domain can be used to explain test performance (Gierl, 2007; Gierl et al., 2000; Gierl et al., 2008; Leighton & Gierl, 2007; Leighton et al., 2004). Measurement methods evaluate whether student responses are consistent with the attribute hierarchy. The psychometric components of this method include identifying the *expected response patterns*, which are the response patterns perfectly consistent with the attribute hierarchy. Actual student responses are then compared to these *expected response patterns* to determine how likely it is that students possess the cognitive attributes in the hierarchy. This type of assessment is used to provide feedback that describes what students do or do not know in terms of the knowledge, skills, or abilities described by the attributes in the hierarchy.

Understanding of Slope is Essential Mathematics

Researchers agree that understanding the concept of slope is essential mathematics. “The slope of a line is one of the most powerful concepts in mathematics” (Andersen & Nelson, 1994, p. 27). Understanding rate of change is vital because this concept is relevant to a wide variety of contexts (Wilhelm & Confrey, 2003). “Slope is one of the most important mathematical concepts that students encounter in middle school and high school” (Joram & Oleson, 2007, p. 261; NCTM, 2006). As such, both national and state level curriculum guidelines contain statements about what students should know and understand about slope.

Common Core State Standards (CCSS)

The Common Core State Standards (CCSSO/NGA, 2010) were developed collaboratively with the intent of being adopted by all states in the USA. The purpose of the document is to guide the teaching, learning, and assessment of students in kindergarten through high school. The CCSS for mathematics describe understanding of slope developing over middle and high school grades. The sixth grade mathematics standards target the understanding of ratio reasoning to solve problems. Such reasoning relies on a student's ability to understand fractions made up of two different types of quantities, such as in a recipe that contains a ratio of 3 cups of flour to 4 cups of sugar. The sixth grade standards describe students interpreting ratios in tables of numbers and computing percents and unit conversions. The seventh grade standards target the growth of ratio reasoning into proportional reasoning, which is present when students can compute and interpret unit rates and constants of proportionality between like or different types of quantities. Seventh grade students should be able to interpret and represent proportional relationships numerically, symbolically, and graphically, and be able to use these representations to solve problems. The eighth grade standards require students to analyze the properties of linear functions, specifically their slopes, represented algebraically, numerically, graphically, or verbally. Students are expected to construct and interpret linear functions, specifically their rates of change or slopes. Students are expected to apply their understanding of linear functions and constant rates of change to their analyses of bivariate data represented in scatter plots and lines of best fit. The high school standards describe students reasoning about functions, where students identify function behavior to be increasing, decreasing, positive, or negative. Students also should be able to calculate, estimate, and interpret average rates of change for functions represented numerically, graphically, or symbolically.

NCTM Standards

The content standards in the *Principles and Standards for School Mathematics* (PSSM) (NCTM, 2000) are intended to help educators to organize high quality curricula and instructional programs that prioritize important mathematics and engaging instruction aimed at teaching for understanding. This document lists several references to the development of understanding of slope and its application to solving problems, particularly in the Algebra standard. In the Algebra standard, elementary students are expected to analyze repeating and growing patterns, and later to represent patterns and functions in tables, in graphs, and in words. These students should be able to explain changing quantities such as how many inches they grow in a year and later explain how a change in one variable relates to a change in a second variable. They should be able to identify, describe, and compare situations that demonstrate constant rates of change. These skills contribute to their preparation to learn slope as they examine changing quantities and relate one changing quantity to another related changing quantity. Middle grades students are expected to identify linear functions presented in tables, graphs, or equations and contrast their properties with non-linear functions presented similarly. Slope is one essential quality that is different for linear and non-linear functions. Students are expected to understand the meaning of slope in different representations. Specifically, they should know what slope means when it is expressed as a number in an equation, and they should know what the slope means when it is represented by the orientation of a line graphed on the coordinate plane. High school students should be able to interpret the meaning of slope as a rate of change for functions given in any representation, to include equations, tables, and graphs.

Concepts and skills related to slope are also present in content standards other than in the Algebra standard. The middle grades Geometry standards (NCTM, 2000) describe students using the coordinate plane to graph and analyze polygons with parallel or perpendicular sides. In order to complete these analyses, students must plot polygons on the coordinate plane and then compute the slopes of the sides of the polygons. The middle grades Measurement standard references problems involving rates, calling attention to the comparison of two varying quantities. The high school Data Analysis and Probability standard lists regression coefficients, regression equations, and correlation coefficients in addition to determining the function that models the data set. Regression not only is a process dependent on determining a slope, but also is a culminating activity for using slope to solve problems and to make data driven decisions such as inferences and predictions.

NCTM Curriculum Focal Points

In an effort to focus educators on essential themes at each grade level from kindergarten through eighth grade, the NCTM (2006) published *Curriculum Focal Points for Pre-kindergarten through Grade 8 Mathematics: A Quest for Coherence* to recommend a coherent curriculum for elementary and middle school students. The concept of slope is introduced in the seventh grade focal points. Seventh grade students should "graph proportional relationships and identify the unit rate as the slope of the related line" (NCTM, 2006, p. 62). Students should relate their work with proportional reasoning and slope when they work with scale factors and similar objects and when converting between different units of measure.

The concept of slope is developed in the eighth grade focal points (NCTM, 2006). Eighth grade students should understand direct variation as a special case of linear functions and recognize the constant of proportionality as represented by the slope of the graphed line. They should be able to interpret slopes in terms of the labels on the axes of a graph, translate among numerical, graphical, symbolic, and verbal representations of linear relationships, and describe how the slope appears in the different representations. They should be able to compare and contrast parallel and perpendicular lines by analyzing the lines' slopes, and apply their understanding of linearity and slope to solve real-world problems involving rates such as motion and constant speed.

Reasoning and Sense Making in High School Mathematics

The NCTM published a series of books titled *Focus in High School Mathematics: Reasoning and Sense Making* (Graham, Cuoco, & Zimmerman, 2010; NCTM, 2006; Shaughnessy, Chance, & Kranendonk, 2009) to advise educators on how to increase the focus of their instruction and how to help students learn important mathematics with understanding. In the first supporting publication focused on statistical reasoning and probability (Shaughnessy et al., 2009), students must explain what the slope of the line means in terms of a graph's axis labels and in terms of a problem's context variables. Students should also be able to evaluate and compare slopes presented in tables, equations, and graphs and explain their meaning within the context of the problem. A second book (Graham et al., 2010) in the reasoning and sense making series focuses on algebraic reasoning. In this book the authors cited "reasoning about slope" (p. 25) to be among the top few areas presenting the greatest challenge to students and educators. The authors characterize slope to be the relationship between the geometric properties of the points lying on a line and the ordered pairs associated with these points, particularly when the line is neither horizontal nor vertical.

Kansas Curricular Standards for Mathematics

The *Kansas Curricular Standards for Mathematics* for eighth grade list several expectations related to an understanding of slope (Kansas State Board of Education, 2003). Eighth grade students should be able to recognize, describe, and analyze constant and linear relationships presented numerically, graphically, symbolically, and in problem contexts (verbally). In particular, they should be able to "explain the concept of slope" (p. 278) and "find the slope of a line" (p. 291). High school students should be able to "recognize how changes in the slope within a linear function change the appearance of a graph" (p. 311) and "interpret the meaning of the slope of a line in the context of a real-world situation" (p. 311). They should be able to "calculate slope given ordered pairs and explain how the graph of a line is related to its slope" (p. 321). Thus students in Kansas should gain conceptual knowledge and procedural expertise with slope in all of its representations, that is, computing slope from numbers, determining slope from graphs of lines, identifying slope from equations, and identifying slope from verbal descriptions of problems. Students also should demonstrate the ability to translate among these representations, which is called representational fluency.

Foundational Concepts of Slope

"Slope involves the construction of a ratio as the measure of a given attribute" (Lobato & Thanheiser, 2002, p. 164). As such, a firm grasp of reasoning with ratios forms the foundation for understanding slope. Two types of reasoning with ratios are covariational reasoning and proportional reasoning (Hoffer, 1988). Several researchers have described their ideas about levels of covariational reasoning (Tourniaire & Pulos, 1985). Other researchers have explored the complex nature of proportional reasoning and how it provides a foundation for learning higher level mathematics (Kurtz & Karplus, 1979).

Covariational Reasoning

Adamson (2005) determined covariational reasoning to be an essential element of understanding slope. He worked with college students taking an intermediate algebra course and found that covariational reasoning occurs when a student is able to understand and explain how one quantity changes as another quantity changes (Adamson, 2005). As Lobato and Thanheiser (2002) described in their studies of adolescents in middle school, students develop this reasoning by first identifying the attribute to be measured, e.g. steepness of a ramp, and the quantities that contribute to its measurement, such as base length and height. Students then require assistance in using the two measurements to form a ratio that represents the attribute of steepness or slope (Lobato & Thanheiser, 2002).

Carlson, Jacobs, Coe, Larsen, and Hsu (2002) investigated how calculus students used covariational reasoning to interpret and represent dynamic situations. Their premise was that covariational reasoning contributes to a student's ability to interpret information represented in graphs, particularly functions. They proposed a framework consisting of five levels of covariational reasoning. For each level, they noted ways students could demonstrate that particular level of understanding. The first level in the framework requires students to coordinate changes in one variable with changes in a second variable. Students can demonstrate this understanding by reading a graph and labeling the axes of a graph with verbal descriptions of what the two varying quantities could represent. The second level in the framework requires students to coordinate the direction of change in one variable with the change in a second variable. Students can demonstrate this understanding by constructing an increasing or decreasing graph to represent a situation described verbally. The third level in the framework requires students to coordinate the amount of change in one variable with the changes in a second variable. Students can demonstrate this understanding by plotting points or constructing a linear graph according to specific values given in a problem. While the first two levels describe qualitative understanding, the third level introduces a quantitative comparison between values of the two variables. The fourth level in the framework requires students to coordinate and interpret average rates of change. Students can demonstrate this understanding by describing the amount a dependent variable changes as the independent variable changes by uniform increments. At this level students are able to work with non-linear functions and variable rates of change. The fifth level in the framework requires students to coordinate and interpret instantaneous rates of change of non-linear functions. Students can demonstrate this understanding by constructing smooth curvilinear graphs or by describing the instantaneous rate of change at any point in the domain of a function. These authors observed students who succeeded in advanced high school mathematics courses despite apparent voids in their understanding of covariation. They suggested these voids in understanding limit students' abilities to reason mathematically.

Another study targeting covariational reasoning specifically suggested a progression of four levels of how students describe information presented in graphs (Moritz, 2005). In this work, Moritz worked with students from third grade through ninth grade and defined covariation to be the correspondence of variation, or equivalently, the correspondence of two variables. He recommended graph production and graph interpretation as important activities by which students develop covariational reasoning. Reading a graph requires a student to be able to understand the axis labels, scales, and relationships between the two variables, and to understand what the data items expressed as ordered pairs mean. As such, the first of the four levels in Moritz's model consists of a student having a visual, holistic image of a graph or being able to describe the meaning of the axis labels. The second level consists of students recognizing univariate data patterns, such as describing the variable on the x-axis as increasing or decreasing. The third level consists of students describing how each of the two variables change, but not necessarily offering a description of any relationship between these changes. The fourth level consists of students being able to describe how both variables change in relation to each other, including descriptions of their directions of change. Moritz used three types of problems to probe students about their conceptions of covariation. Students were required to translate verbal statements into graphs, translate scatter plots into statements, and interpolate data given a graph. Moritz found that students' abilities to verbally and numerically interpret graphs were highly correlated and suggested this may be due in part to the fact that both of these actions require students to read and interpret the axes and labels of a graph. He also found that typical errors stemmed from students using single bivariate data points in isolation rather than in relation to other data. Other errors stemmed from reasoning about one variable at a time rather than about both variables in concert.

Wavering (1989) investigated the reasoning middle school students used when constructing graphs from tables of data. While this study categorized student responses in regard to the procedures they applied to complete tasks, Wavering, like Moritz (2005), found that students apply univariate reasoning to bivariate data and graphs before they are able to conceive of two variables changing in relation to one another. Wavering used nine categories to classify how students responded to tasks in which they were asked to graph data presented in tables. Students in the first category made no attempt to construct a graph. Students in the second category ordered the data for each variable. Students in the third category ordered the data for each variable and plotted the resulting, erroneous ordered pairs. Students in the fourth category ordered the data pairwise and plotted the data, but the axes were not appropriately scaled. Students in the fifth category ordered and plotted the data, ordering one axis and scaling the other axis, thereby accurately graphing linear, positively correlated data. Students in the sixth category, given randomly ordered data representing a linear function with a negative slope, first ordered the values of each variable from least to greatest, and then plotted erroneous, positively correlated ordered pairs. Students in the seventh category, also given data representing a linear function with a negative slope, graphed the data correctly with a negative slope, but were unable to interpret the relationship between the variables. Students in the eighth category graphed and interpreted data representing functions with negative slopes. Students in the ninth category were able to graph and interpret positively and negatively correlated data in linear and non-linear relationships. The types of reasoning described in categories two through four refer to univariate reasoning, whereas the types of reasoning described in categories five through nine refer to increasing levels of covariational reasoning.

Proportional Reasoning

Proportional reasoning "involves a sense of covariation, multiple comparisons" (Cramer, Post, & Behr, 1989, p. 445), and the ability to perform multiple processes (Lesh et al., 1988). Proportions consist of pairs of equal ratios (National Research Council, 2001). Therefore, reasoning about proportions requires students to compare ratios, i.e. to perform operations on operations, which according to Piaget, constitutes formal thought (Hoffer, 1988; Lesh et al., 1988). For this reason, proportional reasoning is regarded as a capstone of elementary mathematics and a cornerstone of secondary mathematics (Lamon, 1993).

Qualitative reasoning about proportions. Proportional reasoning should be learned through both qualitative and quantitative methods (Cramer et al., 1989). Qualitative reasoning is present when a student can perceive and interpret how the proportional quantities relate to one another without necessarily working with specific values, whereas quantitative reasoning is present when a student solves proportions in terms of specific numerical values. In a study aimed at helping middle school students to distinguish proportional relationships from non-proportional relationships, Cramer et al. (1989) developed multistep tasks in which students were presented with descriptions of different situations involving two varying quantities. The students were required to construct tables of data, generalize the data into rules governing the bivariate relationships, and graph their data. The rules were not required to be symbolic or algebraic representations. They could be descriptive of the invariant properties governing the relationship between the two variables concerned. After the students had processed each of the tasks, they were asked to compare and contrast the features of each of the tasks in order to highlight the differences between proportional relationships and non-proportional relationships. The benefit of working through these tasks was that students learned through discovery that what distinguished proportional relationships was that they were always described by a multiplicative rule and were always represented by straight line graphs intersecting the origin.

Proportional reasoning is multifaceted in that it depends on a person's qualitative sense of covariation as well as that person's quantitative ability to work with rational numbers (Heller, Post, Behr, & Lesh, 1990). The conceptual links between rational number concepts and direct proportions complicate a person's ability to reason proportionally. In a study aimed at uncovering how junior high school students reasoned qualitatively about rates, Heller et al. (1990) developed a sample of tasks in which students were asked to reason about how the values of rates represented as fractions would change based on different changes in the numerators or denominators of these fractions. These tasks were presented either numerically in terms of the values in the numerators and denominators of fractions or contextually in terms of different applications of proportions such as speeds, scales, or mixtures. Students were asked to determine whether the rate increased, decreased, or stayed the same based on alterations to the number values or details of the scenario. These authors found that common student errors stemmed from using additive rather than multiplicative reasoning in approaching proportional situations. They also concluded that students' abilities to solve numerical fraction problems appeared to draw on different reasoning than the reasoning used to reason qualitatively about proportional situations. The authors suggested that students who may have memorized the procedures needed to solve numerical problems, such as comparing ratios, lacked the conceptual foundation about proportions to solve problems presented contextually, such as comparing rates. They concluded that developing students' abilities to reason qualitatively about proportional situations may be an important, though overlooked, prerequisite for quantitative reasoning about proportions.

Reasoning with ratios varying in complexity. Several authors have identified an interaction between students' proportional reasoning and the complexity of the values of the ratios given in a problem (Tourniaire & Pulos, 1985). Two types of comparisons are necessary to reason proportionally (Cramer, Post, & Currier, 1993; Vergnaud, 1988). Between-type comparisons examine two different kinds of quantities or two quantities that are measured in different units (Cramer et al., 1993). Within-type comparisons examine two like quantities or two quantities measured in the same unit (Cramer et al., 1993). Regardless of whether the integral value is the result of comparing within or between two different types of values presented in a problem, students prefer to work with ratios that simplify to integers (Cramer et al., 1993; Hart, 1988). Many investigators have explored the progression students make as they develop the ability to compare different values in proportional situations (Tourniaire & Pulos, 1985).

Noelting (1980a) worked with young children to identify how children's thinking progressed and developed into proportional reasoning. In order to reduce any effects introduced by different problem contexts, Noelting maintained the use of a single problem context, namely juice mixture concentration, and varied the numerical comparisons students were asked to make when solving problems. Unity of content permitted the investigator to focus on describing the hierarchy of reasoning strategies used by children as they responded to proportional reasoning tasks. Noelting determined a constant order of successive stages in acquiring proportional reasoning in addition to a rationale for how students passed from one stage to the next. The different stages are related to the number and type of comparisons students make and to the values of the ratios in those comparisons. Students graduate from one to multiple comparisons, and they proceed from correctly comparing values whose ratios simplify to integer values, to those whose ratios simplify to unit fractions, and finally to those whose ratios simplify to fractions that do not simplify to integers or unit fractions.

Noelting's experiment consisted of 23 questions about comparing the relative orange taste of a pair of drinks made up of different amounts of orange concentrate and water (Noelting, 1980b). Each drink was described by an ordered pair of the form (*j, w*), where the number in the place of *j* represented the amount of juice concentrate, and the number in the place of *w* represented the amount of water. The students were asked to describe which of two ordered pairs represented a drink with a stronger orange taste and to explain why they thought so.

Noelting (1980b) described a sequence of eight stages through which students proceed as they acquire proportional reasoning strategies. Students in the first stage only identify the elements, such as the orange concentrate and the water. Students in the second stage compare only the first values in the two ordered pairs. These comparisons are called between comparisons as they reflect comparisons of the amounts of juice concentrate in two different drinks. Students in the third stage identify equal first values between two ordered pairs, and proceed to compare the second values in the ordered pairs. For example, one of Noelting's items presented two drinks in which the first drink had one unit of orange concentrate and two units of water, and the second drink had one unit of orange concentrate and five units of water. Students in stage three noticed the equality of the first terms and proceeded to compare the values of the second terms, whereas students in stage two viewed these ordered pairs as equal, having evaluated only the first elements of the ordered pairs. Students in the fourth stage identify a reverse relation within the ordered pairs. For example, when presented with the ordered pairs (3, 4) and (2, 1), these students can see that the ratio 3:4 is smaller than the ratio 2:1 simply because 3 < 4 and 2 > 1. These comparisons are called within comparisons as they reflect comparisons within juice concentrations rather than between two concentrations.

Students in Noelting's fifth stage are able to recognize the equivalence class of the 1:1 ratio (Noelting, 1980b). Students in the sixth stage are able to recognize the equivalence class of any ratio. Students in the seventh stage are able to compare two ordered pairs in which one of the between comparisons is a multiple. For example, given (1, 3) and (2, 5), 2 is a multiple of 1. Since 2 = 2 times 1, the student compares 5 to 2 times 3. Noticing that 5 < 6, then, the student in this stage correctly concludes that the juice concentration described by (1, 3) is weaker than the juice concentration described by (2, 5). Students in the eighth stage are able to compare any pair of ratios describing different juice concentrations and correctly determine which contains a higher juice concentration.

Another study investigating the effects of numerical values on middle school students' abilities to reason proportionally also presented tasks that varied in numerical complexity (Pulos, Karplus, & Stage, 1981). In this study the authors developed an instrument to compare middle school students' proportional reasoning with problems containing ratios that simplified to equal integer values to students' proportional reasoning with problems containing ratios that simplified to unequal, non-integer values. These two categories produced a main effect demonstrating that working with equal integer ratios is much easier for students than working with unequal, non-integer ratios.

In another study, Karplus, Pulos, and Stage (1983b) examined the number and types of comparisons middle school students made when reasoning about proportional situations. This study used lemonade recipe puzzles for the context of all problems in order to remove potential context effects from student responses. The authors identified six cognitive strategies related to proportional reasoning. The first strategy is to compare or construct equal integer ratios. The second strategy is to compare an integer ratio to a non-integer ratio. The third strategy is to compare or construct equal non-integer ratios. The fourth strategy is to compare unequal non-integer ratios. The fifth strategy is to use one type of comparison either between recipes or within a recipe. The sixth strategy is to use both types of comparison and all four values relevant to the proportion in the problem. The first four strategies formed a Guttman scale, indicating a linear progression from one type of ability to the next. In the case of the fifth and sixth strategies, there was evidence that using one type of comparison preceded using both types of comparisons. These authors found that in general, as the difficulty of the ratios in a problem increased, the use of additive reasoning also increased. They also reiterated the earlier finding that students were drawn to integral ratios, regardless of whether these occurred between ratios or within ratios.

Proportional reasoning in different problem contexts. Proportional reasoning has been investigated primarily by examining how students work with word problems concerning mixtures, such as juice concentration, or rates, such as speed (Tourniaire & Pulos, 1985). These two families of problems differ in how units of measurement are treated within the context and their uses of continuous and discrete quantities, both of which introduce difficulties for students. Rate problems typically include two different quantities measured in different units, and the ratio formed by these two quantities represents a comparison of the two original quantities. Conversely, mixture problems concern two similar quantities using the same unit of measure, and the ratio formed by the two quantities in a mixture context normally forms a new object. An example of a mixture that results in a new object is when red and yellow paint are mixed to form orange paint. Another factor contributing to the difficulty of proportion tasks has to do with whether the quantities referenced in a problem are discrete, countable objects or continuous amounts. Students more easily conceive of discrete, countable sets than continuous amounts. Since mixture contexts often involve continuous amounts, and there is a new quantity formed by the mixture, these problems are more difficult for students.

Karplus, Pulos, and Stage (1983a) investigated how middle school students responded to four different types of proportional reasoning tasks that varied in terms of the content referenced by the task and the numerical values presented in the task. They sought to identify what types of comparisons students made and what strategies they employed in solving proportion problems. They found that the context or setting of a problem and its numerical complexity both affected the types of comparisons the students made and what strategies they used. More familiar contexts, integral ratios, and small numbers were easier for students than less familiar contexts, non-integral ratios, and large numbers. These authors suggested that successful proportional reasoning can be described in terms of three steps to solving problems. The first step is to identify the two variables that are proportionally related in the problem. The second step is to recognize the rate or ratio that relates the two variables and remains constant in the relationship. The third step is to apply this relationship to make predictions or to compare two different rates.

Although there are many applications of proportional reasoning to a wide variety of problem contexts, school learning situations focus on solving missing value problems with a standard algorithm and comparing two ratios given the four values to construct them (Farrell, 1985). While proportional reasoning can be used to solve these problems, these problems can also be solved through rote skill, i.e. by applying the standard multiply then divide algorithm.

A great variety of problems should be used to develop proportional reasoning (Lesh et al., 1988). A useful problem type requires students to anticipate how the relationship between two ratios might change if one of the values changes. In these problems, a proportion is given in terms of four values, such as $\frac{A}{B} = \frac{C}{D}$. The student is asked to explain how the relationship between the two ratios would change if one of the values was increased or decreased. For example, if the value of A was increased and the other three values remained constant, then the relationship between the ratios would change from two equal ratios to unequal ratios, i.e., $\frac{A}{B} > \frac{C}{D}$. Such problems can be used by providing equal or unequal ratios with the assignment of creating a pair of ratios with a specific relationship, i.e. equal, greater than, or less than. A second useful problem type requires students to convert from ratios to rates to unit rates. Unit conversions are particularly relevant to real world applications in science and economics. A third useful problem type, which these authors identified to be most difficult for students, requires students to translate a ratio from one representation to another representation, such as deriving a numerical rate from the information given in a verbal description of that rate.

Proportional relationships are special cases of covariation, in which two quantities vary in such a way that as one quantity changes, the second quantity changes in a precise way (Lamon, 2005). In some of these relationships, namely direct variation situations, the variation can be captured by a single quantity called a constant of proportionality. Students who can reason proportionally not only conceive of the two measures varying together as in all covariational relationships, but they also can abstract this relationship into a single quantity or scale factor. As students prepare for algebra, they should be able to build a table of values given the description of directly proportionally related quantities. They should be able to find the scale factor that relates one column of numbers to the other column of numbers, and eventually derive the function that relates the two sets of values. Proportional thinkers are able to abstract the rates from direct proportions so that they can work with both unit and composite rates, such as 26 miles per 2 gallons of gas and 13 miles per gallon.

Proportional reasoning requires students to be able to think abstractly because they must compare ratios to determine equivalence of an ordered relationship (Post, Behr, & Lesh, 1988). Proportional relationships form a bridge between numerical patterns and linear relationships that can be represented algebraically in the form of direct proportions using equations of the form of $y = mx$, which can represent a wide variety of economics and science problem scenarios. Although proportions made up of two equivalent ratios with one missing value dominate classroom materials and applications, these alone fail to support student understanding of unit rate, particularly when rates are represented numerically or graphically. Proportions represented in tables, graphs, symbols, and stories offer teachers and students multiple modalities for exploring how to translate among representations in order to communicate about phenomena that involve constant rates of change.

In a study investigating the relationship between students' understanding of proportional reasoning and their understanding of steepness, Cheng (2010) found a positive correlation between these two types of knowledge. Her work and interviews with middle school students revealed that students chose different strategies based on their proportional reasoning abilities when they solved steepness problems. Students who earned higher scores on a test of proportional reasoning were more likely to consider two quantities to derive a measure of steepness, perceive unit rates, and effectively use ratios to compare steepness of two or more objects than students who earned lower scores on the same test of proportional reasoning. Additionally, students who reasoned proportionally were more successful solving problems in less familiar contexts and when different unit labels were used for the two varying quantities in a problem.

Sources of Difficulty in Understanding Slope

There are many ways to perceive of slope ranging from real manifestations of rates to a variety of mathematical representations, which makes this concept very complicated to learn (Orton, 1984; Stump, 2001). The concept of slope, or rate of change, is manifested physically in reality by the steepness of various objects such as ramps or hills, by speeds of traveling objects, by changes in states such as temperatures or growth of plants and animals, and by unit costs. These rates and slopes are associated with a variety of mathematical representations such as functions, graphs, tables of numerical data, and verbal descriptions. Calculus students experience difficulty when working with the slope concept in each of its representations and particularly when translating between representations (Stump, 2001).

From middle school throughout high school, students experience difficulty with the numerical aspects of computing slope (Barr, 1980; Seymour & Lehrer, 2006). Common errors are to construct the slope ratio upside down or to compute an incorrect value for the slope of a line passing through two points. These errors extend from an incomplete understanding of slope and often correspond to procedurally focused instruction (Barr, 1980). Students also struggle to abstract a single value for slope from the two values used in the ratio to compute slope (Stump, 1999). In fact, students frequently fail to recognize non-fractional values of slope because instruction often focuses on slope as a ratio rather than a rate (Barr, 1980, 1981). A similar error manifests itself graphically, where students understand slope to be the ratio of $\frac{rise}{run}$ but have trouble interpreting it as a quotient that determines a line's orientation in the plane or a rate that governs the relationship between two variables (Seymour & Lehrer, 2006; Walter & Gerson, 2007).

Students experience difficulty interpreting features of graphs, such as slope, that are not directly readable from the graph (Dugsdale, 1993). They have trouble interpreting any global features of graphs in terms of the context or labels on the axes of graphs (Bell & Janvier, 1981; Dugsdale, 1993; Kieran, 1993). Students readily misinterpret graphs whose appearance does not pictorially represent the information that is graphed (Bell & Janvier, 1981). For example, the shape of a graph whose context is that of climbing a hill is confused with the shape of a hill itself (Bell & Janvier, 1981). Students also confuse the values represented on a graph (Bell & Janvier, 1981). Typical errors involve mistaking the amount of increase (Δy) with the value of the dependent variable (y) or the rate of increase (slope) with the amount of change in the dependent variable (Δy) (Bell & Janvier, 1981). This error is manifested verbally when students erroneously interchange the terms "growing more" and "growing faster" (Bell & Janvier, 1981, p. 38).

Seymour and Lehrer (2006) identified the need for middle school students to learn and use appropriate vocabulary to work with mathematical concepts in different representations. In one example, students used the word "times" when describing the rule depicted in the equation of a line. The teacher had to ask how this "times" was represented in the graph of the equation. The students' responses led the teacher to conclude that they used this word to indicate a size difference, an additive interpretation, rather than as a multiplicative operation. Students persistently favored repeated addition when working with the graph of the line. This was demonstrated by their preference for reading slope from point to point instead of interpreting the functional relationship between the two variables.

Orton (1984) found that high school calculus students, who typically succeeded through at least two years of algebra instruction, struggled with concepts such as constant speed, variable speed, average speed, and instantaneous speed. These students failed to correctly associate described acts or motions with appropriate mathematical representations. More specifically, they were less able to interpret rates of change with tables of data than they were when given graphs of lines. That is, students had difficulty describing a rate of change in terms of the two variables depicted in a table of values, but they had more success describing a rate of change in terms of the two variables depicted in a graph. Orton also found that these students experienced less difficulty interpreting graphs built from tables of data than from graphs accompanied by their corresponding symbolic equations. These calculus students had the greatest amount of difficulty interpreting any mathematical representation of a rate in terms of its context. For example, in studying problems involving cars traveling from one place to another, students failed to consider the implications of traffic lights in how speed changed during travel.

Similarly, Stump (2001) found that high school pre-calculus students needed practice in examining both physical manifestations and mathematical representations of slope. She identified gaps in students’ understanding of slope as rate of change, their ability to form connections between equations of linear functions and their graphs, and their ability to form connections between graphs of linear functions and their corresponding rates of change or slopes. The conclusions from this study suggested that many students did not use ratios or proportional reasoning when examining steepness or associating steepness to slope and rate of change.

In a study aimed at identifying students’ knowledge of steepness, slope, and rate of change, Teuscher and Reys (2010) found that high school calculus students struggled to relate these three qualities to the graphs of linear functions. They identified that many students shared a tendency to view steepness, slope, and rate of change as completely separate concepts. Some students, however, viewed the terms synonymously, likely due to experience with poorly designed classroom application problems. For example, when a student learning about slope was asked to compute the slope of a roof, that student rarely was asked to consider the direction of the pitch. This reinforced the notion that direction did not always matter, and at the same time promoted confusion between the definitions of steepness and slope. Steepness refers only to the pitch or incline, whereas slope and rate of change both carry a directional or signed value. The common problem identified by these researchers was that students had not developed appropriate conceptual connections among the concepts of steepness, slope, and rate of change, and were unable to associate these correctly with mathematical representations, particularly graphs.

The absence of conceptual understanding causes students to make serious errors when solving problems involving slope (Barr, 1980). Students are unable to transfer their knowledge of slope when it is based on rote, mechanical process learning (Barr, 1980, 1981; Kieran, 1993). There must be conceptual grounding for the concept of slope for students to be able to apply their knowledge appropriately to novel problems (Barr, 1980, 1981). Slope should not be a mindless algorithm.

In a study aimed at identifying how students work with slope and graphs, Reiken (2008) found that teenaged algebra students most often performed procedurally but seldom demonstrated conceptual understanding of slope. He found that many students were able to compute slope values by using a numerical formulaic approach, i.e., $\frac{y_2-y_1}{x_2-x_1}$, by performing a similar procedure to measure distances between points on the graph of a line, i.e., *rise over run*, or by picking the slope out of an equation as the number in front of the "*x*." He found that only some students were able to work with slope as a value that related two sets of numbers when they were listed in a table of values, and even fewer students used slope as a measure of the rate of change in descriptions that related how one variable changed in response to the other variable's change. The last two capabilities indicated more conceptual understanding than the procedural behaviors, but detecting how students interpreted slope from information presented in tables or graphs required activities in which students worked with more than one representation.

Adamson (2005) explored student reasoning about slope with college algebra students. He identified three types of reasoning that contributed to students' ability to work with slope or rate. Computational reasoning was required to produce accurate numerical answers, covariational reasoning was required to describe how one quantity changed in relation to another quantity, and proportional reasoning was required to explain how one quantity changed in a constant way relative to another quantity's change. He found that proportional reasoning was present when students were able to accurately apply and explain scalar relationships. Like Lobato and Thanheiser (2002), Adamson (2005) found that students experienced difficulty identifying what quantities contributed to calculating steepness or rate. His students also experienced difficulty when trying to interpret the meaning of slope in real world contexts or when tables and graphs were labeled with units other than time, people, and money.

Descriptions of Instructional Activities for Teaching about Slope

A person's ability to understand the meaning of slope depends upon that person's ability to acknowledge the relationships that exist between different representations of the slope concept (Seymour & Lehrer, 2006; Stump, 1999). Therefore, students should learn about slope in multiple representations (Magidson, 2005; Moschkovich et al., 1993). Instruction should target assisting students to make connections among the different representations of slope (Magidson, 2005) so that students will acquire complete knowledge of the slope concept rather than partial knowledge that is void of essential connections among different representations of slope (Moschkovich et al., 1993).

Common approaches to teaching students about slope rely on graphs of lines (Barr, 1980), tables of values (Bell & Janvier, 1981), or equations of linear functions (Norman, 1993), without an emphasis on using more than one representation to explore slope (Stump, 1999). Teachers' limited content knowledge of slope may contribute to their tendency to confine their instruction to single representations (Stump, 1999). Stump (1999) found that the majority of secondary mathematics teachers she studied described slope in terms of the appearance of the graph of a line, and only a few teachers in her study interpreted slope in terms of the function that described how two variables were related to one another.

Andersen and Nelson (1994) recommended that early instruction about slope should be based on the notion of steepness, which is equivalent to the absolute value of slope. These researchers proposed that students should not be exposed to any definitions of slope until they have explored the characteristics of physical objects such as ramps, handrails, and steps, represented these on dot paper, and calculated their steepness as the vertical height divided by horizontal length. Steepness activities can also be used to demonstrate the properties of horizontal and vertical lines and relate these properties to zero and undefined slopes, respectively.

Similarly, Grocki (1990) introduced his students to the notion of steepness and how to use horizontal and vertical measurements to calculate its value by taking them to a gymnastics studio. The students measured various gymnastics equipment such as folding mats arranged as inclined planes, parallel bars set at different inclines, and the horse set in a non-horizontal, i.e., inclined, position. Grocki encouraged his students to use available measurement tools such as the equal width sections of the mats to form the bases of their measurements. For example, students measured one incline plane to be four folds of a mat in horizontal length and one fold of the same type of mat in height. Thus they estimated the steepness to be $\frac{1}{4}$.

Anderson (2008) introduced the concept of slope to her pre-algebra students by describing two boys and their walking speeds. The students built tables of values listing time as the independent variable and distances walked as two dependent variables, and used these tables to graph the two boys' walking speeds. After students spent time describing how these tabular and graphical representations related to the description of the boys' walking, the teacher introduced the slope-intercept form of a linear equation as another way to represent the situation. From this work, Anderson found that it was very important to retain the idea that slope or steepness is derived from the measures of two quantities, which can be lost when students work with symbolic linear functions in which slopes are represented by integer values.

Bell and Janvier (1981) and Smith (2003) recommended instructional activities that require middle school students to collect and model data using tables and graphs in order to explore the concept of slope. Familiarity with the contexts of problems and experiments allows students to understand the labels of graphs and foster meaning for the appearance of graphs (Bell & Janvier, 1981). Additionally, students form connections between the data they collect, the values in the tables, and the appearance of the graphs, thereby strengthening their representational fluency (Bell & Janvier, 1981). Smith (2003) also recommended this type of instruction because it would increase the focus on the covariation between the quantities in the context of the experiment or problem.

Joram and Oleson (2007) described an instructional activity based on data about the growth and heights of trees over time. They found that change over time was a familiar context that middle school students readily understood, which offered a conceptual foundation for investigations into slope. The instruction described in this research focused on the relationship between the age of a tree and its height and eventually developed the idea that a tree's height was a function of its age. The students interpreted average rates of growth from the tables of numbers, constructed number sentences based on their analyses of the tables, and graphed their data. Many students correctly associated the rates of growth determined from the tables of data to the steepness of the lines on their graphs and identified the rate as the number of feet a tree grew in one year. Like the instruction recommended by Bell and Janvier (1981) and Smith (2003), Joram and Oleson's activities prompted students to consider and form connections among different representations of slope.

Wilhelm and Confrey (2003) recommended activities using technology in which high school algebra students construct an understanding of slope as a rate by accumulating values in the context of saving money in a bank account. Such activities prompt students to model the daily deposits and daily cumulative values of a savings account using mathematical representations such as graphs and equations. By exploring this type of problem using mathematical representations, students learn to associate daily deposits with the slope of a function and the accumulated savings amount to the value of the function. The advantage of this instructional design is that it originates with a context and requires students to build mathematical representations from the context, whereas many other instructional activities reverse this process. While this design might help students to relate mathematical representations to problem contexts, these researchers found that students still had difficulty acquiring a complete understanding of the relationship between the function's rate and value and their corresponding meanings in the context of the problem.

Summary

Concept learning is based on a person's concrete experiences (Dienes, 1971). While concepts are cognitive representations of real objects and concrete experiences (Ausubel, 1968), a person's conceptions mature through education and life experiences (Martorella, 1972). When faced with new information related to an existing conception, a human being attaches the new information to existing knowledge in a process called adaptation (Adler, 1971), and through assimilation, uses this revised conception to interact with future experiences (Martorella, 1972).

Different authors asserted that learning concepts requires passage through sequential stages. Concept formation occurs when an individual internalizes the meaning of an object or idea from concrete, empirical experiences (Ausubel, 1968; Bruner et al., 1956), which is the predominant form of learning known to children from birth to about the age of three (Novak, 1998). Concept attainment involves analysis and evaluation of the criterial attributes that qualify an object or idea for membership into a category with a concept's name (Bruner et al., 1956; Martorella, 1972; Novak, 1998). Concept assimilation is a process whereby a person relates new ideas to existing, structured knowledge, which is the principal way in which people learn and refine their conceptions after early childhood (Novak, 1998).

Misconceptions occur when flawed information or erroneous connections are associated with a concept. One type of misconception occurs when a person possesses incomplete or incorrect knowledge of the components that make up a concept (Glaser, 1986). A second type of misconception occurs when a person has a flawed sense of the requirements for set membership (Henderson, 1970). A third type of misconception occurs when a person's conception is in progress (Klausmeier, 1992; Wilson, 2009), that is, when a person has organized the elementary components of a concept but has not grasped everything associated with a concept.

A variety of models are used to display how concept knowledge might be organized in a person's cognitive structure. Concept maps (Baroody & Bartels, 2001), learning hierarchies (Gagné, 1968), construct maps (Wilson, 1992, 2009), and learning progressions (Wilson, 1992, 2009) are models containing descriptions of the concepts and skills that make up larger learning targets (Popham, 2008). They define what students must learn to understand concepts and map out optimal learning sequences (Wilson, 2009).

Various authors have attempted to describe what it means to understand and document how students learn with understanding. A common theme emphasized by many authors was that conceptual understanding depends on connected knowledge. Specifically, connections between existing knowledge and new information have been and continue to be prominent in the recommendations for how to develop and foster conceptual understanding.

Progressive educators, e.g., Dewey and Pratt, emphasized the importance of helping students to connect new experiences to their existing knowledge and previous experiences (Dewey, 1938; Willis et al., 1994). When children connect their experiences to their knowledge, they are able to form meaningful understanding of newly introduced material.

Piaget asserted that "intelligent activity is always an active, organized process of assimilating the new to the old and of accommodating the old to the new" (Flavell, 1963, p. 17). Building on the work of Piaget and progressive educators, constructivists insist that human beings actively construct new knowledge in response to their personal concrete, perceptual, or abstract experiences (Confrey, 1990; Noddings, 1990; von Glasersfeld, 1990). Learning with understanding must be viewed as a continuous process of reconstructing a person's record of experiences (Dewey, 1938). Conceptual knowledge that is deliberately constructed through constructivist learning experiences is inherently connected and integrated with other concepts

(Confrey, 1990). Consequently, future learning is enhanced by the presence of networked knowledge, supporting the cognitive actions that form links to more sophisticated understandings.

Gagné (1971) advocated the idea that learning new information depended on the connections a person could build to relate the new information to what that person already knew. He developed a model of concept learning based on the notion that in order to acquire a concept, one must have learned all pre-requisite concepts; otherwise it would be impossible for such learning to occur. Gagné's (1968) theory of learning hierarchies required the identification of a hierarchy of subordinate knowledge components for each terminal learning outcome. Gagné (1968, 1971) suggested that by evaluating students' knowledge of the components of a hierarchy, teachers could accurately determine what subordinate knowledge was known and identify what needed to be learned.

Understanding mathematics depends upon richly connected knowledge (Carpenter & Fennema, 1991; Marshall, 1990; NCTM, 1989, 2000), which involves knowing concepts and using procedures as well as deliberately building progressively sophisticated cognitive networks or schema relating concepts and procedures (Beyer, 1993; Hiebert & Carpenter, 1992; Hiebert & Lefevre, 1986; Marshall, 1990; Messick, 1984; National Research Council, 2001; Vergnaud, 1997; Webb & Romberg, 1992). The depth of a person's understanding of a mathematics concept depends upon the number of nodes that person has arranged for the concept, the number and strength of these nodes' interconnections, and the person's acknowledgement of these nodes and their interconnections (Vergnaud, 1997).

Connections that promote understanding of mathematical concepts include (1) ways in which different representations are used to express the meaning of a concept, (2) ways in which concepts are associated with procedures, and (3) ways in which prior knowledge is associated with new information. Connections support fluency with concepts, procedures and representations, which is an important indicator of mathematical understanding. Connections are critical to building mathematical understanding, but they "are frequently overlooked in curriculum planning and assessment" (Boaler, 2002, p. 14).

Educational measurement theory and practice aim to develop means by which educators can describe and interpret what students know and can do. In order for the measurement process to inform valid decisions and appropriate consequences, judgments made at every stage in the measurement process must be valid. The most convincing validity rationales contain evidence from all possible means along with descriptions of the connections between the constructs assessed, the tasks used to measure them, the empirical data collected and scored, the interpretations of the data, the decisions based on these interpretations, and the consequences of these decisions (Messick, 1989, 1994b).

Valid educational measurement should originate with descriptions of the knowledge, skills, and abilities students are expected to learn (Haertel, 1985; Kane, 1992, 2001; Messick, 1984, 1994a, 1994b, 1995b; Mislevy, Almond, et al., 2003; Mislevy & Haertel, 2006; Mislevy et al., 1999; Mislevy, Steinberg, et al., 2003; Mislevy et al., 2000). The measurement process should also take into account the nature of how people come to know things, how they demonstrate their knowledge, and how observations are interpreted in terms of decisions and consequences (Mislevy, Almond, et al., 2003; Mislevy, Steinberg, et al., 2003; Mislevy et al., 2000).

Evidence-centered design (ECD) is an assessment design framework that requires test developers to build assessments by first analyzing the domain of knowledge that is to be measured and having this analysis drive test design, test delivery, scoring, reporting, and interpretation of test scores (Mislevy, Almond, et al., 2003; Mislevy & Haertel, 2006; Mislevy, Steinberg, et al., 2003; Pellegrino & Huff, 2010). Evidence-centered assessment design uses five layers to frame evidentiary arguments and cases for validity (Mislevy & Haertel, 2006). The layers needed to establish the validity argument are domain analysis, domain modeling, conceptual assessment framework, assessment implementation, and assessment delivery.

Test developers who implement ECD are likely to incorporate models of how the knowledge targeted by an assessment is arranged in a person's cognitive structure. Models can be used to represent what knowledge people possess, how knowledge is typically acquired, and how knowledge is measured. Theoretical mental models of the knowledge and cognitive processes people use when they work through problems can inform test item development because mental models delineate the components of the knowledge and processes being targeted by a particular test (Pellegrino, 1988). When mental models are incorporated in assessment design, the same mental models can be used to analyze student test responses, which can help to inform educators more specifically about what students know and can do (Pellegrino et al., 1999).

Learning progressions (Popham, 2008, 2011), construct maps (Wilson, 2009), and learning hierarchies (Gagné, 1968) are three types of models useful for describing the knowledge people possess, how it is typically learned, and common misconceptions. These types of models can be used to guide the development of assessment items and tasks that measure what people know in a manner that is consistent with how knowledge is acquired.

A relatively new type of cognitive model is the cognitive attribute hierarchy, which can be used to model the cognitive processes or skills students need to complete test items (Gierl et al., 2008; Leighton et al., 2004). In a cognitive attribute hierarchy, each component of knowledge, skill, or ability needed to demonstrate proficiency in a domain is represented by an attribute (Leighton et al., 2004). Furthermore, the attributes are arranged hierarchically to depict how people likely acquire domain knowledge, and the dependencies, or cognitive connections, among the attributes are clearly indicated (Leighton et al., 2004).

"Slope is one of the most important mathematical concepts that students encounter in middle school and high school" (Joram & Oleson, 2007, p. 261). As such, both national and state level curriculum guidelines contain statements about what students should know and understand about slope. The CCSS (CCSSO/NGA, 2010) for mathematics describe understanding of slope developing over middle and high school grades. The PSSM (NCTM, 2000) references the development of understanding of slope and its application to solving problems in the standards for middle and high school mathematics. The *Kansas Curricular Standards for Mathematics* for eighth grade list several expectations related to an understanding of slope (Kansas State Board of Education, 2003).

Understanding the slope of a line is one of the most important mathematical proficiencies students should acquire in secondary mathematics courses (Andersen & Nelson, 1994; Joram & Oleson, 2007), but acquiring the ability to reason about slope is difficult for students (Graham et al., 2010). Understanding slope builds on a student's ability to reason with ratios (Lobato & Thanheiser, 2002). Two types of reasoning contribute to a student's ability to work with ratios, namely, covariational reasoning and proportional reasoning (Hoffer, 1988).

Covariational reasoning is evident when a student can describe the relationship between two quantities that vary together (Adamson, 2005). This understanding requires a student to recognize that two varying quantities together produce a third type of quantity whose value is represented as a ratio (Lobato & Thanheiser, 2002). Covariational reasoning supports a person's ability to interpret information presented in graphs (Moritz, 2005).

Proportional reasoning develops after a person has the ability to conceive of covariation and is able to work with rational numbers (Heller et al., 1990). Proportional reasoning involves multiple comparisons, namely, comparisons between the two values that make up a ratio and comparisons between two or more ratios. For this reason, proportional reasoning is regarded as a capstone of elementary mathematics and a cornerstone of secondary mathematics (Lamon, 1993).

There are many ways to perceive of slope ranging from real manifestations of rates to a variety of mathematical representations, which makes this concept very complicated to learn (Orton, 1984; Stump, 2001). Students experience difficulty when working with the slope concept in each of its representations and particularly when translating between representations. Students struggle with the numerical aspects of computing slope (Barr, 1980; Seymour & Lehrer, 2006) and especially with interpreting slopes in terms of problem contexts (Orton, 1984).

Common errors extend from an incomplete understanding of slope and often correspond to procedurally focused instruction (Barr, 1980). Reiken (2008) found that students often performed procedurally but seldom demonstrated conceptual understanding of slope. Reiken (2008) found that only some students were able to work with slope as a value that related two sets of numbers when they were listed in a table of values, and even fewer students used slope as a measure of the rate of change in descriptions that related how one variable changed in response to the other variable's change.

A person's ability to understand the meaning of slope depends upon that person's ability to acknowledge the relationships that exist between different representations of the slope concept

(Seymour & Lehrer, 2006; Stump, 1999). Therefore, students should learn about slope in multiple representations (Magidson, 2005; Moschkovich et al., 1993), and students should form and acknowledge cognitive connections among the different representations of slope (Magidson, 2005). Organized and connected knowledge of essential components of the slope concept provides a foundation for conceptual understanding of slope.

CHAPTER THREE

RESEARCH DESIGN

Introduction

The overall structure of the present study followed evidence-centered design (ECD) (Mislevy & Haertel, 2006), which is a framework that supports educational researchers in accumulating evidence for validity throughout the assessment development and decision-making process. The scope of this study was to analyze the concept of slope, to assemble a model of how understanding of foundational concepts related to slope develops, and to construct an assessment to measure student understanding of selected foundational concepts related to slope. Therefore, this study did not require all of the steps of the ECD process described by Mislevy and Haertel (2006). Instead, this study implemented two layers of ECD, namely, domain modeling and task modeling. These two processes along with their outcomes are each described following a discussion of the participants.

The assessment method used for this study was the attribute hierarchy method (AHM), a method developed by Leighton, Gierl, and Hunka (2004). These researchers described the AHM to be a variation of the rule-space method (Birenbaum & Tatsuoka, 1993). The two methods for designing assessments are similar in that they rely on a theoretical analysis and modeling of the cognitive domain to be assessed, and they both aim to provide diagnostic information about students' knowledge states that can inform teaching and learning. Both models require a construct to be defined in terms of the attributes that collectively describe what students must know and be able to do to demonstrate their knowledge in a particular content domain.

The AHM is distinguished from rule-space because it assumes a dependent, hierarchical relationship among the attributes that describe a body of knowledge (Gierl et al., 2000; Leighton et al., 2004). The hierarchy developed for this study influenced item development, that is, the items were designed to assess whether examinees had acquired the attributes depicted in the hierarchy. The hierarchy also influenced the interpretation of test scores, that is, the scores were used to describe examinees' knowledge of foundational concepts related to slope in terms of the attributes depicted in the hierarchy.

The present study was conducted in two phases. During the first phase, the domain modeling phase, the investigator developed an attribute hierarchy model of five foundational concepts related to slope and collaborated with subject matter experts to refine its contents and structure. During the second phase, the task modeling phase, an instrument was developed to assess student understanding of the five foundational concepts related slope as described by the model defined in the first phase. Then the assessment was administered to students in middle and high school grades through the Kansas Computerized Assessment (KCA) software, which is an online environment.

Purpose of this Study

The goal of this study was to explore the concepts students should possess to demonstrate the ability to consider a problem or a graph depicting a direct variation relationship and to interpret the rate or slope that governs the variation relationship in terms of the problem's context variables. The primary objectives of the study were to determine a model to represent selected foundational concepts related to understanding slopes depicted in graphs or described in problems and to develop a means by which such understanding may be assessed. The study included the development of an instrument to assess the foundational concepts related to understanding slope.

Research Questions

This investigation answered these research questions.

1. What insight is gained about the validity of the proposed cognitive model from an analysis of student data generated from an assessment informed by the model?
2. To what extent did student participants exhibit common misconceptions regarding slope?

The domain modeling phase provided a theoretical description of the attributes contributing to understanding of five foundational concepts related to slope as well as their hierarchy. This model was called the Foundational Concepts of Slope Attribute Hierarchy (FCSAH) and informed the task modeling phase and the resulting assessment, that is, the Foundational Concepts of Slope Assessment (FCSA). Each item on the FCSA was designed to evaluate one precise set of attributes in the FCSAH. The data collected from the FCSA was analyzed to confirm the FCSAH and to provide answers to the first and second research questions.

Participants

The FCSA was administered to students in middle and high school grades studying Pre-algebra, Algebra 1, Geometry, Algebra 2, or courses with similar content taken before Pre-calculus. According to the Kansas curriculum standards (Kansas State Board of Education, 2003), the concept of slope should be taught and is tested on the state assessments for eighth grade and in high school. According to the Common Core State Standards for Mathematics (CCSSM) (CCSSO/NGA, 2010), students should develop an understanding of unit rates beginning in sixth grade, and this understanding should increase in sophistication throughout the middle grades. Eighth grade students should be able to interpret a unit rate as the slope of the graph of a proportional relationship (CCSSO/NGA, 2010). High school students are expected to apply their understanding of constant unit rates to their study of functions and modeling (CCSSO/NGA, 2010). These curriculum standards convey expectations for students in these courses to possess some knowledge of slope and be able to demonstrate their knowledge. Additionally, students in these courses should have varied levels of knowledge of slope, as some have greater experience in mathematics than others.

In May 2011, 30 teachers administered the FCSA to a sample of 1629 students. These students attended 26 different school districts in Kansas. There were 630 students enrolled in Pre-algebra, 492 students enrolled in Algebra 1, 365 students enrolled in Geometry, and 142 students enrolled in Algebra 2. Individual student demographic characteristics were not collected. To provide some description of the students in this sample, district demographic information was collected for every district represented in the student sample. This information was obtained from the publicly available school report card information for the academic year 2009-2010 from the Kansas State Department of Education (KSDE) website. Each district report card listed the percent of students by gender, socio-economic status, and race. The percents of students in these districts that were male, female, economically disadvantaged, not economically disadvantaged, African American, Hispanic, other race, or White are shown in Table 1.

Table 1

Number of Student Participants and Demographic Information for Each Participating District

		Gender		Socio-economic Status		Race			
District	Student Participants (*n*)	Male	Female	Disadvantaged	Not Disadvantaged	African American	Hispanic	Other	White
1	46	54%	46%	55%	45%	0.2%	5%	7%	88%
2	18	54%	46%	52%	48%	0.4%	3%	2%	94%
3	135	52%	48%	23%	77%	6.5%	12%	8%	74%
4	50	53%	47%	65%	35%	0.3%	7%	2%	91%
5	84	53%	47%	43%	57%	2.5%	1%	30%	66%
6	161	51%	49%	15%	85%	3.2%	6%	4%	87%
7	111	51%	49%	40%	60%	2.2%	2%	3%	93%
8	74	51%	49%	64%	36%	1.3%	4%	4%	91%
9	100	50%	50%	57%	43%	0.5%	3%	4%	92%
10	51	51%	49%	57%	43%	0.9%	3%	4%	93%
11	24	58%	42%	36%	64%	0.0%	6%	2%	92%
12	47	52%	48%	57%	43%	7.5%	8%	9%	75%
13	36	49%	51%	56%	44%	0.6%	26%	2%	72%
14	6	56%	44%	48%	52%	0.7%	1%	3%	95%
15	35	50%	50%	46%	54%	0.0%	3%	3%	94%
16	81	52%	48%	33%	67%	0.3%	6%	8%	86%
17	35	51%	49%	24%	76%	0.0%	2%	1%	97%
18	64	51%	49%	51%	40%	2.8%	5%	12%	80%
19	113	53%	47%	40%	60%	0.9%	6%	4%	89%
20	64	51%	49%	51%	49%	3.0%	5%	4%	88%
21	83	52%	48%	35%	65%	0.6%	3%	6%	91%
22	13	51%	49%	28%	72%	0.6%	1%	2%	97%
23	88	53%	47%	64%	36%	12.8%	5%	7%	75%
24	75	51%	49%	6%	94%	2.9%	4%	12%	81%
25	37	51%	49%	49%	51%	0.0%	22%	3%	75%
26	113	53%	47%	40%	60%	0.9%	6%	4%	89%

Domain Modeling Phase

Purpose

The purpose of the domain modeling phase was to analyze the slope concept to identify the attributes, that is, the specific concepts and skills, contributing to a person's understanding of slope. Then the attributes were arranged in a hierarchical model, i.e., the FCSAH, depicting the independent and dependent relationships among the attributes. The modeling process was informed by the work of Leighton and Gierl (Gierl, 2007; Gierl et al., 2000; Leighton & Gierl, 2007; Leighton et al., 2004). These investigators advocated that cognitive analysis should be the starting point of any assessment developed with the goal of diagnosing student knowledge. Cognitive analysis is often accomplished cooperatively by a researcher and subject matter experts and yields a cognitive model that identifies the specific attributes associated with a particular body of knowledge (Gierl et al., 2008). The model displays these attributes to reflect how they are hypothetically arranged and interconnected in the mind of a subject matter expert. Cognitive models can be based on think aloud sessions, during which students say out loud what they are thinking as they respond to a variety of questions or problems (Gierl et al., 2008). Cognitive models may also be based on previous research that identifies the concepts and skills students must possess to be prepared to acquire more sophisticated knowledge.

Norris, Macnab, and Phillips (2007) described a cognitive model as a means to provide plausible links between conceptual understanding and test performance outcomes. Because a cognitive model replicates how concepts and skills are acquired and connected in a person's cognitive structure, the model can be used to represent the specific combinations of attributes students must possess in order to successfully complete different types of questions and problems in a particular content area. When a cognitive model guides assessment development and practices, test users are able to gather information about examinees in terms of what knowledge they do and do not possess. This information strengthens test interpretations by including diagnostic information describing what examinees understand about the material that was assessed and what they still must learn, thereby offering teachers detailed information to guide their instructional planning (Norris et al., 2007).

Foundational Concepts of Slope Attribute Hierarchy (FCSAH)

The FCSAH developed for this study incorporated the work of scholars who have studied student understanding of covariation and proportional reasoning. The FCSAH is displayed in Figure 1 and described in the paragraphs that follow.

Drawing on the work of Adamson (2005), the first two attributes of the FCSAH consisted of a student's ability to perceive of covariation in different problem contexts. The first attribute concerned the ability to detect which quantities in a problem situation vary in correspondence to one another without any reference to their directions of change. The second attribute concerned the ability to identify the direction of change of two covariates in constant rate problem contexts, as depicted in the framework developed by Carlson et al. (2002). The first and second attributes in the FCSAH were positioned in the model as prerequisites for proportional reasoning.

Drawing on the work of Adamson (2005) and Cheng (2010), the third, fourth, and fifth attributes modeled in the FCSAH concerned a student's proportional reasoning abilities. These three attributes concerned the ability to interpret the meaning of the slope ratio in terms of the context of a problem presented either verbally or graphically. Three different types of ratios were modeled in the FCSAH to reflect the work of Noelting (1980a, 1980b) and Karplus, Pulos, and Stage (1983b), who studied how students learned to interpret the meaning of quantities expressed as ratios. The third attribute in the FCSAH concerned a student's ability to interpret the meaning

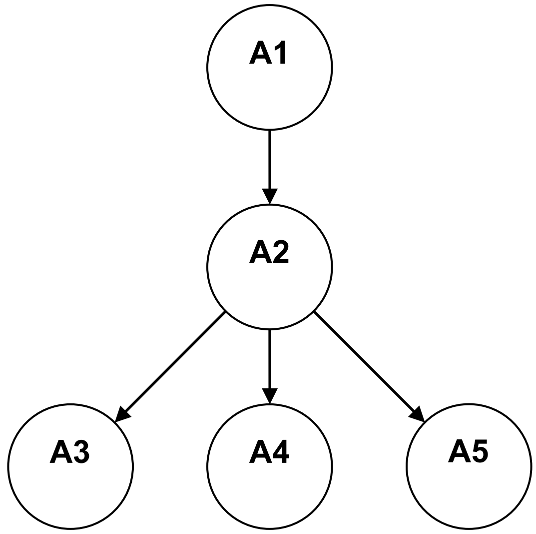


Figure 1. The Foundational Concepts of Slope Attribute Hierarchy (FCSAH) depicts five attributes associated with understanding the concept of slope. Attribute A1 is defined as the ability to identify covariates from a problem scenario. Attribute A2 is defined as the ability to identify covariates and the direction of their relationship. Attribute A3 is defined as the ability to interpret a slope whose value equals a whole number. Attribute A4 is defined as the ability to interpret a slope whose value simplifies to a positive unit fraction. Attribute A5 is defined as the ability to interpret a slope whose value simplifies to a positive rational number that is neither a whole number nor a unit fraction.

of the slope ratio when its value was a whole number. The fourth attribute in the FCSAH concerned a student's ability to interpret the meaning of the slope ratio when its value was a positive unit fraction, that is, a ratio whose simplest form equals one over an integer greater than one. The fifth attribute in the FCSAH concerned a student's ability to interpret the meaning of the slope ratio when its value was a positive rational number value but neither an integer nor a unit fraction.

Students' difficulties in learning about slope are complicated by the fraction representations of different ratios. Studies completed by Noelting (1980a, 1980b) and Karplus, Pulos, and Stage (1983b) revealed that students' abilities to solve proportional tasks are related to the complexity of the numerical values presented in the tasks. Students have the least difficulty working with ratios whose fraction representations simplify to integers, more difficulty with ratios whose fraction representations simplify to unit fractions, e.g. one over an integer greater than one, and the most difficulty with ratios whose fraction representations do not simplify to either of these forms. To account for different numerical complexity, the FCSAH developed for this study included attributes concerning slopes that simplified to whole numbers, positive unit fractions, and positive rational values in neither of these forms.

The slope ratio expresses a single comparison of two quantities (Hoffer, 1988). Each quantity in the slope ratio represents an amount of change in one variable. When interpreting slope, students must compare these two amounts of change. After students are able to conceptualize the slope ratio, they learn that the value of the slope ratio remains constant for a particular linear function. A unit rate represents the amount of change in the dependent variable when the independent variable's value increases by exactly one unit. Thus a unit rate can be any real-number value. For example, if the slope of a line is two-thirds, then the unit rate is the value 2/3 or approximately 0.667, meaning for every increase of one unit in the independent variable, the dependent variable increases by 2/3 of a unit.

Procedure for Developing the FCSAH

The researcher scrutinized the literature and previous investigations into concepts related to slope to identify foundational concepts that contribute to understanding of slope. Using the literature and research findings, the researcher arranged five attributes into the Foundational Concepts of Slope Attribute Hierarchy (FCSAH) to depict the order in which the attributes might optimally be acquired. Drawing on the work of Lesh, Post, and Behr (1988), the first draft of the FCSAH developed for this study included a total of seven attributes. The ability to interpret the slope ratio in terms of two values was distinguished from the ability to interpret the slope ratio as a single value or unit rate. However, with seven attributes in the FCSAH, it was not practical to develop an assessment containing enough items to reliably measure knowledge of each attribute and one that could be completed in one class period. Therefore, the researcher reduced the number of attributes by using one attribute for the ability to interpret the slope ratio in terms of its two values and the ability to interpret the slope ratio as a unit rate. This process affected attributes A4 and A5 in Figure 1.

The FCSAH was shared with two independent subject matter experts (SMEs) in order to gain their opinions on both the collection of attributes identified as contributing to an understanding of foundational concepts related to slope and how the attributes were hierarchically arranged. One expert was from a large Midwestern university. This individual had recent experience working with undergraduate students in remedial mathematics classes at the university. The other expert was a secondary mathematics teacher with experience working with middle school mathematics students in Kansas. After both SMEs reviewed the FCSAH independently, the researcher met with the two experts together to discuss their reactions.

Both SMEs agreed with the overall structure of the FCSAH in that covariational reasoning should precede interpreting slope values in terms of problem contexts. However, both SMEs also expressed concern about the order of attribute A1 and attribute A2. The SMEs thought that students might find the problems for attribute A1 to be more difficult than the problems for attribute A2. The problems for attribute A1 were to be presented in verbal form, whereas the problems for attribute A2 were to be presented using graphs. The SMEs thought that problems containing graphs would be more familiar than problems containing only verbal descriptions. After discussing this with the SMEs, the researcher reviewed the literature about student understanding of covariation. The first level in a framework from Carlson et al. (2002) described the ability to coordinate the changes in one variable with changes in a second variable, whereas the second level described the ability to coordinate the direction of change in one variable with the change in a second variable. After sharing this framework with the SMEs, the researcher and the SMEs agreed to leave attributes A1 and A2 as they were modeled in the FCSAH.

Task Modeling Phase

Purpose

The purpose of the task modeling phase of the study was to use the FCSAH developed in the domain modeling phase to inform the development of an assessment comprised of items designed to evaluate the attributes and their dependent hierarchical combinations. Following the process described by Gierl, Leighton, and Hunka (2000), the researcher determined several matrix representations of the FCSAH, the last of which directly influenced item and test design and subsequently score interpretations.

Test Development Based on the FCSAH

Adjacency matrix (A). The adjacency matrix depicts the direct (one-step) relationships of attributes in an attribute hierarchy with ones and zeros (Gierl et al., 2000). The adjacency matrix for the FCSAH is order 5 × 5 and contains one row and one column for each attribute in the FCSAH. A one in position (*j, k*) indicates attribute *j* is a prerequisite for attribute *k*. Zeros occupy the remaining elements of each row of the adjacency matrix. The adjacency matrix for the FCSAH is shown below.

$$\begin{array}{c|ccccc} & A1 & A2 & A3 & A4 & A5 \\ \hline A1 & 0 & 1 & 0 & 0 & 0 \\ A2 & 0 & 0 & 1 & 1 & 1 \\ A3 & 0 & 0 & 0 & 0 & 0 \\ A4 & 0 & 0 & 0 & 0 & 0 \\ A5 & 0 & 0 & 0 & 0 & 0 \end{array}$$

Reachability matrix (R). The reachability matrix depicts the direct and indirect relationships among the attributes in an attribute hierarchy (Gierl et al., 2000). The reachability matrix for the FCSAH is order 5 × 5, where each row of the R matrix represents one of the paths from the first attribute in the FCSAH to each of the remaining attributes in the model. The jth row specifies with ones all the attributes for which attribute j is a prerequisite, including attribute j. Another way to view the derivation of this matrix is that it was derived from the A matrix and the identity matrix (I) using Boolean algebra, that is, $\mathrm{R} = (A + I)^3$. The reachability matrix for the FCSAH is shown below.

$$\begin{array}{c|ccccc} & A1 & A2 & A3 & A4 & A5 \\ \hline A1 & 1 & 1 & 1 & 1 & 1 \\ A2 & 0 & 1 & 1 & 1 & 1 \\ A3 & 0 & 0 & 1 & 0 & 0 \\ A4 & 0 & 0 & 0 & 1 & 0 \\ A5 & 0 & 0 & 0 & 0 & 1 \end{array}$$

The reachability matrix explicates the dependent relationships among the attributes in the FCSAH. This matrix demonstrates that attribute A1 is a prerequisite for attributes A2, A3, A4, and A5. Attribute A2 is a prerequisite for attributes A3, A4, and A5. Attributes A3, A4, and A5 have no hypothesized relationships among one another.

Incidence matrix (Q). The incidence matrix represents all possible combinations of the attributes in a hierarchy (Gierl et al., 2000). The incidence matrix for the FCSAH is order 5 × 31 and displays no particular relationship to the FCSAH. It contains one row for each attribute and one column for every possible combination of one attribute with one or more of the others in the

FCSAH. Each column thereby represents a combination of attributes for which a test item could be developed. This matrix would be an appropriate template for potential items if all of the attributes in a cognitive model were theoretically independent of one another. However, this is not the perspective of the attribute hierarchy method or this study. The incidence matrix for the FCSAH is shown below. The attributes are indicated by the rows of the matrix.

A1	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1
A2	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1
A3	0	0	0	1	1	1	1	0	0	0	0	1	1	1	1	0	0	0	0	1	1	1	1	0	0	0	0	1	1	1	1
A4	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1
A5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

Reduced incidence matrix (Qr). The reduced incidence matrix contains a column for each combination from the pool of all possible combinations of attributes (Q) that meets the constraints defined by the attribute hierarchy (Gierl et al., 2000). The reduced incidence matrix for the FCSAH contains five rows, one for each attribute in the FCSAH, and includes the five columns from the Q matrix that logically follow from the FCSAH. This matrix was determined by selecting from the Q matrix precisely those columns that were specifically listed in or implied by the R matrix. An additional constraint for this study stemmed from the fact that the test items deemed appropriate to assess knowledge of attributes A3, A4, and A5 were mutually exclusive, which resulted in no item types designed to assess any pair of attributes A3, A4, or A5 in a single item. The five columns selected for the reduced incidence matrix indicate that there are five item types needed to effectively measure the knowledge modeled by the FCSAH. The reduced incidence matrix for the FCSAH is shown below. The attributes are indicated by the rows of the matrix, and the five question types are indicated by the columns and the column labels T1, T2, T3, T4, and T5.

$$\begin{array}{c|ccccc} & T1 & T2 & T3 & T4 & T5 \\ \hline A1 & 1 & 1 & 1 & 1 & 1 \\ A2 & 0 & 1 & 1 & 1 & 1 \\ A3 & 0 & 0 & 1 & 0 & 0 \\ A4 & 0 & 0 & 0 & 1 & 0 \\ A5 & 0 & 0 & 0 & 0 & 1 \end{array}$$

Foundational Concepts of Slope Assessment (FCSA) Test Items

The FCSA was created to measure student understanding of selected foundational concepts related to slope in terms of the attributes in the FCSAH. In order to maintain an emphasis on understanding slope rather than computing slope, the items developed for this assessment targeted the ability to interpret slope ratios in problems presented verbally or graphically. The items specifically were not designed to evaluate whether students were able to compute the slope ratio. This focus was selected in order to further investigate the ability to interpret the slope ratio in terms of a problem's context variables, particularly for problems presented verbally or graphically. For each column in the Q_r matrix, the researcher developed four multiple-choice items. A description of the items developed for each column follows, and a copy of the FCSA is attached as Appendix A.

The items developed for attribute A1 were designed to evaluate the ability to detect which quantities in a problem situation varied in correspondence to one another without any reference to their directions of change. Therefore, the items developed to assess this ability were all presented in the form of word problems and verbal descriptions of the quantities that varied. Of the four items developed for attribute A1, two items presented a constant rate scenario in the problem and required the student to identify one quantity that influenced another specific quantity. The other two items presented a constant rate scenario in the problem and required the student to identify a pair of quantities that together influenced a third specific quantity. The contexts of the items developed for attribute A1 included average driving speed, weekly deposits in a bank account, unit price, and ingredients in a recipe. The two items that required students to select single quantities were presented first and were followed by the items that required students to select pairs of covariates. This sequence was selected in order to place what the researcher anticipated to be the easiest items before the more difficult items.

The items developed for attribute A2 were designed to evaluate the ability to identify the direction of change of two covariates in constant rate problem contexts. The items developed for attribute A2 contained graphs and verbal problem descriptions. Of the four items developed for A2, two items presented a constant rate scenario in a verbally stated problem and required the student to identify the graph that represented the same situation. The other two items developed for attribute A2 presented a constant rate scenario as a graph and required the student to identify the verbal description that represented the same situation. The contexts of the items developed for attribute A2 included ingredients in a recipe, lengths and costs of phone calls, average work times, and average running speeds. The items were presented in an alternating sequence on the assessment, that is, the fifth and seventh items on the FCSA presented verbal problems and required students to select graphs, while the sixth and eighth items presented graphs and required students to select verbal descriptions.

Careful design decisions were made in developing the items for attribute A2. When developing the two items in which a verbal problem was stated and the students were required to select a graph, the same four graphs were used for the answer choices for both items, and none of the graphs contained numbers or scales on the axes. The axis labels for the four graphs shown in an item were the same and named the covariates in the problem context. The similarities among the graphs maintained the items' focus on the ability to detect the direction of covariation. One item concerned a positive slope scenario, and the other problem concerned a negative slope scenario. Likewise, when developing the items in which a graph was given and the students were required to identify the corresponding verbal description, neither graph contained numbers or scales on the axes. Furthermore, the verbal descriptions offered as answer choices were constructed such that pairs of answer choices had the same sentence structure. One item concerned a zero slope scenario, and the other problem concerned a negative slope scenario.

The items developed for attribute A3 were designed to evaluate the ability to interpret the meaning of the slope ratio in terms of the context of a problem presented either verbally or graphically. Items developed for attribute A3 concerned slopes whose ratio values simplified to whole numbers. The items developed for attribute A3 contained graphs and verbal problem descriptions. Of the four items developed for attribute A3, two items presented a constant rate scenario in a verbally stated problem and required the student to identify the graph that represented the same situation. The other two items developed for attribute A3 presented a constant rate scenario as a graph and required the student to identify the verbal description that represented the same situation. The contexts of the items developed for attribute A3 included ingredients in recipes, unit price, and a direct comparison of two rates of saving money. The items were presented in an alternating sequence on the assessment, that is, the ninth and eleventh items on the FCSA presented verbal problems and required students to select graphs, while the tenth and twelfth items presented graphs and required students to select verbal descriptions. Two items developed for attribute A3 contained graphs whose axes were scaled by one, and the other two items contained graphs whose axes were scaled by two, that is, they had scaled graphs.

Careful design decisions were made when developing the items for attribute A3. For the items in which students were required to interpret a graph and choose a verbal description, the answer choices all contained comparisons of the same two values, and the verbal statements were constructed such that pairs of statements shared the same sentence structure. Likewise, for the items in which students were required to interpret a verbally described problem and choose a graph, the incorrect answer choices were constructed such that the graphs represented errors stemming from either incorrect use of univariate reasoning, additive reasoning, or interpreting the slope ratio upside down. For one of these problems two graphs had positive slopes, two graphs had negative slopes, two graphs had slopes using the correct values from the problem, and two graphs had slopes using incorrect values from the problem. For the other of these problems, one graph had an undefined slope, one graph had a slope whose value was zero, one graph had the correct slope according to the problem, and one graph had a slope whose value was the reciprocal of the correct slope's value.

The items developed for attribute A4 were designed to evaluate the ability to interpret the meaning of the slope ratio in terms of the context of a problem presented either verbally or graphically. Items developed for attribute A4 concerned slopes whose ratio values simplified to unit fractions. The items developed for attribute A4 contained graphs and verbal problem descriptions. Of the four items developed for attribute A4, two items presented a constant rate scenario in a verbally stated problem and required the student to identify the graph that represented the same situation. The other two items developed for attribute A4 presented a constant rate scenario in a graph and required the student to identify the verbal description that represented the same situation. The contexts of the items developed for attribute A4 included ingredients in a recipe, average running speed, and unit price. The items were presented in an alternating sequence on the assessment, that is, the thirteenth and fifteenth items on the FCSA presented verbal problems and required students to select graphs, while the fourteenth and sixteenth items presented graphs and required students to select verbal descriptions. Three items developed for attribute A4 contained graphs whose axes were both scaled by one, and the other item contained graphs whose axes were scaled by two. Two items developed for attribute A4 required students to simplify slope values and interpret them as rates, and the other two items required students to interpret slopes in terms of the slope ratios' values.

Careful design decisions were made when developing the items for attribute A4. For the items in which students were required to interpret a graph and choose a verbal description, the verbal statement answer choices contained comparisons of two values extending from correct multiplicative reasoning or two values extending from incorrect additive reasoning, and the verbal statements were constructed such that pairs of statements shared the same sentence structure. Likewise, for the items in which students were required to interpret a verbally described problem and choose a graph, the incorrect answer choices were constructed such that the graphs represented errors stemming from either incorrect use of univariate reasoning, additive reasoning, or interpreting the slope ratio upside down. For one of these problems the four graphs had positive slopes, two graphs had slopes using the correct values from the problem, and two graphs had slopes using incorrect values from the problem. For the other problem, one graph had an undefined slope, one graph had a slope whose value was zero, one graph had the correct slope according to the problem, and one graph had a slope whose value was the reciprocal of the correct slope's value.

The items developed for attribute A5 were designed to evaluate the ability to interpret the meaning of the slope ratio in terms of the context of a problem presented either verbally or graphically. Items developed for attribute A5 concerned slopes whose ratio values simplified to positive rational numbers but neither whole numbers nor unit fractions. The items developed for attribute A5 contained graphs and verbal problem descriptions. Of the four items developed for attribute A5, two items presented a constant rate scenario in a verbally stated problem and required the student to identify the graph that represented the same situation. The other two items developed for attribute A5 presented a constant rate scenario as a graph and required the student to identify the verbal description that represented the same situation. The contexts of the items developed for attribute A5 included ingredients in a recipe, average rate of growth, unit price, and the scale of an enlarged photo. The items were presented in an alternating sequence on the assessment, that is, the seventeenth and nineteenth items on the FCSA presented verbal problems and required students to select graphs, while the eighteenth and twentieth items presented graphs and required students to select verbal descriptions. Two items developed for attribute A5 contained graphs whose axes were both scaled by one, and the other two items contained graphs whose axes were scaled by two. Two items developed for attribute A5 required students to simplify the slope value and interpret it as a rate, and the other two items required students to interpret the slope in terms of the slope ratio's two values.

Careful design decisions were made when developing the items for attribute A5. For the items in which students were required to interpret a graph and choose a verbal description, the verbal statement answer choices contained comparisons of two values extending from correct multiplicative reasoning or two values extending from incorrect additive reasoning, and the verbal statements were constructed such that pairs of statements shared the same sentence structure. Likewise, for the items in which students were required to interpret a verbally described problem and choose a graph, the incorrect answer choices were constructed such that the graphs represented errors stemming from either incorrect use of univariate reasoning, additive reasoning, or interpreting the slope ratio upside down. For these problems the four graphs had positive slopes, two graphs had slopes using the correct values from the problem, and two graphs had slopes using incorrect values from the problem.

The items developed for the FCSA reflected different slope values, different scales in the graphs, different mathematical representations, and different contexts. These aspects of each item on the FCSA are depicted in Table 2. Attribute A1 targeted the ability to detect which quantities in a problem situation varied in correspondence to one another without any reference to their directions of change. Therefore, the items developed to assess this ability were all presented in the form of word problems and verbal descriptions of the quantities that varied. The items developed for the other four attributes in the FCSAH contained graphs and verbal problem descriptions. If graphs were presented in the problems designed for attribute A1, then directional information would be part of the problem context. However, since attribute A1 targeted only whether students were able to perceive of which quantities were related in context, and not how they were related, graphs were not appropriate representations to use to evaluate this attribute. In order to maintain the focus of all the items on the FCSA on the ability to interpret the meaning of the slope ratio in terms of context variables, the slope values were all positive. Furthermore, the linear function underlying each problem had a y-intercept equal to zero. Using only positive slope values and intercepts equal to zero restricted the focus to interpreting slope values rather than confusing slope with other features of linear functions such as varying y-intercepts.

Table 2

Scales of Graphs, Mathematical Representations, and Problem Contexts used in the Items

				Representation			Problem Context		
Item	Attribute	Slope Value	Scale of Graph	Verbal-verbal	Verbal-graph	Graph-verbal	Unit price	Average rate	Recipe or scale
1	A1			X				X	
2	A1			X				X	
3	A1			X					X
4	A1			X					X
5	A2				X				X
6	A2					X	X		
7	A2				X			X	
8	A2					X		X	
9	A3	3/1	1		X			X	
10	A3	5/1	1			X	X		
11	A3	4/1	2		X				X
12	A3	8/1	2			X			X
13	A4	1/7	1		X			X	
14	A4	1/4	1			X			X
15	A4	1/6	2		X		X		
16	A4	1/5	1			X	X		
17	A5	3/5	1		X		X		
18	A5	3/4	1			X			X
19	A5	5/2	2		X			X	
20	A5	5/7	2			X			X

Note: Scale of graph equal to 1 indicates that the graphs were scaled by one. Scale of graph equal to 2 indicates that the graphs were scaled by two. Verbal-verbal indicates that the question stem and the response options were represented verbally. Verbal-graph indicates that the question stem was represented verbally and the answer options were represented by graphs. Graph-verbal indicates that the question stem was represented by a graph and the answer choices were represented verbally.

In developing all of the items for the FCSA, the researcher constructed the answer choice distractors, i.e. the incorrect responses, to reflect common student misconceptions or errors. One common error associated with learning about slope is to construct the slope ratio upside down, that is, as the change in the independent variable divided by the change in the dependent variable (Barr, 1980). A second common error is to confuse the amount of change in one variable with its value at a particular point (Bell & Janvier, 1981). A third common error stems from students using additive reasoning instead of multiplicative reasoning when working with slope, which is manifested by students subtracting values instead of dividing them to produce slope values (Heller et al., 1990). A fourth common error is associated with students failing to identify the quantities or variables in a problem context that are related to the slope or rate in the given context (Moritz, 2005). A fifth common error occurs when students consider only one variable at a time rather than the comparison of two variables (Moritz, 2005). The items on the FCSA contained incorrect answer choices designed to attract students who held specific misconceptions. Table 3 depicts the misconceptions included in item response options for the items on the FCSA.

Table 3

Misconceptions Represented in Assessment Item Response Options

Item	Key	Identify quantity	Single variable strategy	Additive strategy	Total vs. change	Reciprocal slope	Opposite slope
1	D	A, B, C					
2	C	A, B, D					
3	A	B, C, D					
4	D	A, B, C					
5	A		C, D				
6	D				A, B, C		
7	B		C, D				
8	B	C, D					A
9	C					A, B	B, D
10	D				A, C	B	
11	C		A, D			B	
12	B					A, C	C, D
13	A		B		D	C	
14	A			C, D		B	
15	D			B, C		A, B	
16	C		A, B				
17	C			B		D	
18	A					B, C, D	
19	A		C			B	
20	D					C	

Note: Key = the correct response option. Identify quantity indicates the response option represented an error in identifying the quantities related to the slope. Single variable strategy indicates the response option represented consideration of only one variable instead of two variables. Additive strategy indicates the response option represented an additive computation rather than a multiplicative computation. Total vs. change indicates the response option represented the confusion of an amount of change and a variable's value at a point. Reciprocal slope indicates the response option represented the reciprocal of the correct slope. Opposite slope indicates the response option represented the opposite of the correct slope.

Procedure for Developing the Foundational Concepts of Slope Assessment (FCSA)

Item quality assurance reviews. An assessment containing 21 items designed to measure the attributes in the FCSAH was assembled and shared with the two SMEs, who were previously consulted about the FCSAH. The SMEs were asked to blindly identify which attributes they believed were assessed by each test item. The researcher compiled data to reflect the agreement of the SMEs and took actions to improve the FCSA based on the SMEs' responses.

There was perfect agreement among both SMEs and the researcher regarding the alignment of 11 items. The researcher met again with the SMEs to discuss and refine the alignment of the other 10 items. The disagreements about seven items were resolved through discussion and refinement of the attributes' descriptions. The disagreements about two items led to item revisions. One item was deleted from the FCSA.

One item disagreement was associated with attributes A1 and A2. This disagreement was caused by confusion about the definition of attribute A1. This disagreement was addressed by refining the description of attribute A1 so it did not include students' ability to detect the nature of the relationship between variables, but only to target students' ability to identify which variables are related in a problem situation. After refining this description, the researcher and SMEs reached agreement on which attribute was measured by this item.

Three item disagreements were associated with attributes A3 and A4. The disagreements were caused by confusion about the early drafts of the definitions of attributes A3 and A4. First, the researcher used feedback from the SMEs to refine the descriptions of attributes A3 and A4 to more clearly distinguish them according to whole number (A3) or unit fraction (A4) slopes. Then the researcher and SMEs reached agreement on which attribute was measured by each of these three items.

Three item disagreements were associated with attributes A4 and A5. The disagreements were caused by confusion about the early drafts of the definitions of attributes A4 and A5. First, the researcher used feedback from the SMEs to refine the descriptions of attributes A4 and A5 to clearly distinguish them according to unit fraction (A4) or neither integer nor unit fraction (A5) slopes. Then the researcher and SMEs reached agreement on which attribute was measured by each of these three items.

Two item disagreements led to item revisions. In both cases the items were edited so that the scales of the graphs could be simplified to a scale of one or a scale of two. In one case, the context of the problem was changed in order to maintain a realistic situation while reducing the range of values needed in the graph. In the other case, the relationship between the two quantities in the problem was changed so that the graph could have a scale of one on both the x-axis and the y-axis.

Three individuals other than the SMEs participated in the study to insure the technical quality of the test items developed for the FCSA. These individuals reviewed each item to confirm that the language used in each item was clear and contained grade-appropriate vocabulary. They confirmed the accuracy of each item's mathematical content. They also reviewed the answer choices for each item to insure that there was exactly one correct answer and that the distractors represented typical misconceptions held by students or errors students might make when responding to the items.

Expected item response vectors. For each test there exists a unique set of expected item responses corresponding to each combination of attributes possessed by examinees that is consistent with the attribute hierarchy used to design the test (Gierl et al., 2000). Each expected response vector contains a one for each item a hypothetical student should get right, given the set of attributes that student possesses, and a zero for the other items. The responses an examinee gives to test items are dependent upon the attributes that examinee possesses. When a student does not fully possess an attribute assessed by a test item, that student has a reduced probability of answering the item correctly.

The set of expected item response vectors for the FCSA was developed by the researcher to reflect the expected responses students would give depending on which attributes from the FCSAH they possessed. This set of vectors is represented in a matrix in which each row contains the answer pattern of a hypothetical student that possesses a particular combination of attributes from the FCSAH. There are 20 digits in each row, where each digit corresponds to an item response on the FCSA. Recall items one through four on the FCSA assess knowledge of attribute A1. Items five through eight assess knowledge of attributes A1 and A2. Items nine through twelve assess knowledge of attributes A1, A2, and A3. Items thirteen through sixteen assess knowledge of attributes A1, A2, and A4. Items seventeen through twenty assess knowledge of attributes A1, A2 and A5. A one represents an expected correct response to an item, and a zero represents an expected incorrect response to an item.

The first row in the matrix of expected response vectors represents the expected responses of a student who possesses none of the attributes in the FCSAH. This student should have incorrect responses to all of the test items. The second row in the matrix represents the expected answers of a student who only possesses attribute A1, that is, the ability to identify the two covariates in a particular problem scenario. This student should have correct responses to the first four items and incorrect responses to all other items. The third row in the matrix represents the expected answers of a student who possesses attributes A1 and A2, that is, the abilities to identify covariates and to discern the direction of their relationship. This student should have correct responses for the first eight items and incorrect responses to all other items. The fourth row in the matrix represents the expected answers of a student who possesses attributes A1, A2, and A3, that is the abilities to identify covariates, to discern the direction of their relationship, and to interpret a slope ratio whose value is a whole number. This student should have correct responses to the first twelve items and incorrect responses to all other items.

The fifth row in the matrix represents the expected answers of a student who possesses attributes A1, A2, and A4, that is the abilities to identify covariates, to discern the direction of their relationship, and to interpret a slope ratio whose value simplifies to a positive unit fraction. This student should have correct responses to items one through eight and items thirteen through sixteen, and this student should have incorrect responses to items nine through twelve and items seventeen through twenty. The sixth row in the matrix represents the expected answers of a student who possesses attributes A1, A2, and A5, that is the abilities to identify covariates, to discern the direction of their relationship, and to interpret a slope ratio whose value simplifies to a positive rational number that is neither a whole number nor a unit fraction. This student should have correct responses to items one through eight and items seventeen through twenty, and this student should have incorrect responses to items nine through sixteen. The seventh row in the matrix represents the expected answers of a student who possesses attributes A1, A2, A3 and A4. This student should have correct responses to the first sixteen items and incorrect responses to all other items. The eighth row in the matrix represents the expected answers of a student who possesses attributes A1, A2, A3, and A5. This student should have correct responses to items one through twelve and items seventeen through twenty, and this student should have incorrect responses to items thirteen through sixteen. The ninth row in the matrix represents the expected answers of a student who possesses attributes A1, A2, A4, and A5. This student should have correct responses to items one through eight and items thirteen through twenty, and this student should have incorrect responses to items nine through twelve. The tenth row in the matrix represents the expected answers of a student who possesses attributes A1, A2, A3, A4, and A5. This student should have correct responses to all of the test items. The expected response vector matrix is shown below.

$$\begin{bmatrix} 00000000000000000000 \\ 11110000000000000000 \\ 11111111000000000000 \\ 11111111111100000000 \\ 11111111000011110000 \\ 11111111000000001111 \\ 11111111111111110000 \\ 11111111111100001111 \\ 11111111000011111111 \\ 11111111111111111111 \end{bmatrix}$$

Figure 2. The Expected Response Vector Matrix for the FCSA

Assessment Implementation

Recruitment for participation. The FCSA was named “SlopeAssessment” in the testing software and published as a formative assessment by the Center for Educational Testing and Evaluation (CETE). It was made available to all middle and high school mathematics teachers in Kansas schools on May 5, 2011. District leaders were notified about this assessment’s purpose and availability through an email announcement sent by the Kansas State Department of Education (KSDE). This announcement contained a brief description of the goals of the study in addition to instructions about how to access the assessment. A copy of the announcement is attached as Appendix B. In the announcement, teachers were encouraged to use this assessment as part of their regular instructional activities in the areas of proportional reasoning, graphing lines, and interpreting rates of change or slope. The researcher’s contact information was included in the announcement, and teachers were encouraged to request additional information as needed.

After the announcement was shared with districts and teachers, the researcher received several questions about how to access the assessment. The researcher learned that some districts restricted the pool of tests teachers can view when they login to the CETE formative assessment website. In such districts, teachers were unable to see the FCSA as a test they could select for their students until a district or building administrator added this assessment to the pool of tests that was visible to the teachers. The researcher then drafted instructions to teachers and administrators in these schools. This additional information was sent via email to district leaders and teachers who had previously communicated with the researcher and again, on May 13, 2011, to district test coordinators. A copy of this additional information is attached as Appendix C.

Teachers who gave the FCSA to their students were directed to a survey to collect teacher contact information, building and district names and numbers, and the average time students took to take the FCSA. A copy of this survey is attached as Appendix D.

Process used by teachers to access the FCSA and review results. Teachers who chose to administer the FCSA logged in to the CETE website for formative assessment tools. There, they clicked on the tab “Browse Tests” to find the “SlopeAssessment.” Then they added the “SlopeAssessment” to “My Tests.” Then they created new administrations for their classes, which produced a test session number and password their students used when they logged in to the Kansas Computerized Assessment (KCA) software. These codes directed the student to the correct test, that is, the FCSA. Teachers then administered the FCSA by having their students log in to the KCA software and take the test as they would any other formative assessment delivered through this software.

After students completed the FCSA, teachers were able to view their students' item level responses online and download them as text documents or spreadsheets. The class level reports contained a row for each student and 20 columns, with one column corresponding to each item on the FCSA. The row for each student contained a one for each question the student answered correctly. The entry for incorrect responses was the letter of the response the student chose, i.e., A, B, C, or D.

Data collection. Student test responses were captured by the KCA test engine. CETE staff used information collected in the teacher survey to identify the teachers who administered the FCSA. For every administration of the FCSA, a CETE staff member copied the student item level responses into a spreadsheet. Along with the student responses, the CETE staff member recorded the teacher's last name and the course name, both captured in the test administration data, as well as the district number, which was captured by the teacher survey. The data set delivered to the researcher thus contained a row for each student. Each row contained the teacher's last name, the course name, the district number, the student's response to each item on the FCSA, and the student's percent correct.

Variables

The independent variables in this study were the five attributes identified as attributes in the FCSAH, namely, attributes A1, A2, A3, A4, and A5. The dependent variables were the ten latent classes into which students were placed based on their response patterns on the FCSA. The ten latent classes corresponded to the ten expected response vectors and represented the different combinations of attributes a student could have possessed.

Data Analysis

Analyses Using Item Response Theory

Item response theory (IRT) uses information gathered from test responses to estimate the ability of each student in the tested domain, which in this study is the student's knowledge of selected foundational concepts related to slope. Ability estimates are placed on a scale with a midpoint of zero and standard deviation of one. Although the theoretical range of ability estimates extends from negative to positive infinity, the practical range usually spans [-3, 3] or [-4, 4]. Not only does IRT produce ability estimates, but it also produces estimates of one, two, or three aspects of each item on a test. The first aspect of each item is the item's ability to discriminate between students of different ability levels. A larger discrimination parameter value indicates that an item is more discriminating than an item with a smaller discrimination parameter value. The second aspect of each item is the item difficulty, which describes how difficult the item was for the sample of students who responded to the item. Item difficulty estimates are placed on the same scale as student ability estimates and indicate the ability needed for an examinee to have a 50% probability of answering the item correctly. The third aspect of each item estimates the probability of a low ability student answering the item correctly by chance, which is sometimes referred to as the guessing parameter.

The three aspects of each test item are represented by three statistical parameters, which together produce an item characteristic function and corresponding curve for each item on a test. This function is used to estimate the likelihood, p, that a student with a particular ability estimate, Θ, will answer an item correctly. The three parameter logistic (3PL) model is displayed as Equation 1, where i references a particular test item, a represents the item's power to discriminate between students with different ability levels, *b* is the item's difficulty, and *c* is the item's guessing parameter.

$$p_i(\theta) = c_i + \frac{1-c_i}{1+e^{-a_i(\theta-b_i)}} \quad (1)$$

The test characteristic function is the sum of the item characteristic functions for the items on a test. The graph of the test characteristic function illustrates the overall difficulty of a test and its ability to discriminate among students of different ability levels.

Test items with different item parameter values provide different amounts of information for students of different ability levels. Item information functions illustrate the amount of information provided by items for students of different ability levels, where the horizontal axis represents student ability and the vertical axis represents the amount of information provided by items.

The test information function is the sum of the item information functions for the items on a test. The test information function illustrates the amount of information provided by the collection of items on a test for students of different ability levels, where the horizontal axis represents student ability and the vertical axis represents the amount of information provided by the collection of items comprising an entire test. A subtest information function can also be determined for any subset of items on a test.

The purpose of using IRT to measure achievement on the FCSA was to estimate how much knowledge about foundational concepts related to slope each student had. When the FCSA was scored using IRT, the focus was on how each student answered each question rather than the student's composite test score. The IRT model yielded a probability of how likely it was that a student possessed a particular amount of knowledge about selected foundational concepts related to slope. The IRT model also produced three item parameters for each item on the FCSA.

Analyses Using the Attribute Hierarchy Method

In preparation for analyzing student response data, the researcher determined a hypothetical student to represent each combination of attributes that was consistent with the FCSAH. For the proposed FCSAH, there were ten hypothetical students, each of which was represented by a row in the expected response vector matrix displayed in Figure 2. Each expected response vector contained a one for every item the hypothetical student should answer correctly based on the attributes that student theoretically possessed and a zero for the items the student should answer incorrectly based on the attributes that student theoretically did not possess. The expected response vectors were evaluated with IRT calculations to estimate the ability of each of the ten hypothetical students represented by the expected response vectors.

After student responses were collected, they were organized into what are called observed response vectors containing ones for the items a student answered correctly and zeros for the items the student answered incorrectly. Following the work of Leighton, Gierl, and Hunka (2004), each observed response vector was compared to every expected response vector using the formula shown in Equation 2 and the ability estimates of the hypothetical students. These comparisons produced estimates of the likelihood that an observed student response vector matched an expected response vector. The formula used for comparisons is listed as Equation 2, where j corresponds to the expected response vector to which an actual student's responses are compared, k is the number of ones in the difference vector, and m is the number of negative ones in the difference vector. Each comparison followed this procedure:

1. The student response vector was subtracted from the expected response vector to produce a difference vector. Each element of this vector was 1, 0, or -1. An entry of 1 in the difference vector indicated that the student with the knowledge implied by the expected

response vector should have answered the item correctly, but the actual student did not – an error. An entry of 0 in the difference vector indicated that the student with the knowledge implied by the expected response vector and the actual student answered the item in the same way. An entry of -1 in the difference vector indicated that the student with the knowledge implied by the expected response vector should have answered the item incorrectly, but the actual student answered the item correctly.

2. The difference vector was used in the formula shown in Equation 2 with the ability estimate (Θ) corresponding to the expected response vector to yield a likelihood of the actual student having the same ability estimate as that corresponding to the expected response vector.

$$P_{j\ expected}(\theta) = \prod_{k=1}^{k} P_{jk}(\theta)\ \prod_{m=1}^{m}\left[1 - P_{jm}(\theta)\right] \tag{2}$$

3. The procedure described above produced ten likelihood estimates for each student in the data set. These ten likelihood estimates were summed for each student. Then each likelihood estimate was divided by the student's sum to produce a probability corresponding to each likelihood estimate. The highest probability was identified for each student, and the student was classified into the same category as the expected response vector corresponding to the highest probability. In this way each student was classified into a knowledge state according to the responses the student gave and the attributes the student likely possessed. These probabilities also were evaluated to judge the likelihood that the FCSAH was true.

After all students were classified into knowledge states, the ability estimates of the students classified into each knowledge state were examined. The distribution of the ability estimates of the students classified into each knowledge state was displayed in a box-and-whisker plot. Then the ten plots were arranged together on one graph, where the vertical width of each box was proportional to the number of students classified in the knowledge state represented by the box. This graph was used to display the relationships among the distributions of the ability estimates of the students classified into the different knowledge states.

Analyses Using Classical Test Theory

Student responses were analyzed to describe the misconceptions held by students with regard to selected foundational concepts related to slope. The analyses were reported in tables, one for each set of four items assessing one attribute. The data for each item included the percent of students selecting each response choice, the IRT parameter estimates, and the item characteristic curve. These summaries were used to describe the common misconceptions held by students regarding the concept of slope. The percent of students selecting a response indicating the presence of each type of misconception was calculated and displayed in a table containing one row for each item on the FCSA. The values in the table were averaged to describe the overall selection rate of each of the misconceptions identified in the literature and used in the distractors of the items on the FCSA.

Student responses were analyzed to describe how students responded to items with different problem contexts. The analyses were reported in tables summarizing how students answered the items on the FCSA in terms of each item's problem context. The percent of students who answered each item correctly was calculated and displayed in a table containing one row for each item on the FCSA and one column for each of the three problem contexts used for the items on the FCSA. The values in the table were averaged for each of the problem contexts. These averages were used to summarize how these students responded to items with different problem contexts.

CHAPTER FOUR

RESULTS

Introduction

The aim of the present study was to analyze the concept of slope, to assemble a model of how understanding of selected foundational concepts related to slope develops, and to construct and use an assessment to measure student understanding of the selected foundational concepts related to slope. The assessment model used for this study was the attribute hierarchy method (AHM), a method developed by Leighton, Gierl, and Hunka (2004). This model requires a construct to be defined in terms of the attributes that collectively describe what students must know and be able to do to demonstrate their knowledge in a particular content domain. The analysis of the slope concept yielded the Foundational Concepts of Slope Attribute Hierarchy (FCSAH), a hypothetical representation of the knowledge students must possess in order to understand selected foundational concepts related to slope. Five attributes were identified as components of understanding slope. The FCSAH guided the development of an assessment designed to measure student knowledge of the selected foundational concepts related to slope in terms of the five attributes delineated in the FCSAH. Examinees' knowledge states were described in terms of whether they demonstrated knowledge of the attributes in the hierarchy.

Research Questions

This investigation answered these research questions.

1. What insight is gained about the validity of the proposed cognitive model from an analysis of student data generated from an assessment informed by the model?
2. To what extent did student participants exhibit common misconceptions regarding slope?

This chapter presents the quantitative results of this study in three sections. The first section reviews the Foundational Concepts of Slope Attribute Hieraracy (FCSAH) and presents statistical properties about the test items and different combinations of items using item response theory (IRT). These results include an analysis of the items on the Foundational Concepts of Slope Assessment (FCSA), an analysis of the FCSA as a test, and analyses of the five subtests, each targeting one attribute. The second section presents results about the individual student responses. These results begin with a description of the expected student response patterns and their associated ability estimates, that is, the expected latent classes, or knowledge states, used for classifying students in terms of their knowledge about the selected foundational concepts related to slope. Next, an analysis of the actual student response patterns will be given with their associated ability estimates, followed by a description of how each student was classified into one of the knowledge states. Then, an analysis of the distribution of the ability estimates of the students classified into each latent class will be presented. The third section presents results about what student responses indicate about student knowledge of the foundational concepts related to slope. This section will also include findings concerning the misconceptions about the foundational concepts related to slope held by students.

Analysis of the Test Items

Slope Attribute Hierarchy

The FCSAH developed for this study incorporated the work of scholars who have studied student understanding of covariation and proportional reasoning. The FCSAH is displayed in Chapter 3 in Figure 1. The FCSAH was analyzed to develop four matrix representations of the attributes and their dependent hierarchical combinations. These matrices were used to develop test items specifically targeted to the attributes contained in the FCSAH.

An assessment was developed to evaluate the attributes delineated in the FCSAH and their dependent hierarchical combinations. Following the process described by Gierl, Leighton, and Hunka (2000), the first step was to determine several matrix representations of the FCSAH, the last of which directly influenced item and test design.

The adjacency matrix depicts the direct (one-step) relationships of attributes in an attribute hierarchy with ones and zeros (Gierl et al., 2000). The adjacency matrix for the FCSAH is shown below.

	A1	A2	A3	A4	A5
A1	0	1	0	0	0
A2	0	0	1	1	1
A3	0	0	0	0	0
A4	0	0	0	0	0
A5	0	0	0	0	0

The reachability matrix depicts the direct and indirect relationships among the attributes in an attribute hierarchy (Gierl et al., 2000). The reachability matrix for the FCSAH is shown below.

	A1	A2	A3	A4	A5
A1	1	1	1	1	1
A2	0	1	1	1	1
A3	0	0	1	0	0
A4	0	0	0	1	0
A5	0	0	0	0	1

The incidence matrix (Q) represents all possible combinations of the attributes in a hierarchy (Gierl et al., 2000). The incidence matrix for the FCSAH is shown below.

A1	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1
A2	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1
A3	0	0	0	1	1	1	1	0	0	0	0	1	1	1	1	0	0	0	0	1	1	1	1	0	0	0	0	1	1	1	1
A4	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1
A5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

The reduced incidence matrix (Q_r) contains a row for each attribute and a column for each combination from the pool of all possible combinations of attributes (Q) that meets the constraints defined by the attribute hierarachy (Gierl et al., 2000). The reduced incidence matrix for the FCSAH contains five rows, one for each attribute in the FCSAH, and includes the five columns from the Q matrix that logically follow from the FCSAH. The reduced incidence matrix for the FCSAH is shown below. The attributes are indicated by the rows of the matrix, and the five question types are indicated by the columns and the column labels T1, T2, T3, T4, and T5.

	T1	T2	T3	T4	T5
A1	1	1	1	1	1
A2	0	1	1	1	1
A3	0	0	1	0	0
A4	0	0	0	1	0
A5	0	0	0	0	1

Foundational Concepts of Slope Assessment

The FCSA was developed to measure student understanding of the attributes delineated in the FCSAH. In order to maintain an emphasis on the understanding of the meaning of slope, the items developed for this assessment were designed to measure whether students were able to interpret the meaning of slopes presented in word problems or graphs. Four multiple-choice response items were developed for each column of the Q_r matrix.

Item level data analyses. Item response theory (IRT) was applied to the item responses to the FCSA to estimate each student's ability with regard to attributes in the FCSAH. Not only did IRT calculations produce student ability estimates, but it also produced estimates of three aspects that together described each item's power to discriminate among students of different ability levels, relative difficulty, and probability of a low-ability student guessing the correct answer. These aspects were captured as item parameters and were labeled, respectively, the a-parameter, b-parameter, and c-parameter. The a-parameter represented the item's power to discriminate between students of different ability levels. The b-parameter represented the item's relative difficulty. The c-parameter estimated the probability of a low ability student answering the item correctly by chance.

The Foundational Concepts of Slope Assessment (FCSA) was administered to 1629 students in middle and high school grades studying Pre-algebra, Algebra 1, Geometry, Algebra 2, or courses with similar content taken before Pre-calculus. It was important to confirm that the FCSA assessed student ability in only one cognitive dimension in order to meet the required assumptions for IRT analysis. Factor analysis confirmed a single dimension, or factor, in these data, which is illustrated by the scree plot in Figure 3. The plot illustrates that a large eigenvalue was extracted for the first factor contrasted with much lower values for all other factors. Furthermore, the difference between the eigenvalues extracted for the first factor and the eigenvalue extracted for the second factor is very large, whereas the differences between the eigenvalues extracted for the second through the twentieth factors are very small. These values confirmed that the FCSA assessed student ability in a single dimension, which met the assumption of unidimensionality for IRT analysis of the data.

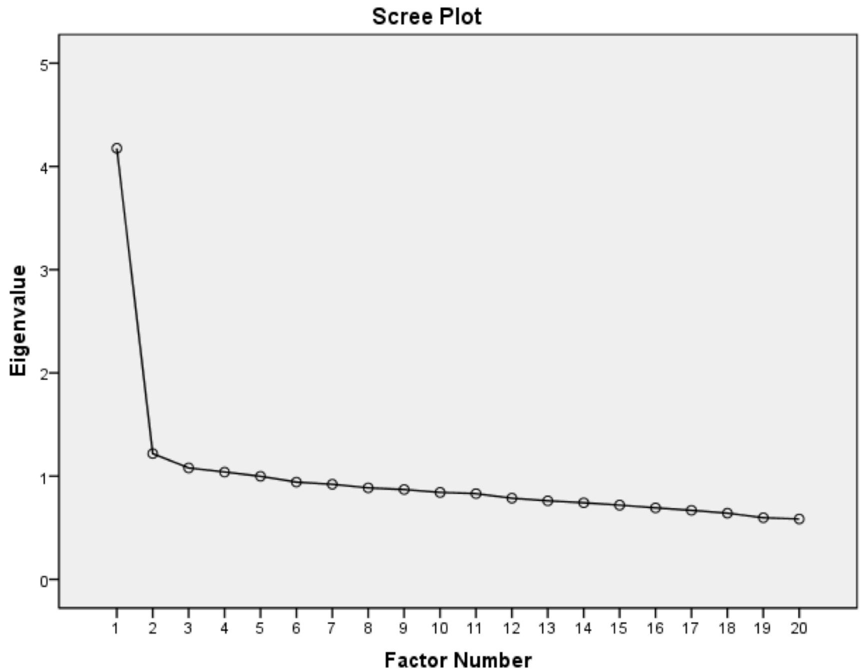


Figure 3. Scree Plot of Student Responses to the FCSA

After unidimensionality was confirmed, students' responses were used to estimate the item parameters using BILOG-MG3 (Zimowski, Muraki, Mislevy, & Bock, 2003). The default settings in BILOG-MG3 (Zimowski et al., 2003) were used except that the number of response options, which by default is set to five, was set to four. This choice was made because each item on the FCSA had four answer choices rather than five. The item parameter estimates are shown in Table 4.

The three item parameters for each item are used together to produce the item characteristic function and corresponding curve for each item on a test. The item characteristic curve (ICC) illustrates the relationship between a population's responses to an item and the amount of ability hypothetically needed to answer the item correctly. The ICC is a monotonically

Table 4

BILOG-MG3 Item Parameter Estimates for the FCSA

Item	a-parameter	b-parameter	c-parameter
1	0.48	-2.35	0.26
2	0.69	-2.95	0.24
3	0.58	-2.03	0.24
4	0.65	-1.47	0.26
5	0.73	-2.06	0.26
6	0.74	-1.17	0.31
7	0.87	0.00	0.27
8	0.67	-1.15	0.28
9	0.71	-0.76	0.25
10	1.41	-0.82	0.23
11	1.00	-0.43	0.19
12	1.24	-1.27	0.19
13	1.01	-0.03	0.14
14	1.10	-0.95	0.15
15	0.89	-0.08	0.18
16	0.65	-1.51	0.23
17	0.76	-1.23	0.17
18	1.36	-1.20	0.22
19	1.03	-0.24	0.20
20	0.82	2.09	0.26

increasing function, where students with less ability in the domain have lower probabilities of answering the item correctly, and students with more ability in the domain have higher probabilities of answering the item correctly. The point of inflection on the ICC represents the point at which students with the ability level equal to the abscissa of the point of inflection have a 50% chance of answering the item correctly. The abscissa of the ICC's point of inflection equals the b-parameter, which is used to describe the difficulty of the item. The slope of the ICC at the point of inflection is proportional to the a-parameter, which is used to determine how efficiently the item discriminates between students with similar ability levels.

The ICCs for the items on the FCSA are shown in Figure 4. The horizontal axis represents student ability, and the vertical axis represents the probability of students with different ability levels answering each item correctly. The plot of the item characteristic curves illustrates that many of the items on the FCSA had relatively low difficulty, and only one item had relatively high difficulty. The plot also shows large slopes for several items, indicating that these items discriminated well between students of higher and lower ability levels.

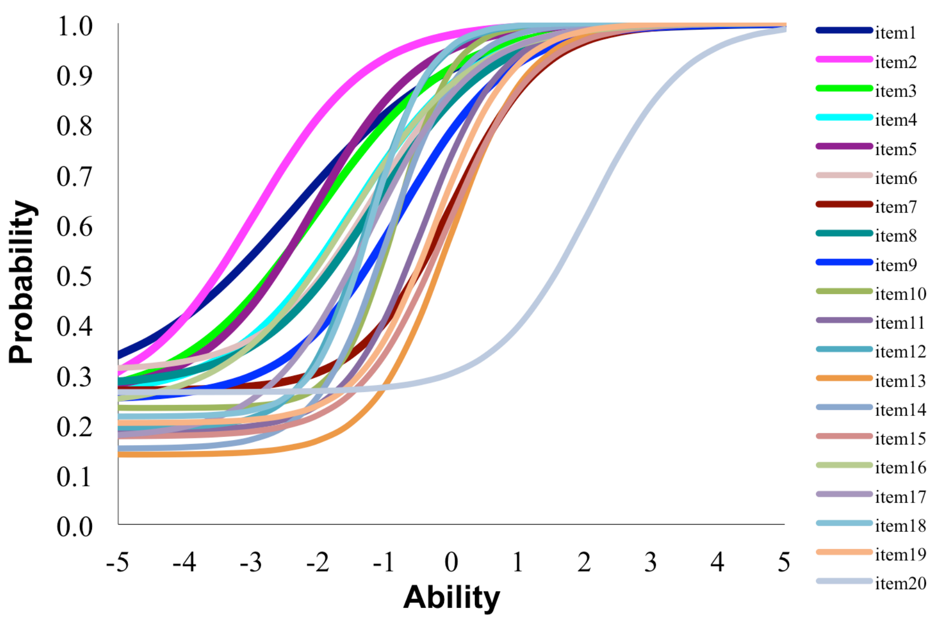


Figure 4. Item Characteristic Curves for the Items on the FCSA

The test characteristic function equals the sum of the item characteristic functions and is shown in Figure 5. The horizontal axis represents student ability, and the vertical axis represents the expected scores for students with different ability levels. This graph illustrates that the FCSA was not very difficult for the students who took it because the point of inflection has an abscissa that is between -1 and 0. The test characteristic function also illustrates that the FCSA did not effectively distinguish between students of different ability levels because the slope of the test characteristic function at the point of inflection is not very steep.

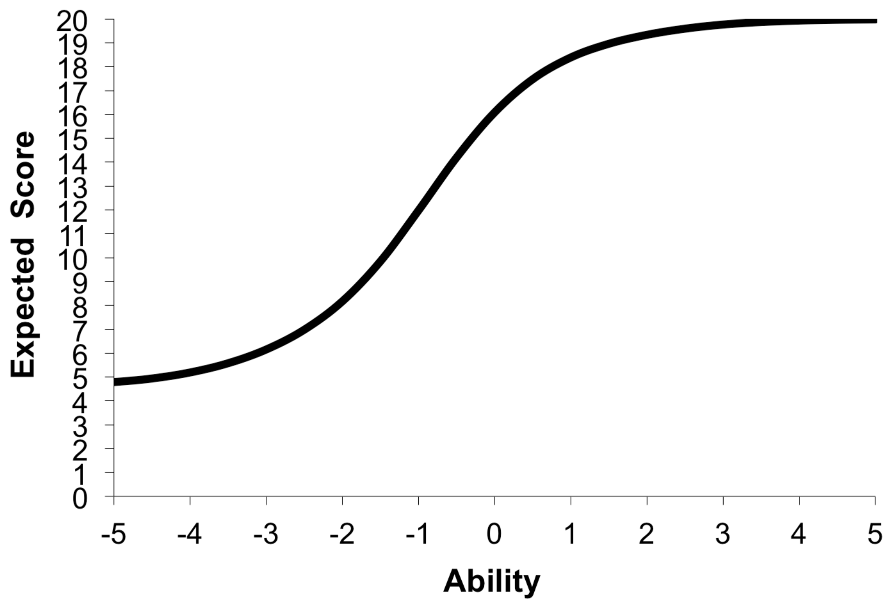


Figure 5. Test Characteristic Curve for the FCSA

The item information function illustrates the contribution an item makes to estimating student ability. The value of the item's contribution varies with the ability of the students. The relationship between item information and student ability is displayed graphically, where the horizontal axis represents student ability, and the vertical axis represents the amount of information the item provides for students of different ability levels. More information is illustrated by larger item information function values, and less information is illustrated by smaller item information values. A plot of the item information functions for the items on the FCSA is shown in Figure 6. This plot illustrates that several items on the FCSA were very informative for students with abilities in the range of about -1.6 to 0. One item on the FCSA was most informative for students with abilities in the range of about 1.5 to 3.5. Several items, many of which were answered correctly by many students, were not very informative. This finding is illustrated by the item information functions with maximum function values of less than 0.3.

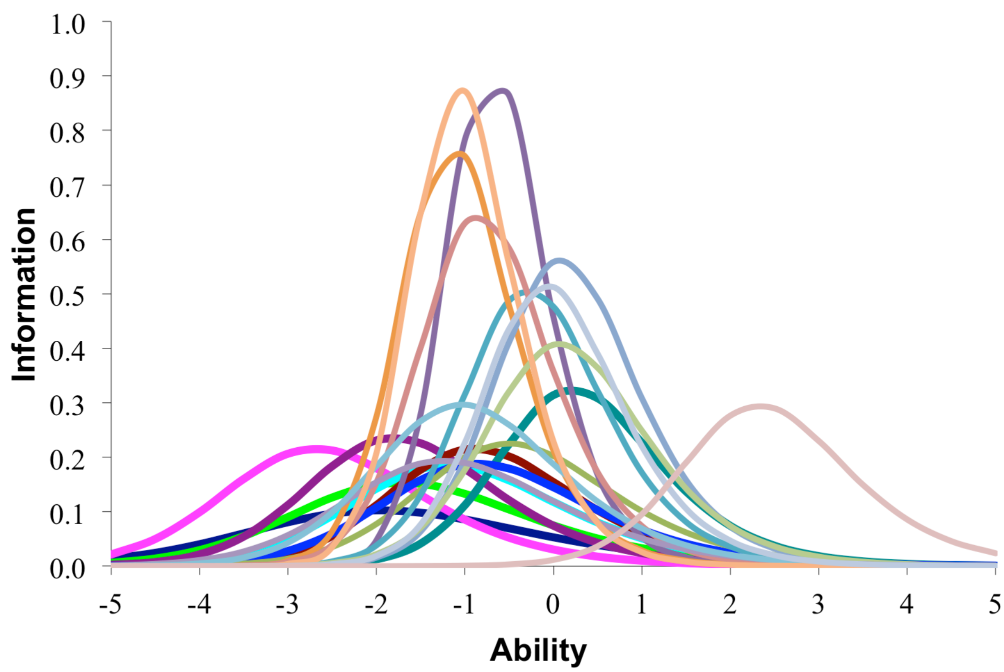


Figure 6. Item Information Functions for the Items on the FCSA

The test information function (TIF) equals the sum of the item information functions. A plot of the TIF and standard error is shown in Figure 7. The TIF shows that the FCSA was most informative for students with abilities in the range of -2.0 to 1.0.

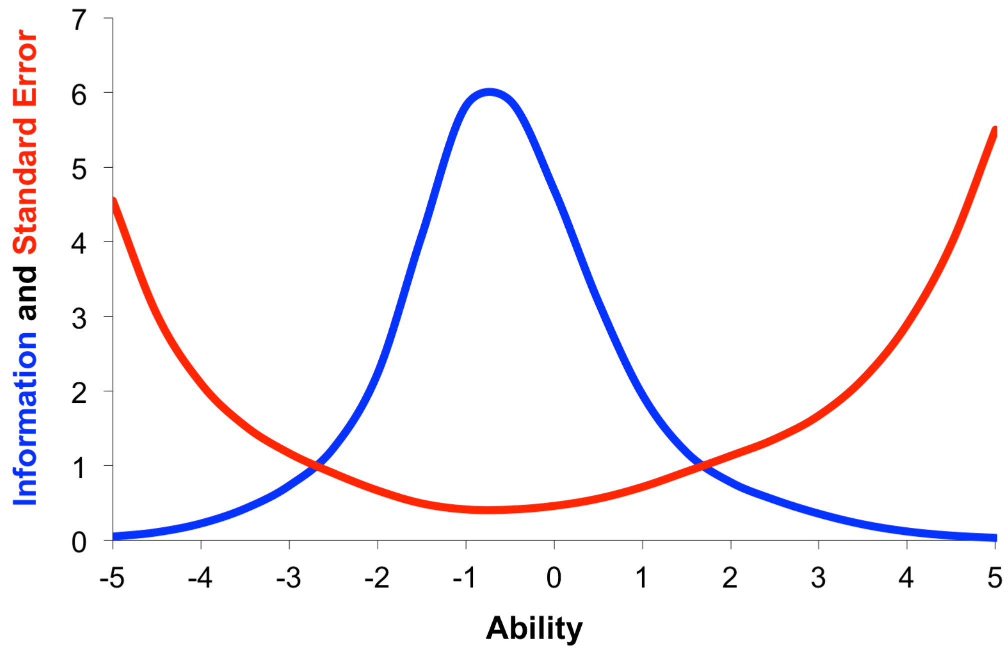


Figure 7. Test Information Function and Standard Error for the FCSA

Analysis of the subtests designed to assess the attributes in the FCSAH. The FCSA contained four items developed for each of the attributes in the FCSAH. These five sets of four items each were analyzed as attribute subtests by determining the test information function for each of the five subtests. A plot of the five subtest information functions is shown in Figure 8. This plot shows that the subtest comprised of items one through four was less informative than the other subtests for the entire population in the study and was most informative for students with ability estimates near the value of -2.0. The subtest comprised of items five through eight was slightly more informative overall than the first subtest and was most informative for students with ability estimates near the value of -0.5. The subtest comprised of items nine through twelve was the most informative for the entire population in the study and was particularly informative for students with ability estimates near the value of -1.0. The subtest comprised of items thirteen

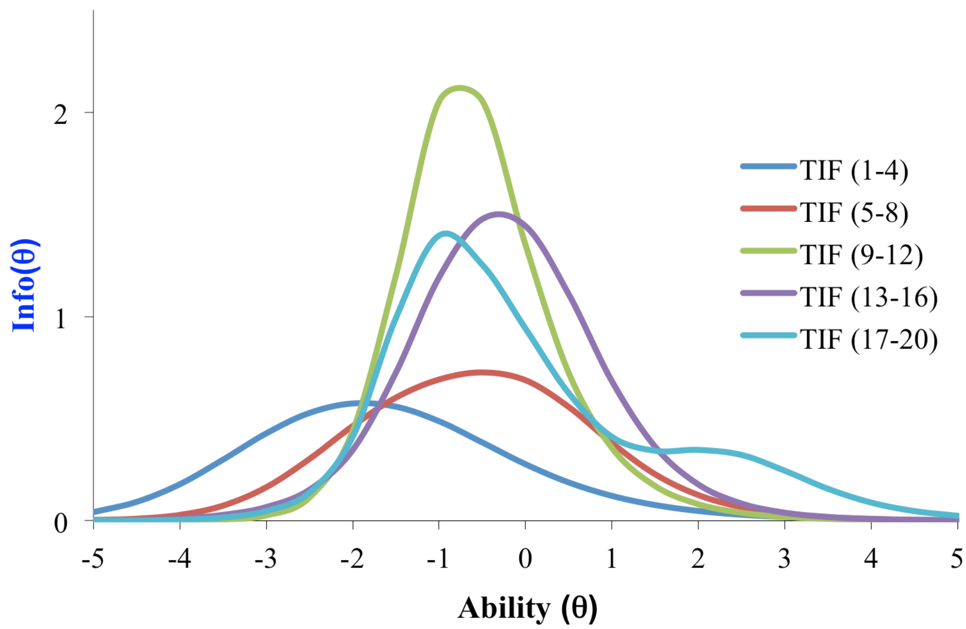


*Figure 8***.** Attribute Subtest Information Functions

through sixteen was the second most informative subtest for the entire population in the study and was particularly informative for students with ability estimates near the value of -0.2. The subtest comprised of items seventeen through twenty was the third most informative subtest for the entire population in the study and was particularly informative for students with ability estimates near the value of -1.0.

Analyses of Individual Student Item Responses

Knowledge States Consistent with the FCSAH

For each attribute hierarchy, there exists a set of hypothesized combinations of attributes that can be possessed by a set of examinees (Leighton et al., 2004). There were ten hypothesized attribute combinations that were consistent with the FCSAH and were used to classify the students who took the FCSA. These ten combinations comprised the rows of the expected response vector matrix displayed in Chapter 3 in Figure 2 and are shown below in the second column of Table 5. Each row in Table 5 is hereafter referred to as a knowledge state. The first knowledge state represents students who possess none of the attributes in the FCSAH. The abbreviation for this category in Table 5 is A0. The second knowledge state represents students who possess only attribute A1 in the FCSAH. The abbreviation for this category in Table 5 is A1. The third knowledge state represents students who possess attributes A1 and A2 in the FCSAH. The abbreviation for this category in Table 5 is A12. The fourth knowledge state represents students who possess attributes A1, A2, and A3 in the FCSAH. The abbreviation for this category in Table 5 is A123. This pattern continues through the tenth knowledge state, which represents students who possess all five attributes in the FCSAH. The abbreviation for this category in Table 5 is A12345.

Expected student responses. Each of the ten knowledge states corresponds to one of the rows of the expected response vector matrix shown in Chapter 3 in Figure 2. Each vector contains 20 digits, that is, a one for each item a hypothetical student should get right, given the set of attributes that student possesses, and a zero for the other items. For the purposes of this study, student knowledge of each attribute was considered to be either present or not present. Therefore, if a theoretical student possessed an attribute, then that student should have answered all four items aligned to that attribute correctly. Likewise, if a theoretical student did not possess an attribute, that student should have answered all four items aligned to that attribute incorrectly. This rationale yielded ten unique expected student response vectors.

Ability estimates for hypothetical students. The knowledge states and expected response vectors shown in Table 5 were used to create hypothetical students whose abilities were estimated using IRT calculations. The item parameters determined previously were used with Equation 1 to estimate these hypothetical students' abilities. Table 5 displays the ten hypothesized student knowledge states in terms of the combination of attributes each student might possess along with each hypothesized student's expected response vector and corresponding ability estimate.

Table 5

Knowledge States, Expected Responses, and Ability Estimates Consistent with the FCSAH

Knowledge State	Expected Response Vector	Ability Estimate
A0	00000000000000000000	-2.92
A1	11110000000000000000	-2.23
A12	11111111000000000000	-1.67
A123	11111111111100000000	-0.95
A124	11111111000011110000	-1.19
A125	11111111000000001111	-1.23
A1234	11111111111111110000	-0.14
A1235	11111111111100001111	-0.21
A1245	11111111000011111111	-0.42
A12345	11111111111111111111	1.45

Individual Student Ability Estimates

Item response theory (IRT) was applied in order to use information gathered from test item responses to estimate each student's ability with regard to the concept of slope. For the 1629 student participants, a plot of the student ability estimate frequencies is shown in Figure 9.

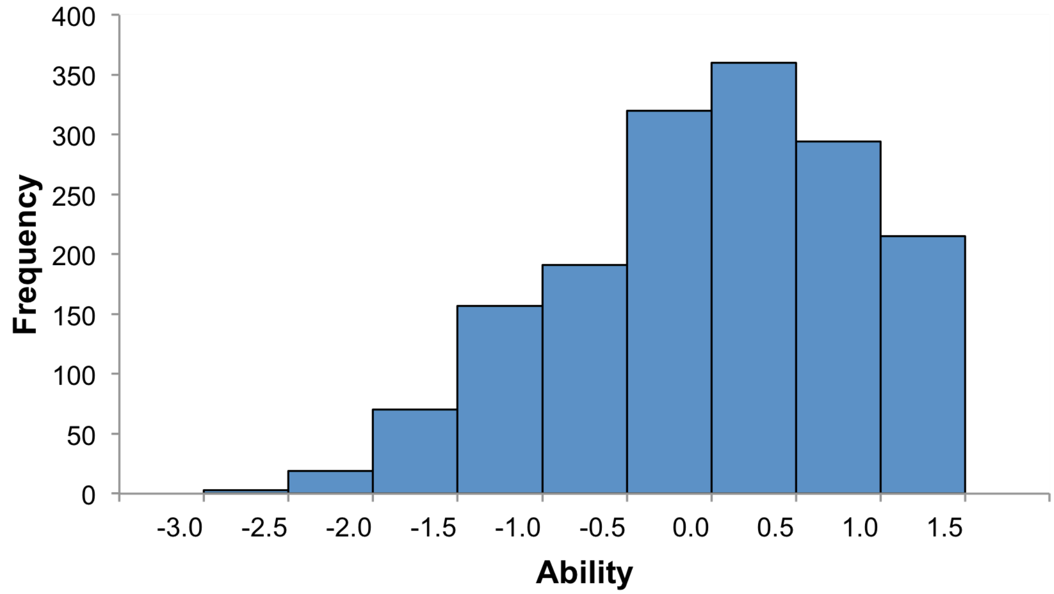


Figure 9. Student Ability Estimate Frequencies

Comparisons between Observed Response Vectors and Expected Response Vectors

Student responses were organized into observed response vectors containing ones for the items students answered correctly and zeros for the items they answered incorrectly. Following the work of Leighton, Gierl, and Hunka (2004), each observed response vector was compared to every expected response vector using Equation 2. The comparison producing the highest likelihood was used to place each student into a knowledge state consistent with the FCSAH. An example of one student's comparisons is shown in Table 6. This table contains one row for each knowledge state consistent with the FCSAH. Each row lists the ability estimate and expected response vector of the theoretical student possessing the attributes and corresponding expected responses of the particular knowledge state. The row also lists the estimated likelihood, $L(\Theta)$, that the actual student being compared shared the same ability estimate as the theoretical student and the probability, $P(\Theta)$, corresponding to this likelihood. The student whose comparison is shown in Table 6 was classified in knowledge state A1235 because this category corresponded to the highest likelihood and probability values in Table 6. This student correctly answered all of the items that assessed attributes A1, A2, and A3. Therefore, this student was classified as possessing attributes A1, A2, and A3. This student incorrectly answered items thirteen and twenty. Item thirteen was a relatively easy item that assessed attribute A4. This student must have answered item thirteen correctly in order to be classified as possessing attribute A4. Since this student missed item thirteen, this student was classified as not possessing attribute A4. Item twenty was by far the most difficult item that assessed attribute A5. Students who answered this item incorrectly but correctly answered the other three items that assessed attribute A5 were classified as possessing attribute A5. This is because this item was so much more difficult than the other three items that assessed A5.

Table 6

Likelihood Estimates, Probabilities, and Knowledge State Classification of a Sample Student with Observed Response Vector 11111111111101111110 and Ability Estimate of 0.64

Ability Estimate	Expected Response Vector	$L_{j\text{Expected}}(\Theta)$	$P_{j\text{Expected}}(\Theta)$	Knowledge State
-2.92	00000000000000000000	0.00	0.00	A0
-2.23	11110000000000000000	0.00	0.00	A1
-1.67	11111111000000000000	0.00	0.00	A12
-0.95	11111111111100000000	0.03	0.04	A123
-1.19	11111111000011110000	0.01	0.01	A124
-1.23	11111111000000001111	0.00	0.00	A125
-0.14	11111111111111110000	0.23	0.34	A1234
-0.21	11111111111100001111	0.27	0.39	A1235
-0.42	11111111000011111111	0.12	0.17	A1245
1.45	11111111111111111111	0.03	0.05	A12345

All 1629 student participants were classified into one of the ten knowledge states that are consistent with the FCSAH. The ability estimates calculated for participants lay on a distribution from -2.67 to 1.45, as illustrated in Figure 9 and shown in Table 7, where -2.67 is the minimum value in the “Minimum Ability” column, and 1.45 is the maximum value in the “Maximum Ability” column. Table 7 shows the ability estimate assigned to the expected response vector for each knowledge state, and the percent of students classified into each knowledge state is also shown. Table 7 also lists the descriptive statistics needed to plot box and whisker plots of the abilities of the students assigned to each knowledge state. These box whisker plots are shown in Figure 10, where the relative height of each box is proportional to the percent of students assigned to that knowledge state.

Table 7

Descriptive Statistics about Knowledge State Classifications

Knowledge State	Knowledge State Ability Estimate	Minimum Ability	Lower Quartile Ability	Median Ability	Upper Quartile Ability	Maximum Ability	Percent of Students	Percentile
A0	-2.92	-2.67	-2.48	-2.25	-1.93	-1.78	1	1
A1	-2.23	-2.32	-1.98	-1.83	-1.65	-1.31	2	3
A12	-1.67	-2.32	-1.81	-1.60	-1.46	-1.19	3	6
A123	-0.95	-1.84	-1.00	-0.61	-0.41	0.01	9	15
A124	-1.19	-1.77	-1.27	-0.94	-0.62	-0.18	9	24
A125	-1.23	-1.93	-1.29	-0.99	-0.64	-0.39	8	32
A1234	-0.14	-0.48	0.06	0.41	0.72	0.93	16	48
A1235	-0.21	-0.50	0.04	0.35	0.58	1.00	19	67
A1245	-0.42	-0.97	-0.20	0.19	0.56	1.10	21	88
A12345	1.45	1.14	1.14	1.14	1.45	1.45	13	100
All Categories	n/a	-2.67	-0.56	0.66	0.10	1.45	100	100

These results indicate that students who were classified into knowledge state A0 appear to have different ability levels than the other students because the box and whisker plot for this knowledge state does not share many of the same values listed on the horizontal axis with the plots of other knowledge states. Students who were classified into knowledge states A1 and A12 appear to have somewhat similar ability levels because the box and whisker plots for these two knowledge states share some of the values listed on the horizontal axis. However, these students appear to have different ability levels than students classified to higher levels because the box and whisker plots for knowledge states A1 and A12 do not share many values listed on the horizontal axis with the box and whisker plots for knowledge states including more than two attributes. Students who were classified into knowledge states A123, A124, and A125 appear to have very similar ability levels, a finding that is illustrated in Figure 10 where the box and whisker plots for these knowledge states, displayed in shades of green, share many of the same values listed on the horizontal axis. The students classified into knowledge states A123, A124, and A125 appear to have very different ability levels than students classified at lower or higher ability levels. These findings suggest that knowledge states A123, A124, and A125 are not very different for these students. Students who were classified into knowledge states A1234, A1235, and A1245 appear to have very similar ability levels, a finding that is illustrated in Figure 10 where the box and whisker plots for these knowledge states, displayed in shades of blue and purple, share many of the same values listed on the horizontal axis. The students classified into knowledge states A1234, A1235, and A1245 appear to have very different ability levels than students classified at lower or higher ability levels. These findings suggest that knowledge states A1234, A1235, and A1245 are not very different for these students. The students classified into knowledge state A12345 appear to have very different ability levels than students classified at lower levels.

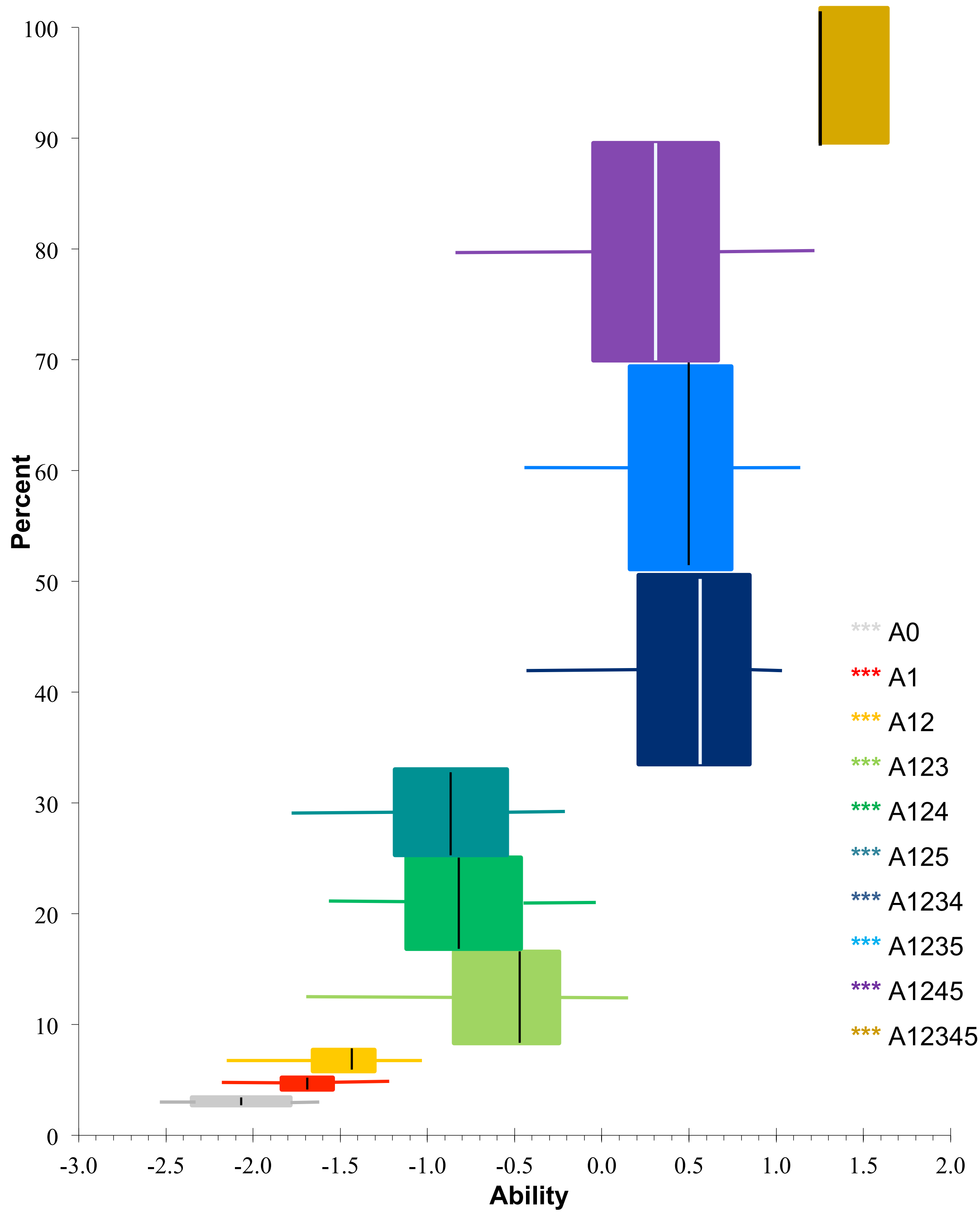


Figure 10. Percent of Students Assigned to each Knowledge State by Ability Estimate

Student Understanding of the Attributes in the FCSAH

Overview of the FCSA

This study was conducted in part to determine what students understand about selected foundational concepts related to slope and to describe typical misconceptions held by students with regard to those foundational concepts related to slope. To that end, the FCSA was developed to measure student understanding of the attributes delineated in the FCSAH. In order to maintain an emphasis on the understanding of the meaning of slope, the items developed for this assessment were designed to measure whether students were able to interpret the meaning of slopes presented in word problems or graphs. Four multiple-choice response items were developed for each combination of attributes that was consistent with the FCSAH. There were five combinations for which four items were designed, for a total of 20 items on the FCSA.

The items developed for attribute A1 were designed to evaluate the ability to identify the quantities that vary together in constant rate problem contexts. The items developed for attribute A2 were designed to evaluate the ability to identify the directions in which two quantities change in constant rate problem contexts. The items developed for attributes A3, A4, and A5 were designed to evaluate the ability to interpret the meaning of a slope ratio in terms of the quantities relevant to a constant rate problem scenario. What distinguished these three attributes' items was the value of the slope in each problem type. Items developed for attribute A3 concerned slopes whose ratio values simplified to whole numbers. Items developed for attribute A4 concerned slopes whose ratio values simplified to positive unit fractions, that is, values that could be expressed in the form of one divided by an integer greater than one. Items developed for attribute A5 concerned slopes whose ratio values simplified to positive rational numbers that were neither integers nor unit fractions.

In developing all of the items for the FCSA, the researcher constructed the answer choice distractors, i.e. the incorrect responses, to reflect common student misconceptions or errors. One common error associated with learning about slope is to construct the slope ratio upside down, that is, as the change in the independent variable divided by the change in the dependent variable (Barr, 1980). In the data analysis, this error will be referred to as the reciprocal of the correct slope. A second common error is to confuse the amount of change in one variable with its value at a particular point (Bell & Janvier, 1981). In the data analysis, this error will be referred to as confusing total values with amounts of change. A third common error stems from students using additive reasoning instead of multiplicative reasoning when working with slope, which is manifested by students subtracting values instead of dividing them to produce slope values (Heller et al., 1990). In the data analysis, this error will be referred to as additive reasoning. A fourth common error is associated with students failing to identify the quantities or variables in a problem context that are related to the slope or rate in the given context (Moritz, 2005). In the data analysis, this error will be referred to as failing to identify the appropriate quantities. A fifth common error occurs when students consider only one variable at a time rather than the comparison of two variables (Moritz, 2005). In the data analysis, this error will be referred to as univariate reasoning.

The items developed for the FCSA reflected one or two mathematical representations, different contexts, and different slope values. The items on the FCSA also contained a variety of incorrect answer choices designed to attract students who held specific misconceptions. The graphs presented in the items were labeled at every value or were scaled by two. The next section presents analyses of the student responses to the individual items on the FCSA. The four items that were written to address the attribute are presented and analyzed together.

Items Intended to Assess Attribute A1

Items one through four assessed whether students could identify the quantities that varied together in constant rate problem contexts. All four of these items presented students with verbal problems, and students had to identify the names of the values that varied together in the problems. The percents of students that chose each response option are shown in Table 8, the IRT item parameters are shown in Table 9, and the item characteristic curves are shown in Figure 11.

Table 8

Student Item Response Percents for Items 1-4 on the FCSA

Item	Key	"A" Responses	"B" Responses	"C" Responses	"D" Responses	Blank
1	D	5	3	4	89	0
2	C	0.4	2	96	2	1
3	A	88	1	10	0.4	0
4	D	1	7	7	84	0

Table 9

BILOG-MG3 Item Parameter Estimates for Items 1-4 on the FCSA

Item	a-parameter	b-parameter	c-parameter
1	0.48	-2.35	0.26
2	0.69	-2.95	0.24
3	0.58	-2.03	0.24
4	0.65	-1.47	0.26

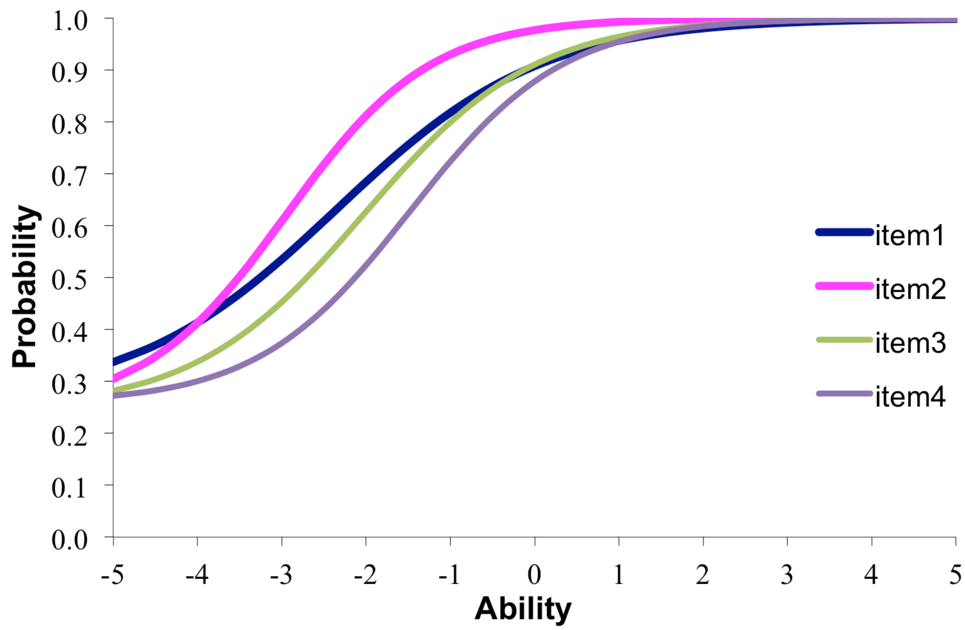


Figure 11. Item Characteristic Curves for Items 1-4 on the FCSA

These four items as a group were relatively easy for this group of students. All four items were relatively non-discriminating, a fact that is illustrated by the gentle slopes of the graphs and by all small slope values in Table 9, that is, $a < 1.0$. The easiest item in this group was item number two, which was answered correctly by 96% of the students who took the FCSA. This item also had the lowest difficulty parameter of the four items, that is, $b = -2.95$, confirming that students with very low ability estimates demonstrated a 50% chance of answering this item correctly. The most difficult item in this group was item number four, which was answered correctly by 84% of the students who took the FCSA. This item also had the highest difficulty parameter of these four items, that is, $b = -1.47$.

Items Intended to Assess Attribute A2

Items five through eight assessed whether students could identify the directions in which two quantities changed in constant rate problem contexts. Items five and seven presented students with verbal problems, and students had to select the graphs that matched the scenarios in the problems. Item seven was the only problem on the FCSA for which the correct response had a negative slope value. Items six and eight presented students with graphs, and students had to select the verbal descriptions that matched the graphs. The percents of students that chose each response option are shown in Table 10, the IRT item parameters are shown in Table 11, and the item characteristic curves are shown in Figure 12.

Table 10

Student Item Response Percents for Items 5-8 on the FCSA

Item	Key	"A" Responses	"B" Responses	"C" Responses	"D" Responses	Blank
5	A	91	2	3	3	1
6	D	2	5	10	83	0
7	B	24	64	7	6	1
8	B	7	81	5	7	0

Table 11

BILOG-MG3 Item Parameter Estimates for Items 5-8 on the FCSA

Item	a-parameter	b-parameter	c-parameter
5	0.73	-2.06	0.26
6	0.74	-1.17	0.31
7	0.87	0.00	0.27
8	0.67	-1.15	0.28

Item five included two distractors designed to attract students who used univariate reasoning instead of bivariate reasoning and one distractor designed to attract students who may have interpreted the opposite direction of variation, that is, the opposite slope. For this item, response options C and D represented univariate reasoning, and option B represented the opposite slope. As shown in Table 10, a total of 6% of the students who took that FCSA chose options C and D for item five, indicating that the students may have used univariate reasoning, whereas only 2% chose option B, indicating that the students chose the opposite of the correct slope.

Item six included three distractors designed to attract students who confused total values with amounts of change. For this item, response options A, B, and C represented this misconception. As shown in Table 10, a total of 17% of the students who took the FCSA chose these options for item six, indicating that the students may have confused total values with amounts of change.

Item seven included two distractors designed to attract students who used univariate reasoning instead of bivariate reasoning and one distractor designed to attract students who interpreted the opposite direction of variation, that is, the opposite slope. For this item, response options C and D represented univariate reasoning, and response option A represented the opposite slope. As shown in Table 10, a total of 13% of the students who took that FCSA chose options C and D for item seven, indicating that the students may have used univariate reasoning when working this item. Table 10 also shows that 24% chose option A, indicating that the students chose the opposite of the correct slope for this item.

Item eight included two distractors designed to attract students who failed to identify the correct quantities that varied in the problem context and one distractor designed to attract students who interpreted the opposite of the correct slope. For this item, response options C and D represented the incorrect quantities, and response option A represented the opposite slope. As shown in Table 10, a total of 12% of the students who took the FCSA chose options C and D for item eight, indicating they may have compared the incorrect quantities. Also shown in Table 10, 7% of the students who took the FCSA chose option A, indicating that they selected the opposite of the correct slope.

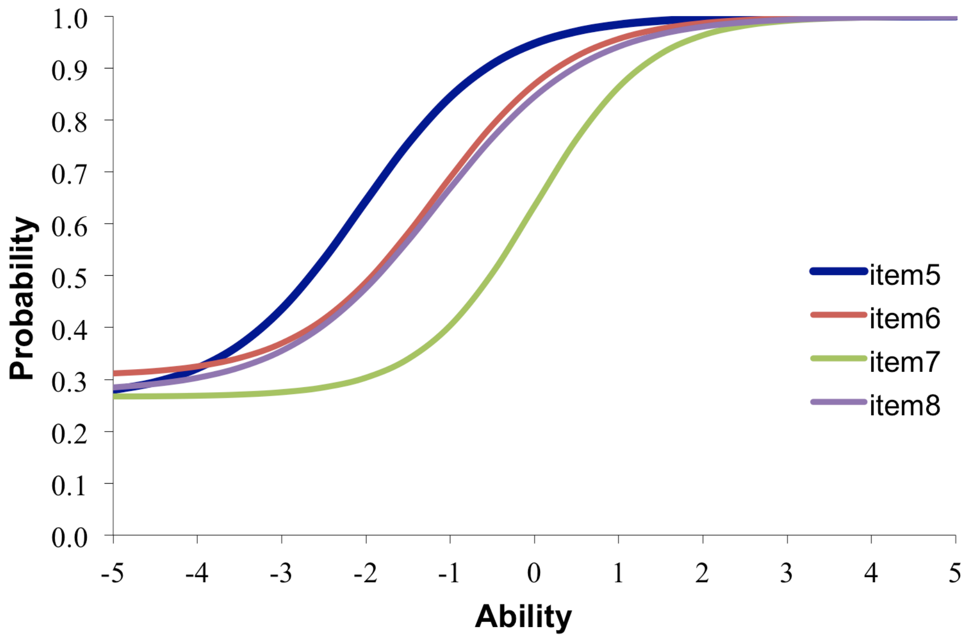


Figure 12. Item Characteristic Curves for Items 5-8 on the FCSA

These four items as a group were somewhat easy for this group of students. All four items were more discriminating than items one through four, but somewhat non-discriminating overall, a fact that is illustrated by the slightly steeper slopes of the graphs and by slope values between a = 0.5 and a = 1.0, as shown in Table 11. The easiest item in this group was item number five, which was answered correctly by 91% of the students who took the FCSA. This item had the lowest difficulty parameter of these four items, that is, b = -2.06, confirming that students with relatively low ability estimates demonstrated a 50% chance of answering this item correctly. The most difficult item in this group was item number seven, which was answered correctly by 64% of the students who took the FCSA. This item also had the highest difficulty parameter of these four items, that is, b = 0.00.

Items Intended to Assess Attribute A3

Items nine through twelve assessed whether students could interpret the meaning of slope in terms of the quantities relevant to a constant rate problem scenario. These items contained slopes whose ratio values simplified to whole numbers. Items nine and eleven presented students with verbal problems, and students had to select the graphs that matched the scenarios in the problems. Items ten and twelve presented students with graphs, and students had to select the verbal descriptions that matched the graphs. The percents of students that chose each response option are shown in Table 12, the IRT item parameters are shown in Table 13, and the item characteristic curves are shown in Figure 13.

Table 12

Student Item Response Percents for Items 9-12 on the FCSA

Item	Key	"A" Responses	"B" Responses	"C" Responses	"D" Responses
9	C	11	4	75	10
10	D	5	8	6	81
11	C	9	14	69	8
12	B	4	87	5	4

Table 13

BILOG-MG3 Item Parameter Estimates for Items 9-12 on the FCSA

Item	a-parameter	b-parameter	c-parameter
9	0.71	-0.76	0.25
10	1.41	-0.82	0.23
11	1.00	-0.43	0.19
12	1.24	-1.27	0.19

Item nine included one distractor designed to attract students to the reciprocal of the correct slope, one distractor designed to attract students to the opposite of the correct slope, and one distractor designed to attract students to the opposite reciprocal of the correct slope. For this item, response option A represented the reciprocal slope, option B represented the opposite reciprocal slope, and option D represented the opposite slope. As shown in Table 12, 11% of the students who took the FCSA chose option A for item nine, indicating that the students chose the reciprocal of the correct slope for this item. Also shown in Table 12, 4% of the students chose option B, indicating that the students selected the opposite reciprocal slope. Table 12 also shows that 10% of the students who took the FCSA chose option D for item nine, indicating that the students chose the opposite slope for this item.

Item ten included two distractors designed to attract students who failed to distinguish total values from amounts of change and one distractor designed to attract students to the reciprocal of the correct slope. For this item, response options A and C represented incorrect total values or amounts of change, and response option B represented the reciprocal of the correct slope. As shown in Table 12, a total of 11% of the students who took the FCSA chose options A and C for item ten, indicating that the students may have confused total values with amounts of change. Table 12 also shows that 8% of the students who took the FCSA chose option B, indicating that the students chose the reciprocal of the correct slope for this item.

Item eleven included two distractors designed to attract students who used univariate reasoning instead of bivariate reasoning and one distractor designed to attract students to the reciprocal slope. For this item, response options A and D represented univariate reasoning, and response option B represented the reciprocal of the correct slope. As shown in Table 12, a total of 17% of the students who took the FCSA chose options A and D for item eleven, indicating that the students may have used univariate reasoning. Also shown in Table 12, 14% of the students who took the FCSA chose option B, which indicated that the students selected the reciprocal of the correct slope.

Item twelve included one distractor designed to attract students to the reciprocal of the correct slope, one distractor designed to attract students to the opposite of the correct slope, and one distractor designed to attract students to the opposite reciprocal of the correct slope. For this item, response option A represented the reciprocal slope, option C represented the opposite reciprocal slope, and option D represented the opposite slope. As shown in Table 12, 4% of the students who took the FCSA chose option A, indicating that the students selected the reciprocal of the correct slope. Also shown in Table 12, 5% of the students who took the FCSA chose option C, indicating the students selected the opposite reciprocal of the correct slope. Also shown in Table 12, 4% of the students who took the FCSA chose option D, indicating the students chose the opposite slope.

These four items as a group were somewhat less easy for this group of students than the items that assessed attributes A1 and A2. Items nine through twelve were more discriminating as a group than items one through eight, a fact that is illustrated by the slightly steeper slopes of the graphs and by slope values between $a = 0.7$ and $a = 1.5$, as shown in Table 13. Items ten through twelve were quite discriminating, with slope values all greater than 1.0. The easiest item in this group was item number twelve, which was answered correctly by 87% of the students who took the FCSA. This item had the lowest difficulty parameter of these four items, that is, $b = -1.27$, confirming that students with somewhat low ability estimates demonstrated a 50% chance of answering this item correctly. The most difficult item in this group was item number eleven,

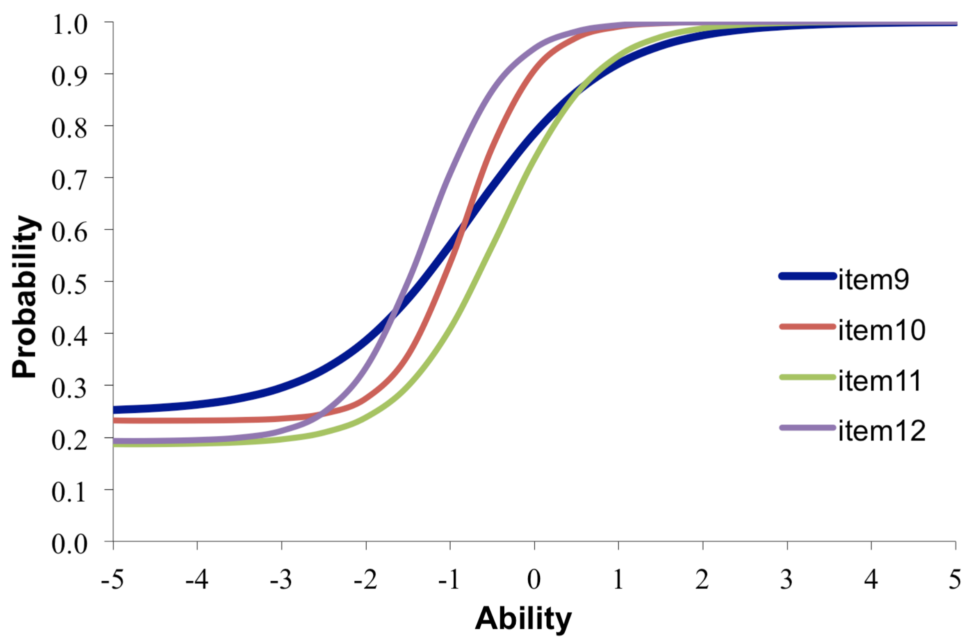


Figure 13. Item Characteristic Curves for Items 9-12 on the FCSA

which was answered correctly by 69% of the students who took the FCSA. This item also had the highest difficulty parameter of these four items, that is, b = -0.43.

Items Intended to Assess Attribute A4

Items thirteen through sixteen assessed whether students could interpret the meaning of slope in terms of the quantities relevant to a constant rate problem scenario. These items contained slopes whose ratio values simplified to positive unit fractions, that is, values that could be expressed in the form of one divided by an integer greater than one. Items thirteen and fifteen presented students with verbal problems, and students had to select the graphs that matched the scenarios in the problems. Items fourteen and sixteen presented students with graphs, and students had to select the verbal descriptions that matched the graphs. The percents of students that chose each response option are shown in Table 14, the IRT item parameters are shown in Table 15, and the item characteristic curves are shown in Figure 14.

Table 14

Student Item Response Percents for Items 13-16 on the FCSA

Item	Key	“A” Responses	“B” Responses	“C” Responses	“D” Responses
13	A	58	26	13	3
14	A	79	12	6	3
15	D	12	8	20	61
16	C	4	7	84	5

Table 15

BILOG-MG3 Item Parameter Estimates for Items 13-16 on the FCSA

Item	a-parameter	b-parameter	c-parameter
13	1.01	-0.03	0.14
14	1.10	-0.95	0.15
15	0.89	-0.08	0.18
16	0.65	-1.51	0.23

Item thirteen included one distractor designed to attract students who used univariate reasoning, one distractor designed to attract students to the reciprocal of the correct slope, and one distractor designed to attract students who confused total amount with amount of change. For this item, response option B represented univariate reasoning, option C represented the reciprocal slope, and option D represented total amount instead of amount of change. As shown in Table 14, 26% of the students who took that FCSA chose option B for item thirteen, indicating that the students may have used univariate reasoning when working this item, whereas 13% of students who took the FCSA chose option C, indicating that the students chose the reciprocal of the correct slope for item thirteen. Only 3% of the students who took the FCSA chose option D, indicating that the students may have confused total amount of change with rate of change.

Item fourteen included one distractor designed to attract students to the reciprocal of the correct slope and two distractors to attract students who used additive reasoning instead of multiplicative reasoning. For this item, response option B represented the reciprocal of the correct slope, and options C and D represented additive reasoning. As shown in Table 14, 12% of the students who took the FCSA chose option B for item fourteen, indicating that the students chose the reciprocal of the correct slope, whereas a total of 9% of the students who took the FCSA chose options C or D, indicating that the students may have applied additive reasoning when working this item.

Item fifteen included one distractor designed to attract students to the reciprocal slope, one distractor designed to attract students who applied additive reasoning, and one distractor designed to attract students who used additive reasoning and then chose the reciprocal slope. For this item, response option A represented the reciprocal of the correct slope, option C represented additive reasoning, and option B represented additive reasoning and the reciprocal of the correct arrangement of the slope fraction. As shown in Table 14, 20% of the students who took the FCSA chose option C for item fifteen, indicating that the students may have used additive reasoning, whereas 12% chose option A, indicating that the students selected the reciprocal of the correct slope. Also shown in Table 14, 8% of the students who took the FCSA chose option B, indicating that the students may have used additive reasoning and selected the reciprocal of the correct slope.

Item sixteen included one distractor designed to attract students who used additive reasoning, and one distractor to attract students who used univariate reasoning. For this item, option A represented additive reasoning and option B represented univariate reasoning. Option D was simply incorrect. As shown in Table 14, 4% of the students who took the FCSA chose option A, indicating that the students may have used additive reasoning, 7% of the students chose option B, indicating that the students may have used univariate reasoning, and 5% of the students chose option D.

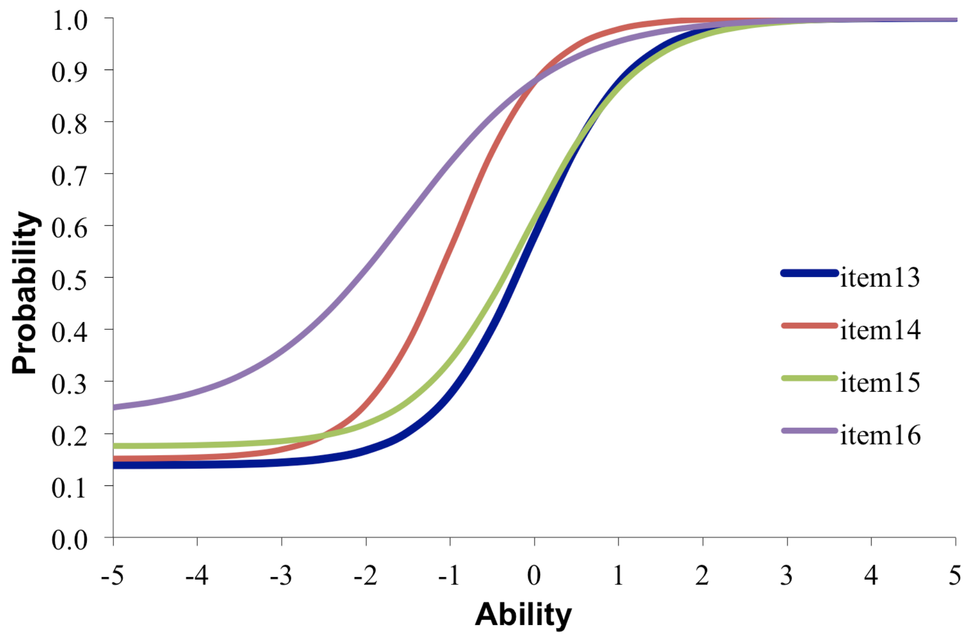


Figure 14. Item Characteristic Curves for Items 13-16 on the FCSA

These four items as a group were somewhat more difficult for this group of students than the items that assessed attributes A1, A2, and A3. Items thirteen through sixteen were more discriminating as a group than items one through eight, but less discriminating than items nine through twelve, a fact that is illustrated by the relatively steep slopes of the graphs and by slope values between a = 0.6 and a = 1.1, as shown in Table 15. Items thirteen and fourteen were quite discriminating, with slope values greater than 1.0. The easiest item in this group was item number sixteen, which was answered correctly by 84% of the students who took the FCSA. This item had the lowest difficulty parameter of these four items, that is, b = -1.51, confirming that students with somewhat low ability estimates demonstrated a 50% chance of answering this item correctly. The most difficult item in this group was item number thirteen, which was answered correctly by 58% of the students who took the FCSA. This item also had the highest difficulty parameter of these four items, that is, b = -0.03.

Items Intended to Assess Attribute A5

Items seventeen through twenty assessed whether students could interpret the meaning of slope in terms of the quantities relevant to a constant rate problem scenario. These items contained slopes whose ratio values simplified to positive rational numbers that were neither whole numbers nor unit fractions. Items seventeen and nineteen presented students with verbal problems, and students had to select the graphs that matched the scenarios in the problems. Items eighteen and twenty presented students with graphs, and students had to select the verbal descriptions that matched the graphs. The percents of students that chose each response option are shown in Table 16, the IRT item parameters are shown in Table 17, and the item characteristic curves are shown in Figure 15.

Table 16

Student Item Response Percents for Items 17-20 on the FCSA

Item	Key	"A" Responses	"B" Responses	"C" Responses	"D" Responses
17	C	3	3	81	13
18	A	86	6	4	4
19	A	66	12	13	9
20	D	29	17	20	33

Item seventeen included one distractor designed to attract students who used additive reasoning, one distractor designed to attract students to the reciprocal of the correct slope, and one distractor that was simply incorrect. For this item, response option B represented additive reasoning, option D represented the reciprocal slope, and option A was simply incorrect.

Table 17

BILOG-MG3 Item Parameter Estimates for Items 17-20 on the FCSA

Item	a-parameter	b-parameter	c-parameter
17	0.76	-1.23	0.17
18	1.36	-1.20	0.22
19	1.03	-0.24	0.20
20	0.82	2.09	0.26

As shown in Table 16, 3% of the students who took that FCSA chose option B, indicating that the students may have used additive reasoning. Also shown in Table 16, 13% of the students chose option D for this item, indicating that the students chose the reciprocal of the correct slope when working this item. Also shown in Table 16, 3% of the students who took the FCSA chose option A.

Item eighteen included one distractor designed to attract students to the reciprocal of the correct slope and two distractors to attract students who used additive reasoning instead of multiplicative reasoning. For this item, response option B represented the reciprocal of the correct slope, and options C and D represented additive reasoning. As shown in Table 16, 6% of the students who took the FCSA chose options B, indicating they chose the reciprocal slope. Also shown in Table 16, 8% of the students who took the FCSA chose options C or D for this item, indicating the students may have used additive reasoning.

Item nineteen included one distractor designed to attract students to the reciprocal slope, one distractor designed to attract students who used univariate reasoning, and one distractor that was simply incorrect. For this item, response option B represented the reciprocal of the correct slope, option C represented univariate reasoning, and option D was simply incorrect. As shown in Table 16, 12% of the students who took the FCSA chose option B for item nineteen, indicating that the students chose the reciprocal of the correct slope. Also shown in Table 16, 13% of the students chose option C, indicating that the students may have used univariate reasoning when working this item, and 9% chose option D.

Item twenty included one distractor designed to attract students to the reciprocal of the correct slope, one distractor to attract students who used univariate reasoning, and one distractor designed to attract students who read the graph incorrectly. For this item, response option B represented univariate reasoning, option C represented the reciprocal of the correct slope, and option A was simply incorrect. As shown in Table 16, 17% of the students who took the FCSA chose response option B for this item, indicating that the students may have used univariate reasoning when working this item. Also shown in Table 16, 20% of the students who took the FCSA chose option C, indicating that the students chose the reciprocal of the correct slope. Also shown in Table 16, 29% of the students who took the FCSA chose option A, which represented an error in reading the graph.

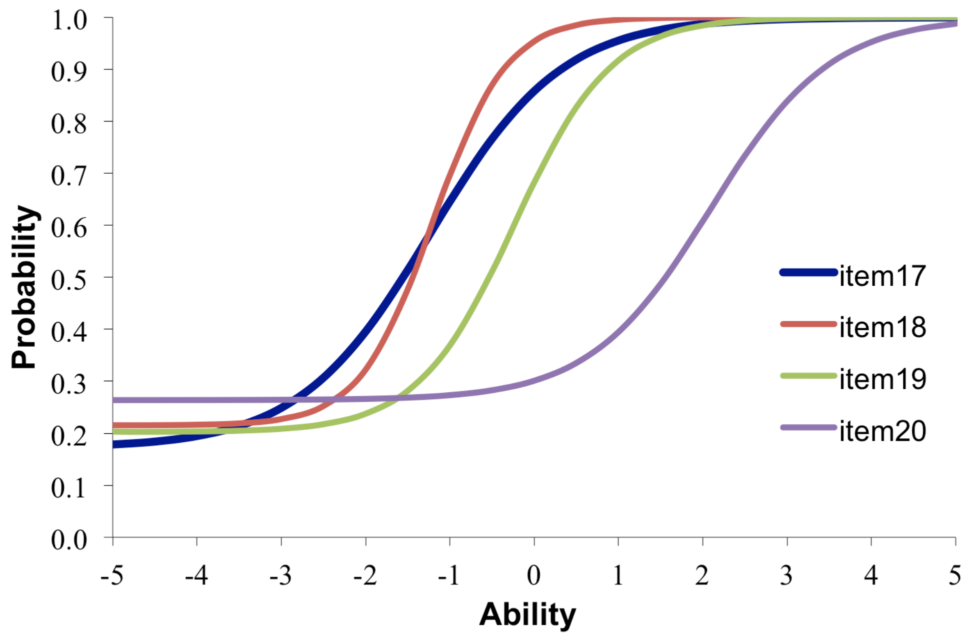


Figure 15. Item Characteristic Curves for Items 17-20 on the FCSA

These four items as a group were somewhat less difficult for this group of students than the items that assessed attribute A4. Items seventeen through twenty were more discriminating as a group than items one through eight, less discriminating than items nine through twelve, and only slightly more discriminating than items thirteen through sixteen, a fact that is illustrated by the relatively steep slopes of the graphs and by slope values between a = 0.75 and a = 1.36, as shown in Table 17. Items eighteen and nineteen were quite discriminating, with slope values all greater than a = 1.0. The easiest items in this group were items seventeen and eighteen, which were answered correctly by 81% and 86%, respectively, of the students who took the FCSA. These items had the lowest difficulty parameters of these four items, that is, item seventeen's difficulty value was b = -1.23, and item eighteen's difficulty value was -1.20. These values confirm that students with somewhat low ability estimates demonstrated a 50% chance of answering these items correctly. While item eighteen was relatively easy, it was also very discriminating for the students who took the FCSA. The most difficult item in this group and on the entire test was item number twenty, which was answered correctly by only 33% of the students who took the FCSA. This item also had the highest difficulty parameter of these four items, that is, b = 2.09, indicating that only students with relatively high ability estimates had a 50% chance of answering this item correctly.

Analysis of Misconceptions and Errors

Observations about the misconceptions and errors made by students in this study were restricted to the erroneous response options offered in the items on the FCSA. Percents of students who selected erroneous responses for each item were presented in the preceding sections. However, it is difficult to assess from single item response data how common different errors are for the population of students who participated in this study. Table 18 displays the FCSA item numbers with the errors presented in the incorrect response options for each item along with the percent of students that chose each of these options. These data illustrate that, on average, less than 15% of the students who took the FCSA selected response options consistent with each of the misconceptions identified in this study.

Table 18

Percent of Students that Chose Different Erroneous Reponses on the FCSA by Item

Item Number	Reciprocal of the Correct Slope	Opposite of the Correct Slope	Total Amount Confused with Amount of Change	Additive Reasoning	Identify (Incorrect) Varying Quantities	Univariate Reasoning
1					11	
2					4	
3					12	
4					16	
5		2				6
6			17			
7		24				13
8		7			12	
9	11	10				
10	8		11			
11	14					17
12	4	4				
13	13		3			26
14	12			9		
15	12			20		
16				4		7
17	13					
18	6			8		
19	12					13
20	20					17
Mean	11	11	10	10	11	14

Analysis of Item Contexts

The items on the FCSA were all contextual problems, that is, each problem contained a situation involving unit price, average rate, or scale factor. For the analysis, problems that involved ingredients in a recipe were classified as scale factor problems. These contexts were chosen because they were found repeatedly in the literature and noted as popular applications of the slope context that likely would be familiar to most students. Table 19 lists the item numbers with their contexts along with the percent of students that answered each item correctly. These data illustrate that students who took the FCSA, on average, demonstrated the most success on problems whose contexts pertained to unit price. Students demonstrated the least success on problems whose contexts pertained to scale factors.

Table 19

Percent of Students that Correctly Answered Items Based on Different Contexts on the FCSA

Item Number	Unit Price	Average Rate	Scale Factor
1		89	
2		96	
3	88		
4			84
5			91
6	83		
7		64	
8		88	
9		75	
10	81		
11			69
12			87
13		58	
14			79
15	61		
16	84		
17	81		
18			86
19		66	
20			33
Mean	80	77	76

Analysis of Mathematical Representations

The items on the FCSA targeting attributes A3, A4, and A5 included graphs and verbal descriptions. Half of the items for these attributes contained graphs in the item stems, and students were required to select verbal descriptions that matched the graphs. The other half of the items for these attributes contained verbal descriptions in the item stems, and students were required to select graphs that matched the verbal descriptions. For the analysis, the problems in which graphs were presented in the stem are labeled as graph-verbal problems, and problems in which verbal statements were presented in the stem are labeled as verbal-graph problems. Table 20 lists the item numbers with their representations along with the difficulty estimates represented by percent of students that answered each item correctly and IRT difficulty estimates (b-parameter values). These data illustrate that students who took the FCSA, on average, demonstrated more success on problems containing graphs in the stems and requiring students to select verbal descriptions that matched the graphs. Students demonstrated less success on problems containing verbal descriptions in the item stems and requiring students to select graphs that matched the verbal descriptions.

Table 20

Difficulty Estimates of the Items on the FCSA Targeting Attributes A3, A4, and A5 when Grouped According to the Mathematical Representations Used in Stem and Answer Choices

		Representation			
		Verbal-graph		Graph-verbal	
Item	Attribute	Percent of Correct Responses	b-parameter	Percent of Correct Responses	b-parameter
9	A3	75	-0.76		
10	A3			81	-0.82
11	A3	69	-0.43		
12	A3			87	-1.27
13	A4	58	-0.03		
14	A4			79	-0.95
15	A4	61	-0.08		
16	A4			84	-1.51
17	A5	81	-1.23		
18	A5			86	-1.20
19	A5	66	-0.24		
20	A5			33	2.09
Mean		68	-0.46	75	-0.61

Note: Verbal-graph indicates that the question stem was represented verbally and the answer options were represented by graphs. Graph-verbal indicates that the question stem was represented by a graph and the answer choices were represented verbally.

CHAPTER 5

SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

Summary

Learning with understanding is important (Novak, 1998) and depends on the human ability to meaningfully relate new information to what is already known and cognitively structured (Ausubel, 1968). Learning with understanding occurs when students actively and consciously connect a new idea to their existing cognitive structure (Ausubel, 1968). Concepts that are learned with understanding are more easily remembered and applied to future experiences than concepts that are not purposefully connected to a person's existing knowledge (Grouws, 1991). Therefore, learning with understanding provides the foundation for future learning.

Various authors have attempted to describe what it means to understand and document how students learn with understanding. A common theme emphasized by many authors is that conceptual understanding depends on connected knowledge. Early in the past century, progressive educators noted the importance of a person's existing knowledge and previous experiences as primary factors in determining what that person was able to learn (Dewey, 1938). Progressive educators, therefore, recommended deliberately designed school activities that built on students' prior knowledge and experiences and provided a foundation for future experiences (Willis et al., 1994). Meaningful knowledge is formed when students relate new experiences to their existing knowledge and previous experiences. Therefore, the knowledge that students understand is directly associated with their experiences.

Piaget also emphasized the importance of existing knowledge and experiences as grounds for intellectual development (Flavell, 1963). He described intelligent activity to be a process in which a person first assimilates new knowledge with existing knowledge, thereby incorporating something new with what is already cognitively organized. The second part of the process occurs when the person adapts the cognitive structure to accommodate the newly acquired information such that this newly revised cognitive structure is applied to all future experiences. The culmination of successive assimilations and adaptations is a sophisticated, networked, highly organized intellectual structure.

Building on Piaget's theory of intellectual development, constructivists emphasized the individual and active nature of the learning process (Confrey, 1990; Noddings, 1990; von Glasersfeld, 1990). Meaningful learning requires students to actively construct new understandings, which may include intermediate, personal representations of ideas or informal conceptions (Battista, 1994; Carpenter & Lehrer, 1999). Over time and experience students form connections among ideas, connections among the different representations of the ideas, and connections among the various symbols or language used to communicate about the ideas (Confrey, 1990).

Mathematics educators and researchers continue to emphasize the importance of understanding and making sense of mathematics (CCSSO/NGA, 2010; NCTM, 2000). In order to understand mathematics, students must develop complex networks of knowledge reflecting conscious organizations of related mathematical facts and processes (Hiebert & Carpenter, 1992; Hiebert & Lefevre, 1986; Klausmeier et al., 1968; Marshall, 1990; Messick, 1984; Skemp, 2006; Webb & Romberg, 1992). Conceptual understanding is, therefore, defined as "an integrated and functional grasp of mathematical ideas" (National Research Council, 2001, p. 118).

Understanding mathematical ideas requires students to construct sophisticated networks of knowledge containing connections among concepts and processes (Beyer, 1993; Carpenter &

Fennema, 1991; Hiebert & Carpenter, 1992; Hiebert & Lefevre, 1986; Marshall, 1990; Messick, 1984; National Research Council, 2001; Vergnaud, 1997; Webb & Romberg, 1992). Connected knowledge develops over time and experience, and the depth of a person's understanding grows with the number of cognitive connections a person constructs among concepts and procedures (Hiebert & Lefevre, 1986). Connections that promote understanding of mathematical concepts include (1) ways in which different representations are used to express the meaning of a concept (National Research Council, 2001; Niemi, 1996), (2) ways in which concepts are associated with procedures (Boaler, 2002; Hiebert & Carpenter, 1992; Mathematical Sciences Education Board, 1993; National Research Council, 2001), and (3) ways in which prior knowledge is associated with new information (Hiebert & Lefevre, 1986; NCTM, 2000; Vergnaud, 1997).

Knowledge of facts and procedures is instrumental for working mathematically, but symbol manipulations are often required as the chief demonstration of mathematical proficiency (Battista, 1999; Vergnaud, 1997), which overlooks the importance of the conceptual underpinnings of mathematical processes. A consequence of this oversight is the dominant view that proficiency in mathematics can be explained by procedural fluency, which conveys little about a student's mathematical understanding (Boaler, 2002). In many cases, high-stakes assessments target small sets of skills and procedures rather than what students understand about mathematics (Neill, 2003), which exerts a negative influence on some instructional situations (Harlen, 2007; Suurtamm et al., 2008). Specifically, teachers focused on test preparation concentrate on skill mastery rather than underlying concepts or mathematical connections, which jeopardizes the potential for students to develop mathematical proficiency (Battista, 1999).

Assessing student understanding of mathematics requires alternative measurement tools that query conceptual understanding in addition to procedural fluency (Webb & Romberg, 1992).

The measurement process used for any assessment should take into account the nature of how people come to know things, how people demonstrate their knowledge, and how observations or test responses are interpreted (Mislevy, Almond, et al., 2003; Mislevy, Steinberg, et al., 2003; Mislevy et al., 2000). Assessments that target understanding should be construct centered (Kane, 2001), that is, they should be developed in response to descriptions of the knowledge, skills, and abilities students are expected to learn.

Evidence-centered design (ECD) is a framework for evaluating the decisions and strategies applied throughout the process of developing and implementing educational assessments (Mislevy, Almond, et al., 2003; Mislevy & Haertel, 2006; Mislevy, Steinberg, et al., 2003). Evidence-centered assessment design uses five layers to frame validity claims (Mislevy & Haertel, 2006). The layers needed to support a validity claim are domain analysis, domain modeling, conceptual assessment framework, assessment implementation, and assessment delivery. Domain analysis consists of gathering information about the domain of knowledge, skills, and abilities to be assessed. Domain modeling involves organizing the information gathered during domain analysis and relating the knowledge, skills, and abilities to be assessed to item types and tasks that will cause students to demonstrate their relevant knowledge, procedures, and strategies. The conceptual assessment framework (CAF) transforms the work done at the domain modeling layer and outlines the technical specifications for an assessment. The assessment implementation layer concerns the preparation of all the operational elements of the assessment that are outlined in the CAF. The assessment delivery layer consists of students taking tests, their responses being scored, and reports being produced and interpreted by test users.

Test developers who implement ECD are likely to incorporate cognitive models of how the knowledge targeted by an assessment is arranged in a person's cognitive structure (Mislevy & Haertel, 2006). Learning progressions (Popham, 2011), construct maps (Wilson, 1992, 2009), and learning hierarchies (Gagné, 1968) are three types of cognitive models useful for describing the knowledge people possess, how it is typically learned, and common misconceptions. These types of models can be used to guide the development of assessment items and tasks that measure what people know in a manner that is consistent with how knowledge is acquired (Pellegrino, 1988). Assessments developed in this way can also provide information about what students know in terms of the elements of the model, thereby directly relating test responses to the learning targets described in the model (Pellegrino et al., 1999).

A relatively new type of cognitive model is the cognitive attribute hierarchy, which can be used to model the cognitive processes or skills students need to successfully complete test items (Gierl et al., 2008; Leighton et al., 2004). In a cognitive attribute hierarchy, each component of knowledge, skill, or ability needed to demonstrate proficiency in a domain is represented by an attribute (Leighton et al., 2004). Furthermore, the attributes are arranged hierarchically to depict how people likely acquire domain knowledge, and the dependencies, or cognitive connections, among the attributes are clearly indicated (Leighton et al., 2004).

Understanding the slope of a line is one of the most important mathematical proficiencies students should acquire in secondary mathematics courses (Andersen & Nelson, 1994; Joram & Oleson, 2007), but acquiring the ability to reason about slope is difficult for students (Graham et al., 2010). Understanding slope builds on a student's ability to reason with ratios (Lobato & Thanheiser, 2002). Two types of reasoning contribute to a student's ability to work with ratios, namely, covariational reasoning and proportional reasoning (Hoffer, 1988).

Covariational reasoning is evident when a student can describe the relationship between two quantities that vary together (Adamson, 2005). This understanding requires a student to recognize that two varying quantities together produce a third type of quantity whose value is represented as a ratio (Lobato & Thanheiser, 2002). One framework suggested that covariational reasoning develops in five successive stages (Carlson et al., 2002). The first level of this framework is the ability to recognize how changes in one variable correspond to the changes in a second variable, given the description of two varying quantities. The second level is the ability to detect or describe the direction of covariation, given the description of an instance of covariation. The third level is the ability to calculate the amount of change in one variable given the amount of change in another variable. The fourth level is the ability to construct a ratio comparing the change in one variable to the change in a second variable. The fifth level is the ability to interpret the ratio comparing the changes in two variables as a rate.

Covariational reasoning supports a person's ability to interpret information presented in graphs. Moritz (2005) developed a four-level model to describe how students acquire the ability to interpret covariation depicted in graphs. The first level in this model is the ability to describe the meaning of the axis labels of a graph. The second level is the ability to describe change in a single variable, either the variable represented on the horizontal axis or the variable represented on the vertical axis of a graph. The third level is the ability to describe changes in two variables displayed on a graph without necessarily noting any relationship between the variables. The fourth level is the ability to describe how two variables displayed on a graph change in relation to each other.

Proportional reasoning develops after a person has the ability to conceive of covariation and is able to work with rational numbers (Heller et al., 1990). Proportional reasoning involves multiple comparisons, namely, comparisons between the two values that make up a ratio and comparisons between two or more ratios. Therefore, several researchers have noted that the ability to reason with proportions signifies formal thinking (Flavell, 1963; Hoffer, 1988; Lesh et al., 1988) and is regarded as a cornerstone of secondary mathematics (Lamon, 1993).

Two components that affect the complexity of proportional reasoning are the qualitative relationships between the quantities compared in a proportion and the quantitative relationships among the numbers in a proportion (Cramer et al., 1989). These two components require two different types of reasoning, which are demonstrated in different ways.

Carlson et al. (2002) proposed definitions for qualitative and quantitative reasoning about proportions. Qualitative reasoning refers to the ability to detect the nature or direction of a proportional relationship and is demonstrated by students who describe straight lines to be steeper or less steep than one another, based on the lines' orientations shown on a graph. Quantitative reasoning refers to the ability to evaluate and compare the numerical quantities in a proportion and is demonstrated by students who calculate the slopes of two lines shown on a graph and associate the steeper line with the larger of the two slopes.

Two additional factors that increase the complexity of proportional reasoning are the two types of comparisons needed to reason with proportions (Cramer et al., 1993). *Between-type* comparisons examine two different kinds of quantities or two quantities measured in different units. *Within-type* comparisons examine two like quantities or two quantities measured in the same unit. Regardless of the type of comparison, however, researchers consistently have determined that students favor comparisons that yield simple numerical values (Tourniaire & Pulos, 1985). Specifically, students tend to work with numerical comparisons that simplify to whole numbers before working with other comparisons. Noelting (1980a) determined that student ability to reason about proportions develops in stages. First, students proceed from the ability to make single comparisons to the ability to make multiple comparisons. Second, students proceed from comparing values whose ratios simplify to integers, to comparing values whose ratios simplify to unit fractions, and finally to comparing values whose ratios simplify to neither of these.

Difficulties students have in learning to reason about covariation and proportions are described in frameworks provided by Carlson et al. (2002), Moritz (2005), Cramer, Post, and Behr (1989), Cramer, Post, and Currier (1993), and Noelting (1980a). Common misconceptions or errors associated with covariation include the failure to identify which quantities are related to one another in a problem (Moritz, 2005) and the failure to detect the direction of covariation of two quantities (Barr, 1980). Common misconceptions or errors associated with proportional reasoning include working with one variable's values instead of two variables' values (Moritz, 2005), applying additive strategies rather than multiplicative strategies (Heller et al., 1990), constructing reciprocal ratios (Barr, 1980), constructing opposite-sign ratios (Barr, 1980), and failing to distinguish total values from amounts of change (Bell & Janvier, 1981).

The present study investigated the knowledge and skills students need to understand the concept of slope. Five components of knowledge were selected and arranged into a hierarchical model, the Foundational Concepts of Slope Attribute Hierarchy (FCSAH), which guided the development of an assessment, the Foundational Concepts of Slope Assessment (FCSA). This study gathered information about student knowledge of selected foundational concepts that contribute to understanding slope and the common misconceptions held by students. The study answered the following research questions:

1. What insight is gained about the validity of the proposed cognitive model from an analysis of student data generated from an assessment informed by the model?
2. To what extent did student participants exhibit common misconceptions regarding slope?

In order to answer these research questions, this study implemented the attribute hierarchy method (AHM) (Leighton et al., 2004). The study was conducted in two phases, a domain modeling phase and a task modeling phase.

The domain modeling phase was conducted in early Spring 2011. Guided by literature on covariational and proportional reasoning, the researcher created a hierarchical model containing five attributes that support a person's understanding of selected foundational concepts related to slope, i.e., the FCSAH. This model was shared with subject matter experts, who helped to refine its structure and to clearly describe the components of knowledge included in the FCSAH. The FCSAH is shown in Chapter 3 in Figure 1.

The task modeling phase was conducted in late Spring 2011. The FCSAH developed in the first phase was used to guide the development of the FCSA, which was administered in May 2011 to gather data about student knowledge of the five attributes included in the hierarchy.

Following the process described by Gierl, Leighton, and Hunka (2000), four matrix representations of the FCSAH were constructed to guide item development. The Adjacency matrix depicted the direct relationships among the attributes in the FCSAH. The Reachability matrix depicted the direct and indirect relationships among the attributes in the FCSAH. The Incidence (Q) matrix depicted all possible combinations of the attributes in the FCSAH. The Reduced Incidence (Q_r) matrix contained the columns from the Q matrix that were consistent with the FCSAH. The Q_r matrix determined five item types necessary to effectively assess the knowledge modeled by the FCSAH. Four items were developed for each of the five item types depicted in the Q_r matrix. The resulting FCSA contained 20 multiple-choice items and was designed to provide a quantitative measure of students' understanding of the attributes modeled in the FCSAH.

The items on the FCSA required students to interpret problem situations presented verbally or in graphs. The items assessed students' ability to determine which quantities varied together in a problem, determine the direction of covariation of two quantities in a problem, and interpret the meaning of a slope ratio in terms of different problem contexts. The FCSA was administered online in May 2011 to 1629 middle and high school students studying Pre-algebra, Algebra 1, Geometry, Algebra 2, or similar courses taken prior to Pre-calculus.

Student responses to the items on the FCSA were analyzed using item response theory (IRT). These analyses were used to describe the items on the FCSA, the FCSA as a test, and the five subtests that each targeted one specific combination of attributes in the FCSAH. Three IRT parameters were determined to describe each item's discrimination, difficulty, and lower bound, and these parameters determined an item characteristic curve (ICC) for each item. A review of the ICCs confirmed that many of the items on the FCSA were relatively easy for the participants in this study, while several of the items on the FCSA discriminated well between students of different ability levels. The test information function illustrated that the FCSA as a test was relatively easy for the participants, and it provided the most information for students with ability levels from -2.0 to 1.0.

The five subtests each targeting knowledge of one specific combination of attributes were analyzed to determine the test information function for each subtest. These analyses revealed that the subtest targeting knowledge of attribute A1 was less informative than the other subtests for the entire population in the study and was most informative for students with ability estimates near the value of -2.0. The subtest targeting knowledge of attributes A1 and A2 was slightly more informative overall than the first subtest and was most informative for students with ability estimates near the value of -0.5. The subtest targeting knowledge of attributes A1, A2, and A3 was the most informative for the entire population in the study and was particularly informative for students with ability estimates near the value of -1.0. The subtest targeting knowledge of attributes A1, A2, and A4 was the second most informative subtest for the entire population in the study and was particularly informative for students with ability estimates near the value of -0.2. The subtest targeting knowledge of attributes A1, A2, and A5 was the third most informative subtest for the entire population in the study and was particularly informative for students with ability estimates near the value of -1.0.

Test items target particular content knowledge, and test item responses are determined by the knowledge possessed by examinees (Gierl et al., 2000). If an examinee possesses the knowledge targeted by a test item, then that examinee should answer the test item correctly. One aim of the present study was to classify each student participant into a particular knowledge state with regard to the attributes modeled in the FCSAH. The different levels of knowledge used for classification were the ten different combinations of attributes consistent with the FCSAH. Each of the ten attribute combinations consistent with the FCSAH was considered to be a different knowledge state, each indicating a different level of knowledge of the attributes modeled by the FCSAH. For each knowledge state, an expected response pattern was determined. Each expected response pattern contained the responses a hypothetical student would give depending on that student's level of knowledge.

Student responses to the FCSA were examined to determine the level of knowledge possessed by each examinee. Using the AHM, each student's observed response pattern was compared to all ten of the expected response patterns. These comparisons yielded a likelihood estimate of how well a particular student's response pattern matched each of the expected response patterns. For each student, the ten likelihood estimates were summed, and each likelihood estimate was divided by the sum to produce a probability. Each student was assigned to the knowledge state corresponding to the expected response pattern with the highest probability.

Student responses to the FCSA and the expected response patterns were analyzed with the IRT model displayed in Equation 1 and the item parameters shown in Table 4 to determine an ability estimate for each student's response pattern and each expected response pattern. These IRT analyses placed each ability estimate on a scale with a mean of zero and a standard deviation of one. The ability estimates of the students classified to each of the ten knowledge states were analyzed to describe the abilities of students within each knowledge state and to compare the abilities of students in different knowledge states. These analyses indicated that some but not all of the ten knowledge states used for classification appeared to be distinct. Specifically, the ability estimates of the students assigned to some categories were relatively distinct, while the ability estimates of the students assigned to other categories were very similar.

Students who demonstrated knowledge of none of the attributes in the FCSAH appeared to have different ability levels than the other students. Students who demonstrated knowledge of attributes A1 and A2 appeared to have somewhat similar ability levels, but these students appeared to have different ability levels than students who demonstrated knowledge of more attributes. Students who demonstrated knowledge of attributes A1 and A2 and either A3, A4, or

A5 appeared to have very similar ability levels, and appeared to have very different ability levels than students who demonstrated knowledge of fewer or more attributes. Students who demonstrated knowledge of attributes A1 and A2 with any pair of attributes from A3, A4 or A5 appeared to have very similar ability levels, and appeared to have very different ability levels than students who demonstrated knowledge of fewer or more attributes. Students who demonstrated knowledge of all five attributes in the FCSAH appeared to have very different ability levels than students who demonstrated knowledge of fewer attributes.

Student responses to the FCSA were analyzed to describe the typical errors or misconceptions held by students. Six common errors or misconceptions were identified from the literature and used during item development to generate the incorrect response options for the items on the FCSA. The percent of students that chose each incorrect response option was calculated to describe the frequency of choices consistent with each misconception or error. Then the percents of students who chose each type of misconception for all the items on the FCSA were averaged. These calculations revealed that, on average, 11% of students who took the FCSA selected responses consistent with constructing the slope ratio upside down, that is, the reciprocal of the correct slope. On average, 10% of the students who took the FCSA selected responses consistent with detecting the opposite of the correct direction of covariation, that is, the opposite of the correct slope. On average, 10% of the students who took the FCSA selected responses consistent with confusing total amounts with amount of change. On average, 10% of the students who took the FCSA selected responses consistent with additive reasoning instead of multiplicative reasoning. On average, 11% of the students who took the FCSA selected responses consistent with identifying incorrectly the quantities that varied together in a problem.

On average, 14% of the students who took the FCSA selected responses consistent with univariate reasoning.

Student responses to the FCSA were analyzed to describe whether different problem contexts were more or less challenging for students. Three item contexts were used for the items on the FCSA. Each item contained a situation involving either unit price, average rate, or scale factor. These contexts were chosen because they were found repeatedly in the literature and noted as popular applications of the slope concept that would likely be familiar to most students. The percent of students who answered each item correctly was calculated. Then the percents for the items having the same context type were averaged. These calculations revealed that on average, students demonstrated most success on problems whose contexts pertained to unit prices, with an average of 80% of students answering correctly. Students demonstrated less success on problems whose contexts pertained to average rates, with average of 77% of students answering correctly. Students demonstrated least success, that is, they demonstrated greatest difficulty, on problems whose contexts pertained to scale factors, with average of 76% of students answering correctly.

Student responses to the items on the FCSA that targeted the ability to interpret slope in terms of a problem's context were analyzed to describe whether the different mathematical representations used in problem stems and answer choices were more or less challenging for students. Two representations were used in these problems. In one set of problems, a graph was shown in the problem stem, and students selected a verbal statement. In the other set of problems, a verbal statement was given in the problem stem, and students selected a graph. The percent of correct responses and IRT difficulty estimates for the items having the same type of stem (graph or verbal statement) were averaged. These calculations revealed that on average, students demonstrated more success on problems in which a graph was given in the stem, with an average of 75% of students answering correctly and average difficulty of -0.61. Students, on average, demonstrated less success on problems in which a verbal statement was given, with an average of 68% of students answering correctly and average difficulty of -0.46.

Conclusions and Discussion

This study provided an example of one way to implement elements of the process articulated in the ECD literature. To date little documentation describes the implementation of the ordered sequence of phases that were completed in this study. In this study each phase was sensitive to and influenced by what was learned in the previous phase. The first phase was the domain analysis, which rendered a theoretical cognitive model. Then that model guided the development of an assessment and the interpretation of student test responses. By proceeding through these phases in the order recommended by the ECD literature, this study illustrated how to build a case for valid inferences throughout the assessment design and development process.

While many mathematics assessments focus on procedural skills, such as computation or problem solving, few mathematics assessments focus on student understanding of the concepts that underlie commonly used procedures. In particular, students are routinely required to demonstrate proficiency with computation of the slope ratio, but they are not as often required to demonstrate their understanding of what the slope ratio means. This study provided an example of how to develop an assessment targeting the cognitive connections students should construct to support their understanding of selected foundational concepts related to slope.

Any application of the AHM requires a cognitive model of the content matter being tested. However, very few studies illustrate the development of a cognitive model that is then used to guide assessment development. More frequently, tests are analyzed after they are developed, and a cognitive model is fit to the items that appear on the test. This study provided an example of how the literature relevant to a content domain can be used to generate a theoretical cognitive model.

Few studies illustrate the application of the AHM to actual student test responses. The present study demonstrated that student test responses can be used to confirm the accuracy of the theoretical cognitive model used by the AHM and also to describe student knowledge in terms of the components in the cognitive model. Assessment models like the AHM that can classify students according to what those students already know and what they still need to learn have the potential to powerfully inform instructional decisions.

Using AHM and the proposed cognitive model, students in this study were classified into different knowledge states with regard to their knowledge of selected foundational concepts related to slope. An analysis of these students' classifications suggested that knowledge of the attributes in the FCSAH, as assessed by the FCSA, develops in a manner that is consistent with the arrangement of the attributes in the FCSAH.

In agreement with the frameworks provided by Carlson et al. (2002) and Moritz (2005), the first attribute in the FCSAH was the ability to identify two quantities that vary in correspondence to one another in a covariation problem. The results from this investigation, using these students' responses to the FCSA, support the theory that the ability to identify the two variables that are associated in a covariation problem is likely acquired before the ability to detect the direction of covariation between two quantities.

Also in agreement with the frameworks provided by Carlson et al. (2002) and Moritz (2005), the second attribute in the FCSAH was the ability to detect the direction of the relationship between the two quantities that vary in correspondence to one another in a covariation problem. The results from this investigation support the theory that the ability to detect the direction of the relationship between two quantities in a covariation problem is likely acquired after the ability to identify which variables are related in a covariation problem and before the ability to interpret the slope ratio in terms of a problem's context variables.

Consistent with the frameworks provided by Carlson et al. (2002) and Moritz (2005), and reflective of Noelting's (1980a, 1980b) work with proportional reasoning, three attributes were identified in the proposed cognitive model used in this study to represent the ability to interpret the slope ratio in terms of a problem's context variables. These three attributes were included in the FCSAH as the third, fourth, and fifth attributes, but this investigation did not hypothesize an ordered relationship among these three attributes. The third attribute in the FCSAH was the ability to interpret slopes whose fractions simplified to whole numbers, the fourth attribute in the FCSAH was the ability to interpret slopes whose fractions simplified to positive unit fractions, and the fifth attribute in the FCSAH was the ability to interpret slopes whose fractions simplified to positive rational numbers that were neither whole numbers nor unit fractions. The results of this investigation support the theory that the ability to interpret a slope ratio in terms of a problem's context variables likely develops after the ability to identify which variables are related in a covariation problem and the ability to detect the direction of the relationship between two quantities in a covariation problem. The results of the present study did not suggest any ordered relationship among the abilities to interpret slopes whose fractions simplified to whole numbers, unit fractions, or other fraction values.

These findings support the theory that there are three main levels of understanding of the selected foundational concepts related to slope. First, students demonstrate the ability to identify quantities that are related as covariates. Second, students demonstrate the ability to identify the direction of covariation in a problem setting. Third, students demonstrate the ability to interpret a slope ratio in terms of a problem's context variables.

On average, test item responses that were consistent with each of the six misconceptions or errors identified in the literature were selected by 11% of the students who participated in this study. Distractors for the multiple-choice items on the FCSA were developed to align with prominent misconceptions described by various mathematics education researchers. None of the misconceptions used in this study were more prevalent among these students' responses to the FCSA than the others.

Recommendations

The following recommendations are relevant to mathematics teaching and learning, mathematics education research, curriculum planning, test development, and educational measurement. These recommendations may be of interest to practitioners, researchers, curriculum developers, and test developers.

1. Mathematics educators should be aware of learning progressions that define optimal sequences for specific learning targets in order to foster learning with understanding. This study's results offer evidence suggesting that students should be expected to learn about foundational concepts related to slope in terms of three levels of understanding. Early instruction should focus on identifying variables in problem contexts presented verbally and in graphs. This should be followed by instruction that focuses on detecting the direction of the relationship between two covariates displayed in graphs or described in verbal problems. Then instruction should focus on interpreting the slope in terms of a problem's context variables when presented on a graph or in a verbal problem.

2. More mathematics concepts and skills should be studied to identify the components needed for understanding and to develop cognitive models. Such studies could be used to identify sources of misconceptions as well as optimal learning sequences that could be used to inform instruction and assessment.
3. Researchers and practitioners should be aware of different types of cognitive models that are available to represent how mathematics concepts and skills are arranged in a person's cognitive structure. Cognitive models provide means by which specific concepts and skills can be studied to identify essential cognitive connections that students should make, potential sources of misconceptions or errors, and optimal learning trajectories, which can inform curriculum planning and sequences of instruction. Cognitive models can also inform assessment development by providing insight to test developers about how different concepts and skills are cognitively organized and what knowledge should be connected.
4. Researchers and test developers should investigate the application of assessments and measurement models that effectively classify students according to their demonstrated levels of knowledge. Although these models have been investigated theoretically, more studies are needed to explore how classification models can be used to evaluate the effectiveness of educational assessments taken by actual students.
5. Further investigations into procedural fluency and conceptual understanding are needed to describe how these two types of knowledge interact to enhance student performance with regard to concepts and skills associated with slope. While the mathematics education community promotes the notion that mathematical proficiency is supported by both conceptual understanding and procedural fluency (CCSSO/NGA, 2010; National Research Council, 2001; NCTM, 2000), a limitation of the present study was that it only investigated

student understanding of the concept of slope and specifically avoided test questions in which procedural fluency would either interfere with or enhance student performance. However, the procedural skills associated with slope must also be part of what students understand about slope, which is an area worthy of further investigation.

6. Further investigation into student understanding of the slope concept should be undertaken using problems presented in tables and in symbolic forms in order to identify essential concepts and misconceptions associated with slope represented in tables and in symbolic forms. The mathematics education community recommends that students should attain representational fluency, that is, the ability to translate between different representations of mathematical ideas (NCTM, 2000). A limitation of this study was that the scope of the assessment implemented was restricted to problems that were presented verbally or in graphs.
7. Further investigations should develop different assessments to target the attributes in the FCSAH. The items used on the FCSA may contain bias or limitations that affected the results of the present study. An assessment comprised of different items may produce different results.
8. The FCSA should be administered to a broader sample of students, including students outside the state of Kansas. The FCSAH and FCSA were guided by national standards documents (CCSSO/NGA, 2010; NCTM, 2000, 2006, 2009) and a variety of mathematics education literature. However, the curriculum standards guiding instruction for students in Kansas describe specific concepts and skills related to slope in terms that may vary considerably from the descriptions used to guide instruction in different states. The grade levels where the concepts and skills related to slope are emphasized in different curricular standards may also

vary from state to state. Analysis of test responses from a population of students whose instruction is guided by a variety of standards documents may improve the generalizability of the results of this study.

9. Further investigation into the different aspects of the items developed to assess the ability to interpret the slope ratio in terms of a problem context should be undertaken to ascertain the differences in student responses to these items. The cognitive model guiding the present study, the FCSAH, included three different attributes for the ability to interpret a slope ratio in terms of a problem's context variables, but the results did not identify an ordered relationship among these three attributes. This finding suggests that there may be another feature of the items developed to assess these three attributes that drew out differences in student abilities.
10. More studies should investigate the implementation of constructed response items and innovative item types such as technology-enhanced items in which students interact with information presented in an online environment. A limitation of the present study was that it implemented only one type of test item, that is, the traditional multiple-choice item type. Mathematics concepts and skills can and should be evaluated using a variety of item types that allow students to respond more freely and to demonstrate the thought processes they use when working with mathematics problems.

REFERENCES

Adamson, S. L. (2005). *Student sense-making in an intermediate algebra classroom: Investigating student understanding of slope.* (Doctoral Dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3166918)

Adler, I. (1971). Piaget on the learning of mathematics. In D. B. Aichele & R. E. Reys (Eds.), *Readings in secondary school mathematics*. Boston, MA: Prindle, Weber & Schmidt, Inc.

Andersen, E. D., & Nelson, J. (1994). An introduction to the concept of slope. *Mathematics Teacher, 87*(1), 27-28.

Anderson, N. C. (2008). Walk the line: Making sense of y = mx + b. In C. E. Greenes & R. N. Rubenstein (Eds.), *Algebra and algebraic thinking in school mathematics: Seventieth yearbook* (pp. 233-246). Reston, VA: NCTM.

Arzarello, F., Robutti, O., & Bazzini, L. (2005). Acting is learning: Focus on the construction of mathematical concepts. *Cambridge Journal of Education, 35*(1), 55-67. doi: 10.1080/0305764042000332498

Ausubel, D. P. (1968). *Educational psychology: A cognitive view*. New York, NY: Holt, Rinehart and Winston, Inc.

Baroody, A. J., & Bartels, B. H. (2001). Assessing understanding in mathematics with concept mapping. *Mathematics in School, 30*(3), 24-27.

Baroody, A. J., & Coslick, R. T. (1998). *Fostering children's mathematical power: An investigative approach to K-8 mathematics instruction*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Barr, G. (1980). Graphs, gradients and intercepts. *Mathematics in School, 9*(1), 5-6.

Barr, G. (1981). Some student ideas on the concept of gradient. *Mathematics in School, 10*(1), 14-16.

Battista, M. T. (1994). Teacher beliefs and the reform movement in mathematics education. *Phi Delta Kappan, 75*(6), 462-470.

Battista, M. T. (1999). The mathematical miseducation of America's youth: Ignoring research and scientific study in education. *Phi Delta Kappan, 80*(6), 424-433.

Bell, A., & Janvier, C. (1981). The interpretation of graphs representing situations. *For the Learning of Mathematics, 2*(1), 34-42.

Benbow, C. P. (2008). Solving America's math problem: National panel offers recommendations. Retrieved 02/27/2011, from http://peabody.vanderbilt.edu/National_Math_Panel_recommendations.xml

Beyer, A. (1993). Assessing students' performance using observations, reflections, and other methods. In N. L. Webb & A. F. Coxford (Eds.), *Assessment in the mathematics classroom: 1993 yearbook* (pp. 111-120). Reston, VA: National Council of the Teachers of Mathematics.

Birenbaum, M., & Tatsuoka, K. K. (1993). Applying an IRT-based cognitive diagnostic model to diagnose students' knowledge states in multiplication and division with exponents. *Applied Measurement in Education, 6*(4), 255-268.

Boaler, J. (2002). Exploring the nature of mathematical activity: Using theory, research and 'working hypotheses' to broaden conceptions of mathematics knowing. *Educational Studies in Mathematics, 51*(1/2), 3-21.

Bond, L., Moss, P., & Carr, P. (1995). Fairness in large-scale performance assessment. In G. W. Phillips (Ed.), *Technical issues in performance assessment* (pp. 117-140). Washington, D.C.: Department of Education.

Bruner, J. S., Goodnow, J. J., & Austin, G. A. (1956). *A study of thinking*. New York, NY: John Wiley & Sons, Inc.

Carlson, M., Jacobs, S., Coe, E., Larsen, S., & Hsu, E. (2002). Applying covariational reasoning while modeling dynamic events: A framework and a study. *Journal for Research in Mathematics Education, 33*(5), 352-378.

Carpenter, T. P., & Fennema, E. (1991). Research and cognitively guided instruction. In E. Fennema, T. P. Carpenter, & S. J. Lamon (Eds.), *Integrating research on teaching and learning mathematics*. Albany, NY: State University of New York Press.

Carpenter, T. P., & Lehrer, R. (1999). Teaching and learning mathematics with understanding. In E. Fennema & T. A. Romberg (Eds.), *Mathematics classrooms that promote understanding* (pp. 19-32). Mahwah, NJ: Lawrence Erlbaum Associates.

Carson, T., & Jordan, B. (2011). *Intermediate algebra* (3rd ed.). Boston, MA: Pearson Education, Inc.

CCSSO/NGA. (2010). *Common core state standards for mathematics*. Washington, D.C.: Council of State School Officers and the National Governors Association Center for Best Practices.

Cheng, D. (2010). *Connecting proportionality and slope: Middle school students' reasoning about steepness.* (Doctoral Dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3405966)

Clement, J. (1989). The concept of variation and misconceptions in Cartesian graphing. *Focus on Learning Problems in Mathematics, 11*(2), 77-87.

Cobb, P., Wood, T., & Yackel, E. (1990). Classrooms as learning environments for teachers and researchers. In R. B. Davis, C. A. Maher, & N. Noddings (Eds.), *Constructivist views of teaching and learning mathematics* (pp. 125-146). Reston, VA: National Council of Teachers of Mathematics, Inc.

Cobb, P., Yackel, E., & Wood, T. (1992). Interaction and learning in mathematics classroom situations. *Educational Studies in Mathematics, 23*(1), 99-122.

Confrey, J. (1990). What constructivism implies for teaching. In R. B. Davis, C. A. Maher, & N. Noddings (Eds.), *Constructivist views on the teaching and learning of mathematics* (pp. 107-124). Reston, VA: National Council of Teachers of Mathematics, Inc.

Cramer, K. A., Post, T., & Behr, M. (1989). Interpreting proportional relationships. *Mathematics Teacher, 82*(6), 445-453.

Cramer, K. A., Post, T., & Currier, S. (1993). Learning and teaching ratio and proportion: Research implications. In D. Owens (Ed.), *Research ideas for the classroom: Middle grades mathematics* (pp. 159-178). New York, NY: Macmillan Publishing Company.

Crawford, A. R., & Scott, W. E. (2000). Making sense of slope. *Mathematics Teacher, 93*(2), 114-118.

de Lange, J., & Romberg, T. A. (2005). Research in assessment practices. In T. A. Romberg, T. P. Carpenter, & F. Dremock (Eds.), *Understanding mathematics and science matters* (pp. 279-305). Mahwah, NJ: Lawrence Erlbaum Associates.

Dewey, J. (1938). *Experience in education*. New York, NY: Kappa Delta Pi.

Dienes, Z. P. (1971). Dienes on the learning of mathematics. In D. B. Aichele & R. E. Reys (Eds.), *Readings in secondary school mathematics* (pp. 222-237). Boston, MA: Prindle, Weber & Schmidt, Inc.

Dugsdale, S. (1993). Functions and graphs - Perspectives on student thinking. In T. A. Romberg, E. Fennema, & T. P. Carpenter (Eds.), *Integrating research on the graphical representation of functions*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Farrell, M. (1985). Adolescents' performance on a sequence of proportional reasoning tasks. *Journal of Research in Science Teaching, 22*(6), 503-518.

Flavell, J. H. (1963). *The developmental psychology of Jean Piaget*. Princeton, NJ: D. Van Nostrand Company, Inc.

Gagné, R. (1968). Learning hierarchies. *Educational Psychologist, 6*, 1-9.

Gagné, R. (1971). Gagné on the learning of mathematics: A product orientation. In D. B. Aichele & R. E. Reys (Eds.), *Readings in secondary school mathematics*. Boston, MA: Prindle, Weber & Schmidt, Inc.

Gierl, M. J. (2007). Making diagnostic inferences about cognitive attributes using the rule-space model and attribute hierarchy method. *Journal of Educational Measurement, 44*(4), 325-340. doi: 10.1111/j.1745-3984.2007.00042.x

Gierl, M. J., Leighton, J. P., & Hunka, S. M. (2000). An NCME instructional module on exploring the logic of Tatsuoka's rule-space model for test development and analysis. *Educational Measurement: Issues and Practice, 19*(3), 34-44. doi: 10.1111/j.1745-3992.2000.tb00036.x

Gierl, M. J., Wang, C., & Zhou, J. (2008). Using the attribute hierarchy method to make diagnostic inferences about examinees' cognitive skills in algebra on the SAT. *The Journal of Technology, Learning, and Assessment, 6*(6), 4-49.

Glaser, R. (1986). *The integration of instruction and testing*. Paper presented at the Redesign of Testing for the 21st Century: Proceedings of the 1985 ETS Invitational Conference, New York, NY.

Gorin, J. S. (2006). Test design with cognition in mind. *Educational Measurement: Issues and Practice, 25*(4), 21-35. doi: 10.1111/j.1745-3992.2006.00076.x

Graham, K., Cuoco, A., & Zimmerman, G. (2010). *Focus in high school mathematics: Reasoning and sense making in algebra*. Reston, VA: NCTM.

Grocki, J. M. (1990). The gymnasium - A dynamic classroom for teaching slope. *Mathematics Teacher, 83*(8), 636-637.

Grouws, D. A. (1991). Improving research in mathematics classroom instruction. In E. Fennema, T. P. Carpenter, & S. J. Lamon (Eds.), *Integrating research on teaching and learning mathematics*. Albany, NY: State Univeristy of New York Press.

Haertel, E. (1985). Construct validity and criterion-referenced testing. *Review of Educational Research, 55*(1), 23-46.

Hancock, L., & Kilpatrick, J. (1993). Effects of mandated testing on instruction. In Mathematical Sciences Education Board & National Research Council (Eds.), *Measuring what counts: A conceptual guide for mathematics assessment* (pp. 149-174). Washington, D.C.: National Academy Press.

Harlen, W. (2007). Criteria for evaluating systems for student assessment. *Studies in Educational Evaluation, 33*(1), 15-28.

Harlen, W., & James, M. (1997). Assessment and learning: Differences and relationships. *Assessment in Education: Principles, Policy & Practice, 4*(3), 365.

Hart, K. (1988). Ratio and proportion. In J. Hiebert & M. Behr (Eds.), *Number concepts and operations in the middle grades* (pp. 198-219). Reston, VA: NCTM.

Hauger, G. S. (1999). High school and college students' knowledge of rate of change. *Dissertation Abstracts International Section A: Humanities and Social Sciences, 59*(10-A).

Heller, P. M., Post, T. R., Behr, M., & Lesh, R. (1990). Qualitative and numerical reasoning about fractions and rates by seventh- and eighth-grade students. *Journal for Research in Mathematics Education, 21*(5), 388-402.

Henderson, K. B. (1970). Concepts. In M. F. Rosskopf (Ed.), *The teaching of secondary school mathematics: Thirty-third yearbook* (pp. 166-195). Washington, D.C.: NCTM.

Hiebert, J., & Carpenter, T. P. (1992). Learning and teaching with understanding. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 65-97). New York, NY: Macmillan Publishing Co, Inc.

Hiebert, J., & Lefevre, P. (1986). Conceptual and procedural knowledge in mathematics: An introductory analysis. In J. Hiebert (Ed.), *Conceptual and procedural knowledge: The case of mathematics*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Hoffer, A. (1988). Ratios and proportional thinking. In T. Post (Ed.), *Teaching mathematics in grades K - 8* (pp. 285-313). Newton, MA: Allyn and Bacon, Inc.

Jaeger, R. M., Mullis, I. V. S., Bourque, M. L., & Shakrani, S. (1995). Setting performance standards for performance assessments: Some fundamental issues, current practice, and technical dilemmas. In G. W. Phillips (Ed.), *Technical issues in performance assessment* (pp. 79-115). Washington, D.C.: Department of Education.

Joram, E., & Oleson, V. (2007). How fast do trees grow: Using tables and graphs to explore slope. *Mathematics Teacher, 13*(5), 260-265.

Kane, M. T. (1992). An argument-based approach to validation. *Psychological Bulletin, 112*(3), 527-535.

Kane, M. T. (2001). Current concerns in validity theory. *Journal of Educational Measurement, 38*(4), 319-342. doi: 10.1111/j.1745-3984.2001.tb01130.x

Kansas State Board of Education. (2003). *Kansas curricular standards for mathematics*. Topeka, KS: Kansas Department of Education. Retrieved from http://www.ksde.org.

Karplus, R., Pulos, S., & Stage, E. K. (1983a). Early adolescents' proportional reasoning on 'rate' problems. *Educational Studies in Mathematics, 14*(3), 219-233.

Karplus, R., Pulos, S., & Stage, E. K. (1983b). Proportional reasoning of early adolescents. In R. Lesh & M. Landau (Eds.), *Acquisition of mathematics concepts and processes* (pp. 45-90). Orlando, FL: Academic Press, Inc.

Kieran, C. (1993). Functions, graphing, and technology: Integrating research on learning and instruction. In T. A. Romberg, E. Fennema, & T. P. Carpenter (Eds.), *Integrating research on the graphical representations of functions*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Klausmeier, H. J. (1992). Concept learning and concept teaching. *Educational Psychologist, 27*(3), 267.

Klausmeier, H. J., Harris, C. W., Davis, J. K., Schwenn, E., & Frayer, D. A. (1968). Strategies and cognitive processes in concept learning: Final report. University of Wisconsin: Research and Development Center for Cognitive Learning.

Kurtz, B., & Karplus, R. (1979). Intellectual development beyond elementary school VII: Teaching for proportional reasoning. *School Science and Mathematics, 79*(5), 387-398.

Lamon, S. J. (1993). Ratio and proportion: Connecting content and children's thinking. *Journal for Research in Mathematics Education, 24*(1), 41-61.

Lamon, S. J. (2005). *Teaching fractions and ratios for understanding: Essential content knowledge and instructional strategies for teachers* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Leighton, J. P., & Gierl, M. J. (2007). Why cognitive diagnostic assessment? In J. P. Leighton & M. J. Gierl (Eds.), *Cognitive diagnostic assessment for education: Theory and applications* (pp. 3-18). New York, NY: Cambridge University Press.

Leighton, J. P., Gierl, M. J., & Hunka, S. M. (2004). The attribute hierarchy method for cognitive assessment: A variation on Tatsuoka's rule-space approach. *Journal of Educational Measurement, 41*(3), 205-237.

Lesh, R., Post, T., & Behr, M. (1988). Proportional reasoning. In J. Hiebert & M. Behr (Eds.), *Number concepts and operations in the middle grades* (pp. 93-118). Reston, VA: NCTM.

Lobato, J., & Thanheiser, E. (2002). Developing understanding of ratio-as-measure as a foundation for slope. In B. Litwiller & G. Bright (Eds.), *Making sense of fractions, ratios and proportions: 2002 yearbook* (pp. 162-175). Reston, VA: NCTM.

Lovell, K., & Butterworth, I. B. (1966). Abilities underlying the understanding of proportionality. *Mathematics Teaching, 37*(1), 5-9.

Macnamara, J. (1982). *Names for things: A study of human learning*. Cambridge, MA: The MIT Press.

Magidson, S. (2005). Building bridges within mathematics education: Teaching, research, and instructional design. *The Journal of Mathematical Behavior, 24*(2), 135-169. doi: DOI: 10.1016/j.jmathb.2005.03.004

Marshall, J. D., Sears, J. T., Allen, L. A., Roberts, P. A., & Schubert, W. H. (Eds.). (2006). *Turning points in curriculum: A contemporary American memoir* (2nd ed.). Upper Saddle River, NJ: Pearson Education, Inc.

Marshall, S. P. (1990). The assessment of schema knowledge for arithmetic story problems: A cognitive science perspective. In G. Kulm (Ed.), *Assessing higher order thinking in mathematics* (pp. 155-168). Washington, DC: American Association for the Advancement of Science.

Martorella, P. H. (1972). *Concept learning: Designs for instruction*. Scranton, PA: Intext Educational Publishers.

Mathematical Sciences Education Board. (1993). *Measuring what counts*. Washington, D.C.: National Academy Press.

McKechnie, J. (Ed.). (1979). *Webster's new twentieth century dictionary of the English language* (2nd ed.). New York, NY: William Collins Publishers, Inc.

Messick, S. (1984). The psychology of educational measurement. *Journal of Educational Measurement, 21*(3), 215-237.

Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (3rd ed.) (pp. 13-103). New York, NY: Macmillan Publishing Co, Inc.

Messick, S. (1994a). The interplay of evidence and consequences in the validation of performance assessments. *Educational Researcher, 23*(2), 13-23.

Messick, S. (1994b). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. Research report RR-94-45. Princeton, NJ: Educational Testing Service.

Messick, S. (1995a). Standards of validity and the validity of standards in performance assessment. *Educational Measurement: Issues and Practice, 14*(4), 5-8. doi: 10.1111/j.1745-3992.1995.tb00881.x

Messick, S. (1995b). Validity of performance assessment. In G. W. Phillips (Ed.), *Technical issues in performance assessment* (pp. 1-18). Washington, D.C.: Dpeartment of Education.

Mislevy, R. J., Almond, R. G., & Lukas, J. F. (2003). A brief introduction to evidence-centered design. Los Angeles, CA: National Center for Research on Evaluation, Standards, and Student Testing (CRESST), Center for the Study of Evaluation (CSE).

Mislevy, R. J., & Haertel, G. D. (2006). Implications of evidence-centered design for educational testing. *Educational Measurement: Issues and Practice, 25*(4), 6-20.

Mislevy, R. J., Steinberg, L. S., & Almond, R. G. (1999). On the roles of task model variables in assessment design. Los Angeles, CA: National Center for Research on Evaluation, Standards, and Student Testing (CRESST), Center for the Study of Evaluation (CSE).

Mislevy, R. J., Steinberg, L. S., & Almond, R. G. (2003). On the structure of educational assessments. *Measurement: Interdisciplinary Research and Perspectives, 1*, 3-62.

Mislevy, R. J., Steinberg, L. S., Almond, R. G., Haertel, G. D., & Penuel, W. P. (2000). *Leverage points for improving educational assessment*. Paper presented at the The Effectiveness of Educational Technology: Research Designs for the Next Decade, SRI International, Menlo Park, CA.

Moritz, J. (2005). Reasoning about covariation. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 227-255). Netherlands: Springer.

Moschkovich, J. N. (1996). Moving up and getting steeper: Negotiating shared descriptions of linear graphs. *The Journal of the Learning Sciences, 5*(3), 239-277.

Moschkovich, J. N., Schoenfeld, A. H., & Arcavi, A. (1993). Aspects of understanding: On multiple representations of linear relations and connections among them. In T. A. Romberg, E. Fennema, & T. P. Carpenter (Eds.), *Integrating research on graphical representations of functions*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Moss, P. (1994). Can there be validity without reliability? *Educational Researcher, 23*(2), 5-12.

National Mathematics Advisory Panel. (2008). Summary of foundations for success*:* The final report. Retrieved 02/27/2011, from http://www.ride.ri.gov/instruction/curriculum/rhodeisland/resources/Math_Summit/National_Math_Panel_Summary_Final_Copy.pdf

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

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

NCTM. (1991). *Professional standards for teaching mathematics*. Reston, VA: Author.

NCTM. (1995). *Assessment standards for school mathematics*. Reston, VA: Author.

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

NCTM. (2006). *Curriculum focal points for prekindergarten through grade 8 mathematics: A quest for coherence.* Reston, VA: Author.

NCTM. (2007). *Mathematics teaching today*. Reston, VA: Author.

NCTM. (2009). *Focus in high school mathematics: Reasoning and sense making*. Reston, VA: Author.

Neill, M. (2003). Leaving children behind: How no child left behind will fail our children. *Phi Delta Kappan, 85*(3), 225-228.

Niemi, D. (1996). Assessing conceptual understanding in mathematics: Representations, problem solutions. *Journal of Educational Research, 89*(6), 351.

Noddings, N. (1990). Constructivism in mathematics education. In R. B. Davis, C. A. Maher, & N. Noddings (Eds.), *Constructivst views on the teaching and learning of mathematics* (pp. 7-18). Reston, VA: National Council of Teachers of Mathematics, Inc.

Noelting, G. (1980a). The development of proportional reasoning and the ratio concept part I: Differentiation of stages. *Educational Studies in Mathematics, 11*(2), 217-253.

Noelting, G. (1980b). The development of proportional reasoning and the ratio concept part II: Problem-structure at successive stages; Problem-solving strategies and the mechanism of adaptive restructuring. *Educational Studies in Mathematics, 11*(3), 331-363.

Norman, F. A. (1993). Integrating research on teachers' knowledge of functions and their graphs. In T. A. Romberg, E. Fennema, & T. P. Carpenter (Eds.), *Integrating research on the graphical representations of functions*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Norris, S. P., Macnab, J. S., & Phillips, L. M. (2007). Cognitive modeling of performance on diagnostic achievement tests: A philosophical analysis and justification. In J. P. Leighton & M. J. Gierl (Eds.), *Cognitive diagnostic assessment for education* (pp. 61-84). New York, NY: Cambridge University Press.

Novak, J. D. (1998). *Learning, creating, and using knowledge: Concept maps as facilitative tools in schools and corporations*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Orton, A. (1984). Understanding rate of change. *Mathematics in School, 13*(5), 4.

Pellegrino, J. W. (1988). Mental models and mental tests. In H. Wainer & H. I. Braun (Eds.), *Test validity* (pp. 49-59). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Pellegrino, J. W., Baxter, G. P., & Glaser, R. (1999). Addressing the “two disciplines” problem: Linking theories of cognition and learning with assessment and instructional practice. *Review of Research in Education, 24*(1), 307-353. doi: 10.3102/0091732x024001307

Pellegrino, J. W., & Huff, K. (2010). *Using evidence-centered assessment design to bridge the gap between summative assessments and classroom practice.* Paper presented at the National Council on Measurement in Education, Denver, CO.

Popham, W. J. (2008). *Transformative assessment*. Alexandria, VA: Association for Supervision and Curriculum Development.

Popham, W. J. (2011). *Transformative assessment in action: An inside look at applying the process*. Alexandria, VA: ASCD.

Post, T., Behr, M., & Lesh, R. (1988). Proportionality and the development of prealgebra understandings. In A. F. Coxford & A. Shulte (Eds.), *The ideas of algebra, K - 12: 1988 yearbook* (pp. 78-90). Reston, VA: NCTM.

Pulos, S., Karplus, R., & Stage, E. K. (1981). Generality of proportional reasoning in early adolescence: Content effects and individual differences. *The Journal of Early Adolescence, 1*(3), 257-264.

Reiken, J. J. (2008). *Coming to understand slope and the Cartesian connection: An investigation of student thinking.* (Doctoral Dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3347606)

Romberg, T. A., Carpenter, T. P., & Kwako, J. (2005). Standards-based reform and teaching for understanding. In T. A. Romberg, T. P. Carpenter, & F. Dremock (Eds.), *Understanding math and science matters* (pp. 3-26). Mahwah, NJ: Lawrence Erlbaum Associates.

Royer, J. M., Cisero, C. A., & Carlo, M. S. (1993). Techniques and procedures for assessing cognitive skills. *Review of Educational Research, 63*(2), 201-243.

Seymour, J. R., & Lehrer, R. (2006). Tracing the evolution of pedagogical content knowledge as the development of interanimated discourses. *Journal of the Learning Sciences, 15*(4), 549-582. doi: 10.1207/s15327809jls1504_5

Shaughnessy, J. M., Chance, B., & Kranendonk, H. (2009). *Focus in high school mathematics: Reasoning and sense making in statistics and probability*. Reston, VA: NCTM.

Skemp, R. R. (2006). Relational understanding and instrumental understanding. *Mathematics Teaching in the Middle School, 12*(2), 88-95.

Smith, E. (2003). Statis and change: Integrating patterns, functions, and algebra throughout the K-12 curriculum. In J. Kilpatrick, W. G. Martin, & D. Schifter (Eds.), *A research companion to principles and standards for school mathematics*. Reston, VA: NCTM.

Steffe, L. P. (1990). On the knowledge of mathematics teachers. In R. B. Davis, C. A. Maher, & N. Noddings (Eds.), *Constructivist views on the teaching and learning of mathematics* (pp. 167-184). Reston, VA: National Council of Teachers of Mathematics, Inc.

Stewart, J. (2003). *Single variable calculus early transcendentals* (5th ed.). Belmont, CA: Brooks/Cole - Thompson Learning.

Stump, S. L. (1999). Secondary mathematics teachers' knowledge of slope. *Mathematics Education Research Journal, 11*(2), 124-244.

Stump, S. L. (2001). High school precalculus students' understanding of slope as measure. *School Science & Mathematics, 101*(2), 9.

Suurtamm, C., Lawson, A., & Koch, M. (2008). The challenge of maintaining the integrity of reform mathematics in large-scale assessment. *Studies in Educational Evaluation, 34*(1), 31-43.

Tall, D., & Vinner, S. (1981). Concept image and concept definition in mathematics with particular reference to limits and continuity. *Educational Studies in Mathematics, 12*(2), 151-169.

Teuscher, D., & Reys, R. E. (2010). Slope, rate of change, and steepness: Do students understand these concepts? *Mathematics Teacher, 103*(7), 6.

Thompson, P. W. (1994). The development of the concept of speed and its relationship to concepts of rate. In G. Harel & J. Confrey (Eds.), *The development of multiplicative reasoning in the learning of mathematics*. Albany, NY: State University of New York Press.

Tourniaire, F., & Pulos, S. (1985). Proportional reasoning: A review of the literature. *Educational Studies in Mathematics, 16*(2), 181-204.

Vergnaud, G. (1988). Multiplicative structures. In J. Hiebert & M. Behr (Eds.), *Number concepts and operations in the middle grades* (pp. 141-161). Reston, VA: NCTM.

Vergnaud, G. (1994). Multiplicative conceptual field: What and why? In G. Harel & J. Confrey (Eds.), *The development of multiplicative reasoning in the learning of mathematics*. Albany, NY: State University of New York Press.

Vergnaud, G. (1997). The nature of mathematical concepts. In T. Nunes & P. Bryant (Eds.), *Learning and teaching mathematics: An international perspective* (pp. 5-28). East Sussex,UK: Psychology Press, Ltd.

von Glasersfeld, E. (1990). An exposition of constructivism: Why some like it radical. In R. B. Davis, C. A. Maher, & N. Noddings (Eds.), *Constructivist views on the teaching and* *learning of mathematics* (pp. 19-30). Reston, VA: National Council of Teachers of Mathematics.

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

Walter, J. G., & Gerson, H. (2007). Teachers' personal agency: Making sense of slope through additive structures. *Educational Studies in Mathematics, 65*, 203-233.

Wavering, M. (1989). Logical reasoning necessary to make line graphs. *Journal of Research in Science Teaching, 26*(5), 373-379.

Webb, N. L. (1992). Assessment of students' knowledge of mathematics: Steps toward a theory. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 661-683). New York, NY: Macmillan Publishing Co, Inc.

Webb, N. L., & Romberg, T. A. (1992). Implications of the NCTM standards for mathematics assessment. In T. A. Romberg (Ed.), *Mathematics assessment and evaluation: Imperatives for mathematics educators* (pp. 37-60). Albany, NY: State University of New York Press.

Wilhelm, J. A., & Confrey, J. (2003). Projecting rate of change in the context of motion onto the context of money. *International Journal of Mathematical Education in Science & Technology, 34*(6), 887-904.

Willis, G., Schubert, W. H., Bullough, R. V., Kridel, C., & Holton, J. T. (Eds.). (1994). *The American curriculum: A documentary history*. Westport, CT: Praeger Publishers.

Wilson, M. (1992). Measuring levels of mathematical understanding. In T. A. Romberg (Ed.), *Mathematics assessment and evaluation* (pp. 213-241). Albany, NY: State University of New York Press.

Wilson, M. (2009). Measuring progressions: Assessment structures underlying a learning progression. *Journal of Research in Science Teaching, 46*(6), 716-730.

Zimowski, M. F., Muraki, E., Mislevy, R. J., & Bock, R. D. (2003). Multiple-group IRT analysis and test maintenance for binary items [Computer software]. Lincolnwood, IL: Scientific Software International, Inc.

APPENDIX A

FOUNDATIONAL CONCEPTS OF SLOPE ASSESSMENT

Note: This test was administered using the Kansas Computerized Assessment software, which is an online environment. During the assessment, each item and its four responses were displayed together in one screen. There was no scrolling required to view each entire test item. However, when the four responses were each graphs, students had the option of enlarging the graphs using a button with the label "click to enlarge." This button activated a pop-up window containing all four graphs. This window automatically closed when the student navigated to a different item on the FCSA. Students selected each response choice by clicking on a radio button beside their choice.

1. Daniel drives his car to and from work at an average speed of 45 miles per hour. Which aspect of this travel is related to how much time he will spend driving to and from work?

   A) the time of day Daniel drives
   B) the direction of Daniel's travel
   C) the price of gas for Daniel's car
   D) X the distance Daniel drives to work

2. Jill deposits the same amount of money into her savings account every time she goes to the bank. She does not withdraw any money. Which fact about Jill's trips to the bank is related to the total amount of money she has in her account?

   A) the time of day
   B) the day of the week
   C) X the number of deposits
   D) the distance to the bank

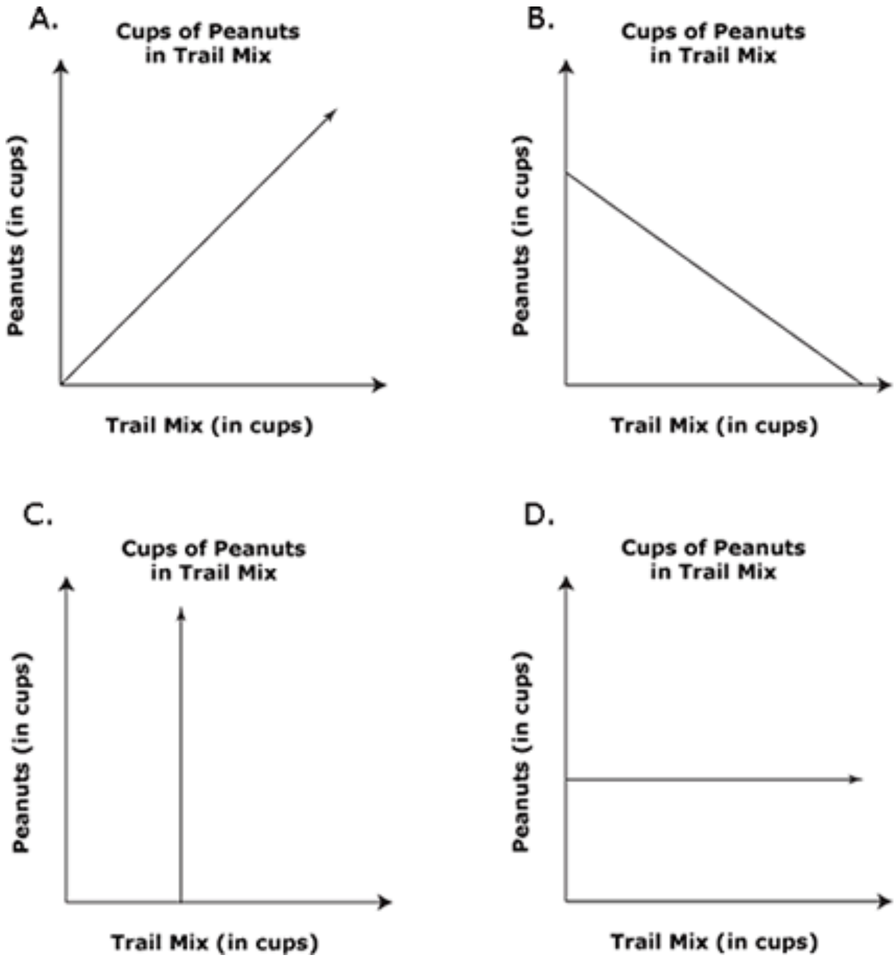
3. A school district is purchasing new math textbooks. Each student will receive one math textbook. Which facts are related to the total cost of the textbooks?

   A) X the number of students and the price per book
   B) the number of teachers and the price per book
   C) the number of students and the number of books
   D) the number of teachers and the number of books

4. Susan is making chocolate chip cookies. Each cookie has the same number of chips. Which facts are related to the total number of chips Susan needs?

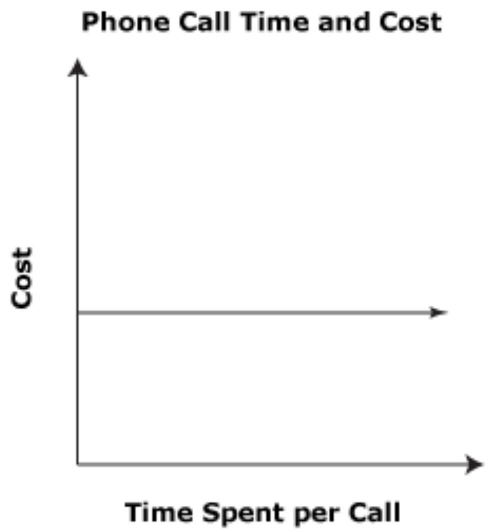
   A) the cost of the chips and the size of the chips
   B) the number of cookies and the size of the chips
   C) the cost of the chips and the number of chips per cookie
   D) X the number of cookies and the number of chips per cookie

5. Greg is making trail mix. The recipe requires peanuts, chocolate candies, pretzels, and cereal. The more trail mix Greg makes, the more peanuts he will need. Which graph shows the correct relationship between the amounts of peanuts and trail mix?



A) X	A	B)	B
C)	C	D)	D

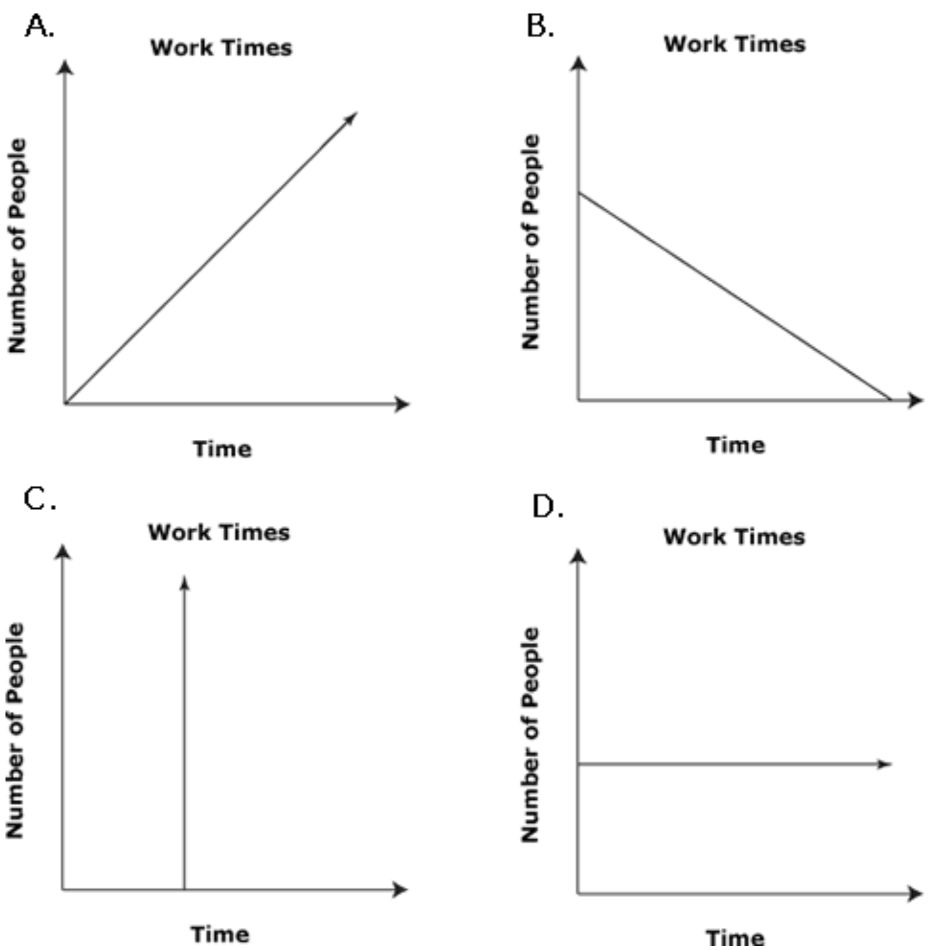
6. A graph about phone call times and costs is shown below.



Which statement could describe this graph?

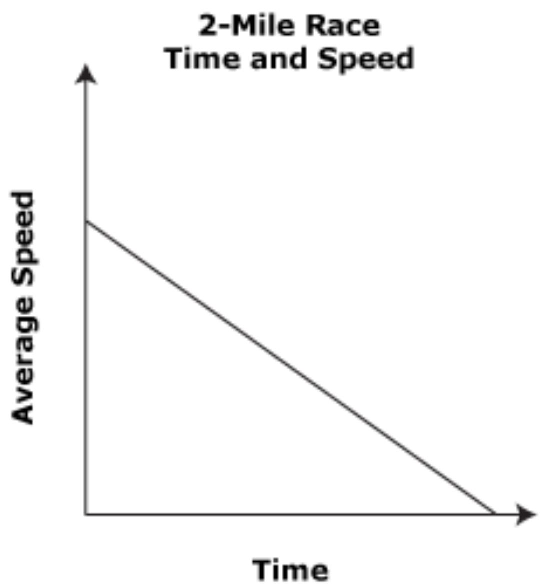
A) More phone calls cost less per call.
B) More phone calls cost more per call.
C) Long phone calls cost more than short phone calls.
D) X Long phone calls cost the same as short phone calls.

7. In many workplaces people work in teams to get big jobs done. When fewer people work, the job takes more time. Which graph shows this relationship between the time and the number of people needed to finish a job?



A) A
B) X B
C) C
D) D

8. The graph below shows the speeds and times of students who ran a 2-mile race.

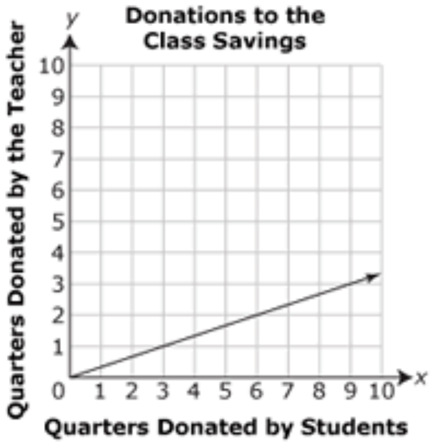


Based on the graph, which statement must be true?

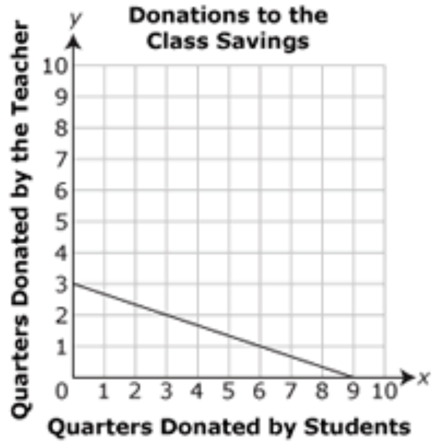
A) A student who runs faster uses more time.
B) X A student who runs slower uses more time.
C) A student who uses more time runs farther.
D) A student who uses less time runs farther.

9. A class is saving for a trip at the end of the year. The teacher puts three quarters in a jar every time a student puts one quarter in the jar. Which graph shows the relationship between the amounts the teacher and the students put in the jar?

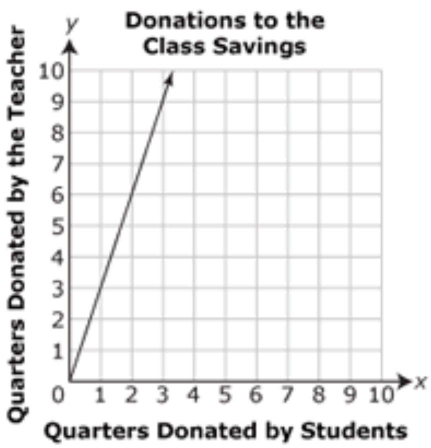
A.



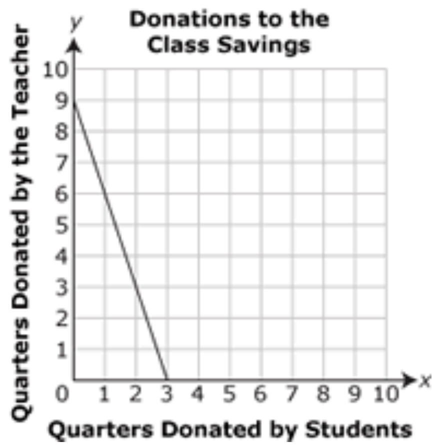
B.



C.

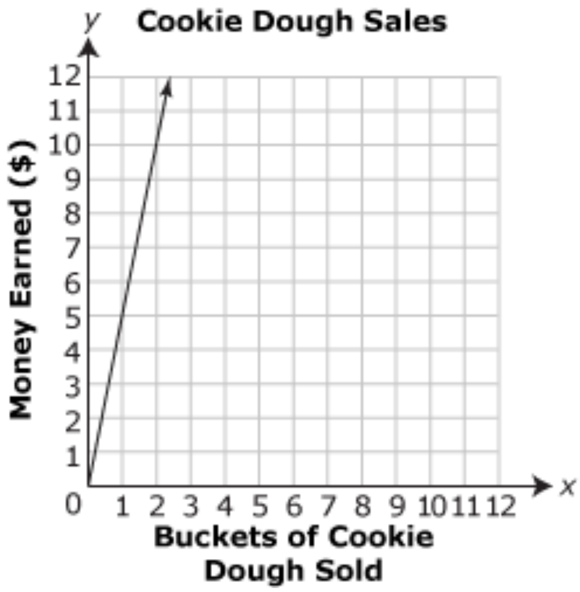


D.



A) A
B) B
C) X C
D) D

10. The graph below shows the amount of money, in dollars, a class could raise by selling cookie dough.

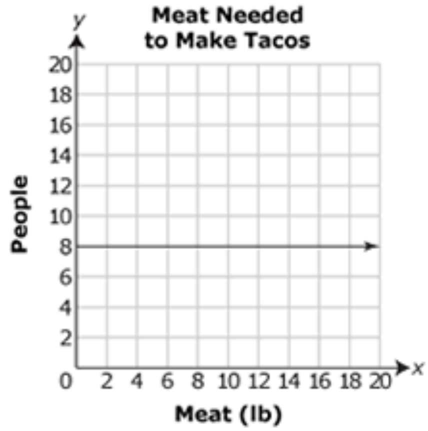


Based on this graph, which statement must be true?

A) For every 1 bucket sold, the class earns \$1.
B) For every 5 buckets sold, the class earns \$1.
C) The class earns \$1 per bucket of cookie dough.
D) X The class earns \$5 per bucket of cookie dough.

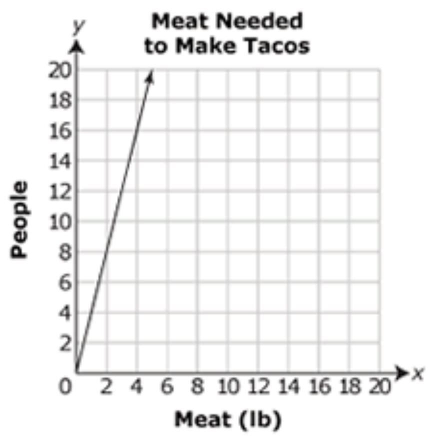
11. Rachel makes tacos for 8 people and uses 2 pounds (lb) of meat. Which graph shows the same relationship between the number of people and the amount of meat?

A.

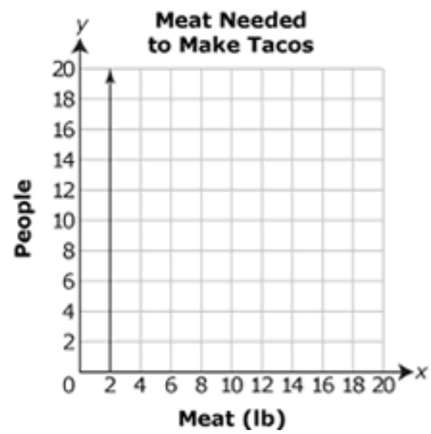


B.

C.



D.



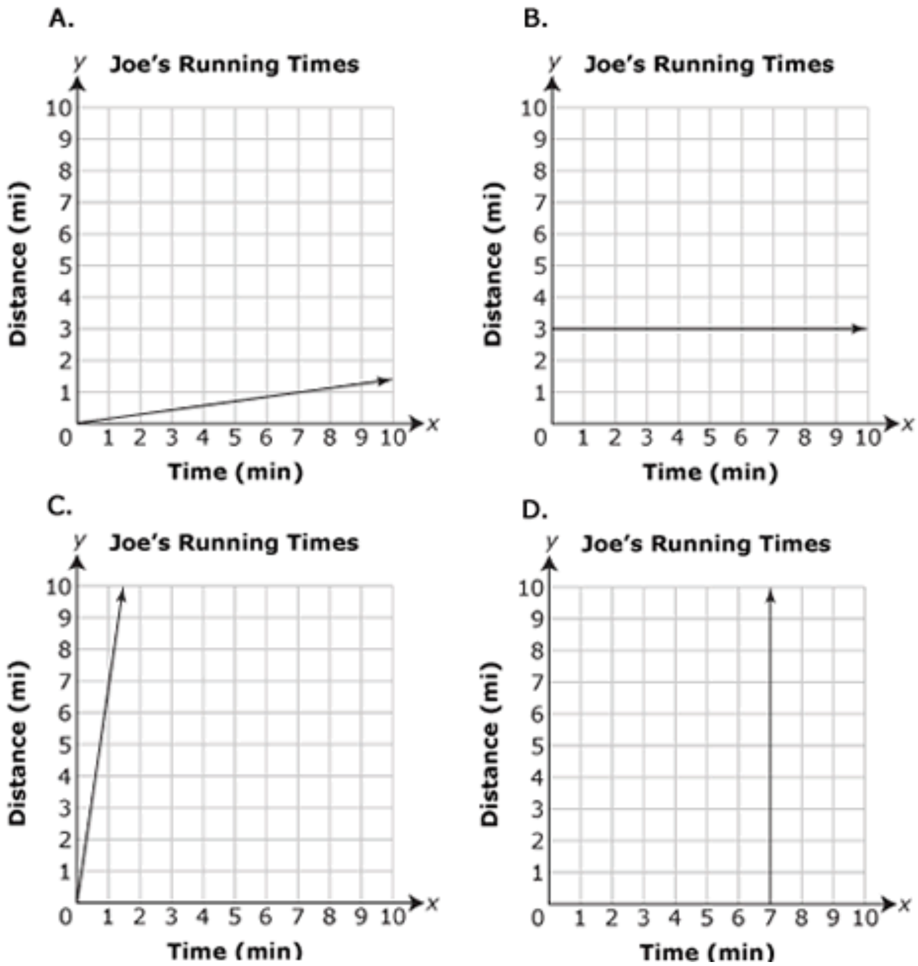
A) A B) B
C) X C D) D

12. The graph below shows the amount of water, in ounces (oz), and the number of scoops of drink mix powder needed to make a sports drink.

Based on the graph, which statement about making this sports drink must be true?

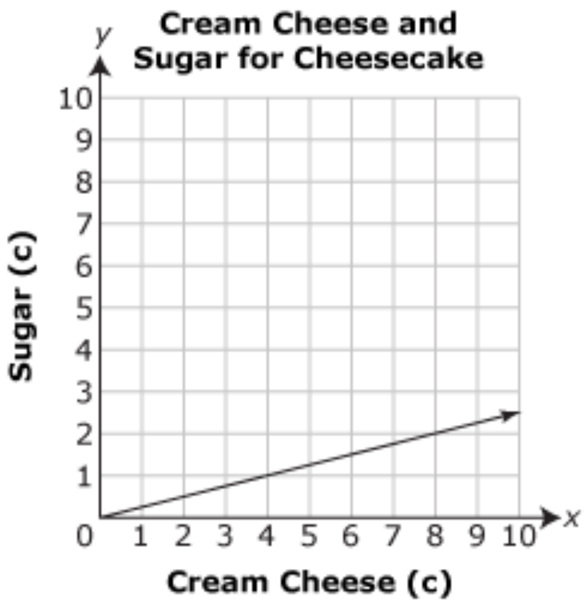
A) Every 2 oz of water are mixed with 1 scoop of powder.
B) X Every 8 oz of water are mixed with 1 scoop of powder.
C) Every 2 scoops of powder are mixed with 1 oz of water.
D) Every 8 scoops of powder are mixed with 1 oz of water.

13. Joe runs at an average speed of 3 miles in 21 minutes. Which graph shows the same running speed?



A) X A
B) B
C) C
D) D

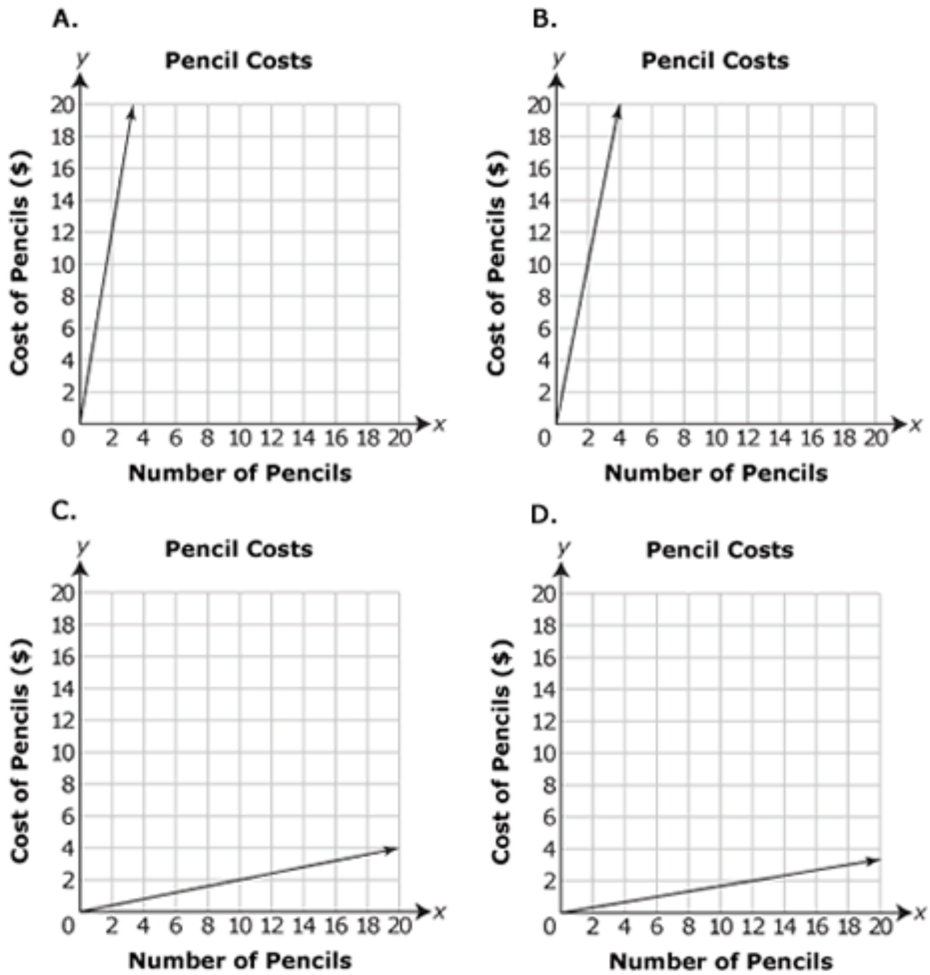
14. The graph below shows the amounts of sugar and cream cheese, in cups (c), needed for a cheesecake recipe.



Which statement is true about this recipe?

A) X For each cup of sugar there are four cups of cream cheese.
B) For each cup of sugar there is one-fourth cup of cream cheese.
C) For every three cups of cream cheese, there is one cup of sugar.
D) For every three cups of cream cheese, there are four cups of sugar.

15. Billy spends one dollar for six pencils. Which graph could be used to find the cost of any number of pencils sold for the same price?



A) A	B) B
C) C	D) X D

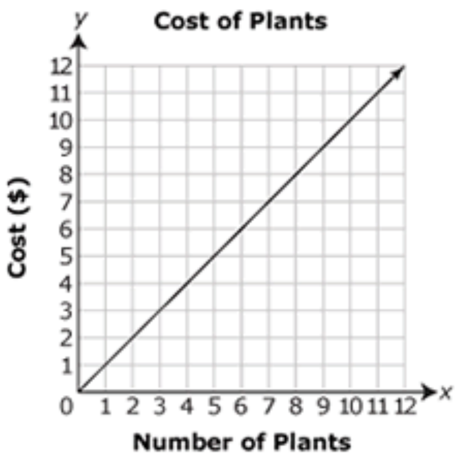
16. The graph below shows the price, in dollars ($), of candy sold at a candy store.

Based on the graph, which statement about the price of this candy is true?

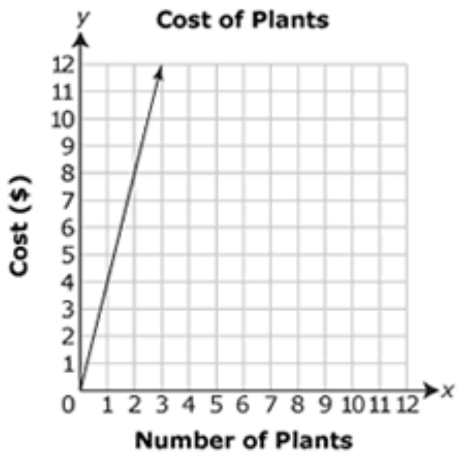
A) The price of 24 pieces of candy is $6.
B) The price of 30 pieces of candy is $5.
C) X The price of 1 piece of candy is $0.20.
D) The price of 1 piece of candy is $0.50.

17. Lisa spent $6 for 10 plants. Which graph shows the cost of these plants?

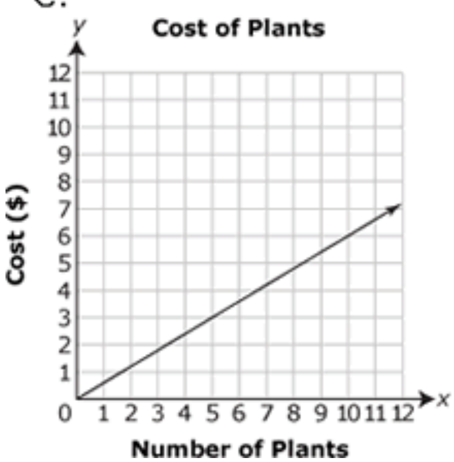
A.



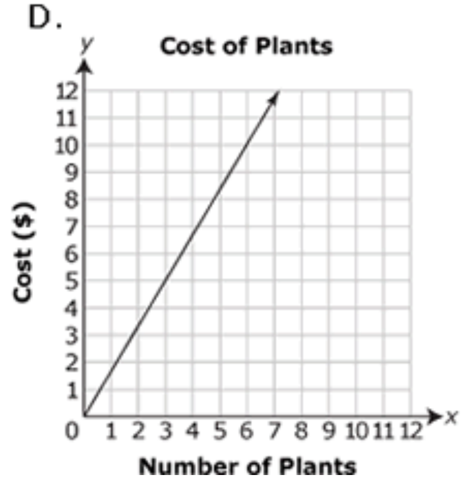
B.



C.

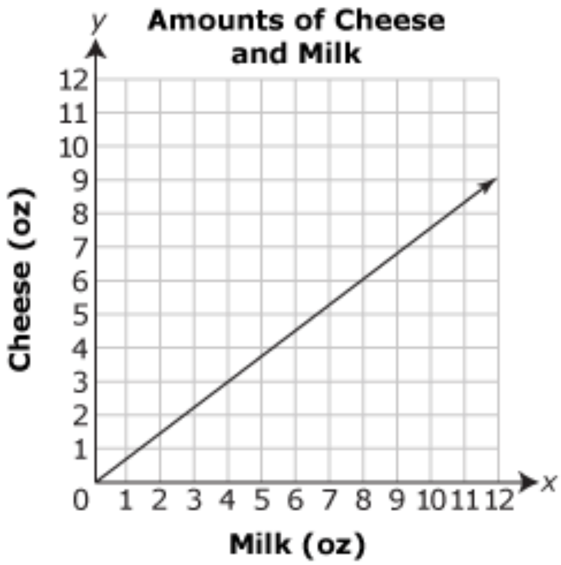


D.



A) A
B) B
C) X C
D) D

18. The graph below shows the amounts, in ounces (oz), of cheese and milk needed to make macaroni and cheese.

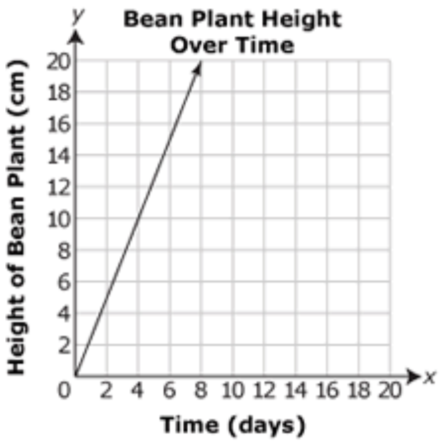


Based on the graph, which statement about the amounts of cheese and milk in the macaroni and cheese is true?

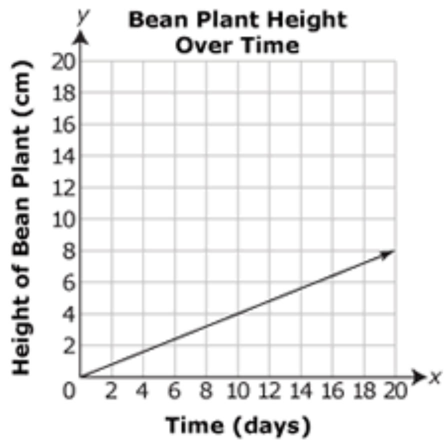
A) X For every 8 oz of milk, there are 6 oz of cheese.
B) For every 8 oz of cheese, there are 6 oz of milk.
C) For every 8 oz of milk, there are 12 oz of cheese.
D) For every 8 oz of cheese, there are 12 oz of milk.

19. A bean plant grows at an average rate of 2.5 centimeters (cm) per day. Which graph shows the growth of this bean plant?

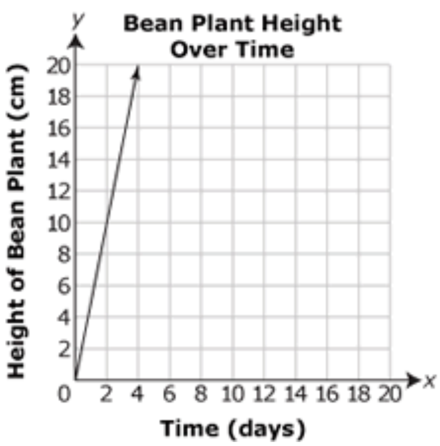
A.



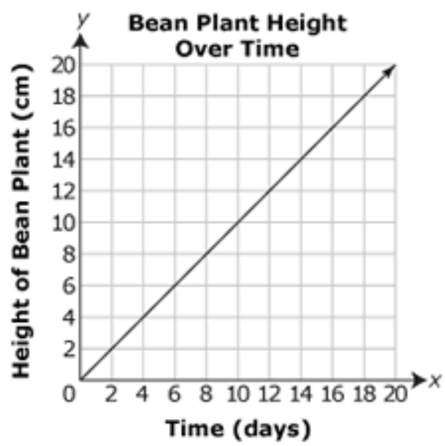
B.



C.

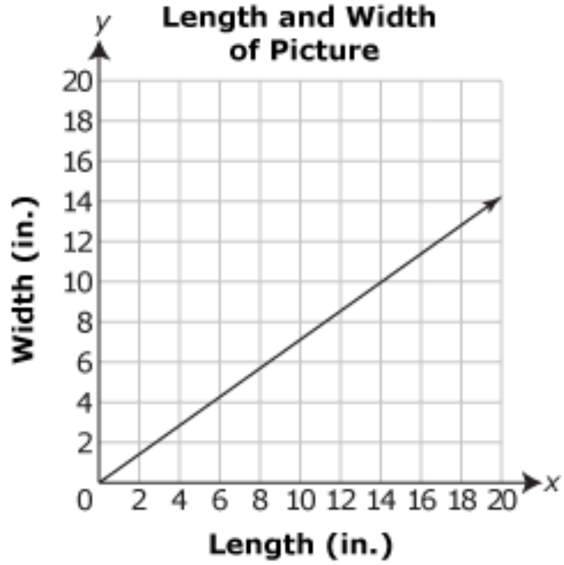


D.



A) X A
B) B
C) C
D) D

20. The graph below shows how the length and width, in inches (in.), of a picture both increase as the picture is enlarged.



Based on the graph, which statement describes the length and width of these pictures?

A) A picture that is 10 in. by 14 in. can be reduced to be 3 in. by 4 in.
B) A picture that is 5 in. by 7 in. can be enlarged to be 10 in. by 12 in.
C) The length of the picture is always as long as the width of the picture.
D) X The width of the picture is always as long as the length of the picture.

APPENDIX B

ANNOUNCEMENT TO RECRUIT TEACHER PARTICIPANTS

What do Your Students Know about Slope?
What: Slope Assessment - A Formative Assessment Targeting the Slope Concept
When: May 2011
Who: Middle and High School Mathematics Teachers and Students
How: CETE TestBuilder Website and KCA Software
Why: This assessment measures student understanding of slope in terms of essential components of prerequisite knowledge. Your participation is valued for moving Kansas assessments forward in the direction of using what we know about how people learn mathematics to influence the nature of formative assessments.

Invitation and Description of Participation Requirements

You are invited to participate in a unique study. The outcome of this study will be to describe what students understand about slope as well as the common misconceptions about slope. The data gathered from this study will also be used to inform the development of future formative assessment tools.

The Slope Assessment measures what students understand about the slope concept and is appropriate for students studying Pre-algebra, Algebra 1, Geometry, and Algebra II, or courses with similar content that are taken prior to Pre-calculus. The assessment contains 20 multiple choice math items and should not take more than 30 minutes for most students to complete.

Teachers are requested to have their students complete the Slope Assessment at their earliest convenience. Teachers are also encouraged to review test items and student responses in the same way they review responses to any other formative assessment delivered via the KCA software.

Every teacher who participates is asked to complete a survey to collect only the teacher's name, contact information, and the building name or number. Click here to access the survey. Furthermore, they must include the **course name** of the students in the **Administration Name** field when selecting the Slope Assessment in the TestBuilder application. This information will be used by CETE to extract from the database the appropriate student responses for this study and categorize them into the correct course levels. CETE will replace individual student identifier information with random numbers prior to delivering the data to the researcher. Thus the researcher will not have any means by which to identify particular students or their teachers from the data.

The data collected and analyzed for this study will be restricted to student responses to the items, i.e. the response choices students made for each of the items on the Slope Assessment, the course name, and the attendance building of each student. Data analyses will be conducted to describe student knowledge of the slope concept. No comparisons will be made between different groups of students, such as those from different school buildings.

For more information contact Angela Broaddus at broaddus@ku.edu or 785-864-2916.

APPENDIX C

INSTRUCTIONS FOR ACCESSING THE SLOPE ASSESSMENT

The Slope Assessment is in the default pool of formative assessments for grades 7, 8, and 10. If your district has its own specialized pool of formative assessments, then we recommend that an administrator place the slope assessment in the district pool.

Once the slope assessment is able to be viewed by teachers, they need only "browse tests," add the Slope Assessment to "my tests," and create a new administration. It is very important for teachers to use the appropriate course name (i.e., pre-algebra, geometry, etc.) in the administration name field and also to complete the teacher survey after administering the assessment. Click here to access the survey.

This assessment will be available through the first week of June.

APPENDIX D

TEACHER INFORMATION SURVEY

THE UNIVERSITY OF KANSAS

Thank you for participating in the Slope Assessment. Please answer the questions that follow to help describe the testing experience of the students who took the Slope Assessment Instrument (SAI).

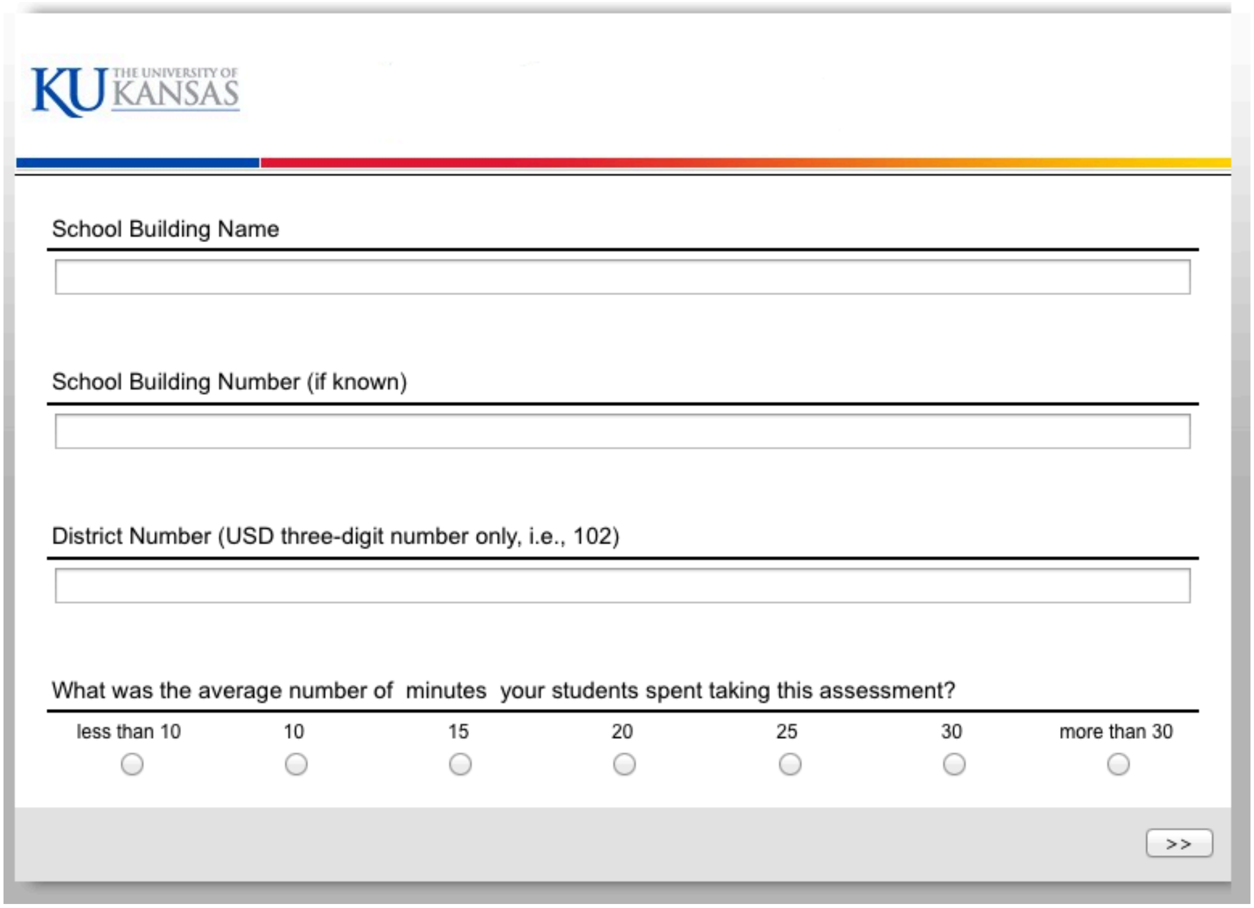
Teacher First Name

Teacher Last Name

Formative TestBuilder Username

Teacher Email
(This will only be used if needed to confirm the correct identification of student responses for the purposes of this study.)

>>



KU THE UNIVERSITY OF KANSAS

School Building Name

School Building Number (if known)

District Number (USD three-digit number only, i.e., 102)

What was the average number of minutes your students spent taking this assessment?

less than 10	10	15	20	25	30	more than 30
○	○	○	○	○	○	○

>>