Developing Instructional Tools to Assist Teachers in Implementing the Common Core State Standards for Mathematics

Angela Broaddus, Ph.D.
Formative and Interim Assessment Programs Coordinator
Center for Educational Testing and Evaluation
1122 West Campus Road
JR Pearson Hall, Room 748
Lawrence, KS 66045
785-864-2916
broaddus@ku.edu

Author note

Angela Broaddus, Center for Educational Testing and Evaluation, University of Kansas.

The author received no financial support for the research or authorship of this article.

Correspondence concerning this article should be addressed to Angela Broaddus, Center for Educational Testing and Evaluation, University of Kansas, 1122 West Campus Road, JR Pearson Hall, Room 748, Lawrence, KS 66045; E-mail: broaddus@ku.edu.

Abstract

Assessments that diagnose student understandings and misconceptions in terms of research-based learning progressions are useful for guiding instructional decisions. This study describes the analysis of a learning progression concerning the slope concept and the development of an instructional tool designed to identify student knowledge levels based on test responses.

Developing Instructional Tools to Assist Teachers in Implementing the Common Core State Standards for Mathematics

Mathematics educators and researchers have worked for decades to improve how American students are taught mathematics. While educators have implemented standards-based instruction and made decisions about student achievement using available data, researchers have continued to explore how to define “mathematical proficiency” and investigated how it is achieved. In 2010, these efforts culminated in the release of the Common Core State Standards for Mathematics (CCSSM)(CCSSO/NGA, 2010). These standards foster high expectations so that students will graduate from high school prepared for college-level coursework or for a variety of careers.

In an effort to articulate the path toward mathematical proficiency, the CCSSM were developed to reflect research-based learning progressions based on current theories of how students acquire mathematics knowledge and skills (CCSSO/NGA, 2010). Teachers who gain familiarity with the learning progressions underlying the CCSSM will improve their ability to provide effective instruction. Additionally, teachers will benefit from access to sophisticated assessment tools that diagnose student understandings and misconceptions in an effort to guide instructional decisions. This study’s purpose is to provide an example of how to analyze the CCSSM and other relevant literature to uncover a learning progression and to develop an instructionally useful assessment tool deliberately designed to be sensitive to the components of knowledge included within that learning progression.

Mathematical Proficiency

Understanding mathematics is complex and develops over years of education and experience (National Research Council [NRC], 2001). People who understand mathematics consciously organize their knowledge to reflect meaningful connections among mathematical ideas, their representations, and the procedures for working with them (Carpenter, 1986; Hiebert & Lefevre, 1986). Mathematical proficiency consists of five strands that support a person's mathematical thinking (NRC, 2001): conceptual understanding, procedural fluency, strategic competence, adaptive reasoning, and productive disposition (NRC, 2001). These strands both represent the content students should learn and suggest behaviors that students should acquire in order to achieve proficiency. The CCSSM underscore the same aspects by equally emphasizing conceptual understanding and procedural fluency and by introducing standards for mathematical practice, which address strategic thinking, reasoning, and disposition (CCSSO/NGA, 2010). These standards articulate the learning targets that comprise the goal of helping all students become mathematically proficient over the course of their K-12 schooling.

Pressure to meet federal accountability requirements has prompted states to implement standardized high-stakes assessments to measure student achievement (Marshall, Sears, Allen, Roberts, & Schubert, 2006). In many cases, such assessments target small sets of skills and procedures rather than what students understand about the content area as a whole (Neill, 2003). Unfortunately, these assessments often influence instruction (Harlen, 2007; Suurtamm, Lawson, & Koch, 2008). In particular, when teachers are focused on students' mastery of assessed material, the classroom instruction often concentrates on performance of assessed skills and routines rather than on promoting understanding of the underlying concepts (Harlen, 2007). The teaching of mathematics is no exception. Teachers are obligated to ensure that their students can carry out specific routines, whether or not these students develop a true understanding of why the routines work and how they arise from related mathematical concepts (Battista, 1999). The attention devoted to preparing for state assessments may undermine classroom teachers' success in promoting rich experiences for students, thereby placing at risk the potential for students to construct a deep understanding of the mathematics they study (Battista, 1999; Haertel, 1985; Hancock & Kilpatrick, 1993) and jeopardizing their potential to develop mathematical proficiency (Battista, 1999).

CCSSM implementation presents educators with the opportunity to change both how students learn mathematics and how their achievement is measured by emphasizing the "Standards for Mathematical Practice" (CCSSO/NGA, 2010, p. 6). Mathematics educators recently urged test developers to embed these practices within new assessments (Krupa, 2011). Previous recommendations have also noted that assessing student understanding of mathematics requires alternative measurement tools (Webb & Romberg, 1992). These recommendations should guide the development of tests used for accountability and the design of instructionally embedded tools such as formative assessments (Krupa, 2011).

Learning Progressions in the CCSSM

For students to acquire a deep understanding of mathematics, their educational experiences should relate their cognitive development, their prior knowledge, and the relevant learning targets. To maintain sensitivity to how students learn, developers of the CCSSM began with "research-based learning progressions detailing what is known today about how students' mathematical knowledge, skill, and understanding develop over time" (CCSSO/NGA, 2010, p. 4). Learning progressions describe the knowledge that students should acquire as they gain expertise in a particular domain (Consortium for Policy Research in Education, 2011).

Specifically, learning progressions contain descriptions of the concepts and skills that make up larger learning targets (Popham, 2008) and map out optimal learning sequences (Wilson, 2009) to delineate how knowledge matures from novice level to expert level (Popham, 2008, 2011; Wilson, 2009).

The learning progressions in the CCSSM extend across grade levels and describe knowledge that increases in sophistication from grade to grade. Becoming familiar with the standards for the surrounding grades will allow teachers to meet the needs of students whose knowledge levels are at different points along the learning progressions. Teachers' abilities to assess student knowledge levels, gaps, and misconceptions will depend on their access to instructional tools that gauge student mastery relative to learning progressions. This study describes the analysis of one learning progression as well as the development of an instructional tool designed to identify student knowledge levels based on test responses.

Method

This study focused on the foundational concepts related to understanding slope. The study included the development of an attribute hierarchy model containing five foundational concepts related to slope, and this model served as a theory for how understanding of slope develops. This model was called the Foundational Concepts of Slope Attribute Hierarchy (FCSAH), and it informed the design of an assessment named the Foundational Concepts of Slope Assessment (FCSA). Each item on the FCSA was created to evaluate one precise combination of attributes in the FCSAH. The FCSA was administered online to students in middle and high school grades.

Development of the FCSAH

The FCSAH developed for this study incorporated the work of scholars who have studied covariation and proportional reasoning. The FCSAH is displayed in Figure 1. The researcher scrutinized the literature as well as previous investigations into ideas related to slope in order to identify foundational concepts that contribute to the understanding of slope. Using the literature and research findings, the researcher arranged five attributes within the FCSAH to depict the order in which the attributes might optimally be acquired. The levels of knowledge represented by the three vertical levels of the FCSAH correspond to a learning progression in the CCSSM for the middle-grade mathematics domains of "Ratios and Proportional Relationships" and "Expressions and Equations." Figure 2 displays the learning progression and the CCSSM standards that describe the knowledge delineated in the FCSAH.

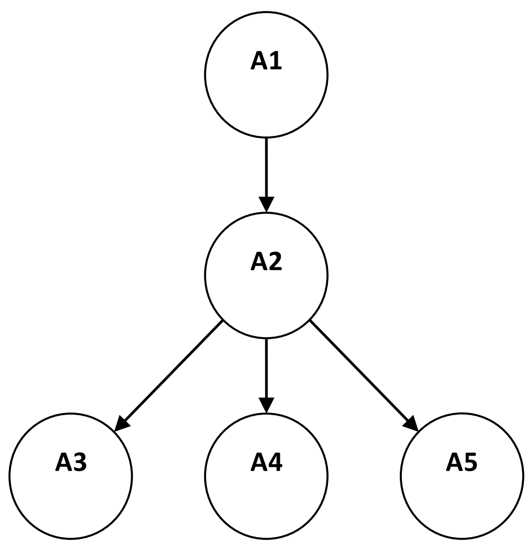

Figure 1. The Foundational Concepts of Slope Attribute Hierarchy (FCSAH) depicts five attributes associated with understanding the concept of slope. Attribute A1 is defined as the ability to identify covariates from a problem scenario. Attribute A2 is defined as the ability to identify covariates and 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.

Drawing on the work of Adamson (2005), the first two attributes of the FCSAH consisted of a student's ability to perceive covariation in different problem contexts. The first attribute concerned the ability to detect which quantities in a problem context 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, Jacobs, Coe, Larsen, and Hsu (2002). The first and second attributes in the FCSAH were positioned in the model as prerequisites for proportional reasoning.

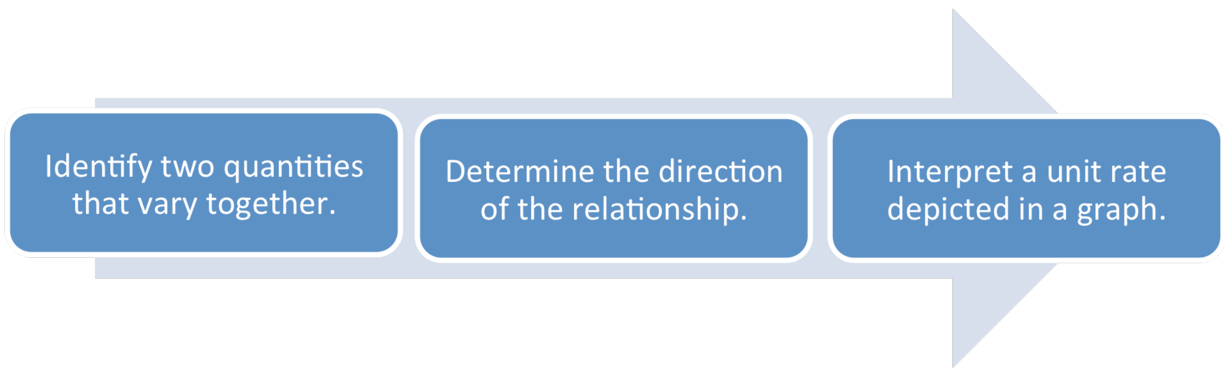

Figure 2. Learning progression with ties to the CCSSM (CCSSO/NGA, 2010).
CC.M.6.RP.1: Understand the concept of a ratio and use ratio language to describe a ratio relationship between two quantities.
CC.M.6.RP.2: Understand the concept of a unit rate a/b associated with a ratio $a:b$ with $b \neq 0$, and use rate language in the context of a ratio relationship.
CC.M.6.RP.3: Use ratio and rate reasoning to solve real-world and mathematical problems.
CC.M.7.RP.2: Recognize and represent proportional relationships between quantities.
CC.M.7.RP.2b: Identify the constant of proportionality (unit rate) in tables, graphs, equations, diagrams, and verbal descriptions of proportional relationships.
CC.M.7.RP.2d: Explain what a point (x, y) on the graph of a proportional relationship means in terms of the situation, with special attention to the points (0, 0) and $(1, r)$ where r is the unit rate.
CC.M.8.EE.5: Graph proportional relationships, interpreting the unit rate as the slope of the graph.
RP = Ratios and Proportional Relationships; EE = Expressions and Equations.

Reflecting 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 the context of a problem presented either verbally or graphically. Three different types of ratios were modeled in the FCSAH. This reflected the work of Noelting (1980a, 1980b) and Karplus, Pulos, and Stage (1983), who studied how students learn to interpret the meaning of quantities expressed as ratios. The third attribute in the FCSAH concerned a student's ability to interpret the meaning 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 (i.e., 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 that was 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 et al. (1983) 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, and the most difficulty with ratios whose fraction representations do not simplify to either of these forms. To incorporate different levels of 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.

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 the way that 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 the SMEs reviewed the FCSAH independently, the researcher met with the two experts together to refine and confirm the structure and contents of the FCSAH.

Development of the FCSA

The FCSAH influenced item development for 20 selected-response items, which evaluated both the targeted attributes in the FCSAH and their dependent combinations. The resulting FCSA contained four items for each of five combinations of attributes.

The FCSA was created to measure student understanding of selected foundational concepts related to slope in terms of the attributes described 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 that were 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 earlier findings (Teuscher & Reys, 2010) suggesting that students experience difficulty relating the slope of a line to its meaning within a problem context, particularly when problems are presented verbally or in graphs.

Each of the items developed for attribute A1 was presented in the form of a word problem, where students were required to select which quantities varied together within the problem context. Each of the items developed for attributes A2, A3, A4, and A5 was presented either verbally or graphically. Students were required to match either a graph to a given verbal problem or a verbal statement to a given graph.

The FCSA was shared with the two SMEs who assisted in the design of the FCSAH. They were asked to judge the alignment of the items to the attributes described in the FCSAH. After the SMEs reviewed the items independently, the researcher met with the two experts together to refine and confirm the alignment of the items to the attributes in the FCSAH.

Three individuals other than the SMEs were consulted to ensure the technical quality of the test items developed for the FCSA. Two of these individuals were from a large midwestern university with experience teaching college mathematics and teacher education courses, and the third individual was from a testing company with expertise in item and test development. These three individuals reviewed each item to confirm that the language used in each item was clear and contained grade-appropriate vocabulary. They also confirmed the accuracy of each item's mathematical content and reviewed the answer choices for each item to ensure that there was exactly one correct answer and that the distractors represented typical misconceptions held by students or common errors that students might make when responding to the items.

The FCSA was administered in May 2011, to 1,629 students in middle and high school grades who were studying Pre-algebra, Algebra 1, Geometry, Algebra 2, or any course with similar content taken before Pre-calculus. While the material being tested appears in the middle-grade standards of the CCSSM, a wide range of student knowledge of slope was needed for the planned analysis. Therefore, courses beyond those in which slope was directly taught were included. Individual student demographic data were not collected for students who took the FCSA. Rather, district demographic information was collected for every district represented in the student sample. Over 27 participating districts, the percent of students identifying as White averaged 86%, the percent of students identifying as Hispanic averaged 6%, the percent of students identifying as African American averaged 2%, and the percent of students identifying as Other averaged 6%. One limitation of this study stems from a lack of racial diversity in the student sample, which restricts the generalizability of the results to similar populations of students.

Data Sources

Student item-level responses were analyzed to determine whether they supported the design of the FCSAH. Item response theory (IRT) was used to analyze the item characteristics and to estimate student abilities. Ten knowledge states reflected the different combinations of attributes that were consistent with the FCSAH. Each knowledge state was associated with an expected set of responses that students should give depending on their level of knowledge about the foundational concepts related to slope. Observed student responses were compared with these expected responses to classify students into knowledge states. The ability estimates of the students assigned to each knowledge state were then analyzed to determine if they represented distinctly different levels of knowledge with regard to the slope concept.

Results

The results provided in this section illustrate that a theoretical learning progression can be tested and confirmed using student test response data. These results are part of a larger study that investigated the application of the Attribute Hierarchy Method (AHM) to actual student test responses to the FCSA. The AHM was implemented to provide information about the accuracy of the FCSAH as a description of how knowledge of slope is organized in a person's cognitive structure. The results are presented here to provide evidence in support of the validity of the learning progression identified in the CCSSM and shown in Figure 2.

The items on the FCSA were analyzed using IRT, which produced three parameter values for each item. BILOG-MG 3 (Zimowski, Muraki, Mislevy, & Bock, 2003) was used to calibrate the items and calculate ability estimates. The item parameters were used to determine ability estimates for the students who took the FCSA as well as the 10 expected response patterns used for classification.

Students' responses to test items depend on both the content of the test items and the nature of the students' knowledge about that content. An important assumption made for the purposes of this study was that if a student possessed the knowledge described by an attribute, then that student would correctly answer all four items targeting that attribute. This assumption was used to develop an expected response pattern for each collection of attributes that was consistent with the proposed cognitive model, that is, the FCSAH. Table 1 lists the expected response patterns along with their corresponding knowledge states and targeted attributes.

Table 1

Knowledge States, Attributes, Expected Responses, and Ability Estimates Consistent With the FCSAH

Knowledge State	Attributes	Expected Response Vector
A0	None	00000000000000000000
A1	A1	11110000000000000000
A12	A1, A2	11111111000000000000
A123	A1, A2, A3	11111111111100000000
A124	A1, A2, A4	11111111000011110000
A125	A1, A2, A5	11111111000000001111
A1234	A1, A2, A3, A4	11111111111111110000
A1235	A1, A2, A3, A5	11111111111100001111
A1245	A1, A2, A4, A5	11111111000011111111
A12345	A1, A2, A3, A4, A5	11111111111111111111

The AHM was implemented to compare observed response patterns to the expected response patterns and to classify each student into an appropriate knowledge state. Initial findings from the AHM analysis suggest that attributes A3, A4, and A5 do not represent distinct knowledge or abilities. Rather, the data suggest that these three attributes may represent a single ability, which students demonstrate to greater or lesser degrees. In response to this finding, the results were organized into six knowledge levels, which are listed in Table 2, along with the descriptive statistics and percent frequency of the student abilities for each knowledge level. Figure 3 displays the distribution of the student abilities for each knowledge level, where each box's height corresponds to the relative percent frequency of the students in each knowledge level.

Table 2

Statistical Characteristics of Students Assigned to Each Knowledge Level

Knowledge Level	Minimum Ability	Lower Quartile Ability	Median Ability	Upper Quartile Ability	Maximum Ability	Percent of Students
No Knowledge	-2.67	-2.48	-2.25	-1.93	-1.78	1
Identify Covariates	-2.32	-1.98	-1.83	-1.65	-1.31	2
Identify Direction of Variation	-2.32	-1.81	-1.6	-1.46	-1.19	3
Interpret Slope Ratio (Low)	-1.93	-1.18	-0.84	-0.55	-0.18	26
Interpret Slope Ratio (Medium)	-0.67	-0.04	0.31	0.61	1.02	55
Interpret Slope Ratio (High)	1.14	1.14	1.14	1.45	1.45	13

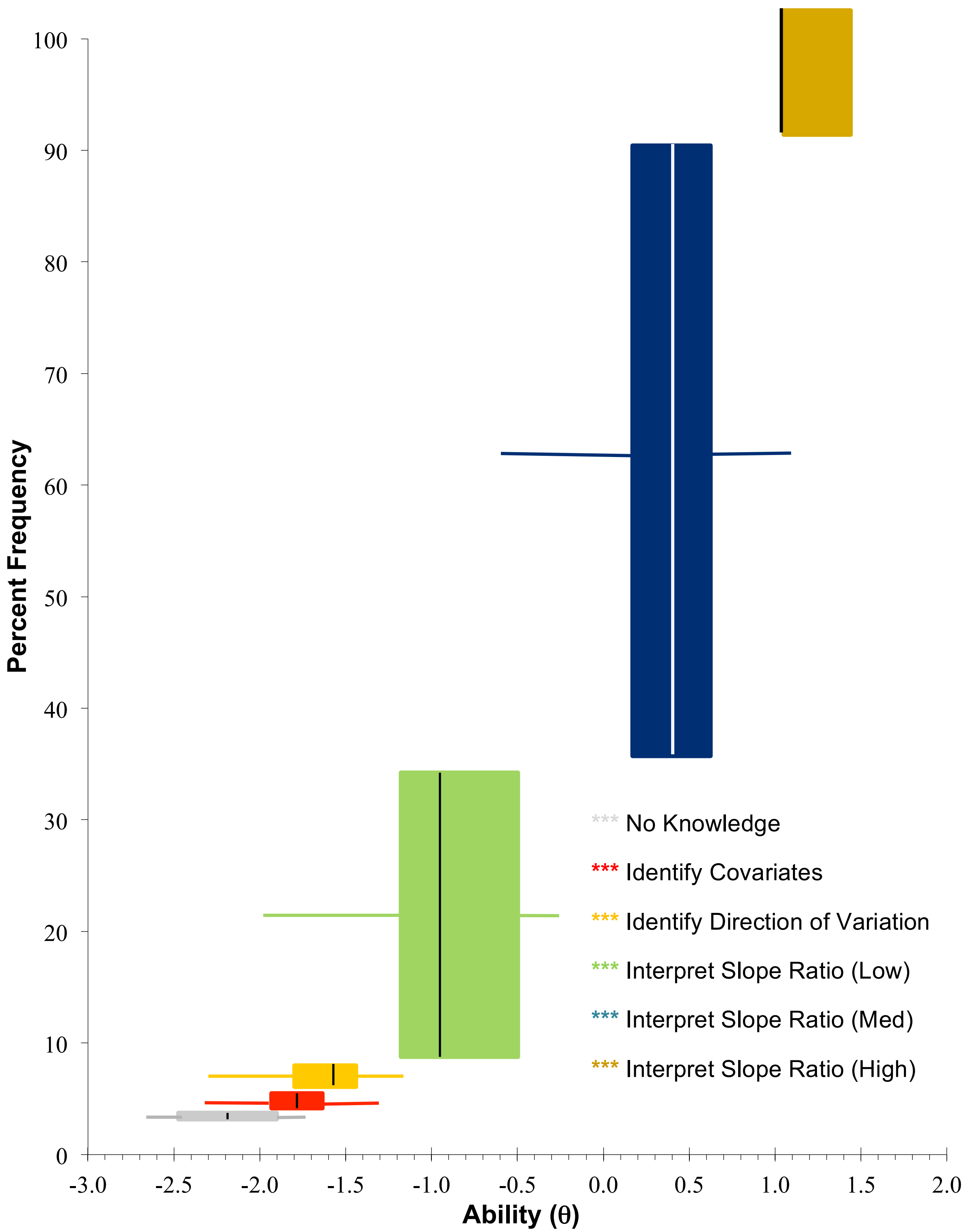

Figure 3. Percent frequencies of student ability estimates by knowledge level.

The box-and-whisker plot graph effectively illustrates the similarities and differences in the abilities of the students classified into different knowledge levels. In cases where adjacent plots share many of the values on the horizontal axis (i.e., ability), there appears to be less difference between the knowledge levels. Conversely, in cases where adjacent plots share very few values on the horizontal axis, there appears to be a greater difference between the knowledge levels.

The results as depicted in Figure 3 indicate that students who were classified as having no knowledge of foundational concepts related to slope had a different ability level than the other students in the sample. Students who were classified as being able to identify covariates, meanwhile, appear to have had a higher ability level than students with no knowledge and a lower ability level than students who were classified as being able to identify the direction of variation. Finally, students who were classified into the three levels of ability to interpret slope ratios appear to have had more overall ability than the students who were classified as being able to identify the direction of variation.

These findings support the theory that there are three main levels of understanding of the selected foundational concepts related to slope, a finding that is consistent with the standards included in the learning progression with ties to the CCSSM and described in Figure 2. 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.

Discussion

The implementation of the CCSSM introduces challenges. Educators will benefit from becoming familiar with the underlying learning progressions and developing instruction that is sensitive to these progressions. Teachers will also need instructional tools to help diagnose student knowledge according to the progressions.

This study demonstrated how to analyze a specific content area in order to identify the components of knowledge that contribute to conceptual understanding. It also described a tool for assessing student knowledge of that content area. In the case of foundational concepts related to slope, the results indicated that the students in this study demonstrated knowledge along a learning progression consistent with the CCSSM.

The presence of learning progressions that cross grade-level boundaries should be the focus of professional development opportunities, bringing teachers from different grade levels together to collaboratively plan their instruction so that it supports these progressions. Such professional development should allow teachers to gain familiarity and skill with analyzing the standards so that they might uncover the relevant learning progressions and plan appropriately sensitive instruction. Teachers who are conscious of the relationships between the learning targets articulated in their grade-level standards and those of surrounding grade levels will be better equipped to teach effectively, both by relating the material they teach to what students have learned previously and by setting the stage for what they will be expected to learn in the future.

Learning progressions have the potential to improve instructional planning by suggesting how learning targets should be sequenced in order to optimize learning. Similarly, as was done with this study, assessment tools should be deliberately designed so that teachers may collect instructionally relevant information about students in terms of the learning targets described in a given learning progression. More studies are needed to clarify the learning progressions implied by the CCSSM and to develop instructional tools that are sensitive to these progressions.

References

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

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.

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. (1986). Conceptual knowledge as a foundation for procedural knowledge: Implications from research on the initial learning of arithmetic. In J. Hiebert (Ed.), *Conceptual and procedural knowledge: The case of mathematics* (pp. 113–132). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

CCSSO/NGA. (2010). *Common core state standards for mathematics*. Washington, D.C.: Council of Chief 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). Retrieved from ProQuest Dissertations and Theses database. (UMI No. 3405966)

Consortium for Policy Research in Education. (2011). Learning trajectories in mathematics: A foundation for standards, curriculum, assessment, and instruction.

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.

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* (pp. 1–27). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

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

Karplus, R., Pulos, S., & Stage, E. K. (1983). 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.

Krupa, E. (2011). *Summary report.* Paper presented at the Moving Forward Together: Curriculum & Assessment and the Common Core State Standards for Mathematics, Pentagon City, Arlington, VA.

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.

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.

Neill, M. (2003). Leaving children behind: How No Child Left Behind will fail our children. *Phi Delta Kappan, 85*(3), 225–228.

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.

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.

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.

Teuscher, D., & Reys, R. E. (2010). Slope, rate of change, and steepness: Do students understand these concepts? *Mathematics Teacher, 103*(7), 519–524.

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.

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). BILOG–MG 3: Multiple-group IRT analysis and test maintenance for binary items [Computer software]. Lincolnwood, IL: Scientific Software International, Inc.