Scaling Up SimCalc Project

Design and Development of the Student and Teacher Mathematical Assessments

Technical Report 05 | March 2010

Nicole Shechtman, Geneva Haertel, Jeremy Roschelle, Jennifer Knudsen, Corrine Singleton

Report Series Published by SRI International

Research collaborators:

University of Massachusetts, Dartmouth; Virginia Polytechnic Institute and State University;
The University of Texas at Austin; and the Charles A. Dana Center at the University of Texas at Austin

SRI International
Center for Technology in Learning
333 Ravenswood Avenue
Menlo Park, CA 94025-3493
650.859.2000

Design and Development of the Student and Teacher Mathematical Assessments

Prepared by:

Nicole Shechtman, SRI International

Geneva Haertel, SRI International

Jeremy Roschelle, SRI International

Jennifer Knudsen, SRI International

Corrine Singleton, SRI International

Acknowledgments

This report is based on work supported by the National Science Foundation under Grant No. 0437861. Any opinions, findings, and conclusions or recommendations expressed in this report are those of the authors and do not necessarily reflect the views of the National Science Foundation.

We thank G. Estrella, K. Rafanan, P. Vahey, A. Stroter, S. Carriere, L. Gallagher, H. Javitz, M. Robidoux, S. Empson, S. Hull, L. Hedges, S. Goldman, H. Becker, J. Sowder, G. Harel, P. Callahan, F. Sloane, B. Fishman, K. Maier, J. Earle, R. Schorr, and B. McNemar for their contributions to this research. We thank the participating teachers and Texas Educational Service Center leaders from regions 9, 10, 13, 17, and 18; this project could not have happened without them. We thank and remember Jim Kaput, who pioneered SimCalc as part of his commitment to democratizing mathematics education.

Design and Development of the Student and Teacher Mathematical Assessments

This technical report documents the design and development of the student and teacher mathematical content assessment instruments for the Scaling Up SimCalc Project. In this project, which is described in Technical Reports 01 and 02 (Roschelle et al., 2007a; Roschelle et al., 2007b) and other publications (e.g., Roschelle et al., in press), we conducted three large-scale studies to address the broad research question, Can a wide variety of teachers use an integration of technology, curriculum, and professional development to increase student learning of complex and conceptually difficult mathematics? The studies examined the SimCalc approach, which uses highly interactive software to help students learn the central middle school mathematical concepts of rate, proportionality, and linear function. The studies took place in Texas during the 2005–06 and 2006–07 school years.

The Assessments and Their Purposes

Two SimCalc studies addressed seventh-grade students, teachers, mathematical content, and curriculum, and the third study was on eighth-grade students, teachers, mathematical content, and curriculum. We developed 3-week curriculum units for each grade level using the SimCalc approach to replace teachers' usual mode of instruction. Across the three studies, four assessments were developed. For each grade-level, we developed (1) an assessment for students of the content of the unit and (2) an assessment for teachers of the mathematical knowledge necessary to teach the unit.

The student assessments, administered pre-replacement unit and post-replacement unit, were the main outcome measure in each of the studies. We found that standardized tests (such as the TAKS [Texas Assessment of Knowledge and Skills]) did not capture the conceptual depth students could reach using the SimCalc technology and curricula; therefore, using such tests for outcome measures would cause us to overlook potentially important effects of the intervention. Thus, the research team decided to build its own assessments. Because student achievement was the primary dependent measure for all the studies, we directed attention and resources to developing assessments that would meet rigorous standards for reliability and validity (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 1999).

The teacher assessments, administered at the study baseline and other time points, were used to examine whether teachers learned mathematical content through participation in this project and to clarify the role of teacher mathematical knowledge in teaching and learning with these technology-enhanced units (for detailed analyses, see Shechtman, Roschelle, Haertel, & Knudsen, in press). Drawing on the work of Ball, Hill, and other researchers in mathematics teaching and learning (Ball, 1990; Ball, Hill, & Bass, 2005; Hill, Rowan, & Ball, 2005; Ma, 1999; Shulman, 1986), the team used the construct of *mathematical knowledge for teaching* (MKT). Because prior research had not assessed MKT at the middle school level, the research team also needed to develop these assessments. Given that findings about teacher knowledge were important but not of primary concern, fewer resources were put into the development and validation of this instrument.

Overview of Assessment Development Processes

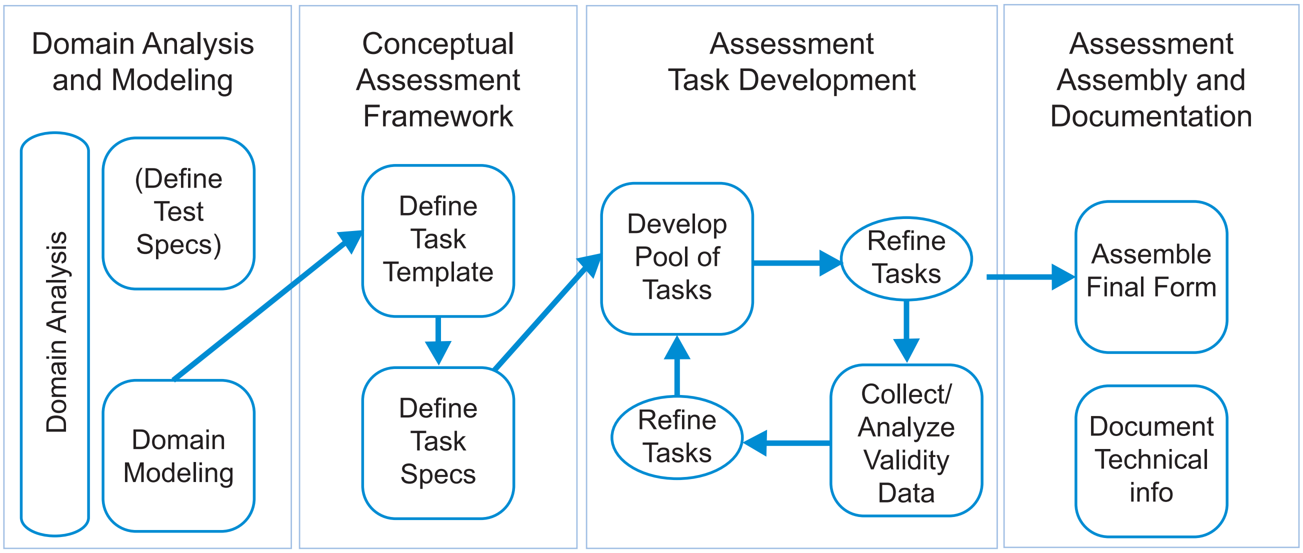
For each grade level, the curriculum, student assessment, and teacher assessment were developed in tandem to ensure they were in alignment with one another. The assessment development process for the seventh-grade studies drew on a pilot study that we had conducted before the large-scale research of the Scaling Up SimCalc Project. As described elsewhere (Tatar et al., 2008), prototype versions of the curriculum, student assessment, and teacher assessment were used in this experiment. Some items were reused with revisions, but the assessments were almost completely redesigned in the fully scaled up work.

We followed best practices in assessment development (e.g., American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 1999) and in the conceptualization of the constructs being assessed. The corresponding test structure and content were grounded in the tenets of Evidence Centered Design (ECD) (Almond, Steinberg, & Mislevy, 2002; Mislevy, Almond, & Lukas, 2003; Mislevy & Haertel, 2006; Mislevy, Steinberg, & Almond, 2002). ECD emphasizes the evidentiary base for specifying coherent, logical relationships among (1) the complex of knowledge, skills, and abilities that are constituents of the construct to be measured; (2) the observations, behaviors, or performances that should reveal the target construct; (3) the tasks or situations that should elicit those behaviors or performances; and (4) the rational development of construct-based scoring criteria and rubrics (Messick, 1994). This evidentiary base supports both the construct and content validity of the assessment.

Figure 1 illustrates the progression of ECD processes followed to build an assessment with a strong evidentiary base. The progression begins with domain analyses and domain modeling activities and ends with assessment assembly and documentation.

In the initial ECD processes, domain analysis and domain modeling, the assessment's conceptual foundation is established. In domain analysis, experts in the content domain articulate the important core knowledge, skills, and abilities to be assessed. During domain modeling, the experts elaborate the structure and content of the assessment tasks to be developed. These processes provide input into the development of a test specification that served as the blueprint for the overall assessment.

Figure 1. ECD processes used to design, develop, validate, and document the final version of each assessment.



In the second process, conceptual assessment framework, the types of assessment items and their properties are specified, including technical issues associated with the scoring model and evaluative rules to be used in scoring the items. The third process, assessment task development, is an iterative cycle of developing a pool of potential assessment items, refining them, collecting and analyzing validity data, using the data to refine the items, and perhaps developing more items. Finally, in the fourth process, assessment assembly and documentation, the validity data are used as a guide to assemble the items to meet the test specifications in the assessment blueprint as closely as possible, and the technical documentation describing the assessment is prepared.

In the following sections, we detail the development of each of the assessments across each of the processes outlined in Figure 1.

Process 1: Domain Analysis and Modeling

The goal of the domain analysis is to establish and articulate important core knowledge, skills, and abilities (KSAs) to be assessed. For each grade level, we performed two major domain analyses. In the first, we developed a conceptual framework that specified the mathematical content that we would focus on in the curriculum and student assessment. Building on this, we developed a conceptual framework for the mathematical knowledge that would be necessary to teach this content.

Here we describe the two sets of domain analyses and the test specifications that were derived from them.

Domain Analysis for the Curriculum and Student Assessment

Because this domain analysis was to serve as the foundation for both our interventions and primary outcome measures, we considered several perspectives in specifying the focal mathematical content for each grade level. One perspective was of the mathematical concepts that students could potentially learn from the SimCalc approach. In this approach, which is described in detail elsewhere (e.g., Roschelle et al., in press), both software and paper curriculum booklets are used. The software anchors students' efforts to make sense of conceptually rich mathematics in their experiences with computer animations of motion. The paper-and-pencil curricula allow students to connect their mathematical understanding across familiar representations (narrative stories and animations of motion) with key mathematical representations (algebraic expressions, tables, graphs). Another perspective was from the Texas seventh- and eighth-grade standards (Texas Essential Knowledge and Skills [TEKS]) and our analysis of the content as actually covered in current Texas mathematics instruction. Our third perspective was to consider the recommendations of the national standards and focal points (i.e., the National Council of Teachers of Mathematics [NCTM]). A fourth perspective was based on the research knowledge about student cognition in the learning sciences and best pedagogical practices for supporting student learning of conceptually difficult mathematics. Finally, we considered the emphasis of the funder of the research, the National Science Foundation (NSF), on "complex and conceptually difficult mathematics."

We started by identifying the broad concepts at the intersection of these perspectives. This led to the identification of proportionality and linear function as the target mathematics. Among middle school mathematical concepts, proportionality ranks high in importance, centrality, and difficulty (Hiebert & Behr, 1988; National Council of Teachers of Mathematics, 2000; Post, Cramer, Behr, Lesh, & Harel, 1993). For example, the NCTM describes proportionality and related concepts as "focal points" for learning in seventh and eighth grade (National Council of Teachers of Mathematics, 2007). Mathematically, proportionality is closely related to the concepts of rate, linearity, slope, and covariation. In addition, proportionality offers an opportunity to introduce students to the concept of a function through the constant of proportionality, *k*, that relates *x* and *f(x)* in the functional equations of the form $f(x) = kx$. A sufficiently deep understanding of function as it relates to rate, linearity, slope, and covariation is central to progress in algebra and calculus. These concepts are also central to students' science learning. Without understanding rate and proportionality, students cannot master important topics and representations in high school science such as laws (e.g., $F = ma$, $F = -kx$), graphs (e.g., of linear and piecewise linear functions), and tables (e.g., interpolating between explicit values relating the width and length of maple leaves). Mathematics education research has identified persistent difficulties in mastering these concepts (i.e., misconceptions) and has theorized that proportionality is at the heart of the conceptually challenging shift from additive to multiplicative reasoning (Harel & Confrey, 1994; Leinhardt, Zaslavsky, & Stein, 1990; Vergnaud, 1988).

In conjunction with our project's mathematics advisory board, which included three mathematicians and three mathematics educators, we developed a mathematics framework for the seventh and then for the eighth-grade intervention that abstracted the mathematical concepts to be used in the curricula and assessments. These are summarized in Table 1, and the more detailed frameworks are in Appendix A. Appendix B shows the subset of TEKS skills and knowledge that were targeted specifically in each grade. In Table 1, we use the symbol M_1 to refer to the mathematics that is measured on the tests used for accountability in Texas. We use the symbol M_2 to refer to mathematics that goes beyond what is tested in Texas.

Table 1. Mathematical conceptual frameworks for the seventh-grade and eighth-grade curricula and assessments.

Framework	M_1 Component *Foundational concepts typically covered in the grade-level standards, curricula, and assessments*	M_2 Component *Building on the foundations of M_1, essentials of concepts of mathematics of change and variation found in algebra, calculus, and the sciences*
Rate and proportionality for the seventh-grade studies	• Simple a/b = c/d or y = kx problems in which all but one of the values are provided and the last must be calculated • Basic graph and table reading without interpretation (e.g., given a particular value, finding the corresponding value in a graph or table of a relationship)	• Reasoning about a representation (e.g., graph, table, or y = kx formula) in which a multiplicative constant k represents a constant rate, slope, speed, or scaling factor across three or more pairs of values that are given or implied • Reasoning across two or more representations
Linear function for the eighth-grade study	• Categorizing functions as linear/ nonlinear and proportional/ nonproportional • Within one representation of one linear function (formula, table, graph, narrative), finding an input or output value • Translating one linear function from one representation to another	• Interpreting two or more functions that represent change over time, including linear functions or segments of piecewise linear functions • Finding the average rate over a single multirate piecewise linear function

Proportionality can be taught both as a formula ($\frac{a}{b}$ - $\frac{c}{d}$) and a function $f(x) = kx$, where k is the constant of proportionality. The analysis of the latter function across algebraic, graphical, tabular, animated, and verbal forms can be the starting point for the learning progression that leads to calculus. In particular, emphasizing the conceptual links among different expressions of "rate" brings coherence to instruction that promotes an ever-deepening understanding of the mathematics of change and variation across many years of material. The particular opportunity in seventh-grade instruction is to connect the multiplicative constant k in the algebraic expression $y = kx$, the slope of a graphed line, the constant ratio of differences in a table comparing y and x values, and the experience of rate as "speed' in a motion. In eighth grade these connections expand to the more complex model implied by the linear function $y = mx + b$.

The NCTM has a perspective that is consistent with, but not identical to, the Texas and SimCalc perspectives. We initially geared our content toward the NCTM's central document, the Principles and Standards (National Council of Teachers of Mathematics, 2000). During our research, the NCTM added a recommended set of three "focal points" to be emphasized at each grade level (National Council of Teachers of Mathematics, 2007). Our mathematics aligned with one of the major focal points in each of seventh and eighth grade. The seventh-grade focal point for algebra emphasizes proportionality. The eighth-grade focal point is described as follows:

> Students use linear functions, linear equations, and systems of linear equations to represent, analyze, and solve a variety of problems. They recognize a proportion ($y/x = k$, or $y = kx$) as a special case of a linear equation of the form $y = mx + b$, understanding that the constant of proportionality (k) is the slope and the resulting graph is a line through the origin. Students understand that the slope (m) of a line is a constant rate of change, so if the input, or x-coordinate, changes by a specific amount, a, the output, or y-coordinate, changes by the amount ma. Students translate among verbal, tabular, graphical, and algebraic representations of functions (recognizing that tabular and graphical representations are usually only partial representations), and they describe how such aspects of a function as slope and y-intercept appear in different representations. Students solve systems of two linear equations in two variables and relate the systems to pairs of lines that intersect, are parallel, or are the same line, in the plane. Students use linear equations, systems of linear equations, linear functions, and their understanding of the slope of a line to analyze situations and solve problems. (National Council of Teachers of Mathematics, 2007, p. 20)

As recommended, the SimCalc content addresses linear and proportional functions, emphasizing that the slope of a line represents a constant rate of change and engaging students in translating among verbal, tabular, graphical, and algebraic representations. Although SimCalc can include systems of linear equations, our research did not address this concept because it is not considered appropriate for eighth-grade mathematics in Texas. (This illustrates one of the conflicts among perspectives.)

Finally, we continued to attempt to be responsive to the NSF's notion of "complex and conceptually difficult" mathematics. Overall, we found this notion ill defined and difficult to apply. The phrase appeared to be attributed to Trends in International Mathematics and Science Study research in the NSF request for proposals to which we responded, but we were unable to find a source for it. We found the phrase ill defined because students can experience mathematics as difficult or complex for many reasons. Classic sources of difficulty and complexity are multistep problem solving, imprecisely framed questions, awkward quantities, and overly formal or abstract presentations. Further, conceptual difficulties may vary with the approach taken: A historical example is that it is easier to add with Roman than Arabic numerals (no place value system is needed) but harder to multiply with Roman than Arabic numerals (because of the lack of place value system).

In sum, Table 1 represents the resolution of the differences in perspectives and provides the conceptual framework for the assessments.

Domain Analysis for the Mathematics Necessary to Teach the Curricula

Because the purposes of the teacher mathematics assessments were to examine teacher mathematics learning and the role of teacher knowledge in the implementation of the curricula, in this domain analysis we established a framework for assessing not only teachers' knowledge of the KSAs in Table 1, but also the specialized mathematical knowledge necessary to support students' learning of these KSAs. Drawing on the work of Ball, Hill, and other researchers in mathematics teaching and learning (Ball, 1990; Ball, Hill, & Bass, 2005; Hill, Rowan, & Ball, 2005; Ma, 1999; Shulman, 1986), the team used the construct of *mathematical knowledge for teaching* (MKT) and distilled seven types of knowledge necessary to support students' formative understandings of the concepts in the units (Table 2).

The domain for the MKT assessment therefore encompassed both general knowledge of the concepts in Table 1 and the specialized knowledge necessary to teach these concepts as structured by the categories in Table 2.

Table 2. Core types of MKT necessary for teachers to support students' formative understanding of the concepts in the seventh- and eighth-grade units.

Type of Knowledge	Example
Interpreting unconventional forms or representations that students are likely to make as they construct their understanding (such representations may be mathematically correct but not conventional)	Interpret the slope of a line on a position graph with time on the y-axis and distance on the x-axis
Generating, choosing, and evaluating problems and examples that can illustrate key curricular ideas	Choose a word problem that illustrates proportional reasoning
Differentiating between colloquial and mathematical uses of language and evaluating student statements for their mathematical precision	Know that "The slope of the line is 0" is more mathematically precise than "the line is flat."
Linking precise aspects of representations and translating a function between multiple representations (i.e., story, graph, table, algebra)	Know that the m in $y = mx + b$ corresponds to the slope of graphed line
Understanding implications of models and representations	Understand that rate and time are inversely proportional, even though the curriculum focuses on the linear relationship between time and distance in $d = rt$
Evaluating the validity of mathematical conjectures	Evaluate under what conditions this student's statement is true: "The shorter the line the faster they run."
Connections to important advanced mathematics beyond the unit	Make connections between changes in position and velocity

Table 3. Test specifications.

Dimension	Student Assessments	MKT Assessments
Mathematical content	• Items aligned with each of the focal KSAs as outlined in Appendix A (such that they create reliable M_1 and M_2 subscales) • Alignment with Texas state standards (TEKS), as outlined in Appendix B	• Items aligned with each of the focal KSAs in Table 1 • Items aligned with categories in Table 2 that support the understanding of KSAs in Table 1 • In seventh-grade studies only, to assess knowledge of mathematics beyond the unit that was covered in the professional development, knowledge of connections between representations of changes in position and velocity
Task types	• Varied across contexts (i.e., motion, money) • Diversity of task types (about one-third each of multiple choice, short response, construction of multiple mathematical representations)	• Multiple choice, following the model of prior MKT work in the field (e.g., Hill, Rowan, & Ball, 2005) • In the seventh-grade studies only, velocity items assessed through multiple choice and constructed response

Test Specifications

As part of the domain analysis process, we developed test specifications (Table 3). All assessments were paper and pencil. The student assessments were designed to be administered by teachers in their own classrooms within one class period (about 45 minutes). The teacher assessments were designed to be administered by workshop leaders or self-administered by teachers at home within about 1.5 hours.

Process 2: Conceptual Assessment Framework

The goals of this stage were to establish the types of assessment items and their properties. For the student assessments, we used existing items from released standardized tests, previously validated instruments, the research literature, the SimCalc pilot, and the SimCalc curriculum to populate the blueprint test specifications as outlined in Table 3.

For the MKT assessments outlined in our test specifications, few items of this type existed, so we had to generate our own pool of items. A common approach in many large-scale assessment development processes is to create *item templates* that form the structure for assessment items and can be filled in with variable content. Drawing on previous examples of MKT assessments (e.g., Ball, 1990; Ball, Hill, & Bass, 2005; Hill, Rowan, & Ball, 2005), we created a set of item templates that could be filled in with new MKT content. Each template could be used to produce a set of multiple-choice questions situated within the context of mathematics classrooms and teaching. Each template addressed one of the key facets of MKT presented in Table 2 and had three parts. The first part was a particular teaching situation that evoked one of the facets of knowledge in Table 2, such as a teacher presenting a problem to the class, grading papers, examining errors students made on a particular problem, or attending a professional development workshop. The second part was a mathematical question about this situation, and the third part was a list of distracters, including the correct answer.

Process 3: Assessment Task Development

Described here are the development and validation processes. Appendix C provides examples of assessment items.

Development of the Item Pool

We collected candidate items for the student assessments from a variety of sources, guided by the assessment blueprint.

- Released standardized tests (seventh-grade TAKS, eighth-grade Trends in International Mathematics and Science Study [TIMSS], eighth-grade National Assessment of Educational Progress [NAEP], California High School Exit Examination [CHSEE], the eighth- and tenth-grade Massachusetts Comprehensive Assessment System [MCAS])
- Items used in early SimCalc design research and the SimCalc pilot
- The rate and proportionality literature (e.g., Kaput & West, 1994; Lamon, 1994; Lobato & Thanheiser, 2002)
- The math of change and variation literature (e.g., Carlson, Jacobs, Coe, Larsen, & Hsu, 2002)
- Items adapted directly from the SimCalc unit

For the seventh-grade and eighth-grade assessments, this produced initial pools of 59 and 58 items, respectively.

To develop our initial pool of items for the MKT assessments, we held a 1.5-day item camp, a workshop in which individuals with various types of expertise came together to collaboratively generate new assessment items. In addition to the SimCalc curriculum designer and core research team, members of the item camps included an experienced middle school math teacher, math education researchers, mathematicians, and assessment experts. Participants were provided with the test specifications, item templates, an outline of the core mathematical KSAs (Tables 1 and 2), the SimCalc curriculum, and various resources such as the Texas middle school mathematics standards, assessments, and textbooks. They were then asked to use these resources to generate items that addressed all the important mathematics teachers should know to support student learning during the unit. In addition, in the seventh-grade assessment, to test mathematics beyond the unit we incorporated into the pool items from previous SimCalc research that assessed knowledge of connections between representations of changes in position and velocity.

For the seventh-grade and eighth-grade MKT assessments, this generated initial pools of 45 and 57 items, respectively.

Collection and Analysis of Validity Data

In item validation, evidence is accumulated to provide a scientifically sound argument that the assessment items measure the constructs they are intended to measure to support the intended interpretation of test scores (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 1999). Table 4 outlines the steps of assessment validation that we followed for each assessment. Here we present data collected with each step.

Formative and Summative Expert Panel Reviews

For each assessment, we conducted both a formative and summative review. Each formative review occurred early in the development process after the initial pool of items had been developed. Each

Table 4. Methods used for assessment validation. (Items were iteratively refined with each step.)

Method	Validity Issue
Formative and summative expert panel reviews (alignment and ratings of items)	• Alignment with the intended content (Table 1) • Alignment with the TEKS • Grade-level appropriateness
Cognitive think-alouds	• Does the task make sense to the respondents? • Does the task elicit the cognitive processes intended? • Can the task be completed in the available time? • Can respondents use the diagrams, charts, tables as intended? • Is the language clear?
Field-testing for psychometric information	*Individual Items* • Range of responses from students representing different levels of mathematical understanding • Amount of variation in responses sufficient to support statistical analysis • Distribution of responses by distractor • Presence of ceiling or floor effects • Discrimination among students at different levels of the construct being assessed *Overall Form and Subscale Analyses* • Internal reliability • Biases among subgroups

summative review was the final step in the assessment development process. In each review, experts were provided with the assessment items and asked to make specific ratings and/or recommendations for each item.

The two formative panels for the seventh-grade and eighth-grade student assessments took place in person in Austin, Texas, in a 1-day workshop. There were two subpanels. For both types of panels, training consisted of an orientation to (1) the SimCalc project, (2) the mathematical content, and (3) the assessment items. The first subpanel was composed of mathematics education researchers and assessment experts (four for the seventh-grade assessment and three for the eighth-grade assessment). This subpanel made two types of categorical concurrence ratings for each item in the item pool: (1) alignment with the SimCalc conceptual framework through classification of each item with respect to KSAs outlined in Appendix A and (2) alignment with a focused subset of the TEKS, as shown in Appendix B. They were also asked to make recommendations for improving the items. The second subpanel for each assessment was composed of two local curriculum and instruction experts (e.g., math supervisor, textbook contributor). This subpanel rated items relative to three aspects of grade-level appropriateness: (1) reading load, (2) computation load, and (3) graphics load. They rated each item on a 3-point scale *(appropriate, somewhat inappropriate, inappropriate)* for each aspect of grade-level appropriateness. They were also asked to recommend modifications that would make the items more grade-level appropriate.

For each assessment, data were compiled at the item level and used to classify and refine the items. Items that did not have a majority of agreement on categorical concurrence classification, were not aligned with the target TEKS, and/or were rated as inappropriate for the grade level were eliminated from the item pool or modified to be suitable. The categorical concurrence data were also used to classify items to determine coverage in our field test instrument across the KSAs in Appendix A.

The summative panels for the student assessments took place after field-testing (see below) and were conducted by mail. The experts were members of our advisory board who had worked with us to develop the mathematical conceptual frameworks. They were provided with the fully refined items and the categorical concurrence classifications determined during the formative review. For each item, they checked off whether they agreed or disagreed with the classification. If they did not agree with it, they were to explain their decision. The summative panel members agreed with each other and with the prior ratings in almost all cases. This summative classification was used to determine content coverage in the final instruments.

Because findings about teacher knowledge were lower stakes in this project than those about student knowledge, to conserve resources, we conducted only a formative review of the MKT assessments and had senior members of the SimCalc research team serve as the experts. For each assessment, two experts aligned each item in the initial item pool with the frameworks in Appendix A and Table 2 and recommended refinements to enhance clarity and alignment. These data were used to determine coverage of the content in these frameworks.

Cognitive Think-Alouds

The second validity study was conducting using a cognitive think-aloud methodology. Cognitive think-alouds were conducted on all the items remaining in the item pool after the formative review. In a cognitive think-aloud, the test-taker speaks out loud his or her thinking as he or she works through assessment items. An interviewer records this monologue and asks a minimal number of probing questions as necessary but does not interfere with the test-taker's engagement with the mathematics. For each student assessment, we conducted think-aloud interviews with eight middle school students who each did a subset of the items. The students represented the full range of achievement levels, and any given item was done by 2-3 high-, 2-3 medium-, and 2-3 low-achieving students. The think-alouds were conducted by a middle school mathematics teacher trained in the think-aloud protocol. These interviews were audiotaped for the seventh-grade assessment and videotaped for the eighth-grade assessment and analyzed by a member of the research team. For each MKT assessment, we conducted think-aloud interviews with three teachers known to represent a range of MKT levels. These interviews were conducted and analyzed by a member of the research team.

For each participant, the interviewer documented the time needed for completion, the mathematical strategies the test-taker used, the mathematical mistakes made, difficulties in comprehending the problem because of ambiguous or unclear language, unfamiliar terminology, or confusing calculations (when applicable). This information was then summarized for each item and used to eliminate and/or modify items in the pool. Items were eliminated if they were too easy or too difficult for the test-takers or if the test-takers could use a construct-irrelevant strategy to solve them (e.g., counting to solve a problem intended to measure proportional reasoning). Item instructions, text, and graphics were modified as necessary to increase clarity and refine mathematical logic.

Field-Testing for Psychometric Information

The third method was field-testing the assessments with a large sample. For each assessment, we used the categorical concurrence ratings of the remaining refined items in the item pool to assemble a field test instrument that met the requirements of the test specifications in Table 3.

We field-tested each student assessment with a representative sample of middle school students (230 for the seventh-grade assessment and 309 for the eighth-grade assessment). For each MKT assessment, we conducted field testing through a mass mailing to a national random sample of 1,000 middle school mathematics teachers (the names and addresses were purchased from an educational data service). The response rates were 17.9% and 12.8%, yielding 179 and 128 teachers for the seventh- and eighth-grade assessments, respectively. On key demographic variables (gender, age, teaching experience, ethnicity, region type, and first language), the samples were representative of the population of teachers we expected to participate in the Scaling Up SimCalc Project; suburban regions, relative to rural and urban regions, were slightly oversampled.

We used classical test theory (CTT) and item response theory (IRT) with the field test data to examine three critical evidentiary concerns. First, we examined individual items for the range of possible responses, statistical variation, ceiling and floor effects, and the capacity of the items to discriminate among test-takers at different ability levels (using IRT parameters for a two-parameter logistic model). Second, we examined the internal consistency of the relationship of total and subscale scores to individual items to the test (i.e., scale reliability). Third, we examined possible biases among population subgroups.

We used these data to refine the instrument for a final version. Items with low discrimination parameters (i.e., items that could not discriminate among individuals of differing ability) or ceiling/floor effects were eliminated or modified. Items were kept that were likely to contribute the most information about the test-taker's ability and to maintain representative coverage of the assessment conceptual framework. Using the IRT data, we calibrated the MKT assessments to be relatively difficult so that the average score would be about 50%.

Table 5. Summary of basic test statistics for each assessment. (MKT subscale statistics not reported because scores always reported in aggregate.)

	Whole Form			M_1 Subscale			M_2 Subscale		
Assessment	Items	Internal Reliability	Pre-test mean *(SD)*	Items	Internal Reliability	Pre-test mean *(SD)*	Items	Internal Reliability	Pre-test mean *(SD)*
Seventh grade									
Student	30	0.86	12.9 (5.7)	11	0.73	7.3 (2.6)	19	0.82	2.6 (3.7)
MKT	24	0.80	10.0 (4.5)						
Eighth grade									
Student	36	0.91	12.1 (7.4)	18	0.79	7.2 (3.7)	18	0.87	4.9 (4.3)
MKT	28	0.80	16.3 (5.0)						

Process 4: Assessment Assembly and Documentation

As summarized in Table 4, the validation methods were used to both iteratively refine the items and forms, as well as provide evidence for validity.

Form and item statistics are documented here using data from the research participants. The sample characteristics can be found in Appendix D. Table 5 summarizes the form statistics, and detailed item data are in Appendices E–H.

Conclusion

These processes led to the development of the four assessments used in the Scaling Up SimCalc Project. The student assessments revealed statistically significant main effects, with student-level effect sizes of .63 and .56 in the Seventh-Grade and Eighth-Grade studies respectively (Roschelle et al., in press). We found mixed support for the relationship between student learning and MKT, as there was a significant relationship between M2 learning gains and teacher MKT in the Seventh-Grade Study ($\beta = .13$, $p<.01$) but not in the Eighth-Grade study ($\beta = -.01$, n.s.) (Shechtman et al., in press).

References

Almond, R. G., Steinberg, L. S., & Mislevy, R. J. (2002). Enhancing the design and delivery of assessment systems: A four-process architecture. *Journal of Technology, Learning, and Assessment, 1*(5).

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (AERA, APA, NCME). (1999). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.

Ball, D. L. (1990). The mathematical understandings that prospective teachers bring to teacher education. *The elementary school journal, 90*(4), 449–466.

Ball, D. L., Hill, H. C., & Bass, H. (2005). Knowing mathematics for teaching. *American Educator, 29*(3), 14.

Carlson, M. P., 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.

Harel, G., & Confrey, J. (1994). *The development of multiplicative reasoning in the learning of mathematics*. Albany, NY: State University of New York Press.

Hiebert, J., & Behr, M. (1988). Introduction: Capturing the major themes. In J. Hiebert, & M. Behr (Eds.), *Number concepts and operations in the middle grades* (pp. 1–18). Hillsdale, NJ: Lawrence Erlbaum.

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

Kaput, J., & West, M. M. (1994). Missing-value proportional reasoning problems: Factors affecting informal reasoning patterns. In G. Harel & J. Confrey (Eds.), *The development of multiplicative reasoning in the learning of mathematics*. Albany, NY: State University of New York Press.

Lamon, S. (1994). Ratio and proportion: Cognitive foundations in unitizing and norming. In G. Harel, & J. Confrey (Eds.), *The development of multiplicative reasoning in the learning of mathematics*. Albany, NY: State University of New York Press.

Leinhardt, G., Zaslavsky, O., & Stein, M. (1990). Functions, graphs, and graphing: Tasks, learning, and teaching. *Review of Educational Research, 60,* 1–64.

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

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

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

Mislevy, R. J., Almond, R. G., & Lukas, J. F. (2003). *CRESST Technical Paper Series*. Los Angeles: Center for Research on Evaluation, Standards, and Student Testing (CRESST).

Mislevy, R. J., & Haertel, G. (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. (2002). On the structure of educational assessments. *Measurement: Interdisciplinary Research and Perspectives, 1,* 3–67.

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

National Council of Teachers of Mathematics. (2007). *Curriculum focal points for prekindergarten through grade 8 mathematics: A quest for coherence.* Reston, VA: National Council of Teachers of Mathematics.

Post, T. R., Cramer, K. A., Behr, M., Lesh, R., & Harel, G. (1993). Curriculum implications of research on the learning, teaching, and assessing of rational number concepts. In T. Carpenter, E. Fennema, & T. Romberg (Eds.), *Rational numbers: An integration of research* (pp. 327–362). Hillsdale, NJ: Erlbaum.

Roschelle, J., Tatar, D., Shechtman, N., Hegedus, S., Hopkins, B., Knudsen, J., & Stroter, A. (2007a). *Can a technology enhanced curriculum prove student learning of important mathematics? (Scaling Up SimCalc Project Technical Report No. 01).* Menlo Park, CA: SRI International.

Roschelle, J., Tatar, D., Shechtman, N., Hegedus, S., Hopkins, B., Knudsen, J., & Stroter, A. (2007b). *Extending the SimCalc approach to grade 8 mathematics (Scaling Up SimCalc Project Technical Report No.02).* Menlo Park, CA: SRI International.

Roschelle, J., Tatar, D., & Kaput, J. (2008a). Getting to scale with innovations that deeply restructure how students come to know mathematics. In A. E. Kelly, R. Lesh, & J. Y. Baek (Eds.), *Handbook of design research methods in education* (pp. 369–395). New York: Routledge.

Roschelle, J., Tatar, D., Shechtman N., & Knudsen, J. (2008b). The role of scaling up research in designing for and evaluating robustness. *Educational Studies in Mathematics, 68,* 149–170.

Roschelle, J., Knudsen, J., & Hegedus, S. (2009a). From new technological infrastructures to curricular activity systems: Advanced designs for teaching and learning. In M. J. Jacobson, & P. Reimann (Eds.), *Designs for learning environments of the future: International perspectives from the learning sciences.* New York: Springer.

Roschelle, J., Shechtman, N., Tatar, D., Hegedus, S., Hopkins, B., Empson, S., Knudsen, J., & Gallagher, L. (in press). Integration of technology, curriculum, and professional development for advancing middle school mathematics: Three large-scale studies. *American Educational Research Journal.*

Shechtman, N., Roschelle, J., Haertel, G., & Knudsen, J. (in press). Investigating links from teacher knowledge, to classroom practice, to student learning in the instructional system of the middle school mathematics classroom. *Cognition and Instruction.*

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

Tatar, D., Roschelle, J., Knudsen, J., Shechtman, N., Kaput, J., & Hopkins, B. (2008). Scaling up technology-based innovative mathematics. *Journal of the Learning Sciences, 17*(2), 248–286.

Vergnaud, G. (1988). Multiplicative structures. In J. Hiebert & M. Behr (Eds.), *Number concepts and operations in the middle grades* (pp. 141–161). Reston, VA: National Council of Teachers of Mathematics.

Appendix A

Detailed Conceptual Frameworks for the Mathematical Content of the Seventh- and Eighth-Grade Interventions and Student Assessments

Table A1. Seventh-Grade Core Mathematical Constructs. Proportionality as Preparation for Algebra and Calculus.

M_1 - Conceptually Simple Proportionality	
Solving for a specific value	
A. Solving problems using the formula *a/b* = *c/d*	Simple *a/b* = *c/d* problem in which three of the values are provided and the fourth must be calculated or the proportion must be recognized
B. Solving unit rate problems with *y* = *kx* or *d* = *rt*	Simple *y* = *kx* or *d* = *rt* problem in which two values are provided and a third must be calculated (even if it is based on own prior work)
Reading a specific value	
C. Basic graph reading of linear relationships	• Reading values at specific points without interpreting their meaning as a rate • Using the labels of axes to determine the meaning of a given pair of *x, y* coordinates • Sketching or plotting a given pair of *x, y* coordinates
D. Basic table reading of linear relationships	Given a particular value, find the corresponding value in a table of a relationship
M_2 - Complex and Conceptually Difficult	
E. Solving problems that invoke the function *y* = *kx*	Reasoning about a representation (e.g., graph, table or *y* = *kx* formula) in which a multiplicative constant k represents a constant rate, slope, speed, or scaling factor across many pairs of values (three or more pairs) that are given or implied
Within representations	
F. Algebraic expression	Interpreting the behavior of a proportional function represented by an algebraic expression or constructing an algebraic representation of a proportional function
G. Table	Filling in cells of a table with many (three or more pairs) values that are related by the same constant of proportionality
H. Graph	Interpreting or constructing the graph of a proportional or linear function
I. Graph with a piecewise linear function	Interpreting or constructing a piecewise linear graph (e.g., releative to narrative description of change over time)
Making connection(s) or comparison(s)	
J. Across two or more functions	Interpreting, comparing, or constructing two or more linear or piecewise linear functions
K. Across multiple representations	Reasoning about the same proportional relationship across at least two of the following representations: graph, table, formula

Table A2. Eighth-Grade Core Mathematical Constructs. Linear Function as Preparation for Algebra and Calculus.

M_1 - Math Addressed Typically in Eighth Grade
1. Problem is within **<u>one representation</u>** of **<u>one linear (but not piecewise linear) function</u>** A. Categorize the function as: i. Linear vs. nonlinear ii. Proportional vs. nonproportional B. Use a linear representation to find an input or output value within: i. Symbolic expression. Given an input, find the output (or given an output, find the input). ii. Table. Given at least two ordered pairs, complete the table. iii. Graph. Given an *x,* find corresponding *y* (or given a *y,* find corresponding *x*). iv. Narrative description. Given a verbal description of an input, find an output.
2. Problem requires **<u>translation</u>** of **<u>one linear (but not piecewise linear) function</u>** from one representation to another (use and/or interpret *m* and *b* as key characteristics or use a few points) A. Graph ↔ Table B. Graph ↔ Symbolic C. Graph ↔ Narrative D. Table ↔ Symbolic E. Table ↔ Narrative F. Symbolic ↔ Narrative
M_2 - Beyond Math Addressed Typically in Eighth Grade
3. Problem requires **<u>interpretation</u>** of **<u>two or more functions</u>** that represent change over time, including linear functions or segments of piecewise linear functions A. Compare: i. Different segments in a piecewise function a) Duration of different segments b) Distance traveled represented by different segments c) Direction of change (e.g., forward/backward, increasing/decreasing) of different segments d) Rate of change (e.g., faster/slower) of different segments ii) Two or more different linear functions a) Time at which two different functions reach a given position b) Given a time, the corresponding position in two linear functions c) Duration of different functions d) Distance traveled represented by two different functions e) Direction of change (e.g., forward/backward, increasing/decreasing) of two different functions f) Rate of change (e.g., faster/slower) of two different functions B. Find the average rate over a single multirate piecewise linear function

Appendix B
TEKS Knowledge and Skills Covered by the SimCalc Student Assessments

Table B1. Seventh-Grade TEKS Knowledge & Skills.

TEKS Standard (K&S)	Standard Component
(Introduction a) Students use algebraic thinking to describe how a change in one quantity in a relationship results in a change in the other, and they connect verbal, numeric, graphic, and symbolic representations of relationships.	
(2) **Number, operation, and quantitative reasoning.** The student adds, subtracts, multiplies, or divides to solve problems and justify solutions.	The student is expected to: (2D) use division to find unit rates and ratios in proportional relationships such as speed, density, price, recipes, and student-teacher ratio (2F) select and use appropriate operations to solve problems and justify the selections
(3) **Patterns, relationships, and algebraic thinking.** The student solves problems involving proportional relationships.	The student is expected to: (3B) estimate and find solutions to application problems involving proportional relationships such as similarity, scaling, unit costs, and related measurement units
(4) **Patterns, relationships, and algebraic thinking.** The student represents a relationship in numerical, geometric, verbal, and symbolic form.	The student is expected to: (4A) generate formulas involving conversions, perimeter, area, circumference, volume, and scaling (4B) graph data to demonstrate relationships in familiar concepts such as conversions, perimeter, area, circumference, volume, and scaling
(7) **Geometry and spatial reasoning.** The student uses coordinate geometry to describe location on a plane.	The student is expected to: (7A) locate and name points on a coordinate plane using ordered pairs of integers
(14) **Underlying processes and mathematical tools.** The student communicates about Grade 7 mathematics through informal and mathematical language, representations, and models.	The student is expected to: (14A) communicate mathematical ideas using language, efficient tools, appropriate units, and graphical, numerical, physical, or algebraic mathematical models

Table B2. Eighth-Grade TEKS Knowledge & Skills.

TEKS Standard (K&S)	Standard Component
(8.3) **Patterns, relationships, and algebraic thinking.** The student identifies proportional relationships in problem situations and solves problems.	The student is expected to: (A) compare and contrast proportional and non-proportional linear relationships; and (B) estimate and find solutions to application problems involving proportional relationships such as similarity and rates.
(8.4) **Patterns, relationships, and algebraic thinking.** The student makes connections among various representations of a numerical relationship.	The student is expected to generate a different representation of data given another representation of data (such as a table, graph, equation, or verbal description).
(8.5) **Patterns, relationships, and algebraic thinking.** The student uses graphs, tables, and algebraic representations to make predictions and solve problems.	The student is expected to: (A) predict, find, and justify solutions to application problems using appropriate tables, graphs, and algebraic equations
(8.15) **Underlying processes and mathematical tools.** The student communicates about Grade 8 mathematics through informal and mathematical language, representations, and models.	The student is expected to: (B) evaluate the effectiveness of different representations to communicate ideas.
(8.16) **Underlying processes and mathematical tools.** The student uses logical reasoning to make conjectures and verify conclusions.	The student is expected to: (A) make conjectures from patterns or sets of examples and nonexamples; and (B) validate his/her conclusions using mathematical properties and relationships.

Appendix C
Sample Items

Figure C1. Seventh-grade student M_1 item.

If $\frac{2}{25} = \frac{n}{500}$ then n =

- ☐ A. 10
- ☐ B. 20
- ☐ C. 30
- ☐ D. 40
- ☐ E. 50

Figure C2. Seventh-grade student M_2 item.

Annie is selling raffle tickets for a trip to Hawaii. Each ticket costs the same amount.

a. Complete the table to find the cost for different numbers of tickets.

Number of Tickets t	**Total Cost ($)** c
10	15
20	30
30	
40	
50	75
	150
	450

b. Write a <u>formula</u> that represents the relationship between number of tickets (t) and total cost (c).

c. On the axes below, sketch a <u>line graph</u> to represent the relationship between any number of tickets and total cost. Label the axes.

Figure C3. Seventh-grade teacher MKT Item (generating, choosing, and evaluating problems and examples that can illustrate key curricular ideas).

Mrs. Covey wants to use story problems during a unit on proportional functions. Which story(ies) **can** be modeled as a proportional function?
(*Mark [X]* ***ALL*** *that apply.*)

☐ A. It costs $5 to enter the airport parking lot, plus $2 every 15 minutes. What does it cost for 1 hour of parking?

☐ B. At Johnson's Market, it costs $1.49 for each pound of peppers. How much do 2.5 lbs. cost?

☐ C. Mark wants to paint a 2200 sq. ft surface. Each gallon of paint covers 500 sq. ft. How many gallons of paint does he need?

☐ D. Yesterday, Maria filled up her 15-gallon gas tank and took a 150 mile trip on I-10. She used 10 gallons of gas. Today, she has to go 50 miles on I-10. Will she make it without filling up again?

☐ E. None of the above.

Figure C4. Eighth-grade student M_1 item.

In the sequence below, which expression can be used to find the value of the term in the *n*th position?

Position	Value of Term
1	0.25
2	0.5
3	0.75
4	1.0
5	1.25
n	

A. $n - 0.75$

B. $\frac{n}{4}$

C. $4n$

D. $n - 1.5$

Figure C5. Eighth-grade student M_2 item.

Annie and Bonnie are running on the same track. They practice several **45-meter** races. For each race, make a line graph that represents their position by time.

a. Annie and Bonnie start at the starting line (0 meters) at the same time, and each runs at a constant speed. Annie finishes the 45-meter race 2 seconds before Bonnie.

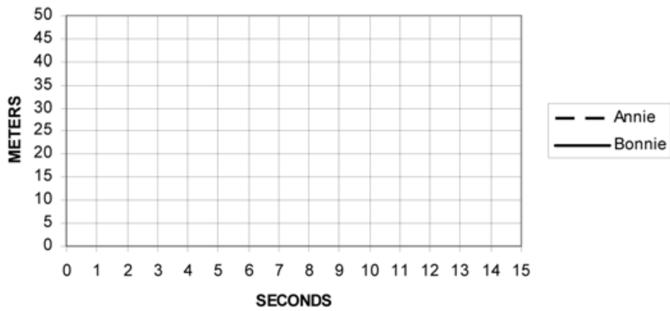
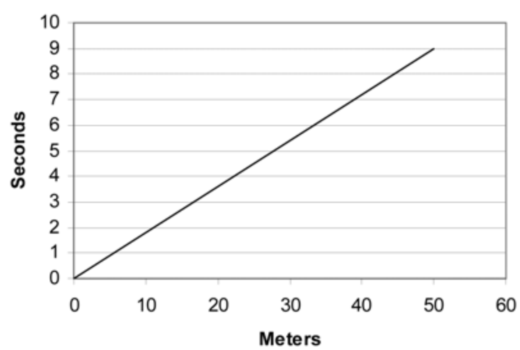


Figure C6. Eighth-grade teacher MKT item (Interpreting unconventional forms or representations that students are likely to make).

Here is a graph of a 50-meter dash that a student made. Notice that distance is on the x-axis and time is on the y-axis.



Which are true statements about the relationship between the line graph and the speed of the runner? *(Choose all that apply)*

A) The slope of the line is 9/50 or .18 as was the average speed in meters per second of the runner during the dash.

B) The slope of the line is 50/9 or 5.56 as was the average speed in meters per second of the runner during the dash.

C) The slope of the line is 9/50 or .18, and the average speed of the runner was 50/9 or 5.56 meters per second.

D) The slope of the line is 50/9 or 5.56, and the average speed of the runner was 9/50 or .18 meters per second.

E) None of the above

Appendix D
Sample Characteristics for Item Data

Table D1. Student Characteristics.

	Seventh Grade Year 1		Eighth Grade	
Variable	Control	Treatment	Control	Treatment
Total count of students	825	796	303	522
Female (%)	50.6	48.9	45.1	47.9
Individual ethnicity (%)				
White	38.7	48.5	65.6	50.0
Hispanic	54.1	44.3	22.7	40.7
Asian	2.0	1.5	1.1	1.3
African American	4.7	4.2	9.5	6.9

Table D2. Teacher Characteristics.

	Seventh Grade Year 1		Eighth Grade	
Variable	Control	Treatment	Control	Treatment
Total count	47	48	23	33
Female (%)	81	77	82.6	84.8
Years teaching total				
Mean	10.5	12.4	9.6	7.9
Range	1–29	1–40	0–27	0–31
Years teaching mathematics				
Mean	9.5	11.0	9.9	8.2
Range	1–29	1–40	0–27	1–32
Teacher ethnicity (%)				
White	70.2	77.1	87.0	78.8
Hispanic	25.5	20.8	8.7	15.1
Asian	4.3	0	0	0
African American	0	2.1	4.3	6.0
Master's degree (%)	17.0	18.8	26.1	6.0

Appendix E
Item Characteristics of the Seventh-Grade Student Assessment

Table E1. Items ordered by difficulty using pretest percent correct (N = 1621, Seventh Grade Year 1). Sample characteristics are reported in Appendix D.

Item Description by Subscale		*p* Pretest All	*p* Posttest Treatment	*p* Posttest Control
M_1				
9a	Basic table reading	.857	.896	.887
2	Unit rate problem with $d = rt$	.754	.810	.821
6	Basic table reading	.718	.832	.778
4	Unit rate problem with $y = kx$	.695	.742	.707
1	Basic graph reading locating a point	.694	.721	.792
7a	Basic graph reading of linear relationships	.675	.936	.770
3	Use the formulat $a/b = c/d$	.622	.672	.652
5	Unit rate problem with $d = rt$	.564	.667	.581
7b	Basic graph reading determining the meaning of a point	.456	.683	.588
M_2				
9c	Translate a table to a graph	.816	.922	.891
7e	Interpret a graph of a proportional function	.722	.830	.764
7c	Basic graph reading determining the meaning of a point	.684	.843	.738
7d	Interpret a graph of a proportional function	.637	.779	.707
9b	Translate a table to an equation	.540	.692	.645
12c	Interpret direction in two piecewise position graphs	.526	.680	.572
8c	Construct a position graph	.469	.813	.537
13	Interpret speeds in a piecewise position graph	.357	.720	.464
14a	Interpret speed and direction in a piecewise position graph	.301	.693	.402
14b	Interpret speed and direction in a piecewise position graph	.252	.661	.338
12b	Interpret direction in two piecewise position graphs	.247	.422	.285
12a	Interpret speed in two piecewise position graphs	.234	.549	.385
14c	Interpret speed and direction in a piecewise position graph	.197	.616	.286
8a	Determine speed from a position graph	.177	.504	.239
10a	Construct values in a table	.179	.362	.244
11	Translate an equation to a graph	.163	.430	.207
14d	Interpret speed and direction in a piecewise position graph	.135	.648	.208
10c	Construct graph from table	.133	.270	.167
10b	Construct equation from table	.108	.241	.175
8b	Determine slope from a position graph	.039	.275	.087
15	Construct narrative about two piecewise position graphs	.028	.138	.052

Appendix F
Item Characteristics of the Eighth-Grade Student Assessment

Table F1. Items ordered by difficulty using pretest percentage correct (N = 825). Sample characteristics are reported in Appendix D.

Item Description by Subscale		*p* Pretest All	*p* Posttest Treatment	*p* Posttest Control
M_1				
8a	On a graph, given an x, find the corresponding y	.834	.866	.911
1	Find an output for an input to a linear equation	.765	.818	.848
4	Translate a table to a graph	.678	.757	.766
8b	Find speed in position graph (b = 0)	.618	.797	.667
2	Translate a narrative to an equation	.584	.611	.587
10a	Translate a graph to a table (motion)	.564	.672	.634
3	Translate a table to a symbolic expression	.545	.653	.653
13	Construct a graph from a narrative (motion)	.392	.563	.492
15a	Construct a position graph from a narrative	.384	.682	.528
11b	Given a table, find the rate (money)	.362	.444	.432
8c	Find speed in position graph (b > 0)	.342	.454	.432
11a	Given several ordered pairs, complete the table (money)	.274	.437	.360
6	Identify a narrative describing a proportional relationship	.258	.481	.373
7	Categorize a function as linear vs. nonlinear	.247	.351	.330
11d	Translate a graph to an equation (money)	.170	.278	.244
10b	Translate a graph to an equation (motion)	.119	.285	.271
11c	Translate a table to a graph (money)	.093	.159	.168
5	Identify a graph of a proportional function	.052	.222	.165
M_2				
8d	In position graphs, compare durations of two functions	.656	.716	.663
14a	Find distance traveled in one segment of a position graph	.507	.613	.597
9	Construct a position graph representing a faster speed than another one	.484	.753	.528
16b	Construct two comparative functions on a position graph	.297	.661	.436
16a	Construct two comparative functions on a position graph	.293	.579	.370
14c	Describe motion in one segment of a position graph (slope = 0)	.287	.625	.436
15b	Construct a two-segment position graph from a narrative	.270	.644	.446
14b	Find speed in one segment of a position graph	.267	.404	.363
16d	Construct two comparative functions on a position graph	.259	.502	.347
14e	Compare distance and time traveled of two position graphs	.256	.437	.376
14d	Describe motion in one segment of a position graph (negative slope)	.251	.540	.363
15c	Construct a two-segment position graph from a narrative	.216	.571	.350
15d	Construct a two-segment position graph from a narrative	.195	.609	.304
17b	Construct a position graph representing the average speed of a multirate graph	.148	.452	.241
16c	Construct two comparative functions on a position graph	.145	.527	.241
17a	Find average speed over a multirate piecewise position graph	.142	.393	.162
12b	Compare rates in a piecewise graph (money)	.079	.188	.145
12a	Compare rates in a piecewise graph (money)	.068	.188	.125

Appendix G

Item Characteristics of the Seventh-Grade MKT Assessment

Table G1. Items ordered by difficulty using percentage correct on the pretest (N = 117). Sample characteristics are reported in Appendix D.

Item Description by Subscale	*p*	Standard Error
Rate and proportionality – M_1		
Story problems to model a proportional function	.368	0.04
Narrative description of a proportional function	.282	0.04
Why cross-multiplication works	.120	0.03
Rate and proportionality – M_2		
Compare proportional functions across representations	.709	0.04
Narrative description of a piecewise linear graph of pay rates	.675	0.04
Evaluate student graphical representation of comparative functions	.675	0.04
Understanding slope when time is on the y-axis	.658	0.04
Representing multiplicative reciprocals	.624	0.04
Evaluate graph conjecture, The shorter the line, the faster the run.	.624	0.04
Mathematically precise language for slope of 0	.624	0.04
Evaluating an unconventional algebraic representation of slope	.590	0.05
Tabular representation of a proportional function	.538	0.05
Story for a piecewise linear graph of changing speeds	.530	0.05
Multiple representations of changing speeds	.530	0.05
Distinguishing multiplicative from additive reasoning	.504	0.05
In d = rt, the relationship between speed and time	.359	0.04
Connecting aspects of a table, graph, and formula of the same function	.188	0.04
Algebraic representation of a proportional function	.154	0.03
Velocity		
Given a velocity graph, choose the position graph	.282	0.04
Narrate motion represented in position and velocity graphs	.256	0.04
Given a velocity graph, draw the position graph	.197	0.04
Given a complex velocity graph, choose the position graph	.188	0.04
Given a complex position graph, draw the position graph	.120	0.03
Given a position graph, draw the velocity graph	.077	0.02

Appendix H
Item Characteristics of the Eighth-Grade MKT Assessment

Table H1. Items ordered by difficulty using percentage correct on the pretest (N = 56). Sample characteristics are reported in Appendix D.

Item Description by Subscale	*p*	Standard Error
M_1		
Model motion on a graph given story	.875	.04
Identifying linearity in a table	.750	.06
Translating linear functions from one representation to another	.679	.06
Distinguish proportional relationships from quadratic and exponential	.518	.07
Proportional reasoning with proportional and linear functions	.482	.07
Identify a graphical representation of a function family for y = kx	.446	.07
Compare different representations of the same function	.321	.06
Proportional reasoning about size and price of object	.304	.06
Representing proportional, linear, and nonlinear functions	.179	.05
M_2		
Distinguishing visual and mathematical properties of slope	.875	.04
Calculate slope of a segment that does not go through the origin	.857	.05
Given a story of motion, choose a piecewise position graph	.821	.05
Evaluate student conjectures about average rate on a piecewise graph of motion	.821	.05
Evaluate student misconceptions about modeling motion	.714	.06
Understanding slope when time is on the y-axis	.679	.06
Model motion that varies in rate and/or starting position with linear function	.661	.06
Given a story of motion, choose a piecewise position graph	.661	.06
Find average rate in a piecewise position graph	.661	.06
Compare two linear functions representing motion (algebraic)	.643	.06
Compare two graphical models of money accumulation	.589	.07
Understanding rate when time is on the y-axis	.554	.07
Given a story, choose a piecewise position graph (distracters are common misconceptions)	.518	.07
Compare two linear functions representing motion (in a story)	.464	.07
Compare two models of motion on a position graph	.429	.07
Identify average rate in a piecewise algebraic function	.411	.07
Describe connections between position and velocity graphs	.375	.07
Identify average rate in a piecewise position graph	.143	.05

Sponsor: The National Science Foundation, Grant REC - 0437861

Prime Grantee: SRI International. Center for Technology in Learning

Subgrantees: Virginia Polytechnic Institute and State University; University of Massachusetts, Dartmouth;
The University of Texas at Austin; and The Charles A. Dana Center at The University of Texas at Austin

SRI International
Center for Technology in Learning
333 Ravenswood Avenue
Menlo Park, CA 94025-3493
650.859.2000

www.ctl.sri.com
www.math.sri.com