Scaling Up SimCalc Project

Extending the SimCalc Approach to Grade 8 Mathematics

Technical Report 02 | December 2007

Jeremy Roschelle, Deborah Tatar, Nicole Shechtman,
Stephen Hegedus, Bill Hopkins, Jennifer Knudsen, Margie Dunn

Report Series Published by SRI International

Research collaborators:

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®

This report is intended to serve as timely means to share findings with Texas teachers and educators, which is necessary for the diffusion phase of our research beginning in June 2007. The research team is simultaneously preparing detailed, scholarly articles for researchers and policy makers, to be submitted to peer-reviewed journals. We prefer that researchers and policy makers wait for and cite the forthcoming peer-reviewed articles. Contact Jeremy.Roschelle@sri.com for more details.

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

Extending the SimCalc Approach to Grade 8 Mathematics

Findings from a second randomized experiment

Prepared by:

Jeremy Roschelle, SRI International
Deborah Tatar, Virginia Polytechnic Institute and State University
Nicole Shechtman, SRI International
Stephen Hegedus, University of Massachusetts at Dartmouth
Bill Hopkins, Charles A. Dana Center at University of Texas at Austin
Jennifer Knudsen, SRI International
Margie Dunn, Rutgers University

Acknowledgments:

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

We thank G. Haertel, 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.

The Effects of Extending the SimCalc Approach to Grade 8 Mathematics

In a second randomized experiment SRI International (SRI) and its partners replicated the approach of their seventh-grade experiment, extending the SimCalc approach to eighth-grade mathematics. We found that eighth-grade students learned more mathematics when their teachers used SimCalc curriculum and software in place of their traditional curriculum.

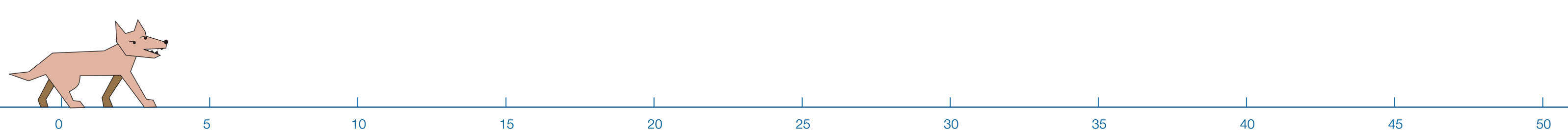
In a series of large-scale randomized experiments in Texas public schools, our research has focused on the question: "Can a wide variety of teachers use an integration of curriculum, software, and professional development to increase student learning of complex and conceptually difficult mathematics?"

In our first report in this series (Roschelle et al., 2007), we reported on an experiment in seventh grade mathematics. We found that wide variety of teachers was able to implement a technology-enhanced replacement unit for the seventh-grade mathematics topic of rate and proportionality and that the students of these teachers learned the topic deeply. Because we randomly assigned teachers to use either a SimCalc replacement unit or their existing curriculum, our experiment provided strong evidence that the SimCalc intervention caused students to learn more mathematics. We argued that the learning gains were important by describing the central role of rate and proportionality in the middle school mathematics curriculum, and we showed that students gained in particular on the more advanced aspects of this topic. Given this positive result, a question naturally arose: "How can we take this to the next level of scale?"

Two aspects of scaling up are covering more grade levels and providing more content. Our second experiment, reported here, implemented an eighth-grade intervention based on the SimCalc approach. We extended that approach by using the same SimCalc MathWorlds™ software but with a different replacement unit curriculum. The new replacement unit targeted the eighth-grade topic of linear functions, which is the next step in the natural sequence from our seventh-grade focus on rate and proportionality. Although the curriculum and teacher professional development were necessarily different for the eighth-grade, they were designed according to the same general principles and in the same style and format as for the seventh grade experiment. By once again comparing an integration of curriculum, software, and teacher professional development with a business-as-usual condition, we accumulated further evidence for the robustness of our approach in an additional grade level. Further research at the James J. Kaput Center for Research and Innovation in Mathematics Education at the University of Massachusetts, Dartmouth is extending the curricular scope of SimCalc to Algebra I and Algebra II.

Another aspect of scaling up is building capacity to train more teachers. In the eighth-grade experiment, we extended the SimCalc approach to a "train-the-trainers" model whereas in the seventh-grade experiment, the curriculum developer led all the teacher workshops. In the eighth-grade extension, the curriculum developer trained teacher educators from regional Education Service Centers[1] (ESCs) throughout Texas; these teacher educators then led the teacher professional development workshops. This train-the-trainers model is common in Texas and can more easily reach large numbers of teachers.

[1] Educational Service Centers are public regional organizations that offer educational support programs to districts throughout the state of Texas.



Below, we report on a main effect in the eighth-grade experiment that was consistent with the main effect in the seventh-grade experiment: in both cases, students of teachers who used SimCalc materials learned more. We also report on two similar moderating variables—the cognitive complexity of teaching goals and the allocation of classroom time to allow students to work with the software. The consistency of the results of our eighth-grade and seventh-grade experiments should increase educators' confidence in the value of the SimCalc approach for enabling teachers in a wide variety of contexts to enhance the mathematics learning of their students.

Addressing a Focal Point of Eighth-Grade Mathematics

The National Council of Teachers of Mathematics (NCTM) has elaborated its earlier recommendations for learning and teaching mathematics to give teachers additional guidance on the most important ideas and major themes of mathematics at each grade level. This elaboration fits the NCTM's curricular principle, which states "a curriculum is more than a collection of activities: it must be coherent, focused on important mathematics, and well articulated across the grade" (NCTM, 2000, p. 14). In particular, the new recommendations list "focal points" for each grade level. The first of three focal points for the eighth-grade is in the algebra strand and states:

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 (NCTM, 2007, p. 20).

The eighth-grade replacement unit we tested aligned with this focal point. Our unit focused on linear and proportional functions, emphasizing that the slope of a line represents a constant rate of change. In the replacement unit, students translated among verbal, tabular, graphical, and algebraic representations, and were introduced to how slope and intercept appeared in different representations. Our unit did not address the concept of "systems of linear equations", which is not considered appropriate for eighth-grade mathematics in Texas. Although students did not use systems of linear equations, they did analyze and compare pairs of linear and piecewise linear functions—a step toward using such equation systems. The replacement unit similarly aligned with relevant portions of the Texas standards for eighth-grade mathematics. In general, the concept of linear function was seen as a crucial transitional topic between earlier mathematics and algebra.

Like the seventh-grade experiment, the replacement unit intervention comprised a paper-based curriculum unit (in both student and teacher versions), the SimCalc MathWorlds software with custom software documents, and teacher professional development.

The eighth-grade replacement unit used the theme of designing cell phone games to address the target mathematics. Students were given roles as designers of electronic games who must use mathematics to make the games operate. Linear functions of the form y = mx + b were developed as models of motion and accumulation. Amounts and rates in the materials were realistic for the specific contexts within this theme. Students learned to use different representations of these functions for problem solving and the connections among representations were emphasized. The mathematics in the unit went beyond standard eighth-grade content to include complex average rate problems. For example, the familiar class of difficult problems in which trains or other vehicles leave a station and travel at differing rates can be more easily conceptualized using graphs.

The unit was designed to be used daily over a 2- to 3-week period, replacing regular lessons on linear functions. The teacher guide provided lesson plans that teachers could adapt, as well as guidance about possible student responses to questions and tasks. Daily access to a sufficient number of computers was required—enough so that students could interact directly with the software individually or in small groups. A student workbook sequenced all the problems and questions and provided space for student responses. The workbook linked to specific software documents that preloaded activities for the students. Activities made use of important SimCalc MathWorlds features, including:

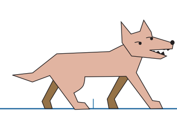
- Visualizing mathematical functions as motions.
- Showing more than one function at a time to allow comparison.
- Editing mathematical functions in both graphical and algebraic forms.
- Linking the verbal, graphical, tabular, and algebraic representations of a function.
- Using piecewise linear functions to model situations with changing rates.

Teachers were trained to use the materials in two workshops. In a 3-day summer workshop, teachers worked through the SimCalc materials as learners, practicing with the software and discussing implementation issues. The facilitator demonstrated appropriate pedagogy. Then, in the fall, teachers gathered for a 1-day planning workshop in which they adapted the lesson plans to fit their classroom contexts but without modifying the materials or the activity sequence of the materials.

Scaling Up via a "Train-the-Trainers" Approach

In the interest of scaling up, we used a train-the-trainer model to deliver preparatory workshops. This is closer to normal practice than the training used in the seventh-grade experiment. In fact, the train-the-trainer model is widely used when teachers are introduced to a new curriculum. In previous work in Texas, our partners at the Charles A. Dana Center at the University of Texas at Austin had used train-the-trainer workshops to introduce their TEXTEAMS materials to middle school mathematics teachers across Texas.

In implementing the train-the-trainer model, we were able to take advantage of our existing relationships with educators in Texas. We gathered a group of experienced professional development leaders, each of whom had responsibility for serving teachers in a particular region of the state. The leaders were already familiar, to some extent, with the SimCalc approach because they had recruited teachers for the previous experiment. Over 2 days, the curriculum developer and a well-known professional development leader taught the soon-to-be trainers the content of the unit. The workshop focused on what we wanted them to do in teacher training workshops, but their experience with the materials was compressed so that more time could be devoted to discussing how teachers might react to the materials and how to deepen their understanding of the mathematical



goals of the program. Over the summer, the trainers led workshops for the teachers in their regions.

Though widely used in a variety of fields, information about the train-the-trainer model and its validity as a training method in any discipline is scant (Orfaly et al., 2005). The train-the-trainer model is endorsed in the field of Public Health as a means of sustainability because the trainers are members of the community rather than experts without community ties. Although the model has the potential for dilution of content or lack of fidelity to the initial premise, studies in the Public Health field demonstrate that such defects do not generally occur (Orfaly et al., 2005). Literature addressing the train-the-trainer model seems to be even sparser in the field of education. A literature review found studies relying on surveys and interviews, but no studies used data collected through direct observation of all levels of training and subsequent posttraining enactments (Bahr et al., 2006; Griffin, 1997; Wildy, Wallace, & Parker, 1996).

In a survey of teacher professional development research, Borko (2004) describes studies that look at a single professional development program enacted by more than one facilitator at more than one site as "Phase 2 research." The central goal of Phase 2 research, according to Borko, is to determine whether or not a program can be enacted with integrity, especially as it becomes further removed from the original professional development providers. "Integrity" in this case does not imply rigid implementation of required activities; rather, it is important to investigate the balance and tradeoffs between fidelity and adaptation. Her literature review yielded no research that provided adequate evidence of such integrity in a Phase 2 project. She did find a small number of Phase 2 projects with widespread enactments, but research on those projects focused almost exclusively on the professional development conducted by the original design teams.

Research Design

Our research question concerned two key aspects of extending our research from the seventh to the eighth-grade:

Can a wide variety of teachers increase student learning of mathematics when (a) they are prepared to use SimCalc materials via a train-the-trainers model, and (b) they implement a SimCalc replacement unit that integrates curriculum and software to address focal topics of eighth-grade mathematics?

To address these questions, SRI and its partners led a randomized experiment whose rigorous design was largely parallel to that of the seventh-grade experiment. The experiment began in summer 2006. Teachers were randomly assigned to participate in either a Treatment or a Control group. The Treatment group received the SimCalc intervention, which began with a 3-day teacher professional development workshop in which teachers learned to teach using the eighth-grade SimCalc unit. Treatment teachers were then asked to teach the replacement unit in place of their usual unit on linear function.

The counterfactual or Control condition was designed to allow comparison between classrooms using the SimCalc unit and classrooms using their regular linear functions materials. In addition, to make sure participation in the study was fair for both groups, teachers in the Control group received professional development on the integration of technology into mathematics teaching, but pertaining to different eighth-grade content. These teachers were assigned to the Teaching Mathematics TEKS Through Technology (TMT3) workshop, a high-quality professional development program with an emphasis on statistics that is offered throughout Texas. Each teacher in both groups received a $650 stipend for his or her completed participation.

The designs of the seventh- and eighth-grade studies differed in two key respects:

1. Whereas in the seventh-grade experiment teachers were asked to participate for 3 years, the eighth-grade study took place over 1 year.
2. Whereas in the seventh-grade experiment Control teachers were offered training in the mathematics that would be assessed but no training in integrating technology, teachers in the eighth-grade experiment received training in integrating technology into mathematics teaching, but with respect to a different mathematical topic.

For each teacher, data were collected for one "target class," which the research team randomly chose from the teacher's roster.

Participants

With the cooperation of five Texas ESCs, we recruited teacher volunteers whose students reflected the regional, ethnic, and socioeconomic diversity of the state. Recruitment for the eighth-grade study did not occur in schools already participating in the seventh-grade experiment; therefore, none of the students participating in the eighth-grade unit had studied the seventh-grade unit. If the replacement units had been offered to the same students in the seventh and eighth-grades, students might have been better prepared to learn from the more difficult eighth-grade unit, making it difficult to tease apart the effect of the eight-grade unit.

Teachers were randomly assigned by school either to the Treatment or Control group. After receiving invitations, 63 teachers from our applicant pool were able to attend and complete a summer workshop. Eventually, 56 teachers (and their 825 students) returned complete data. At intake, the Treatment group (33 teachers) and the Control group (23 teachers) did not differ in any important way (e.g., with respect to teaching experience, ethnicity, gender, mathematical content knowledge, or percent of students eligible for free or reduced-price lunch in school). The greater number of teachers in the Treatment group was an artifact of teachers' scheduling conflicts with the workshops to which they were assigned. Because teachers were not informed about the workshop type until the workshop actually occurred, the consequences for randomization and thus the validity of the experiment are minimal. The attrition rate was comparable to other large experiments with educational technology (Dynarsky et al., 2007) and we have no evidence that would suggest differential attrition, which would be the principal threat to validity.

Assessment of Student Learning

The primary outcome measure in this experiment was student learning of core mathematical content. Our team designed an assessment that was administered in a single class period to students before and after their linear function unit was taught.

Working with a panel of mathematicians and mathematics education experts, we developed an assessment blueprint that encompassed both simple and more complex aspects of linear function. The content which is outlined in Table 1, was aligned to the eighth-grade Texas state standards (the Texas Essential Knowledge and Skills—TEKS). The simple aspects addressed content pertaining to linear function covered in the Texas state test (the Texas Assessment of Knowledge and Skills—TAKS). The complex aspects addressed content aligned with the NCTM's Focal Points for Grade 8.

	M1—Conceptually Simple	M2—Conceptually Complex
Overview of Concepts	Concepts are typically covered in the grade-level standards, curricula, and assessments.	Building on the foundations of M1 concepts, the concepts constitute more complex building blocks for the mathematics of change and variation found in algebra, calculus, and the sciences.
Seventh-Grade Study (focus on rate and proportionality)	• 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 of more pairs of values that are given or implied. • Reasoning across two or more representations.
Eighth-Grade Study (focus on linear function)	• 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 piecewise linear function.

Table 1. Topics found in the student assessments. These were also the topics covered in the replacement unit.

The simpler items were based on those used on the TAKS for the eighth-grade; these items typically ask students to calculate using a linear relationship stated either as a word problem or more mathematically. For example, one question was:

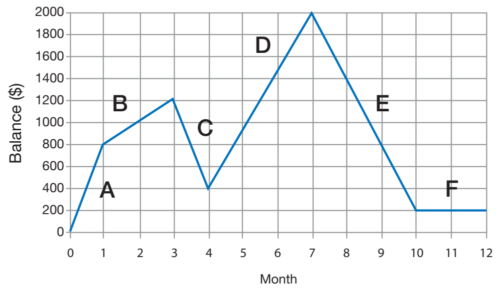
Deondra is saving to buy a new bike. She starts off with $10 in her account and saves $20 a month babysitting. She writes an equation that represents the total amount of money (y) she has saved in any given month (x): y = 20x + 10. How much money has Deondra saved after 6 months?

Students could answer this question by substituting 6 for "x" in the given equation and computing the result, $130. Many students were also able to compute the answer in their heads without using the equation by noticing that 6 times 20 is 120, and then adding 10 more. In either event, the calculation drew on skills aligned with our framework for conceptually simple math.

Other simpler items asked students to choose graphs that represented a proportional relationship or to choose equations that did not represent a linear relationship (e.g., one equation had an x^3 term in it and so was not linear). Other simple items required only simple cognitive tasks, like finding the graphed line that contained a set of points given in a table.

More complex items on our test required comparing multiple rates or finding average rates. One item showed a line graph of the balance in a bank account by month and asked in which month the balance increased at the highest rate (see Figure 1). This item required students to make sense of the "highest rate" within a graph that displayed conceptual distracters, such as the highest balance. Thus the students had to distinguish rate from amount. Further, the rate information was not provided directly; students had to infer it from the slopes in the graph. In addition, the item had both positive and negative slopes; students had to understand the meaning of these directions in the graph.

The graph below shows the balance in Eva's bank account over the course of a year. Use the graph to answer the following questions.



a. In which period did Eva's balance INCREASE at the highest rate?

☐ A ☐ B ☐ C ☐ D ☐ E ☐ F

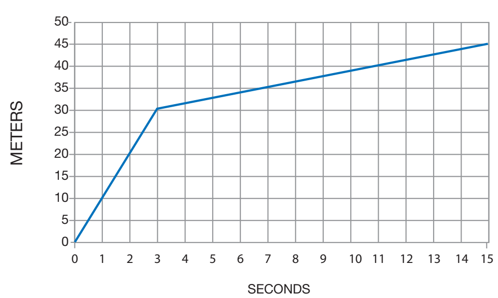
a. In which period did Eva's balance DECREASE at the highest rate?

☐ A ☐ B ☐ C ☐ D ☐ E ☐ F

Figure 1. Advanced item requiring students to identify the fastest rate.

Another, more advanced item showed a graph with two piecewise linear segments, depicting a faster rate followed by a slower rate (see Figure 2). Students were asked to compute average speed and draw a graph representing another motion moving at the average speed for the same length of time.

Juanita ran a 45-meter race as represented in the graph below.



a. What was Juanita's average speed?

b. Vanessa ran the same race. She ran the entire race at Juanita's average speed. Sketch Vanessa's race on the graph above.

Figure 2. Advanced item requiring students to represent the average speed.

The overall 36-item test included 18 simple and 18 complex items. After completing these items, students were asked to answer a few demographic and attitude questions. We carried out rigorous validation processes for the test. This included expert panel reviews for grade-level appropriateness and alignment with Texas state standards and our target content. We also conducted cognitive interviews with students, as well as an item response theory analyses of field test data collected from a large sample of students.

Teacher Measures

In addition to the student assessment, key teacher measures included: (1) an assessment of teachers' knowledge of the mathematics necessary to teach the unit; (2) a questionnaire about each teacher's background, attitudes, and beliefs; (3) a daily log in which teachers provided a structured report of their implementation of the unit; and (4) a teacher retrospective log about the unit as a whole. In addition, demographic data about each participating school were drawn from a database maintained and published by the state of Texas.

Procedure

The timeline for the experiment follows: During summer 2006, Treatment (SimCalc) and Control (TMT3) workshops were conducted in each of the participating regions and all teachers attended a 3-day workshop. Then, in fall 2006, Treatment teachers attended a 1-day planning workshop. At the beginning of the school year, teachers received a box that contained all of the materials they would need, both for teaching and research. Then, during the 2006-2007 school year, teachers taught their assigned units. Student assessments were administered before and after teaching the unit. Measures to assess teachers' mathematics knowledge were administered at the beginning of the summer workshop and after the teachers completed teaching the unit. Teachers also filled out their daily logs during the unit. After completing data collection, they mailed the required materials to the research team. On receipt of the completed materials, teachers received their stipends.

In executing the design, the research team took particular care to avoid any suggestion that one group might be advantaged or perform “better” than the other, and our scientific advisory board reviewed the design to ensure that such suggestions had been avoided. With this concern in mind, the PIs and co-PIs also conducted detailed reviews for bias of each video presentation—the method we used, rather than in-person presentations, to inform teachers about the research design—for the Treatment and Control group teachers and the recruitment partners. The two teacher groups were shown essentially the same video presentations, although the videos for the two groups did differ in regard to a few details to reflect the slightly different research procedures.

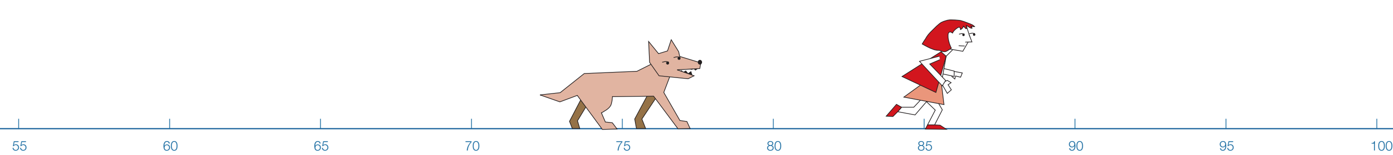
Data Entry, Cleaning, and Analysis

The complex data entry process included careful tracking for completeness and multiple crosschecks for accuracy. Student assessments were scored by a team who had been trained to score to almost perfect agreement. To maintain scorer accuracy throughout the process, a random sample of 10% of the student assessments was double-scored. Computer scanning was used to input scorers’ records for each test. Computer scanning was also used for each of the teacher daily log pages. All other data were entered by hand. On completion of the collection and entry of all data (i.e., student assessments, teacher assessments, teacher questionnaires, teacher logs), a random sample of 5–10% of each data set was compared with the original data to check for accuracy. All data were also subjected to thorough checks for reasonableness of range and distribution. The error rate across all of these checks was negligible.

Because sampling was at the level of intact classrooms (clusters of students), classic statistical models such as the t-test or multiple regression models would have been inappropriate without modification. Accordingly, hierarchical linear modeling (HLM) was employed to estimate the effects of the treatment (Raudenbush & Bryk, 2002). HLM accounts for measurement and sampling error at both the student level and the classroom level, resulting in correctly adjusted standard errors for the treatment effect. Once a main effect is established, HLM enables analysis of the impact of additional student-level or classroom level factors.

Characteristics of the Sample

Tables 2 and 3 (on the next page) show how the teachers and schools varied in this study. In any randomized experiment, it is important to verify whether the assignment procedure resulted in any differences between groups. We found no significant differences between groups on any of these variables.



	Control	Treatment	Total Sample
Total count	23	33	56
Teacher gender, percent			
Female	82.6	84.8	83.9
Male	17.4	15.2	16.1
Years teaching total (mean)	9.57 Range: 0–27	7.91 Range: 0–31	9.59 Range: 0–31
Teacher ethnicity, percent			
Caucasian	87.0	78.8	82.1
Hispanic	8.7	15.2	12.5
African-American	4.3	6.1	5.4
Teacher age (mean)	42.0 Range: 25–62	41.0 Range: 27–64	41.0 Range: 25–64
Percent with master's degrees	26.1	6.0	14.3
Percent by Texas ESC region			5
Wichita Falls (Region 9)	9.1	17.4	12.5
Dallas (Region 10)	24.2	17.4	21.4
Austin (Region 13)	39.4	43.5	41.1
Lubbock (Region 17)	18.2	21.7	19.6
Midland (Region 18)	9.1	0.0	5.4

Note: There were no statistically significant differences between groups for any of these variables.
Table 2. Teacher-level characteristics of the sample.

	Control	Treatment	Total Sample
Total count of schools	19	24	53
Percent free or reduced-price lunch (mean)	43.0	42.4	42.2
School size (mean)	643 Range: 104–2,245	634 Range: 121–1,375	638 Range: 104–2,245

Note: There were no statistically significant differences between groups for any of these variables.
Table 3. School-level characteristics of the sample.

Research Findings

The design of this experiment allowed a comparison of student learning gains in two conditions. Teachers of students in the Treatment group were prepared to use SimCalc materials via a train-the-trainers model and implemented a SimCalc replacement unit that integrated curriculum and technology. Teachers of students in the Control group went to a workshop on integrating technology in their teaching but taught with their existing linear function curriculum materials. The main effect in the eighth-grade experiment was statistically significant; students in the Treatment group learned more than students in the Control group. The overall effect size was 0.79 (see Figure 3), which is considered large in education studies ($z = 5.38$, $p<0.0001$, using a two-level hierarchical linear model with students nested within teacher). The difference between the groups occurred mostly on the complex portion of the test. The effect size of treatment on this portion was 1.27 ($z = 7.62$, $p<0.0001$). The effect size of the treatment on the simple portion was 0.20 ($z = 1.6$, $p = 0.11$, n.s.).

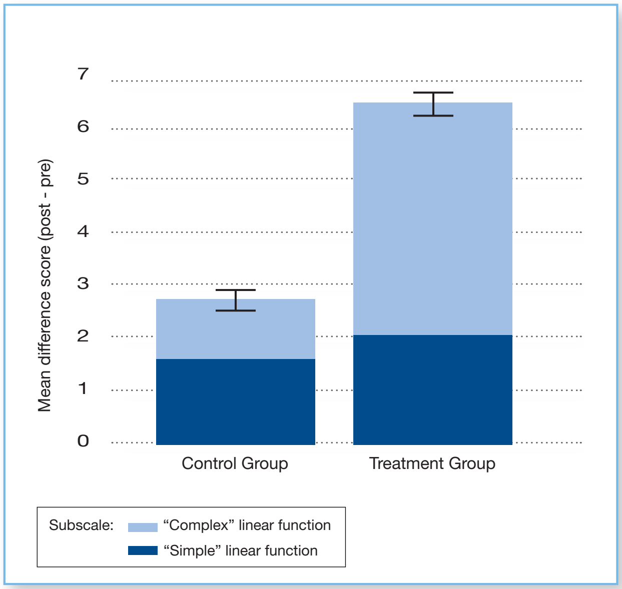


Figure 3. Student gain scores aggregated by teacher. Mean difference scores (± SE of total) on the 36-item student assessment.

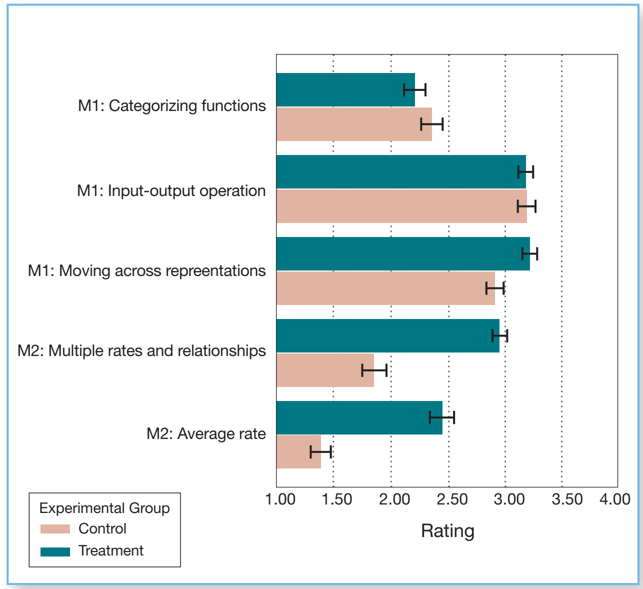


Figure 4: Teacher self-report of topic coverage. Average of daily ratings on a 4-point Likert scale (± SE).

As was the case in the seventh-grade unit, we found that the teachers reported more cognitively complex teaching goals in the Treatment group. Further, teacher report of more complex teaching goals was related to student learning gains. In the daily log, teachers rated their degree of focus for the day’s class on goals of low-cognitive complexity (memorization and use of routine procedures) and high-cognitive complexity (communicating conceptual understanding; making mathematical connections and solving nonroutine problems; and conjecturing, generalizing, or proving). For high-order goals, Treatment teachers (versus Control) reported a stronger daily focus ($t(54) = 4.1$, $p<0.0001$) and the overall statistical association with classroom mean student gain on the complex mathematical subscaleswas positive ($ß = 1.7$, $p<0.01$).

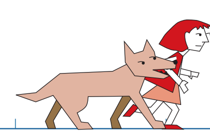
We also collected teacher self-report data on topic coverage. The daily log asked teachers “To what extent did you and your class focus on the following topics?” and listed the following topics aligned with the conceptually simple content of linear function:

- **Categorizing Functions** as nonlinear or linear, proportional or nonproportional.
- **Using Input-Output Operations:** using equations to find unknowns, completing tables, or finding points on a graph.
- **Moving across Representations:** generating one representation (e.g. equations, tables, graphs), given a different representation.

In addition, the daily log listed the following topics aligned with the conceptually complex content of the unit:

- **Multiple Rates and Relationships:** comparing varying rates within the motion of one object or comparing rates across different speeds or prices.
- **Average Rate:** using graphs to find the average rate of an object that changes speed while moving.

Teachers responded on a 4-point Likert scale from “not at all” (1) to “a major focus” (4), and the score for each teacher on each scale was computed by computing the average across their daily reports (see Figure 4). Teachers who used SimCalc covered the first three topics to the same degree as Control group teachers but covered the latter two topics more often. These data suggest that the SimCalc approach achieves its main effect in part by covering



more content within the same unit. This is consistent with SimCalc program philosophy that dynamic representation software allows a greater depth and range of mathematics to be coherently addressed within each classroom activity.

As shown in Figure 5, the teacher daily log also asked teachers to report on the activities in which they and the class engaged during the class period. These activities included whole-class lectures, teacher demonstrations, whole-class discussion, individual student work, student pair work, and student small group work. Each day the teachers checked off which of the activity types they used with their target class. We later counted the number of days students engaged in each activity.

In the Treatment group, there was a negative relationship between the number of days in which whole-class lecture was used and how much complex mathematics students learned ($r(33) = -0.35, p<0.05$), and a positive relationship between the number of days in which students engaged in individual work and how much complex mathematics students learned ($r(33) = 0.64, p<0.0001$). For all the other activity types (teacher demonstrations, whole-class discussions, student pair work, and student small group work), the relationships between the number of days and student gains were not statistically significant. These findings show that students using the SimCalc unit learned more complex mathematics when they engaged more frequently in individual work, whereas learning more complex mathematics was reduced when more time was spent on lectures (see Figure 5). This finding is consistent with the SimCalc program philosophy that individual student work on activities, using both workbooks and software, is important to learning. In forthcoming case studies, we may be able to more fully identify the roles of these different activities in learning.

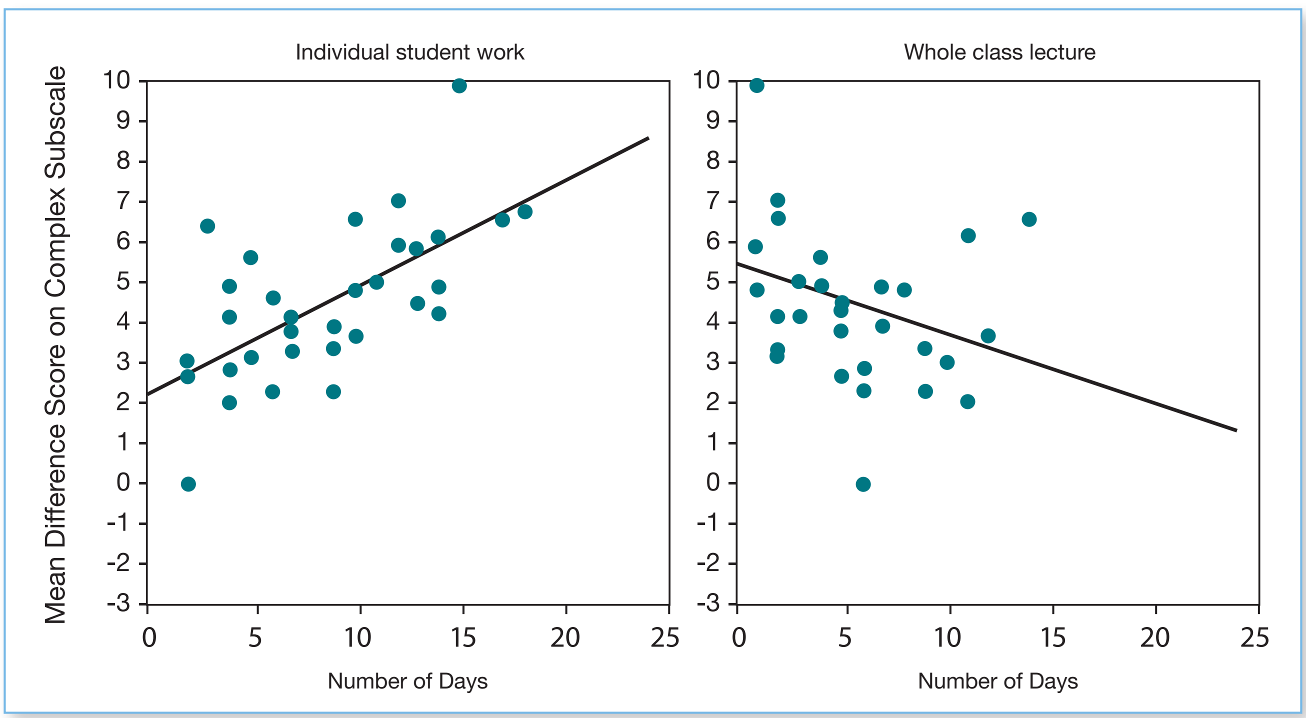


Figure 5. For Treatment teachers only, teachers' report of how many days students engaged in each activity structure, in relation to overall mean student gains on the complex subcale in their classroom.

Discussion

In our eighth-grade experiment, we found that students learned more from a SimCalc replacement unit than they learned from their existing curriculum. Our findings parallel the findings of our seventh grade experiment in several notable ways. Students learned more of the advanced aspects of the target concepts; for the simpler aspects, students of teachers who used the SimCalc replacement unit showed a nonsignificant trend toward greater gains. Teachers who used SimCalc reported more complex teaching goals and students of teachers who reported such goals learned more. Teachers who used SimCalc covered some topics to the same degree as Control teachers, but covered advanced topics more frequently. These findings are consistent with the SimCalc program philosophy, which aims to democratize access to advanced mathematics.

Our eighth-grade experiment went beyond the seventh grade experiment in two notable ways. First, we addressed content that is appropriate for eighth-grade. This content aligns closely to a recognized "focal point" for eighth-grade instruction and sets the stage for Algebra. The clear importance of this content makes our findings more noteworthy. Second, we employed a train-the-trainer model for delivering training to teacher. Relative to our seventh grade experiment, the use of the train-the-trainer model brings an additional realistic element of "scaling up" into consideration in our research. The robustness of the results under a procedure where different teacher educators delivered teacher training adds credibility to the case that SimCalc materials can be used effectively at scale. Dunn (2007) is further analyzing similarities and differences among participating teachers, both in terms of their perceptions and their actual classroom enactments. Dunn will seek to determine whether or not a relationship exists between these elements and the workshop attended.

As in any experiment, these findings should be interpreted with caution. First, the gains applied to more complex and conceptually difficult mathematics and were enhanced when teachers reported more complex teaching goals. Consequently, schools may not see benefits unless teachers have more complex teaching goals and unless they assess more complex reasoning. Second, the results were obtained in Texas, a state with a long record of a stable standards-based educational system. Results may vary in states with different contexts. Third, although we view replacement units as a good strategy to fit within school constraints, the tested replacement units occupied only a modest amount of instructional time. We do not yet know the consequences of more extended uses of such units and do not necessarily recommend using software every day; software use may be most useful when targeted specifically at the most complex and conceptually difficult aspects of mathematics learning. Fourth, our sample lacked a majority African-American school. Fifth, we worked with volunteer teachers and do not know how well nonvolunteer teachers would fare. Sixth, we required schools to have a computer laboratory; however, not all schools have suitable computer facilities.

Although these studies were not designed to isolate particular features of the intervention (e.g., the software, the student printed text, the teacher professional development), two features of this experiment encourage us to conclude that the software—and in particular student use of the software—are important. We gained confidence that the software was important because it was a salient common element in both the seventh- and eighth-grade studies (although the principles and approach to the design of the other materials were also common between the two interventions). Further, the teacher self-report data suggest that when teachers lecture less and allow the students to use the software more in individual work, students learn more. However, care should be taken in interpreting this finding. We are not arguing for a flawed discovery-learning paradigm in which students figure out difficult mathematics concepts on their own (Kirschner, Sweller, & Clark, 2006). Instead, we view student use of the software and teacher explanations and teacher-led discussions as complementary activities (Lobato, Clarke, & Ellis, 2005). We suspect that students can learn more from teacher-led presentations and discussions when they have had direct experience with the software, as in a preparation for future learning paradigm (Bransford & Schwartz, 1999).

In conclusion, the results of this second experiment encourage further scaling up of the SimCalc approach. Several kinds of further efforts might be contemplated. As mentioned earlier, additional research is under way to continue the learning progression into high school Algebra I and II courses. Scaling across grade levels is also possible by using related dynamic mathematics software products such as The Geometer's Sketchpad™, Cabri Geometry™, TinkerPlots™, and Fathom™.In such efforts, educators should take note that our research has found effects from the integration of three elements—curriculum, teacher training, and software—and not just from the software alone. Finally, scaling up to implementations that provide the opportunity for students to experience the SimCalc approach in several consecutive years is also a possibility and one that might result in considerable cumulative gains.

References

1. Bahr, M., Walker, K., Hampton, E., Buddle, B., Freeman, T., Ruschman, N., et al. (2006). Creative problem solving for general education intervention teams: A two-year evaluation study. *Remedial and Special Education,* 27(1), 27-41.

2. Borko, H. (2004). Professional development and teacher learning: Mapping the terrain. *Educational Researcher,* 33(8), 3-15.

3. Bransford, J. D., & Schwartz, D. L. (1999). Rethinking transfer: A simple proposal with multiple implications. In A. Iran-Nejad & P. D. Pearson (Eds.), *Review of research in education* (Vol. 24, pp. 61-100). Washington, DC: American Educational Research Association.

4. Dunn, M. (2007). Effects of employing a train-the-trainer model of professional development in the scale-up of a technology-rich innovation. Unpublished doctoral dissertation in preparation, Rutgers University, New Brunswick, NJ.

5. Dynarski, M., Agodini, R., Heaviside, S., Novak, T., Carey, N., Campuzano, L., et al. (2007). *Effectiveness of reading and mathematics software products: Findings from the first student cohort.* Washington DC: National Center for Educational Evaluation.

6. Griffin, M. L. (1997). A case study of transfer of learning activities in a trainer of trainers program. Unpublished doctoral dissertation, University of Wisconsin-Madison.

7. Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist*, 41(2), 75-86.

8. Lobato, J., Clarke, D., & Ellis, A. B. (2005). Initiating and eliciting in teaching: A reformulation of telling. *Journal for Research in Mathematics Education,* 36(2), 101-136.

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

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

11. Orfaly, R. A., Frances, J. C., Campbell, P., Whittemore, B., Joly, B., & Koh, H. (2005). Train-the-trainer as an education model in public health preparedness. *Journal of Public Health Management and Practice, 11* (Suppl.), S123-S127.

12. Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods.* Thousand Oaks, London, New Delhi: Sage Publications.

13. Roschelle, J., Shechtman, N., Hegedus, S., Pierson, J., McLeese, M., & Tatar, D. (2007). Cognitive complexity in mathematics teaching and learning: Emerging findings in a large-scale experiment. Manuscript submitted for publication.

14. Wildy, H., Wallace, J., & Parker, L. (1996). Decentralising curriculum reform: The link teacher model of in-service training. *School Organisation,* 16(1), 17-28.

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