Helping Students Make Sense of Graphs: An Experimental Trial of SmartGraphs Software

Andrew Zucker, Rachel Kay & Carolyn Staudt

The Concord Consortium
25 Love Lane, Concord MA 01742

October 2013

Accepted by the *Journal of Science Education and Technology*

(the final publication is available at link.springer.com)

Corresponding Author:
Andrew Zucker (azucker@concord.org, 978-405-3231)

Abstract: Graphs are commonly used in science, mathematics, and social sciences to convey important concepts, yet students at all ages demonstrate difficulties interpreting graphs. This paper reports on an experimental study of free, Web-based software called SmartGraphs that is specifically designed to help students overcome their misconceptions regarding graphs. SmartGraphs allows students to interact with graphs and provides hints and scaffolding to help students, if they need help. SmartGraphs activities can be authored to be useful in teaching and learning a variety of topics that use graphs (such as slope, velocity, half-life, and global warming). A two-year experimental study in physical science classrooms was conducted with dozens of teachers and thousands of students. In the first year teachers were randomly assigned to experimental or control conditions. Data show that students of teachers who use SmartGraphs as a supplement to normal instruction make greater gains understanding graphs than control students studying the same content using the same textbooks, but without SmartGraphs. Additionally, teachers believe that the SmartGraphs activities help students meet learning goals in the physical science course, and a great majority reported they would use the activities with students again. In the second year of the study, several specific variations of SmartGraphs were researched to help determine what makes SmartGraphs effective.

Keywords: computers, software, physical science, motion, scaffolding, graphs

Background

In recent decades "graphs and other mathematical representations have come to play an increasingly important role in mathematical activities in schools" (Monk, 2003). Graphs are commonly used in science, mathematics and social science classrooms to convey important concepts. The Common Core State Standards for Mathematics and the Next Generation Science Standards emphasize the importance of students learning to use and understand graphs. Literate adults are also expected to understand and use graphical representations of data (Murray, Kirsch, & Jenkins, 1997). Nonetheless, teaching students to understand graphs is a significant challenge.

Graph Comprehension

The research literature about graph comprehension is voluminous. However, a few key findings, presented here, provide a rationale for the development of new graph-related software, SmartGraphs, as well as for the study of that software reported in this paper.

Compared to tables of numbers, or text, graphs summarize a large quantity of information in a compact manner. Moreover, graphs are able to represent continuous change and co-variation visually and topographically in ways that tables cannot. Knowing that graphs communicate data and relationships so economically, scientists, mathematicians, and others who need to represent relationships between or among variables use them often. Graphs are commonly included in scientific journal articles, for example (Roth, Bowen, & McGinn, 1999), and in lectures in college science classes (Bowen & Roth, 1998).

Yet in spite of the central role of graphs, students at all ages demonstrate difficulties interpreting graphs (Woolnough 2000; van Zee & McDermott 1987). Understanding graphs is not a cognitive ability, per se; instead it is a skill learned through practice (Roth & McGinn, 1997). As is the case with learning to read or to write, students need to repeatedly explore, create, manipulate, and explain graphs in order to learn how to use them well (Monk, 2003). Understanding graphs and other visual representations is called "graphicacy" by some researchers (Friel & Bright, 1996; Roth, Pozzer-Ardenghi, & Han, 2005). Graphicacy must be taught to students.

Graphs are paradoxical because research shows that "although the visual aspect—the visuality—of graphs makes them an extremely rich and powerful medium for generating meaning, it is also the source of many of the incorrect responses students have to graphs" (Monk, 2003). For example, two lines on a graph representing the *speed* of two moving objects over time may intersect, and students may erroneously conclude that the graph represents position, and thus that the two objects are in the same place at the time represented by the intersection. In effect, people often assume that graphs are literal pictures (so that a rising line may be interpreted as a hill) (Leinhardt et al., 1990). There are other common misconceptions about graphs, too, such as a preference for linear relationships, which seem to some students to be more "exact" than non-linear relationships (see Leinhardt et al., 1990 for a discussion of common misconceptions).

Contributing to the challenge of learning to understand graphs, teachers and textbooks often provide inadequate explanations of graphs. For example, a comparison of high school science textbooks and scientific journal articles (Roth, Bowen, & McGinn, 1999) found that "scientific journals provided more resources [e.g. captions and callouts] to facilitate graph reading and more elaborate descriptions and interpretations of graphs than the high school textbooks." The study concluded, "Scientists not only have more training and experience than high school students but, ironically, are provided with more resources for constructing specific meanings from graphs than high school students receive while reading their biology textbooks."

Graphs used in science classes can be viewed as an intermediate form between situations (such as scientific experiments) and the relationships between variables in those situations. Yet teachers and textbooks too often focus on the situations, or on the relationships, without specific reference to the graphical representations as intermediate forms, as though the meaning of graphs were transparent. This presents science learners with a "double problem" because "they neither know the natural phenomenon, nor have they constructed the graph as a sign object" (Roth, 1998). As a result, researchers have found that even undergraduate science students experience difficulty understanding graphs—graphs that instructors have learned to understand through years of practice, but which have become transparent (obvious) to them (Bowen & Roth, 1999).

Because understanding graphs is difficult, many students' knowledge of graphs is brittle and easily compartmentalized by school subject. As one early synthesis of research reported, "often, students who can solve graphing or function problems in mathematics seem unable to access their knowledge in science" (Leinhardt Zaslavsky, & Stein, 1990).

Better instruction, curricular resources, and instructional tools are needed to help students learn to understand and use graphs. Although this conclusion has been clear for decades, growth in the use of computers, tablets, and the Internet in classrooms now provides teachers and students with a different context than their predecessors faced a generation ago.

Web-Based Graphing Software

There are many computer-based tools used for graphing. However, the widespread use of technological tools that help people *create* graphs—including graphing calculators, and graphs generated by spreadsheet software—does not necessarily help students *understand* new graphs they encounter in textbooks, newspapers, or elsewhere. As marvelous as they may be, few technological tools "know" about the topic of a graph unfamiliar to a student and are able to use such knowledge to help the student generate meaning from a particular graph.

At the same time, the growing use of laptop and tablet computers in schools is fueling a movement of instructional materials and assessments onto the Web. For example, the federal government is spending hundreds of millions of dollars for two assessment consortia to develop tests for the Common Core State Standards that will be delivered entirely online. One of the consortia, the Partnership for Assessment of College and Career Readiness (PARCC), will provide an online graphing calculator similar to the TI-84 (Robelen, 2013); using personal graphing calculators will not be allowed when taking the PARCC tests.

Free online software applications that students can use to create graphs include: Create-a-Graph from the National Center for Education Statistics, Google spreadsheets, and the Desmos graphing calculator. These and other Web-based graphing tools—as well as graph-related applets such as those provided by the National Library of Virtual Manipulatives, or by NCTM—are widely used in schools and offer many features useful to students. For example, some of these Web-based tools automatically mark function maxima, minima, and axis-intercepts on a graph (e.g., the Desmos calculator). The use of sensor hardware and software (such as "motion probes" that allow students to graph one-dimensional motion in real time) has become more common in science classes, and has been shown to help students understand specific types of graphs (e.g., Mokros & Tinker, 1987).

Nonetheless, there are few, if any, general-purpose graphical tools specifically designed to help students *understand* the wide range of graphs they encounter in school. As Web-based instructional materials proliferate, and in light of the need to help students better understand graphs and concepts represented in graphs (such as slope, velocity, half-life, and global warming), the Concord Consortium developed SmartGraphs software under a National Science Foundation grant (DRL-0918522).

SmartGraphs Software

Web-based SmartGraphs activities allow students to interact with graphs, for example by clicking on a point in a graph to answer a question, or by labeling a point on a graph using words. Activity authors can also place labels on graphs to attach meaning to a point or region (e.g., "the car accelerated here"). A variety of instructional approaches used in SmartGraphs—such as visualizations (e.g., colored highlights on graphs), and multiple representations of concepts (such as a table or an animation linked to a graph)—have been demonstrated to show promise in science education (Rangel & Linn 2007). These and other features of SmartGraphs are used to create activities in multiple disciplines, including mathematics, science, and social science, always with an emphasis on making sense of graphs and concepts represented in graphs.

Because the software is programmed using HTML5 (Javascript), SmartGraphs activities run directly in a Web browser, without any special installations or download. One exception is that some activities make use of motion sensors connected to the computer, and in those cases an invisible Java applet is used, which requires that Java be installed on the student's computer. Otherwise, SmartGraphs activities do not use Java, Flash or any other plug-in or software besides a modern web browser. This helps assure that these activities are usable in any almost any school with an Internet connection.

A separate HTML5 software application allows authors who are not computer programmers to create their own multi-page lessons or activities, or modify existing activities, by filling in forms. Using the authoring tool, an unlimited number of activities may be created.

As noted above, SmartGraphs software provides students with multiple ways of looking at data or information, including through words, graphs, tables, or animated icons. In addition, when using motion sensors, representations are also made through physical motion of an object in front of the sensor. Many prior studies have investigated how multiple representations can be useful for student learning (e.g. Ainsworth, 2008; Friel, Curcio, & Bright, 2001).

Another important feature of SmartGraphs is the incorporation of scaffolding within activities. The scaffolding includes targeted hints that can respond to specific student answers. The scaffolding effectively creates lessons customized to the needs of the students, because only students who need help see scaffolding on any given page. The scaffolding can be in the form of written hints, equations, or visual markers on the graph or table. De Jong (2006) found that this type of guided inquiry is effective in promoting student learning, specifically stating, "The most effective learning results are found with tools that structure the learning process." Scaffolding can be particularly important when using technology. "Overall, the literature indicates that scaffolding is critical for students' ability to regulate their learning and can therefore enhance the learning of complex topics while using hypermedia" (Azvedo, Cromley, & Seibert. 2004, p 348).

Furthermore, students are more likely to achieve solid understanding of a topic through practice and reinforcement, such as in a multi-page SmartGraphs activity. The concepts taught can be presented in multiple ways within a single activity. Also, by creating multiple activities intentionally designed with overlapping content, students will engage with the content more than once and be able to apply their previous knowledge. For example, our research relied on four activities related to the linear motion of objects and a fifth related to acceleration due to gravity. These lessons began with position-time graphs, using them to help students understand position, direction, and velocity. From there, activities moved on to more difficult velocity-time graphs, matching them with position-time graphs. Lastly, activities related velocity to acceleration.

The Research Study

A two-year study of SmartGraphs began in August 2011. Teacher volunteers were recruited from Pennsylvania, selected because of the common availability of laptops throughout the state as a result of the state's large-scale Classrooms for the Future program (CFF). All of the research teachers were teaching eighth or ninth grade physical science using one of four common physical science textbooks. Tens of thousands of Pennsylvania students in these grades complete the physical science course every year.

A Description of the SmartGraphs "Treatment"

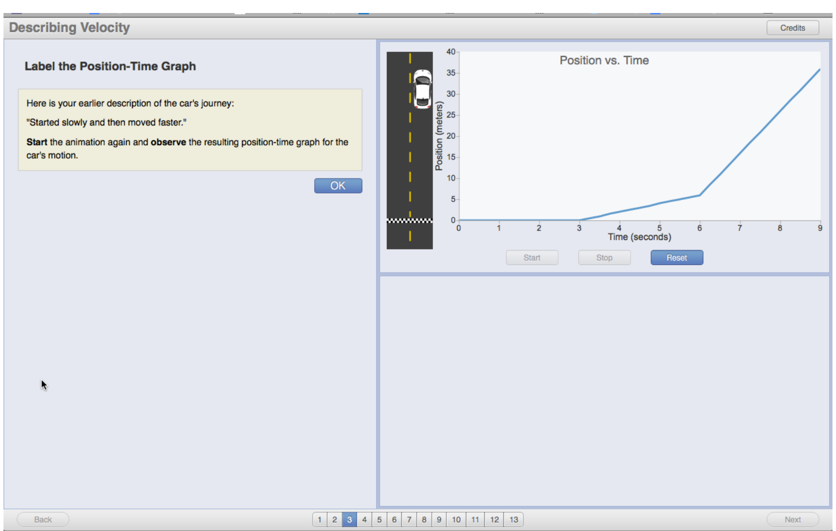
Although SmartGraphs can be used in various disciplines, a controlled study is best focused on only one particular topic. For the study we conducted, four SmartGraphs activities were developed focusing on the motion of objects. These activities were designed to supplement normal textbook-based instruction during the fall semester, when students are taught about the motion of objects. Motion was selected because it is an important topic covered in all physical science curricula and because learning the topic depends on graphs. Studying the motion of objects also allows for incorporation of a motion sensor. An additional SmartGraphs activity about acceleration due to gravity, a topic that usually follows linear motion topics, was developed as well, and made available during the fall of 2011. The five activities (all available online at http://smartgraphs.org) are as follows:

1. **Maria's Run**: Shows that the motion of an object can be described by its position, direction of motion, and speed. (A motion sensor is used in this activity.)
2. **Motion Toward and Away**: Explores different ways of describing motion on a graph. (A motion sensor is used in this activity.)
3. **How Fast Am I Moving?** Uses the position of an object at several times to determine the direction and velocity traveled during different time intervals. (A motion sensor is used in this activity.)
4. **Describing Velocity**: Connects the motion of an object to the corresponding position-time and velocity-time graphs to determine the velocity during different intervals.
5. **Was Galileo Right?** Explores effects of gravity on light and heavy objects as they fall.

According to data collected from teachers who participated in the research during the fall of 2012, on average they spent one month teaching the unit about the motion of objects and acceleration due to gravity. In almost all cases, teachers and students used one class period or less per SmartGraphs activity. The use of SmartGraphs therefore accounted for only a fraction of total instructional time.

To provide an idea of what students were asked to do, Figure 1 shows one page of the activity called Describing Velocity. On this page students can start and stop an animated car moving along a straight path, whose motion is simultaneously revealed on a graph next to the animation. Students are also asked to add labels to portions of that motion graph corresponding to the car staying still, moving slowly, or moving quickly.

Figure 1. Page 3 of Describing Velocity



Year 1 Research Design

For the first part of this study (fall 2011), 35 teachers were randomly assigned to experimental and control conditions. The experimental teachers used four SmartGraphs activities in addition to their usual curriculum, while the control teachers used their usual curriculum alone. All teachers received two days of professional development during August 2011. For the experimental teachers, the two days focused on using SmartGraphs to teach about the motion of objects. The control teachers were instead given professional development about activities related to teaching about heat and temperature. Teachers in both groups received three motion sensors to use with their classes, thereby ensuring that experimental findings would be due to the use of SmartGraphs, not to the availability of sensors. During the first year, nearly 2000 students in 91 classes participated in the study.

The first year of the research study aimed to answer the following research questions (which were also further investigated in Year 2):

RQ1. Do students who use SmartGraphs activities learn more than comparison students studying the same topic from the same textbooks, but who do not use SmartGraphs activities?

RQ2. What do teachers using SmartGraphs physical science activities believe about the software, including its match to important learning goals for the motion unit of study?

Year 2 Research Design

For the second year of the research (fall 2012), 31 of the teachers returned to the study. The small number of teachers who were unable to continue from Year 1 to Year 2 cited as reasons changes in their curriculum, teaching assignments, or schools. Of the 31 participating, two were not able to complete the research during the fall due to technical issues at their schools, leaving 29 teachers from 26 schools teaching 72 sections of Physical Science to nearly 1700 students.

All teachers in the second year used SmartGraphs activities. However, in order to understand more about how SmartGraphs can help students learn about the motion of objects, the activities and the treatment conditions were modified to allow different comparisons. The four research groups of teachers for the second year, and the treatment conditions, are shown in Table 1, which also includes related research questions, as well as comparison data sources.

Teachers in Cohorts A and C, who together formed the experimental group in Year 1, used the same four activities again in Year 2. For Cohort C, the activities were modified so that only one sensor was needed per classroom. This would allow the teacher to use the sensor and a computer projector as a demonstration, while each student could use a computer individually with a student version of the activity that did not use a sensor. During the first year of the research, when teachers could choose their method of delivering activities, we found that most activities with sensors were completed in small groups. It was hypothesized that the benefit to the students in these situations would depend on the size and the dynamics of the group. We also recognized that for many schools having a sufficient number of sensors would be a challenge, but having one sensor would be considerably more affordable. Thus the research for Cohort C attempted to examine the results when those factors were removed, and each student could use an individual computer (without a sensor). These teacher and student versions of the sensor activities had the same number of pages but the student version did not use a sensor.

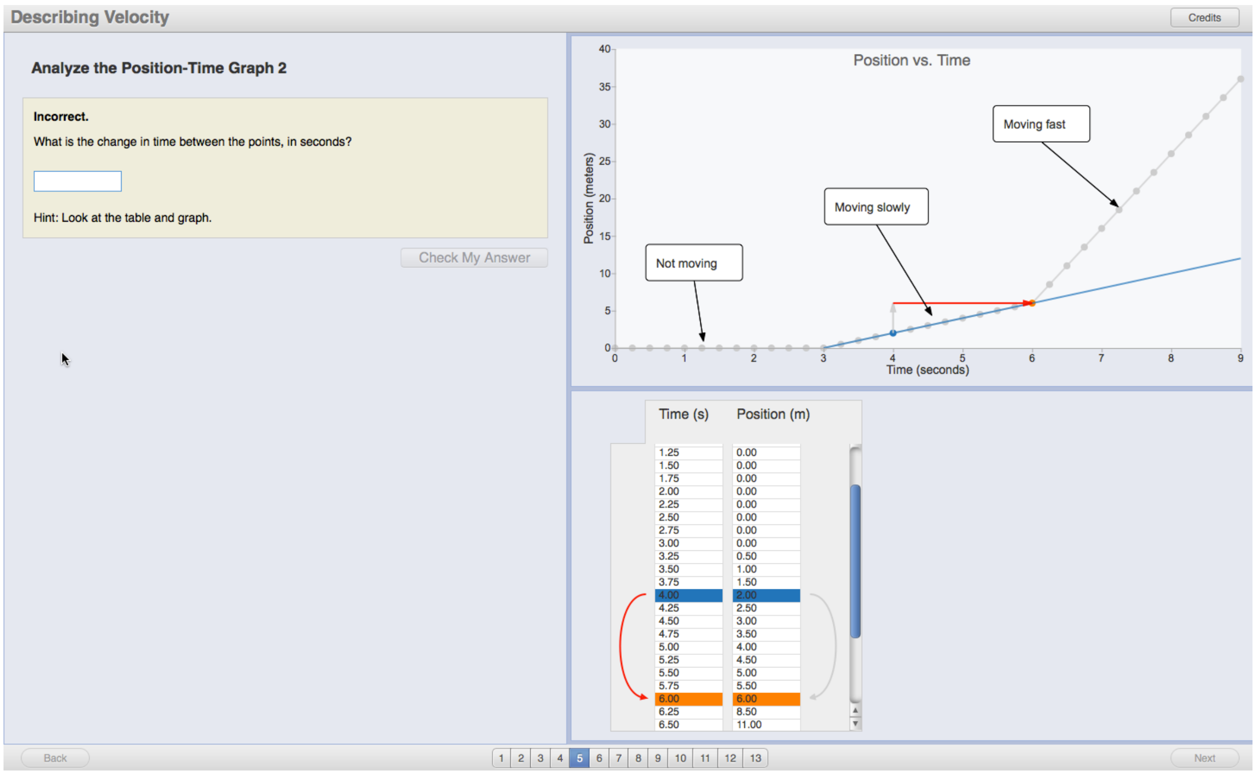
Table 1

Research Groups for Years 1 and 2

Group	**Fall 2011 (Year 1)**	**Fall 2012 (Year 2)**	**Research Question, 2012**	**Comparison Group**
Cohort A	Experimental	Year 1 activities (n = 7 teachers, 20 classes, and 392 students)	RQ3. Does teacher experience with SmartGraphs affect learning gains?	Results from the same teachers, Year 1
Cohort C	Experimental	Modified activities, with whole class (teacher) and individual (student) versions (n = 9 teachers, 21 classes, and 418 students)	RQ4. Do students learn differently when using SmartGraphs on individual computers compared to students working in small groups?	Cohort A
Cohort B	Control	Modified activities, without the SmartGraphs slope scaffolding tool (n = 4 teachers, 8 classes, and 177 students)	RQ5. Does the scaffolding provided by the slope tool lead to different learning gains than more generic scaffolding?	Cohort D
Cohort D	Control	Year 1 activities (n = 9 teachers, 23 classes, and 416 students)	RQ6. Do students of the *same teachers* learn more when using SmartGraphs than without SmartGraphs?	Results from the same teachers in Year 1

Three activities incorporate the Smart Graphs slope tool, which is a specific set of scaffolding steps designed to help students understand and compute the slope of lines or line segments. The slope tool includes an automated series of hints and visual cues on a graph and a table, such as arrows indicating “rise” and “run” on the graph (as shown in Figure 2). Students only see the slope scaffolding if they answer an initial slope-related question incorrectly, and only see all the hints and scaffolds if they need them. It was hypothesized that this tool would be valuable to students who had difficulty understanding or calculating slope. In order to test this hypothesis, a set of activities was created that removed the slope tool from the original activities and replaced the graph-specific hints with generic text hints. The activities without the slope tool were given to Cohort B in order to test this hypothesis by comparing outcomes with Cohort D students, who used the regular activities with the slope scaffolding included.

Figure 2. Examples of scaffolding provided by the SmartGraphs slope tool



The initial goal had been to divide teachers into four approximately equal size groups, but there were notably fewer teachers in Cohort B. Unfortunately, the small number of teachers who left the study between Years 1 and 2 disproportionately affected Cohort B, but there were still a sufficient number of students to warrant analysis of the results from this group.

Instruments

Two primary instruments were used in the study. One was used to measure learning outcomes, and the other was used to understand teachers' experience using SmartGraphs.

Assessment of Student Outcomes

To examine student understanding, identical pre- and post-tests were developed. To develop the test, we first developed a set of nine learning goals for the physical science motion unit, based on analyses of Pennsylvania's science standards and the physical science textbooks used in the state. (See Appendix A for a list of the learning goals.) Eighth and ninth grade physical science teachers reviewed these learning goals to confirm that they were, in fact, the targets of instruction. An example of a learning goal used for this project is, "Identify constant, positive, negative, and 0 rates of change in position with respect to time from a position-time graph." Then an expert in item development was hired to construct test items linked to the learning goals. For content validation, physical science teachers examined the items to confirm they matched the learning goals, and that the items were grade appropriate. Tests with these items were piloted, and items that performed best in terms of reliability and validity, and that covered all of the learning goals, were compiled into one test. The test was composed of eight multiple-choice items and twelve open-response items based on a knowledge-integration format (Lee & Liu 2009). (See Appendix B for sample test items.)

The test was expected to cover five SmartGraphs activities, but since the fifth activity (about acceleration due to gravity) was not available to teachers until the middle of the semester in the fall of 2011 (Year 1 of the research), many teachers had already given the post-test before that activity was available. Therefore, a first pre-post-test analysis examined all 20 test items, while a second analysis examined only the 16 items related to the first four SmartGraphs activities. For the second year of research (fall 2012), two of the twenty items were removed from the test; one multiple-choice item was removed because its content was not adequately covered in the five SmartGraphs activities, and one open-response item was removed because it detracted from reliability and was inaccessible to most students, even on the post-test. One of these items was part of the 16 items used to evaluate the first four activities, so analyses looking at only these four activities across the two years only include 15 items.

Weekly Logs

The other key source of data was teacher logs. After each week in which teachers taught motion, they were instructed to complete an online log that documented what content was covered, what technology was used, and any special circumstances that week. In Year 1, nearly 500 teacher logs were completed, each accounting for one week for a particular section during the time teachers taught motion and gravity. In Year 1, the experimental teachers were also asked which SmartGraphs activities they used, how the activities worked in the classroom and how well activities matched the physical science curriculum. In Year 2, over 300 teacher logs were completed. In Year 2, all teachers were asked about their experience with SmartGraphs. Cohort A was also asked how their experience compared with their first year using SmartGraphs, while Cohort C was asked how well the format using only one sensor worked for that activity.

Findings

Results from Year 1 are briefly summarized below. Then Year 2 results are presented.

Year 1 Results

The first year of the study included 1,686 students who took both the pre and post-tests. Data were analyzed pertinent to research questions 1 and 2, and are summarized here. (More detailed results from the first year of the study can be found in Kay, Zucker, & Staudt 2012.)

Research Question 1

After eliminating items related to the fifth activity on gravity, results showed improvement pre- to post- for students in both experimental and control groups, but the gain scores were significantly greater for students in the experimental group. However the effect size difference was small (0.13 for the complete test).

Research Question 2

Results from teachers for the log items about the SmartGraphs activities showed that teachers felt the content was appropriate, addressed the class's learning goals, and helped students meet those goals. For 99% of the instances reported on the logs, teachers said they would use that activity again, either as is (63%) or with minor changes (36%). Thus while the difference in learning gains was small between the two groups they were statistically significant, and the teachers' responses to the activities were positive. These findings warranted further investigation of the strengths of SmartGraphs in Year 2.

Year 2 Results

In the second year of the study, there were 1,403 students who completed both the pre- and post-tests. The test was composed of 7 multiple-choice items (worth 1 point each) and 11 open-response items (worth up to 4 points each) for a total of 51 possible points on the test. Because of the nature of the rubric for the open-response items, receiving 4 points on any given item was extremely difficult; therefore, the maximum scored on the pre-test was 43 points with a median of 16, and the maximum scored on the post-test was 45 points with a median of 23. Results of the pre and post-tests by cohort are shown in Table 2. (When computing multiple-choice scores, the multiple-choice portions of open-response items were included, so there are a total of 13 multiple-choice items evaluated for this score.)

Table 2

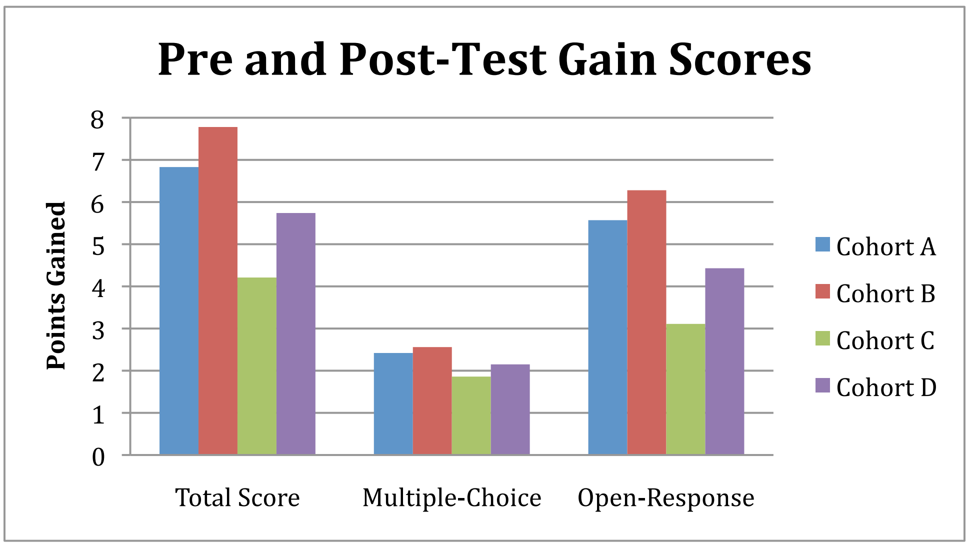
Results of Pre and Post-Tests

	Cohort A		**Cohort B**		**Cohort C**		**Cohort D**	
	n = 392		**n = 177**		**n = 418**		**n = 416**	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post
Total Score	13.94	20.77	18.02	25.80	18.33	22.54	15.38	21.12
Multiple-Choice Score	5.68	8.10	6.71	9.27	6.80	8.66	6.02	8.18
Open-Response Score	11.04	16.61	14.54	20.82	14.89	18.00	12.41	16.84

For all question types, students in Cohorts A and D started with pre-test scores significantly below students in Cohorts B and C, as measured by analysis of variance (ANOVA). Also on all three measures (total score, multiple-choice score, and open-response score), students in all cohorts significantly improved from pre-test to post-test.

Figure 3 shows the gain scores for each cohort of students. ANOVA showed significant differences among cohorts for all three scores. For all three scores, Cohort C (which used one sensor for the teacher, and individual computers without sensors for the students) showed significantly lower gains than Cohorts A and B on all three measures, and significantly lower gains than Cohort D for total score and open-response score. For the total test score gain, Cohort B (whose activities did not include the slope tool) showed significantly greater gains than Cohort D, the group that used the slope tool. Both Cohorts A and B showed significantly greater gains than Cohort D for open-response score. Further discussion of the meaning of these data is included below.

Figure 3. Results of Pre and Post-Test: Gain Scores



Also, in the second year of the study 328 logs were completed across all teachers in all cohorts. Each teacher completed a log for each section participating in the research (up to three sections) starting the first week they began teaching the motion unit and ending the week they administered the post-test to that class.

Research Question 2

Table 3 shows teacher responses to two questions asked in Year 2 about each activity they completed with students. In the first column, the amount of agreement with the statement "The activity helped my students meet the learning goals" is shown. As can be seen from the table, the great majority of teachers felt that the activity did meet the learning goals of the class. The last column of Table 3 shows teacher responses to the question "Would you use this activity in your classroom again?" Again, the responses were overwhelmingly positive.

Table 3

Selected Responses from Weekly Logs in Year 2 (2012)

	The activity helped my students meet the learning goals.	**Would you use this activity in your classroom again?**
	Strongly Agree or Agree	Yes or Yes with minor changes
Cohort A	100% (n=90)	100% (n=91)
Cohort B	95% (n=40)	95% (n=40)
Cohort C	96% (n=81)	99% (n=81)
Cohort D	86% (n=114)	92% (n=115)
TOTAL	94% (n=325)	96% (n=327)

The most positive responses came from Cohorts A and C. You will recall that these two cohorts of teachers comprised the experimental group that used the SmartGraphs activities the previous year. The least positive responses came from Cohort D. It should also be recalled that Cohorts B and D used five activities, including the fifth activity (Was Galileo Right?) about acceleration of objects due to gravity. Cohorts A and C did not use the gravity activity in order that their Year 2 results could be compared based to results based on the same activities they had completed the previous year. Therefore, it made sense to examine if responses for "Was Galileo Right?" (the fifth activity) were different from the results for all activities taken together. In fact, for that activity only 73% of the teachers in Cohorts B and D agreed or strongly agreed with the statement, "The activity helped my students meet the learning goals." Also, only 77% of teachers would use this activity in their classroom again, either as is or with minor changes. It is likely that differences in teachers' attitudes about the activities among the cohorts (A and C compared to B and D) were a result of use of this fifth, less popular activity.

Cross-year Analyses

Four research questions were answered using a combination of Year 1 and Year 2 data. Findings about each of them are discussed below.

Research Question 3: Does teacher experience with SmartGraphs affect learning outcomes?

To answer this research question, results from students of Cohort A teachers were compared between Year 1 (2011) and Year 2 (2012). Because many of the teachers in Cohort A were unable to use the fifth activity during the first year, they were instructed to only complete the first four activities in the second year, so that Year 2 results could be compared to Year 1. Students' scores on the fifteen items unrelated to the fifth activity were compared across years. There were 5 multiple-choice questions and 10 open-response questions, for a possible total of 45 points. Across all 15 questions there were 11 multiple-choice items (using the multiple-choice parts of open-response items as part of this score) for 11 possible multiple-choice points. The open-response score was composed of scores from the 10 open-response items, worth a total of 40 points. The mean scores for Cohort A students in the two years are shown in Table 4.

Table 4

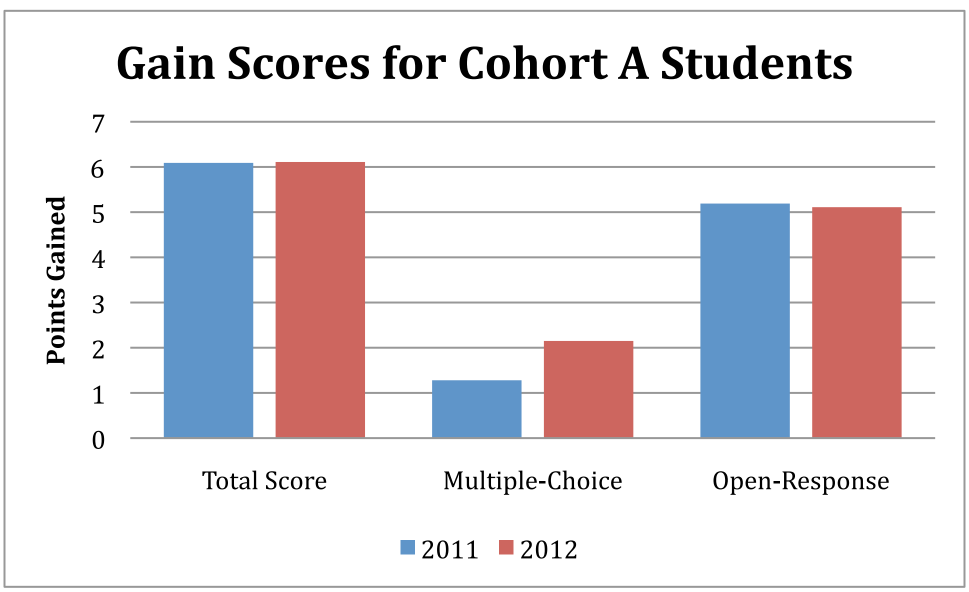
Pre and Post-Test Results by Year for Cohort A Students

	2011 (n = 416)		2012 (n = 392)	
	Pre-Test	**Post-Test**	**Pre-Test**	**Post-Test**
Total Score (45 pts max)	12.32	18.42	11.34	17.46
Multiple-Choice (11 pts max)	4.96	6.24	4.52	6.67
Open-Response (40 pts max)	10.34	15.53	9.61	14.72

There was no significant difference between years for total test or open-response scores on either the pre-test or post-test. However, there was a significant difference between years for both the multiple-choice pre-test and multiple-choice post-test, with the 2011 students outperforming the 2012 students on the pre-test, but 2012 students outperforming the 2011 students on the post-test.

Gain scores (Figure 4) confirm this difference, showing that the gains for total score and open-response score were nearly identical between years, but the difference in gains for multiple-choice scores was significantly different by year ($t = -7.348$, $df = 806$, $p < .0005$), with greater gains in 2012 than in 2011. The multiple-choice score indicates a factual understanding of content, in contrast to the open-response items, which reflect a deeper ability to explain the items. Thus the greater increase in multiple-choice gains in Year 2 indicates that students using SmartGraphs had greater gains in factual knowledge about the motion of objects during the second year than the first year.

Figure 4. Pre and Post-Test Gains by Year for Cohort A Students



Weekly logs continued to show satisfaction with the activities from 2011 to 2012, with the ceiling for agreement eliminating the possibility for improvement. For example, in 2011, 100% of Cohort A teachers agreed or strongly agreed with the statement "The activity helped my students meet the learning goals," and again 100% agreed with the statement in 2012. Similarly, in 2011, 100% of Cohort A teachers said they would use the activity again as is or with minor changes, and in 2012, 100% of teachers still would use the activity again. Teachers' comments, however, supported the idea that experience with the activities was helpful in their implementation of the activities. For example:

> *I prepped more for the activity this time – I went through all of the key terms, checked for understanding and demonstrated how to do the activity before the students actually did it. When they started it, the process ran smoother.*
>
> *I am more confident, and know more what to expect to help students.*
>
> *I found the experience using this activity this year as compared to last year much easier. I was able to troubleshoot issues and felt more prepared.*

Overall, while the greater experience of the teacher only lead to small increases in learning gains on multiple-choice items, an increase in teacher comfort and confidence was evident in their use of SmartGraphs activities with students.

Research Question 4: Do students learn differently when using SmartGraphs on individual computers compared to students working in small groups?

For this question, Year 2 students in Cohort A were compared to Year 2 Cohort C students. Both groups had used four SmartGraphs activities in Year 1. As described above, the activities for Cohort C were modified into teacher and student versions, with only one sensor required (to use with the teacher version). By contrast, Cohort A teachers could assign the activities however worked best with their classes. The great majority of teachers in Cohort A had students work in small groups, with an average of five sensors per class (i.e., five groups, each using one sensor and computer).

To examine the data pertaining to this research question, the weekly logs from the three activities which required sensors were compared for Cohort A and Cohort C teachers. In Cohort A, there were 69 instances of sections using one of the three motion sensor activities. In 86% of the cases using these three activities the teachers *strongly agreed*, "the activity helped my students meet the learning goals." In 97% of the cases teachers reported they would use the activities again "as is" while in 3% of the cases they reported they would use the activities again "with minor changes."

In Cohort C there were 60 instances of sections using one of the three motion sensor activities. In 52% of the cases using these three activities the teachers *strongly agreed*, "the activity helped my students meet the learning goals." In 70% of the cases teachers reported they would use the activities again "as is" while in 28% of the cases they would use the activities again "with minor changes." In the other 2% of cases teachers reported they would not use the activity again.

There was one activity (Describing Velocity) that did not require use of a motion sensor. In Cohort A 78% of the time that teachers used that activity with a section they reported they strongly agreed, "the activity helped my students meet the learning goals," and 100% of the time they would use it again "as is." When Cohort C teachers used that activity, 48% of the time they strongly agreed, "the activity helped my students meet the learning goals," and 48% of the time they reported they would use the activity again "as is."

Explaining these differences is challenging. It is possible that Cohort C teachers' experiences using only one sensor per class for the first three activities biased their perceptions of the fourth (non-sensor) activity. It is also possible that the teachers differed in their opinions about the value of the activities as early as last year. To test the latter proposition, we used last year's data set and retrospectively divided the experimental teachers into the same two cohorts, A and C. Indeed, data show that they did differ in their perceptions of the value of activities even in 2011, as shown in Table 5.

Table 5

Opinions of Cohort A and C Teachers About SmartGraphs Activities in 2011 and 2012

	SENSOR ACTIVITIES		DESCRIBING VELOCITY	
	Helped students (strongly agree)	Use again as is	Helped students (strongly agree)	Use again as is
Group A 2011	86%	77%	74%	70%
Group A 2012	86%	97%	78%	100%
Group C 2011	60%	64%	64%	52%
Group C 2012	52%	70%	48%	48%

These data tend to show that Cohort A teachers' second year of experience using the activities in the same manner as the first year *increased* their confidence in the activities, especially their opinion that they would use an activity again "as is." In contrast, Cohort C teachers, who used the sensor activities in a different way the second year than the first year, did *not* exhibit growing confidence in using the activities—even the non-sensor activity.

In response to the question, "Was this an effective method to use SmartGraphs activities with your class? Please explain," there were many comments to the effect that the single sensor was an effective method. For example, one teacher wrote, "Yes, it went very well with our class discussions and lessons." Another comment was, "Yes. I think overall the students learned better by completing the activity together [and] then, trying the rest of the lab on their own."

However there were also indications some teachers were not enthusiastic about using a single sensor in the teacher version of the activity (and no sensors in the student version of the same activity) and teachers provided reasons why. Here are some of the negative comments teachers in Cohort C made:

> *No. … I will never do this activity as a full class again. The small group work was so much better in terms of student collaborating and time wise, where you don't have to wait for the slower students to finish answering questions and before you know it the class period is over. I feel that I have wasted so much time this unit waiting.*
>
> *I actually did not like doing this activity as a whole class. I found it difficult to keep track of what page the students were supposed to be on versus what page the projector was on. I didn't like that not all students were able to take a turn creating a graph with the motion sensor. I also didn't like how the whole class was stuck waiting for a few slower students to finish before we could move on to the next page and the activity does not allow for collaboration among students (students teaching other students).*
>
> *I prefer having them work in lab groups since the computers don't all move at the same rate sometimes students get behind where we are working.*

As seen previously (in Figure 3), gain scores from students in Cohorts A and C differed significantly, with Cohort A showing greater gains than Cohort C. This was true for the complete test as well as when the items related to the fifth activity were removed (because Cohorts A and C did not use the fifth activity). In the first year of the study, gains were also greater for Cohort A students, but the difference was significant only when looking only at items related to the first four activities. Thus the differences between the gains of students in Cohorts A and C may reflect a difference between these two groups of teachers more than a result of the difference in use of the sensors. Thus the effect of presentation mode (one sensor or many sensors per class) likely does not affect learning gains.

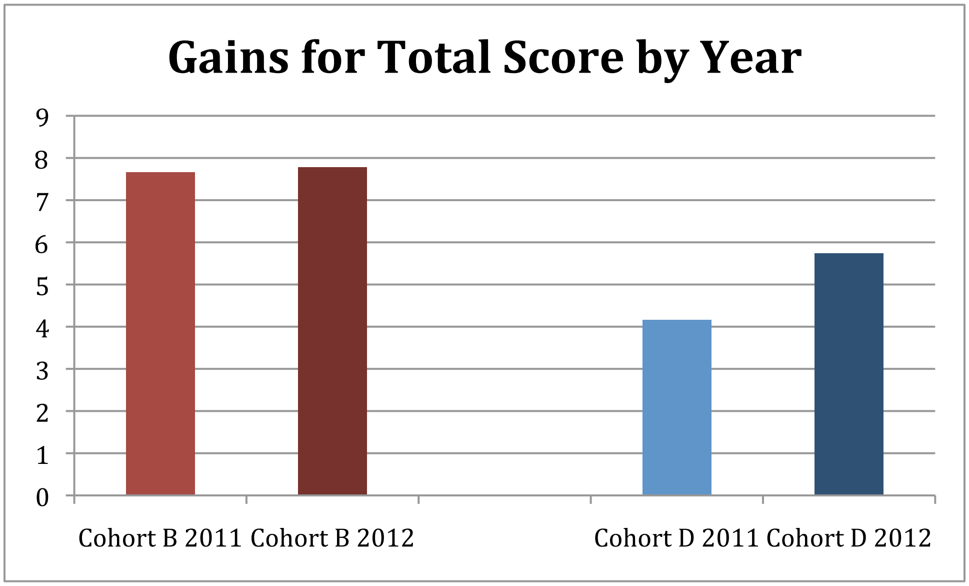
In summary, although the evidence is ambiguous, and not all teachers agreed, it appears that teachers show greater enthusiasm for using multiple sensors with small groups of students, rather than a single sensor projected for the whole class to see.

Research Question 5: Does the scaffolding provided by the slope tool lead to different learning gains than more generic scaffolding?

In order to specifically assess differences in learning gains by access to the slope tool, lessons were created with the slope tool removed, as described previously. Cohort B used these revised activities whereas Cohort D used the activities with the slope tool. As previously noted, Cohort B students had significantly higher gains than Cohort D students on both the total test score and the open-response test score. This would imply that the use of the slope tool leads to *decreased* learning gains. In order to examine this further, learning gains for the same 18 items were examined from Year 1 of the study. In Year 1, the gains of Cohort B were significantly greater than the gains of Cohort D in all three categories. However the increase in gains from

2011 to 2012 was less for Cohort B than for Cohort D, indicating that Cohort D may be "catching up." This difference in gains for the total score is shown in Figure 5. Results for multiple-choice and open-response scores are similar.

Figure 5. Gain Scores of Cohorts B and D by Study Year



Items from the 2012 assessment that specifically included computation of slope were examined further. There were three multiple-choice items and two open-response items. Results for these five items are shown in Table 6. As seen in the table, three of the slope items (3, 6, and 10) do not show significant differences in gains for Cohorts B and D, but two items (2 and 16) do show significant differences. But again, these differences favor Cohort B.

Table 6

Gains on Items Involving Computation of Slope

		Cohort B (n = 177)			**Cohort D (n = 416)**			
		Pre	Post	Mean Gain	Pre	Post	Mean Gain	Significant difference in gains?
Multiple-Choice	Item 2	0.37	0.78	0.41	0.30	0.60	0.30	t = 2.202, df = 591, p = .028
	Item 3	0.65	0.82	0.17	0.56	0.72	0.16	t = 0.212, df = 591, p = .832
	Item 6	0.40	0.75	0.34	0.36	0.61	0.25	t = 1.561, df = 591, p = .119
Open-Respo nse	Item 10	1.34	1.80	0.46	1.04	1.47	0.44	t = 0.252, df = 591, p = .801
	Item 16	0.82	1.89	1.07	0.70	1.45	0.75	t = 2.925, df = 591, p = .004

That the group without slope scaffolding performed significantly better than the group with the slope tool could be an artifact of the particular teachers in these cohorts, as students of the same teachers in Cohort B also had greater gains than students of Cohort D in the first year of the study when neither group was using SmartGraphs. However, the size of the sample (593 students in 31 classes) suggests that this is an unlikely explanation.

These results disprove our original hypothesis, that the scaffolding included in the slope tool—providing connections between text, graph, and table—helps students more than generic scaffolding (not graphically tied to the particular graph shown to students).

Research Question 6: Do students of the same teachers learn more when using SmartGraphs than without SmartGraphs?

While the first year of the study compared learning gains between students taught by teachers with or without SmartGraphs software—a classic experimental trial—one advantage of a two-year study is the ability to compare students of the same teachers teaching *without* and then, a year later, *with* the SmartGraphs software. While the students between years changed, there are enough constants between the years—notably school, population characteristics, curriculum, and teaching style—that a two-year comparison between students of the same teachers has strong credibility. In effect, this research question is a variant of research question 1 (Do students who use SmartGraphs activities learn more than comparison students studying the same topic from the same textbooks, but who do not use SmartGraphs activities?).

In order to make this comparison, the 18 items that were common between the assessments in both years were used to compute scores and gain scores. Results are shown in Table 7 and Figure 6 for students of teachers in Cohort D, who used SmartGraphs in 2012 but not in 2011.

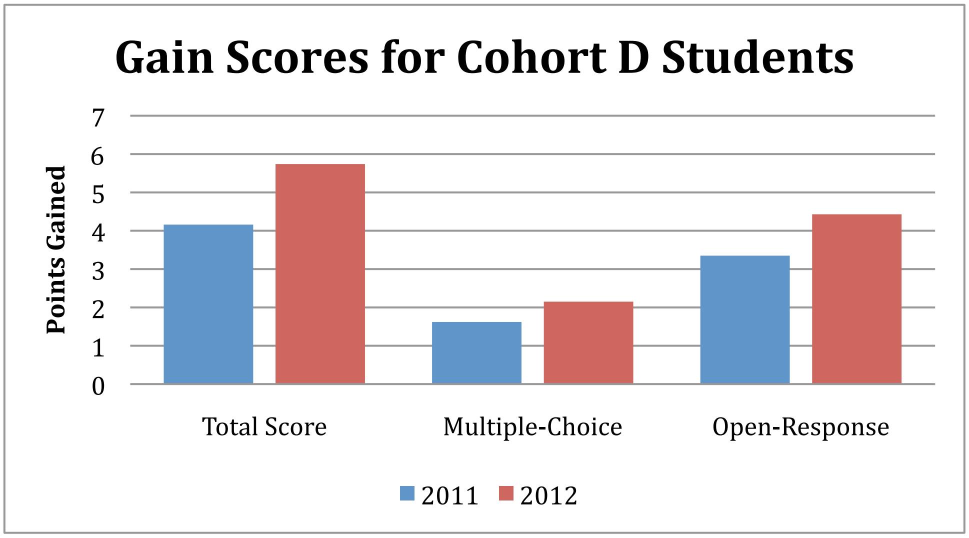
Table 7

Pre and Post-Test Results by Year for Cohort D Students

	2011 (n = 365)		**2012 (n = 416)**	
	Pre-Test	**Post-Test**	**Pre-Test**	**Post-Test**
Total Score (51)	14.01	18.17	15.38	21.12
Multiple-Choice (13)	5.68	7.30	6.02	8.18
Open-Response (44)	11.14	14.49	12.41	16.84

In the second year of the study, students of Cohort D teachers had significantly higher pre-test scores (for total score and open-response score), but even with this difference in starting point, the gains were also significantly higher for all three categories (total score: t = -3.907, df = 779, p < .0005; multiple-choice score: t = 03.257, df = 779, p = .001; open-response score: t = -2.911, df = 779, p = .004).

Figure 6. Gain Scores for Cohort D by Year



Based on this information one can conclude that use of SmartGraphs leads to greater learning gains than instruction by the same teachers using the same textbooks but without SmartGraphs. Although these effect sizes are usually considered small, the differences between years are greater for this cross-year comparison than the effect size comparing outcomes for the experimental and control groups in Year 1. Specifically, comparing 2012 and 2011 results for Cohort D, the effect size for the total score gain was $d = 0.28$, for the multiple-choice gain was $d = .23$, and for the open-response gain was $d = 0.31$, showing higher student outcomes in 2012 when SmartGraphs was used. In terms of learning gains, an effect size of 0.28 corresponds to moving a student from the 50th percentile on the total score to the 61st percentile.

HLM Analyses

Because students are clustered within sections, which are clustered within teachers, it makes sense to look at this data with statistical techniques that account for this nested data structure. Specifically, hierarchical linear modeling (HLM) uses this structure within the analysis in order to account for similarity between students within a certain group. As the experiment for Research Question 4 was the most straightforward, initial HLM analyses focused on students of Cohort D teachers in both the first and second year of the study. The level-1 unit of analysis was the student. For level-2, the unit of analysis was section. While there are certainly similarities between sections of the same teacher, there were enough differences between sections to justify this grouping. For example, some teachers had an honors section and a section of students with special needs both included within this study, and the results for these two sections were very different. A third level of "teacher" could be added to the analyses, but that would not allow sufficient power for results to be detected.

Multiple models were completed using HLM7 software. There were three outcome variables used: total test score gain, multiple-choice gain, and open-response gain. For these analyses, Level 1 is student-level and will have the outcome variable of gains (total, multiple-choice, or open-response). Level 2 is section-level. For all of the models, there are no level-1 variables. Therefore, Level 2 is only predicting the intercept of the Level 1 equation (β_0).

For each outcome variable, three models were examined. The first model is the null or empty model, in which no variables are added. The second model adds the variable Year, which is coded 0 or 1. Year coded as 0 is the first year of the study, when teachers did not have access to SmartGraphs software. Any sections from Year = 1 (the second year of the study) had access to SmartGraphs software. The third model included variables to account for how the teacher used SmartGraphs. Specifically, each section was classified as using SmartGraphs mainly in whole class mode, mainly in small group mode, or mainly with students using computers individually. In many cases, classes did a combination of modes across the five SmartGraphs activities, but the most prominent mode was selected for the sake of classifying the section. In most, but not all cases, teachers were classified with the same mode in all sections they taught. Overall there were 7 sections classified as “whole class,” 14 classified as “small group,” and 2 classified as “individuals.” To use these classifications in HLM, three dummy variables were developed, with the selected classification coded as “1” while the others were coded as “0” for that section. In all cases, sections from the first year of the study are coded as 0. Thus, the results from the dummy variables show that classification as compared with students with no access to SmartGraphs.

The results for total test gains are shown in Table 8. The intraclass correlation in model 1 was 0.07 indicating that 7% of the variance in gain scores is between sections. While this is small, it is sufficient to justify using HLM.

Table 8

Effects of SmartGraphs Use on Total Gain Scores Using 2-level Hierarchical Linear Regression

Variable	Model 1 coeff.	Model 2 coeff.	Model 3 coeff.
Intercept (β_0)			
Intercept (γ_0)	4.9*	4.0*	4.0*
Year	---	1.7*	---
Dummy: Whole class	---	---	2.7*
Dummy: Small groups	---	---	1.2
Dummy: Individuals	---	---	1.3
Variance components			
Level-1 variance (σ^2)	30.86	30.85	30.87
Intercept (τ_{00})	2.49	1.89	1.73

*Significantly different from 0 at $\alpha = .05$

In the second model, year is a significant predictor of total test gains. Specifically, students in the second year of the study would expect to gain an additional 1.7 points on the post-test. An effect size was computed by dividing the coefficient of the treatment variable (year) by the between-section standard deviation from the null model. (Rethinam, Pyke, & Lynch, 2008). In this case, the effect size for year was .30, meaning that the gain scores for students in the second year of the study were 0.30 standard deviations greater than the gain scores for students in the first year of the study (without SmartGraphs). This is very close to the 0.28 effect size found without HLM.

For the third model that examines results by mode of use for SmartGraphs, only the whole class mode is a significant predictor, with a mean additional gain of 2.7 points over students from the first year of the study. This effect size is .48, meaning there is a medium effect of using SmartGraphs in whole class mode versus not using SmartGraphs (from the study's first year). This is especially interesting in light of the lack of difference in score gains between Cohort A and Cohort C (see Question 3 above), in which teachers in Cohort C were required to use the activities in whole class mode. However, it should be noted that teachers in Cohort C had a variation of the regular SmartGraphs activities with separate versions for students and teachers, so students could follow along with the demonstration. In Cohort D, it is unclear whether students even had computers in front of them as their teacher demonstrated the activity.

Results for multiple-choice gains and open-response gains were similar to total gains. In both cases, year was a significant predictor of gains, indicating the use of SmartGraphs led to greater gain scores on both multiple-choice and open-response measures. For the multiple-choice outcome, there is a .26 effect size for year, while for the open-response outcome there is a .23 effect size for year.

Overall, the HLM analyses support the claim that students of the same teachers performed better with SmartGraphs than without. Because the statistically significant outcome from the initial t-tests still holds with HLM analyses, the result is robust. It appears that teachers who used SmartGraphs activities in whole class mode had students with the greatest gains, while students of teachers who used the activities in small group mode did not have significant gains at all compared with non-SmartGraphs users. This difference in activity delivery mode is worth investigating further.

Discussion

Based on the two years of data collection and analysis, we reached a number of conclusions about SmartGraphs. In addition, we began to look for explanations of data that were puzzling.

Five Conclusions

Multiple analyses of data from this two-year experimental study allow us to reach five conclusions. We have a high degree of confidence in these conclusions.

1. Using SmartGraphs activities that focus on the motion of objects as a supplement to normal instructional activities in physical science classes results in statistically significant learning gains for students. This conclusion was reached in Year 1, and then confirmed with a new set of comparisons in Year 2. Replicating the finding makes us more confident that the conclusion is correct, as does the large number of schools, teachers, and students who participated in the research. HLM analyses support this finding as well. While the effect size was small, it should be recalled that on average teachers spent a month on this unit, only four or five days of which were devoted to using SmartGraphs. Thus, on average less than 25% of class time was devoted to using SmartGraphs.

2. Teachers overwhelmingly reported that the activities helped students achieve learning goals for this unit of study, and that they would use the activities in the future. Once again, the Year 1 finding was replicated in Year 2. For example, on 94% of the weekly logs submitted in Year 2 teachers agreed or strongly agreed that, "the activity helped my students meet the learning goals" (Table 3).

In both years teachers felt that SmartGraphs activities were successful for their students. The following are examples of teacher comments on the activities:

> *It is very good for making them consider different motions with a graph; many of them hadn't considered starting away from the origin.*
>
> *The students love this one! They get to play around with the motion sensor and experiment with different movements and the learning is so intuitive they don't realize they're learning!*
>
> *This is one of my favorite activities because it really helps students relate the motion to the graphs of objects moving forward and backward.*

3. Teachers who used SmartGraphs for a second year had greater confidence in their use of the activities. Although this finding is not surprising, it is a reminder that changing instructional practice takes time. Teachers need time to learn and become comfortable with new ways of teaching. Although there was not a greater gain on students' total test score in Year 2, the score on the multiple-choice portion of the test was statistically higher in 2012 than in 2011, suggesting that teachers' greater confidence using the activities also had a favorable impact on student learning (Figure 4).

4. The scaffolding tool we developed to help students understand slope apparently did not add value to the activities. This is a surprising finding to the project team because we expected the slope tool would be valuable, or we would not have created it. Yet both an analysis of total test score and an analysis of test items specifically focusing on slope did not show an advantage for students in the cohort using the slope tool, as compared to students receiving generic hints about slope (Table 6). One possible explanation for this finding is that many more students had difficulty calculating the slope of a line on a position-time graph than we expected—an hypothesis we investigated further after the second year classroom trials (see discussion in the next section).

5. Teachers using motion sensors as part of SmartGraphs prefer having small groups of students use a motion sensor, rather than using a single sensor and a computer projector. We realized that a single sensor per classroom would be less expensive for schools, but were not sure whether it would be as effective. For this reason, we designed an experiment to compare a single-sensor version of these activities with the 2011 version in which students work in small groups, each group using a sensor. This arrangement also required every student to answer questions on his or her computer (rather than possibly watch while other students in a small group used a computer). Teachers apparently did not agree with this arrangement, in part because coordinating the pages of the activity that students were on with the page the teacher was on was hard to do, and in part because they felt that having more students walk in front of the sensor (in the small groups) was more effective for student learning. As a result we will not further disseminate the version of the sensor activities in which teachers use a sensor but students do not.

Additional Information about Teaching and Learning Slope

SmartGraphs activities running on students' computers did not save student data back to a Concord Consortium server. As a result, we did not generate a data file showing whether particular students answered any particular question correctly in each activity, or how many hints students needed. However, in Year 2 we discovered that we could use Google Analytics to explore students' detailed responses because visiting each page in a SmartGraphs activity, and each step on a page (such as a particular hint), can be tracked, via Google Analytics, as an "event" to be logged.

The Analytics data showed that many students had great difficulty calculating the speed of an object based on a graph (in other words finding the slope of a straight-line segment on a position-time graph). For example, in the fall of 2012 60% of students saw a page in Describing Velocity noting that their answer to a certain slope question was incorrect—even though the slope they were asked to provide was a small whole number (2 meters per second) and therefore easily computed if the concept of speed as slope on a graph were understood.

Once we realized that calculating slope was more challenging for eighth and ninth graders than expected, we surveyed the Physical Science teachers in early 2013 and asked them about teaching and learning slope. About three-quarters of the 23 Physical Science teachers who responded to this survey believe that the mathematics department in their school has greater responsibility than the science department for teaching students to calculate the slope of a line. About one-third of the teachers do not expect students to learn to calculate speed as slope in the Physical Science class—but about two-thirds do. Asked to estimate the percentage of students who would answer a simple speed/slope question correctly from a graph *after* completing the motion unit, the teachers' average was 63% and the median was 70%. Expectations are that students will understand steepness of a line as an indication of greater speed, but for some teachers there is not an expectation that students will be able to calculate slopes. (Note that the learning goals we used in this study, which are listed in Appendix A, accurately reflect participating teachers' greater emphasis on developing students' qualitative understanding of slope rather than students' ability to perform quantitative calculations.)

Teachers expressed concern about the difference in vocabulary used in math and science classes. These particular science teachers (who teach in schools across Pennsylvania) believe that the math teachers in their schools teach the process of calculating rise over run without sufficient attention to what slope represents. For example, one teacher responded "I don't think students make the connection between meters per second and speed. They get slope in math but learn rise over run and don't connect it to speed." Science teachers typically use units of measure, such as meters and seconds, and the survey respondents suggested that math teachers often do not use units. The teachers pointed out that variables in science are not usually designated X and Y but rather by other names, such as Time and Distance, and these Physical Science teachers suggest that students are confused by the different contexts in math and science.

Assuming we were to repeat this research, we believe that we would see greater gains for students using SmartGraphs, compared to students not using the software, if we were to add another activity narrowly focused on calculating speed as the slope of a line on a position-time graph. The slope scaffolding we developed (see Figure 2) still seems appropriate for students who already know how to calculate slope and need reminders, but it is less appropriate for students who are just learning to connect science and math vocabulary, the use of graphs, the concept of speed, and arithmetic calculations. One teacher specifically responded, "I like the idea of having a specific activity for calculating slope. This may help bridge the gap and make connections between what they learn in math and science classes." What we learned from the teachers about teaching slope will inform development of another short SmartGraphs activity, which will focus only on calculating speed viewed as the slope of a line on a position-time graph, and carefully connect science and mathematics concepts and vocabulary.

Concluding Remarks

Reviewing prior research about students' understanding of graphs was critical in the development of new graph-related software, SmartGraphs. Once the software was developed, choices we made about the topic of instruction to study, and the outcome measures, enabled our research to focus on students' understanding of graphs important in physical science courses. The large-scale randomized study of SmartGraphs reported here shows evidence that classroom use of research-based software tools can significantly improve students' understanding of graphs and concepts represented in graphs.

Appendix F of the Next Generation Science Standards describes the "practices" of science that students need to learn (Achieve, Inc., 2013). One of these practices is "analyzing and interpreting data." The appendix explains that as they mature, "students are expected to improve their abilities to interpret data by identifying significant features and patterns." We believe this expectation includes making sense of new and unfamiliar graphs. But to achieve this goal, teachers will need to spend more time helping students understand graphs. Individual teachers could read and apply research directly to their instructional practice (which is unlikely to happen on the scale of tens of thousands of teachers), or—much more likely—teachers need to make use of instructional materials and digital tools that are based on science and mathematics education research.

Although this two-year study has concluded, work on SmartGraphs software, new activities, and the authoring system continues (further information can be found at http://smartgraphs.org). Given the enthusiastic response of teachers, and the statistically significant student gain scores we found in both years of the study, we believe that there is potential for large-scale use of the SmartGraphs software. Discussions are under way with publishing companies to explore their interest in integrating SmartGraphs into digital textbooks.

No study is flawless, nor is any one technological tool used for teaching and learning able to help teachers resolve all student misconceptions about graphs. Yet despite these facts, we hope that SmartGraphs software and the research reported here contribute to both research and practice in STEM education.

Acknowledgements

This research was conducted under National Science Foundation Grant No. DRL-0918522, awarded to the Concord Consortium (Carolyn Staudt, PI). Any opinions, findings, and conclusions or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References

Achieve, Inc. (2013). *The next generation science standards, appendix F: Science and engineering practices*. Retrieved from http://www.nextgenscience.org/next-generation-science-standards.

Ainsworth, S. (2008). The educational value of multiple-representations when learning complex scientific concepts. In J.K. Gilbert, M. Reiner, & M. Nakhleh (Eds.), *Visualization: Theory and Practice in Science Education Models and Modeling in Science Education,* (Vol 3, pp. 191-208). Springer Science+Business.

Azevedo, R., Cromley, J. G., & Seibert, D. (2004). Does adaptive scaffolding facilitate students' ability to regulate their learning with hypermedia? *Contemporary Educational Psychology*, *29*, 344-370.

Bowen, G.M., & Roth, W.-M. (1998). Lecturing graphing: What features of lectures contribute to student difficulties in learning to interpret graphs? *Research in Science Education*, 28, 77-90.

de Jong, T. (2006). Technological advances in inquiry learning. [Education Forum]. *Science*, *312*, 532-533.

Friel, S. N., Curcio, F. R., & Bright, G. W. (2001). Making sense of graphs: Critical factors influencing comprehension and instructional implications. *Journal for Research in Mathematics Education, 32*(2), 124-158.

Friel, S. N., & Bright, G. W. (1996, April). *Building a theory of graphicacy: How do students read graphs?* Paper presented at the annual meeting of the American Educational Research Association, New York. (ERIC Document Reproduction Service No. ED 395 277)

Kay, R., Zucker, A., & Staudt, C. (2012). *Being smart about SmartGraphs: Findings from an experimental trial in physical classrooms*. Concord, MA: The Concord Consortium. Retrieved from http://concord.org/sites/default/files/pdf/smartgraphs-research-fall-2011.pdf.

Lee, H-S. & Liu, O. L. (2009). Assessing learning progression of energy concepts across middle school grades: The knowledge integration perspective. *Science Education*, *94*(4), 665-688.

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

Monk, S. (2003). Representation in school mathematics: Learning to graph and graphing to learn. In J. Kilpatrick, W. G. Martin & D. Schifter (Eds.), *A research companion to Principles and Standards for School Mathematics* (pp. 250-262). Reston, VA: National Council of Teachers of Mathematics.

Mokros, J., & Tinker, R. (1987). The impact of microcomputer-based labs on children's ability to interpret graphs. *Journal of Research in Science Teaching, 24*(4), 369-383.

Murray, T. S., Kirsch, I. S., & Jenkins, L. B. (1997). *Adult literacy in OECD countries: A technical* report *for the first international adult literacy survey*. Washington, DC: National Center for Education Statistics.

Rangel, E. S., & Linn, M. C. (2007). *Science education that makes sense* (Research Points #5:1). Washington, DC: American Educational Research Association.

Rethinam, V., Pyke, C. & Lynch, S. (2008). Using multilevel analyses to study the effectiveness of science curriculum materials. *Evaluation and Research in Education, 21*(1), 18-42.

Robelen, E. (2013, August). Calculator use on exams to shift with the Common Core. *Education Week*, 33:1, 1.

Roth, W.-M. (1998). Unspecified things, signs, and "natural objects": Towards a phenomenological hermeneutic of graphing. In S.B.Berenson, K.R.Dawson, M.Blanton, W.N.Coulombe, J.Kolb, K.Norwood, & L.Stiff (Eds.), *Proceedings of the 20th Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (pp. 291-297). Columbus, OH: ERIC Clearinghouse for Science, Mathematics, and Environmental Education.

Roth, W.-M., & McGinn, M.K. (1997). Graphing: Cognitive ability or practice? *Science Education*, 81, 91-106.

Roth, W.-M., Bowen, G. M. and McGinn, M. K. (1999), Differences in graph-related practices between high school biology textbooks and scientific ecology journals. *J. Res. Sci. Teaching*, 36: 977-1019.

Roth, W.-M., Pozzer-Ardenghi, L., & Han, J. H. (2005). *Critical Graphicacy: Understanding Visual Representation Practices in School Science*. Dordrecht, The Netherlands: Springer-Kluwer.

van Zee, E., & McDermott, L. C. (1987). Investigation of student difficulties with graphical representations in physics. In J. Novak (Ed.), *2nd Int. Seminar "Misconceptions and Educational Strategies in Science and Mathematics", Ithaca, NY, 1987* (Vol. III, pp. 531-539): Cornell University.

Woolnough, J. (2000). How can students learn to apply their mathematical knowledge to interpret graphs in physics? *Research in Science Education*, *30*(3), 259-268.

Appendix A

SmartGraphs' Learning Goals for the Physical Science Motion Unit

Understand direction of motion

1. Determine direction of motion from a position-time graph.
2. Determine direction of motion from a velocity-time graph.
3. Distinguish between speed and velocity when reading a velocity-time graph.

Understand rate of change

4. Identify constant, positive, negative, and 0 rates of change in position with respect to time from a position-time graph.
5. Identify constant, positive, negative, and 0 rates of change from a velocity-time graph.
6. Estimate and predict through interpolation and extrapolation, an object's speed in different time intervals on a position-time graph.
7. Estimate, predict, and determine an object's average speed over a given time interval on a position-time graph.
8. Interpret and compare position or velocity of an object from two graphs showing different rates of change, as well as interpreting the intersection of the two graphs in real context.

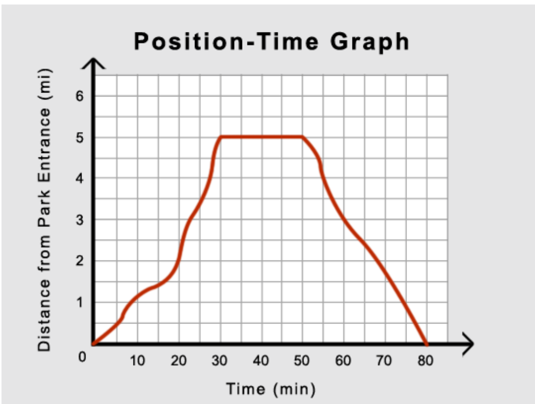
Extend learning to graphs other than motion of objects, e.g. temperature, money, algebra.

9. Interpret characteristics (rate of change etc) such as above in new contexts.

Appendix B

Two Sample Test Items

A Multiple-choice item: Position-Time Graph

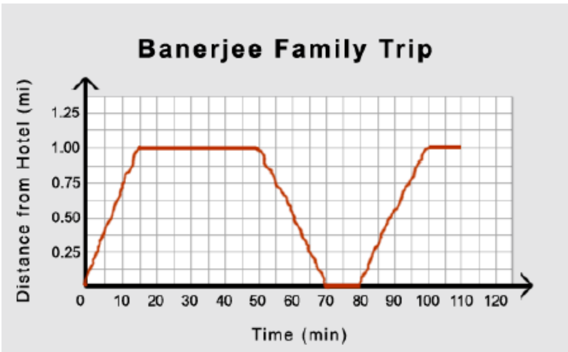


Eric roller-skated from the entrance of a park to a juice bar. He stopped at the juice bar for a smoothie. Then he went back to the park entrance. The graph above represents his entire trip.

According to the graph, what did Eric do between minutes 30 and 50?

- ☐ Eric skated at a constant speed.
- ☐ Eric skated along a straight path.
- ☐ Eric stopped and drank a smoothie.
- ☐ Eric skated back to the park entrance.

An Open-response item: Banerjee Family Trip



The Banerjee family decided to spend Monday at the beach. The graph above shows the distances they walked to and from the beach during the morning. They begin at time zero at the hotel.

Which section or sections of the graph show a time during which the family walked from the beach to the hotel? Check all that apply.

- ☐ The segment between (0,0) and (15,1)
- ☐ The segment between (15,1) and (50,1)
- ☐ The segment between (50,1) and (70,0)
- ☐ The segment between (70,0) and (80,0)
- ☐ The segment between (80,0) and (100, 1)
- ☐ The segment between (100, 1) and (110,1)

Explain in detail how you chose your answer(s) and why you did not select the other choices.

____

____

____

____