RESEARCH REPORT

Comparative Effectiveness of TI-84 Graphing Calculators on Algebra I and Geometry Outcomes:

A Report of Randomized Experiments in the East Side Union High School District and the San Diego Unified School District

Gloria I. Miller
Andrew Jaciw
Brandon Hoshiko
Empirical Education Inc.

Xin Wei
Stanford University

August 2007

Empirical Education Inc.
www.empiricaleducation.com
425 Sherman Avenue, Suite 210
Palo Alto, CA 94306
(650) 328-1734

Acknowledgements

We are grateful to the people in the East Side Union High School District and San Diego Unified School District for their assistance and cooperation in this research and for providing access to their data under an agreement between Empirical Education Inc. and the districts. The research was sponsored by Texas Instruments Inc. through a contract with Empirical Education.

About Empirical Education Inc.

Empirical Education Inc. was founded to help school districts, publishers, and the R&D community assess new or proposed instructional and professional development programs through scientifically based pilot implementations. The company draws on the expertise of world-class researchers and methodologists assuring that the research is objective and takes advantage of current best practice in rigorous experimental design and statistical analysis. The company's findings let educators quantify the value of programs and help them partner with providers to implement those most effective for their students.

Comparative Effectiveness of TI-84 Graphing Calculators on Algebra I and Geometry Outcomes:

A Report of Randomized Experiments in the East Side Union High School District and the San Diego Unified School District

Table of Contents

Introduction

Texas Instruments has undertaken a research program with the goal of producing scientifically-based evidence of the effectiveness of graphing calculators and the *TI-Navigator™* classroom networking system in the context of a professional development and curriculum framework. The program includes a two-year longitudinal study. The research is conducted as a set of randomized experiments working with teachers of Algebra I and Geometry in the East Side Union High School District (ESUHSD) and the San Diego City Schools (SDCS). This report addresses the findings from ESUHSD and SDCS for the first year of the two year study.

The research measures the impact of important components of the Texas Instruments (TI) offering, in the context of a field experiment. In this case, East Side Union High School District is a large urban school district with a large variation in the student population with teachers within the normal variation in schools, such as faculty experience, pedagogy, classroom environments, and scheduling. San Diego City Schools is a large urban school district with a diverse student population with teachers set in school environments representative of the variations found in schools, such as faculty experience, instructional practices, classroom environments, and scheduling.

The Texas Instruments intervention includes teacher professional development and materials designed to supplement standard textbooks with activities that provide students with additional pathways to explore concepts. Specifically, the TI hardware components consist of the TI-84 Silver Edition Plus graphing calculator, Cabri Jr. Dynamic Geometry application, TI-Navigator 2.0, and T^3 Professional Development. Additionally, every student in the study was provided a TI-84 Silver Edition Plus graphing calculator for home or classroom use. Findings from previous studies provide some evidence to indicate that when students have more access to calculators, both during class time and at other times, students score higher on end-of-course test scores (Heller, Curtis, Jaffe, and Verboncouer, 2005). This was the rationale that TI used when providing calculators for every student in the study.

The outcomes of interest are student achievement on standardized tests of Algebra and Geometry. The results are measures of the incremental effect of introducing advanced TI products and services into textbook-based college preparatory Algebra I and Geometry instruction. We have analyzed the effect of each component as well as important interactions among the components and interactions between the components and characteristics of the students and teachers such as background and incoming math expertise. We also address the effect of student and teacher familiarity with the technology over time since the full impact of technology introduction may not be seen within a single academic year.

This experiment in SDCS and ESUHSD will inform the current literature on the benefit of calculator use in Algebra I. Our review of experimental research (Khoju, Jaciw & Miller 2006) has shown a consistent impact of calculator use as applied to Algebra I. Integrating results across three studies that addressed graphing calculators in Algebra I, we found a standardized effect size of 0.85. This effect size metric is used throughout this report and is explained later in more detail. For the current purpose, it is important to know that an effect size this large is substantial and almost always has practical value in education. The experiments reviewed, however, were not well controlled and measured results from relatively small implementations where it is often easier to produce strong results. The TI experiments will contribute to this literature in two ways. First, we are operating in a context of use that is relatively large and represents realistic level of teacher professional development and support. Second, the program consists of a coherent set of elements including professional development, materials, and technologies all working together. In order to understand the use of the training and these materials in the classrooms, we have collected a wealth of qualitative data through regular surveys, observations and interviews of teachers.

The design of our experiment reflects the requirements of the No Child Left Behind Act, which directs schools to consult reports of rigorous research in making adoption decisions about instructional programs. This research is a randomized control trial (RCT) comparing student achievement outcomes for classes being taught using the TI graphing calculator technology and classes being taught with other materials. An RCT eliminates the variety of biases that regularly compromise the validity of research. While pretest and outcome measures were taken at the student level, randomization occurred at the teacher level. Participating teachers were paired and a coin was tossed to determine which teacher would use the graphing calculator program and which teacher would continue to use their existing math program.

Random assignment to experimental conditions does not assure that we can generalize the results beyond the districts where they were conducted. We designed our study to provide useful information to support local decisions that take into account the specifics of district characteristics and their implementation of the program. The results should not be considered to apply to school districts with practices and populations different from those in this experiment. The individual reports provide a rich description of the conditions of implementation in order to assist the district in strengthening its program and to provide the reader with an understanding of the context for our findings.

Methods

Research Design

Our study is a comparison of outcomes for classes taught using the TI graphing calculator technology and materials (*GC* group) and classes taught with the current materials in use in the district. Teachers volunteered for participation and from the pool of volunteers the researchers assigned equal numbers to the *GC* and control groups. The outcome measures are student level test scores. Within a multi-level model (students clustered in classes, classes within teachers), analyses of covariance are used to control for the influence of multiple factors both at the class level and student level, as well as to identify interactions with the experimental conditions.

Introduction of TI-Navigator™

We report on the first phase of the two year experiment, conducted during the 2005-2006 school year. This Phase I is divided into two parts. After one semester, we divided the *GC* group into teachers who continued to use the calculators and those using the calculators with TI-Navigator 2.0 (*GC* + *Nav*) as a further enhancement of the instructional capabilities of the technology. In this way we apportioned the first year into two semester-long experiments which also combine to form a year-long experiment. Besides allowing us to get an initial understanding of this enhancement, which was to be used systematically in the second phase of the experiment, this also allowed us to put all the elements into practice in the first year so that implementations could be refined as necessary for subsequent phases.

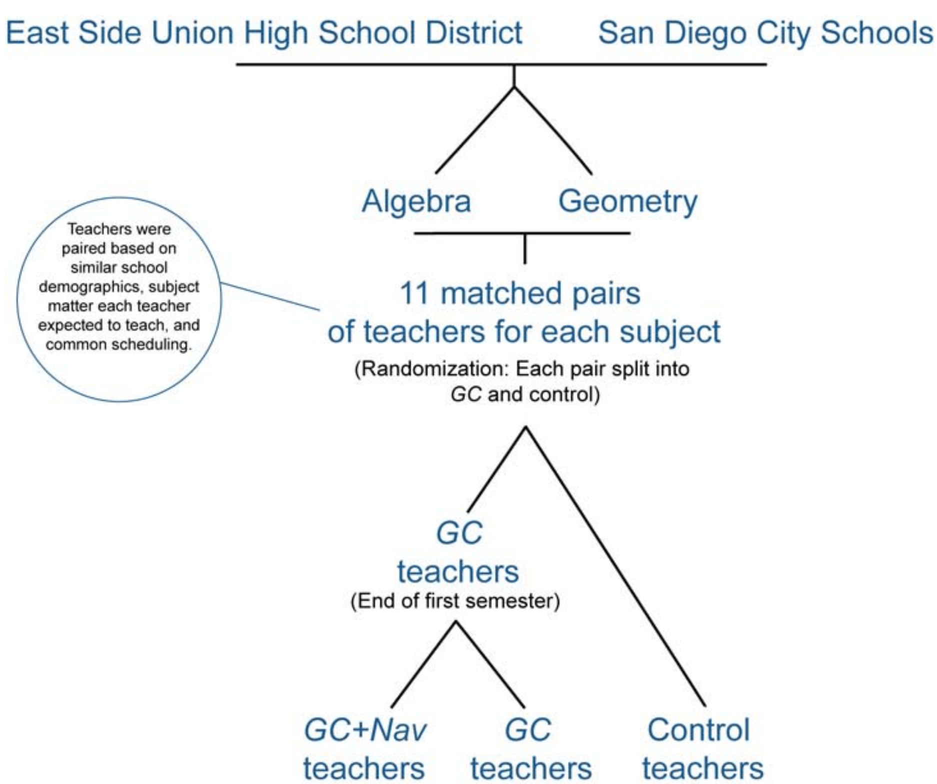

Figure 1. Research Design

Analysis of Sample Size Needed

Our research design assumed that we would report the results for the two districts independently as well as with the two districts combined. With the combined data we estimated that 40 teachers would be sufficient to detect an effect far smaller than our meta-analysis revealed for research on graphing calculators applied to Algebra I. We designed for an effect size of 0.3. An effect size is arrived at by dividing the effect by a measure of how dispersed the data points are (called the standard deviation). An effect size of .3 is three-tenths of a standard deviation.

The determination of the minimum detectable effect size involves making educated assumptions about design parameters. We assumed that we would be working with a fairly substantial correlation between the pre and post tests (.64). We also had to be concerned with how much of the variability in student outcomes was due to average differences at the teacher level. This intra class correlation (ICC) is important in designs that involve more than one level (in this case, randomization was done with teachers, but the outcome measures came from the students). Intuitively, the ICC is the proportion of the variability in student scores that can be accounted for by differences in teacher-level averages of the student scores. When the ICC is very large, for instance, much of the variation in student scores is accounted for by differences among teachers in their students' scores. If the differences among teachers are large and/or the differences within classes are small, then the sample size that matters the most for the experiment, is the number of teachers. If the differences among teachers are negligible then the sample size that matters most is the number of students. In general we need larger samples to detect smaller effects, and the ICC allows us to calculate how small an effect we can detect given available numbers of students and teachers. In this experiment we assumed a fairly conservative intraclass correlation of .22. This is a value that has been computed in other studies of math outcomes at the high school level.

In our calculation of a .3 minimum detectable effect size we also assumed conventional levels of tolerance for false-positive and false-negative outcomes, setting them at .05 and .20, respectively.

We also believed that the pairing of teachers prior to randomization would give our experiment additional power to detect effects in the 0.3 range.

Our experiment spanned two districts, and can be regarded as a multi-site trial. However, randomization was done at each site so we can consider results at each location separately. In other words, a separate experiment was performed at each site. The sample size at each location was smaller than the combined sample size. The minimum detectable effect size at each location will therefore be larger than .3. However, we believe that with the use of a matched pairs design, with the willingness to set tolerance for false-positive outcomes above conventional levels, and given the relatively large effect sizes for similar interventions in the meta-analysis we performed, that we are able to draw valid inferences about the impact of the intervention within each site.

We did not design our experiment specifically to detect results for subgroups of teachers at the same 0.3 level. We do, however, examine Algebra I and Geometry results separately recognizing that the effect would have to be larger to be detectable with the same level of confidence. *We caution that in the case where we are looking at results within-sites for subgroups of teachers, the minimum detectable effect size may be quite large, and failure to find an effect may be the result of not having adequate statistical power.* We also plan to examine the moderating effects of teacher characteristics as well as the impact of *GC + Nav* condition on student outcomes, but leave these to the analysis of both districts together. With a small number of teachers there can easily be chance imbalances of teacher characteristics that affect the outcome and thereby compromise our conclusions.

Examination of subgroups of students, such as students whom are not English proficient or students with different levels of incoming skill, is possible because each teacher will have some of each subgroup among their students. The power of an experiment where the intervention and randomization is conducted at the teacher level is largely dependent on the number of teachers rather than on the total number of students.

Questions Addressed

Our design allows us to address the following questions in the quantitative aspect of this research:

- Does the program consisting of graphing calculators and related technologies and professional development result in higher achievement in tests of Algebra I or Geometry?
- Is the impact of the program different for students with different incoming achievement levels or for English proficient?

The impact of teacher characteristics and the impact of different implementations will be addressed in the report of the combined data set that includes both districts. The impact of the TI-Navigator system that was introduced in the second semester will likewise be addressed in the context of the combined data.

Materials

The resources provided by Texas Instruments fall into three categories: training, equipment, and activity materials. The deployment and use of these resources is detailed more fully in the Results section of this report.

There were two sets of equipment deployed in this experiment, the TI-84 Silver Edition Plus graphing calculator (graphing calculator) system technology and the TI-Navigator system.

The graphing calculator was deployed to the treatment teachers during the first semester (October to January). The TI-Navigator system was deployed in January at the start of the second semester to a randomly selected subset of the treatment teachers.

According to Texas Instruments the graphing calculator system has several features that make it an appropriate Algebra I and Geometry learning tool:

- Connectivity with a variety of presentation tools that allow opportunities for demonstrations of graphing and analysis techniques, data collection and analysis, and problem solving methods
- Applications for Algebra I including linear equations, functions, and inequalities
- Cabri® Jr. Dynamic Geometry is an interactive Geometry application for constructing, exploring, and analyzing of a variety of geometric objects
- Applications in probability theory and statistics
- Data acquisition through the use of sensors (CBL/CBR)
- Students can share data and graphical displays via the presentation connector

The TI-Navigator system is designed to work with the TI graphing calculators and adds the following capabilities:

- Wireless communication between students' graphing calculators and the teacher's PC
- Activity center, quick polling, and screen capture activities.

Additional hardware provided by TI to all study teachers included a standard notebook computer with TI SmartView software, a data projector, and calculator-based laboratory (CBL/CBR) units. The text materials provided by TI for the trainings are discussed in the Implementation section, as are the workshops attended by the teachers.

Site Descriptions

Texas Instrument personnel initially introduced the researchers to the East Side Union High School District (ESUHSD) and the San Diego Unified School District (SDCS). After many discussions with the districts, several high schools expressed interest in participating in the research.

An introductory meeting was held at each school site and teachers who teach Algebra I and/or Geometry were invited to attend on May 31, 2005 in ESUHSD and June 15, 2005 in SDCS.

City of San Jose

San Jose, California is a large city located approximate 50 miles south of San Francisco. The total population of San Jose is 873,882 persons. The East Side area of San Jose has a population and ethnic breakdown as shown in Table 1.

Table 1. Demographics of the East Side Area

Ethnicity	Total population > 434,000
White	13.1%
Black	4.5%
Hispanic	42.4%
Asian	28.8%
Pacific Islander	1.1%
Filipino	9.7%
American Indian	0.4%

Source. East Side Union High School District, 2006

East Side Union High School District

The East Side Union High School District (ESUHSD) is located in San Jose, California. The district's General Fund (ESUHSD, 2006) is approximately $205 million. The Unrestricted portion of the General Fund accounts for about $127 million, the Special Education portion about $19 million and Categorical/Grants about $59 million. ESUHSD spends approximately $7,548 per student (U.S. Census Form F-33, 2004).

ESUHSD has 21 schools serving grades 9, 10, 11, and 12, one school serving grades K-12 and one school serving grades 10-12. Total enrollment for ESUHSD is 25,496 students. Table 2 provides information about the entire district including the high schools that participated in the study.

Table 2. Demographics of the Entire East Side Union School District

East Side Union High School District	
Total schools	23
Total teachers	1,148.5
Student to teacher ratio	22.2
Grades	KG - 12
Student population	25,496
Migrant students	911
ELL students	7,012
White	13.3%
Black	4.6%
Hispanic	43.8%
Asian	27.3%
Pacific Islander	0.9%
Filipino	9.4%
American Indian/ Native Alaskan	0.4%
Multi racial	0.2%
Source: CCD Public school data 2004-2005 school year	

East Side Union Schools

The research was conducted at nine schools sites in ESUHSD serving grades 9, 10, 11, 12. The specific demographics of each of the schools are reported in Table 3.

Change in Leadership

During the first year of the study, the district and four schools experienced a change in leadership. The East Side Union superintendent resigned from office after the 2004-2005 school year. In August of 2005, the ESUHSD school board announced the selection of an interim superintendent and installed him as the permanent superintendent in early 2006.

At the school level, one school (school ID #26) in the study officially received a new principal appointment in August of 2004. During the 2005-2006 school year, that same school also saw a new principal. Additionally, three schools (school IDs #23, #24, #27) in the study officially received new principal appointments in January of 2005 (ESUHSD website). Although the change in leadership did not impact the communication with the teachers in the study, it did impact the collection of the NWEA-MAP test data. Unlike the previous superintendent's agenda that emphasized the use of the NWEA-MAP for freshmen placement, the current superintendent did not explicitly continue that practice. Several principals stated that it was not part of the ongoing effort to gather placement data. Additionally, the newly appointed principals were not aware of the district's commitment to gather the data for the study.

Table 3. Participating High Schools' Demographics

	Participating East Side Union High Schools[a]								
School ID	**23**	**24**	**25**	**26**	**27**	**28**	**29**	**30**	**31**
Student population	2,744	2,090	1,123	4,005	1,563	1,509	2,006	2,124	2,181
Grades	9-12	9-12	9-12	9-12	9-12	9-12	9-12	9-12	9-12
Scheduling[b]	Daily	Daily	Daily	Daily	Daily	Daily	Daily	Block	Block
Migrant students	53	12	60	139	121	82	55	9	4
Student to teacher ratio	23.8	24.9	20.1	24.0	20.5	21.0	22.7	23.2	23.7
Title I	No	No	Yes	Yes	Yes	Yes	No	No	No
Free lunch	600	257	317	1018	691	695	496	212	209
Reduced lunch	170	57	79	153	87	159	107	49	75
Free/reduced	28.1%	15.0%	35.3%	29.2%	49.8%	56.6%	30.1%	12.3%	13.0%
White	23.9%	16.6%	9.7%	7.3%	2.2%	2.1%	9.8%	15.8%	49.8%
Black	8.6%	5.2%	2.0%	4.0%	2.0%	2.8%	5.6%	4.7%	5.1%
Hispanic	38.4%	19.4%	73.8%	33.8%	56.9%	73.6%	54.1%	22.5%	24.8%
Asian	23.4%	47.5%	6.9%	34.2%	30.3%	9.6%	17.6%	43.5%	15.1%
Pacific Islander	1.0%	0.4%	0.5%	1.0%	1.2%	2.4%	0.6%	1.2%	0.7%
Filipino	3.9%	10.4%	6.3%	18.9%	7.3%	9.4%	11.8%	12.1%	3.1%
American Indian/ Native Alaskan	0.8%	0.4%	0.5%	0.4%	0.1%	0.2%	0.4%	0.1%	0.6%
Multi racial	0.1%	0.1%	0.3%	0.4%	0.0%	0.0%	0.0%	0.2%	0.8%

[a] CCD Public school data 2004-2005 school year

[b] East Side Union High School District, 2006

Existing Math Programs

For the control classes, teachers used a variety of Algebra I and Geometry textbooks. Since the intervention did not include a textbook or a set curriculum, the graphing calculator intervention classes also used textbooks from the following publishers:

- *Algebra I, Exploration and Application* – McDougal Littell
- *Algebra I* – Prentice Hall
- *Geometry* – Holt
- *Geometry* – McDougal Littell
- *Algebra* – Addison-Wesley

The majority of teachers indicated that they use the McDougal Littell series. Calculators and graphing calculators were in use in mathematics classroom but not as a systematic part of the algebra or geometry programs. Control classrooms were not prohibited from using calculators so we expected to find some usage as a part of the existing programs.

City of San Diego

San Diego, California is a large city located approximate 130 miles south of Los Angeles. The total population of San Diego is 1,208,331 persons. San Diego has a population and ethnic breakdown as shown in Table 4.

Table 4. Demographics of San Diego

Ethnicity	Total population 1,208,331
White	63.2%
Black	6.8%
Hispanic	25.9%
Asian	15.8%
Pacific Islander	0.5%
American Indian	0.5%
Other	9.7%
Two or more races	3.5%

Source: United States Census Bureau, 2005

San Diego City Schools

The San Diego City Schools (SDCS) is located in San Diego, California. The total district budget (SDCS, 2006) is approximately $2.2 billion. The General Fund accounts for about $1.22 million dollars. SDCS spends approximately $8,007 per student (U.S. Census Form F-33, 2004).

SDCS has 221 schools serving grades Kindergarten through 12. Total enrollment for SDCS is 134,709 students. Table 5 provides information about the entire district including the high schools that participated in the study.

The research was conducted at 12 schools sites in SDCS serving grades 6 through 12. The specific demographics of each of the schools are reported in Table 6.

Table 5. Demographics of the Entire San Diego City Schools

San Diego City Schools	
Total schools	221
Total teachers	7,189.7
Student to teacher ratio	18.7
Grades	KG – 12
Student population	134,709
Migrant students	142
ELL students	37,076
White	25.8%
Black	14.2%
Hispanic	42.6%
Asian	8.6%
Pacific Islander	1.0%
Filipino	7.2%
American Indian/ Native Alaskan	0.5%
Multi racial	0.0%

Source: CCD Public school data 2004-2005 school year

Change in Leadership

SDCS experienced a district leadership change at the beginning of the 2005-2006 school year. On June 30, 2005, the SDCS superintendent resigned from office and a new superintendent was selected on July 23, 2005 (SDCS website, 2006).

The math department also experienced a change in leadership. Officially in July 2006, the director of the math department retired. The SDCS math department is currently directed by an interim director until a full-time replacement is found.

Table 6. Participating High Schools' Demographics

	Participating San Diego City Schools [a]												
School ID	**10**	**11**	**12**	**13[b]**	**14**	**15**	**16**	**17**	**18**	**20**	**21**	**22**	**35[b]**
Student Population	2350	2920	434	471	465	1462	2001	1685	2568	491	1036	2474	462
Grades	9-12	9-12	9-12	9-12	7-12	9-12	9-12	9-12	9-12	9-12	6-8	9-12	9-12
SchedulinGC	Block	Daily	Block	Block	Block	Daily	Daily	Block	Daily	Block	Daily	Daily	Block
Migrant Students	0	0	0	0	0	0	0	1	0	0	0	0	0
Student to Teacher Ratio	24.2	25.8	22.0	22.5	28.4	22.8	25.9	22.9	25.5	23.5	24.1	25.1	24.8
Title I	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	No	No	Yes
Free Lunch	349	1103	320	261	287	498	638	854	560	321	244	581	318
Reduced Lunch	116	392	33	45	35	119	259	186	264	48	134	222	44
Free/ Reduced	19.8%	51.2%	81.3%	65.0%	69.2%	42.2%	44.8%	61.7%	32.1%	75.2%	36.5%	32.5%	78.4%
White	51.3%	5.0%	3.0%	8.9%	13.3%	43.6%	31.6%	23.7%	25.4%	5.5%	47.3%	51.1%	1.9%
Black	8.8%	20.4%	29.7%	15.1%	17.4%	6.0%	17.0%	17.2%	9.9%	15.9%	13.0%	10.5%	3.5%
Hispanic	14.8%	28.0%	49.8%	74.3%	64.3%	43.7%	33.7%	48.2%	13.4%	77.0%	26.8%	27.6%	93.9%
Asian	16.6%	4.2%	16.8%	0.8%	1.9%	4.1%	10.5%	8.4%	19.0%	0.4%	7.8%	7.8%	0.6%
Pacific Islander	0.4%	2.5%	0.5%	n.a.	0.9%	0.8%	1.6%	0.9%	1.5%	0.0%	1.3%	0.8%	n.a.
Filipino	7.6%	39.7%	0.0%	n.a.	1.5%	1.0%	4.7%	0.8%	30.4%	1.0%	3.3%	1.1%	n.a.
American Indian/ Native Alaskan	0.5%	0.2%	0.2%	0.8%	0.6%	0.8%	0.8%	0.8%	0.3%	0.2%	0.5%	1.0%	0.0%
Multi-Racial	0.0%	0.0%	0.0%	n.a.	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	n.a.

[a] CCD Public school data 2004-2005 school year

[b] Asian ethnic category includes Pacific Islander and Filipino

[c] San Diego Unified School District, 2006

Existing Math Programs

For the *GC* and control classes, teachers used a variety of Algebra and Geometry textbooks. Since the intervention did not include a textbook or a set curriculum, all classes used textbooks from the following publishers:

- *Discovering Geometry* – Key Curriculum Press
- *Geometry* – McDougal Littell

- *Discovering Algebra* – Key Curriculum Press
- *Algebra Concepts and Skill* – McDougal Littell

The majority of teachers indicated that they use the Key Curriculum Press series. Calculators and graphing calculators were in use in mathematics classroom but not as a systematic part of the algebra or geometry programs. Control classrooms were not prohibited from using calculators so we expected to find some usage as a part of the existing programs.

Sample and Randomization

Recruiting

Texas Instrument personnel initially introduced the researchers to the ESUHSD and SDCS districts. We initially met with district staff members and principals from each site to explain the details and procedures of the study. Principals identified eligible teachers, who were then invited to after-school meetings. The initial meeting for the research experiment in ESUHSD occurred on May 31, 2005 with 22 teachers who teach Algebra I and/or Geometry. The initial meeting for the research experiment in the SDCS took place on June 15, 2005 with 22 teachers who teach Algebra I and/or Geometry.

Randomization

The randomization scheme required first finding pairs of teachers who were maximally similar on important factors. From each pair, one teacher is randomly assigned to the new program and the other becomes the control. This scheme avoids bias in the distribution of important characteristics between the group to be trained and to use the new materials and technologies (the *GC* treatment group) and the group to continue to use their normal practices (the control group).

After a question and answer period regarding the research methods and participant duties, 22 teachers at each site volunteered to participate. During the discussion, several important factors that teachers believe will have impact on the results in their district were identified. The sorting criteria for SDCS were similarity between school demographics, the subject matter each teacher expected to teach, and common scheduling (semester courses versus year-long courses). The sorting criteria for ESUHSD were years of teaching experience, Title I status, and non-English proficient population.

In some SDCS cases, the pairs were made within schools, but in most cases pairs were formed across schools with teachers of similar school demographics and expected subject taught. After we identified 11 matched pairs of teachers, each pair tossed a fair coin to determine which teachers was assigned to the graphing calculator (*GC)* group and which was assigned to the control group.

At ESUHSD, teachers were asked to identify the subject matter, Algebra I or Geometry, which they would teach for the 2005 – 2006 school year. Although assignments had not been finalized by their schools, teachers had reasonable expectation as to the classes they were likely to teach. Teachers who taught both subjects were grouped with other teachers teaching both. Within the groups, teachers lined up by years of teaching experience and were paired with the teacher most similar to their years of teaching experience. For several teachers with the same experience (mostly new teachers) they were paired based on school or on demographic similarities of their school, such as Title I status and non-English proficient population. In some cases, the pairs were made within schools, but in most cases pairs were formed across schools with teachers of similar years of teaching experience. After we identified 11 matched pairs of teachers for each site, each pair tossed a fair coin to determine which teachers was assigned to the *GC* group and which was assigned to the control group.

Of the original 22 teachers that agreed to participate at ESUHSD, one asked to be excused from the study because she was reassigned to work at the district level and could no longer be a considered a classroom teacher for the purposes of the study.

Of the original 22 teachers who agreed to participate at SDCS, six teachers had to be excused from the study. Before the pretest administration and after the randomization (October 2005), two control teachers and two treatment teachers left the study. One control teacher was reassigned to a new school and no longer taught eligible classes. The second control teacher had to leave due to personal reasons unrelated to the study. One treatment teacher was reassigned to a different school district and no longer taught eligible classes. The second treatment teacher asked to leave the study because the perceived responsibilities of implementing the treatment would be overly demanding. Shortly after the start of the second semester (February 2006), one treatment and one control teacher asked to leave the study. Both teachers cited the requirements for participation in the study were overwhelming.

Sample Size

Sample size is one of the factors that determine how precisely we can measure the magnitude of an effect. With smaller samples we are usually able only to detect larger effects. We usually measure the size of an effect in terms of standard deviation units, which tells us how big the effect is, controlling for the spread in observed scores.

Our research design assumed that we would report the results for the two districts independently as well as with the two districts combined. With the combined data we estimated that 40 teachers would be sufficient to detect an effect far smaller than our meta-analysis revealed for research on graphing calculators applied to algebra. We designed for an effect size of 0.3. An effect size is arrived at by dividing the effect by a measure of how dispersed the data points are (called the standard deviation). An effect size of .3 is 3 tenths of a standard deviation.

The determination of the minimum detectable effect size involves making educated assumptions about design parameters. We assumed that we would be working with a fairly substantial correlation between the pre and post tests (.64). We also had to be concerned with how much of the variability in student outcomes was due to average differences at the teacher level. This intra class correlation (ICC) is important in designs that involve more than one level (in this case, randomization was done with teachers, but the outcome measures came from the students). Intuitively, the ICC is the proportion of the variability in student scores that can be accounted for by differences in teacher-level averages of the student scores. When the ICC is very large, for instance, much of the variation in student scores is accounted for by differences among teachers in their students' scores. If the differences among teachers are large and/or the differences within classes are small, then the sample size that matters the most for the experiment, is the number of teachers. If the differences among teachers are small then the teacher level contributes relatively less to the uncertainty in the outcome and variation among students within classes becomes more important. In general we need larger samples to detect smaller effects, and the ICC allows us to calculate how small an effect we can detect given available numbers of students and teachers. In this experiment we assumed a fairly conservative intraclass correlation of .22. This is a value that has been computed in other studies of math outcomes at the high school level.

In our calculation of a .3 minimum detectable effect size we also assumed conventional levels of tolerance for false-positive and false-negative outcomes, setting them at .05 and .20, respectively.

We also believed that the pairing of teachers prior to randomization would give our experiment additional power to detect effects in the 0.3 range.

Our experiment spanned two districts, and can be regarded as a multi-site trial. However, randomization was done at each site so we can consider results at each location separately. In other words, a separate experiment was performed at each site. The sample size at each location was smaller than the combined sample size. The minimum detectable effect size at each location will therefore be larger than .3. However, we believe that with the use of a matched pairs design, with the willingness to set tolerance for false-positive outcomes above conventional levels, and given the relatively large effect sizes for similar interventions in the meta-analysis we performed, that we are able to draw valid inferences about the impact of the intervention within each site.

We did not design our experiment specifically to detect results for subgroups of teachers at the same 0.3 level. We do, however, examine algebra and geometry results separately recognizing that the effect would have to be larger to be detectable with the same level of confidence. *We caution that in the case where we are looking at results within-sites for subgroups of teachers, the minimum detectable effect size may be quite large, and failure to find an effect may be the result of not having adequate statistical power.* We also plan to examine the moderating effects of teacher characteristics but leave these to the analysis of both districts together. With a small number of teachers there can easily be chance imbalances of teacher characteristics that affect the outcome and thereby compromise our conclusions.

Examination of subgroups of students, such as students not proficient in English or with different levels of incoming skill is possible because each teacher will have some of each subgroup among their students. The power of an experiment where the intervention and randomization is conducted at the teacher level is largely dependent on the number of teachers rather than on the total number of students.

Data Sources and Collection

The data were collected over the entire period of the experiment beginning with the randomization meeting and ending with the academic calendar of the district in June 2006. Training observations, classroom observations, informal and formal interviews, multiple teacher surveys, email exchanges, and phone conversations were used to provide both descriptive and quantitative evidence of the implementation.

The randomization meeting gave us an initial set of teachers who would be in the project. The initial “intake sheets” provided us with contact information and some background on each teacher which is being supplemented through surveys and interviews.

Interview Data

Interview data were used to provide a description of the program implementation from the perspectives of participating teachers. We also used the interview data to develop more focused case studies on selected teachers’ use of graphing calculators. The case studies are presented in the Implementation section and inform the question: *what do teachers need to know in order to plan a series of lessons using the graphing calculator for demonstration, data gathering, and analysis?* Informal interviews occurred throughout the study in conjunction with classroom visits, phone conversations, and emails. Formal interviews occurred between May and June 2006 and lasted approximately 30 minutes each over the phone.

Student Demographics

Since the meeting was in the spring, the class rosters and final assignments were not available until the school year had begun. At that point, we requested the class rosters and from that compiled a complete list of student ID numbers. This list was the basis for requesting additional demographic information and prior test scores from the district. We integrated the class roster information and all the data from the district into a comprehensive data warehouse.

California schools have a general policy of not providing data regarding socioeconomic levels (free and reduced lunch status) for individual students and only reported this data at the school level. Hence, no analyses were conducted using this information as a covariate.

ESUHSD and SDCS provided Empirical Education with class roster data in September 2005. The 2005-2006 rosters consisted of student demographics, assessment scores, and teacher and classroom assignments. The relevant demographic data was used as a student level covariate in the analyses and to identify NCLB student subpopulations. SDCS has a general policy of not providing socioeconomic data (free and reduced lunch status) for individual students. No analyses could be conducted using socioeconomic status as a covariate.

Survey Data

Surveys were deployed on a bi-weekly basis to both *GC* and control group teachers beginning on January 16, 2006 and concluding on May 19, 2007. Through the surveys we measured and documented the intervention implementation to provide qualitative descriptions of the program. We also used the survey data to quantify the extent of exposure to the materials (opportunities to learn with the technology) as a function of the time teachers spent using the graphing calculator as a demonstration/instructional tool and used as a function of the time students spent using the graphing calculator. Teacher self-report data on each of these indicators was averaged. Nine surveys were administered between January and May to all of the participating teachers. Survey response rates for ESUHSD and SDCS are presented in Appendix A1 and A2 respectively.

Observational Data

Technology Training

Training observation data were minimally coded using the protocols developed by researchers. The data were used to inform researchers how the intervention was expected to be used in class. There were two observed training events involving SDCS. The initial graphing calculator training was held at SDCS from August 9 – 15, 2005. The initial graphing calculator training was held at ESUHSD from August 8 – 12, 2005. The initial TI-Navigator training was held at ESUHSD from January 19 – 21, 2006.

Classroom Observations

Classroom observational data were minimally coded using the protocols found in Appendix C. In general, observational data is used to further inform the description of the learning environment, instructional strategies employed by the teachers, and student engagement. All of the TI-Navigator classrooms and a sample of control classrooms were observed in April 2006.

Achievement Measures

The primary achievement measures are student-level scores on the Northwest Evaluation Association (NWEA) test and the California Standards Test of mathematics (CST-math). The NWEA was administered for this study as a pretest, mid-year test, and posttest. The CST-math was given to all students as part of the California's state assessment system.

CST-Math

The California Standards Test of mathematics achievement (CST-math) is given to all students in grades 2 through 11 as part of California's assessment system. This is criterion referenced and, for students in this experiment, was differentiated into general math, Algebra I and Geometry, depending on the course they were completing during the testing year. The test does not provide vertical alignment and the three versions of the test cannot be compared on the same scale. CST-math scores could not be used a pretest. All analyses that use CST scores as the outcome also use the fall NWEA as the pretest covariate.

NWEA

The NWEA Measures of Academic Performance (MAP) used at ESUHSD and SDCS is a state-aligned, computer-adaptive assessment program that provides educators with information designed to improve teaching and learning. This adaptive test reflects the instructional level of each student and measures growth over time. The test goal structures are created through an alignment process that links state standards documents to the NWEA item bank.

The NWEA Achievement Level Test (ALT) used at SDCS is a state-aligned, leveled, paper-pencil assessment that provides educators with information designed to improve teaching and learning. The ALT test is a two step process where a student initially takes a short locator test that indicates a student's math ability and determines which leveled test booklet is appropriate for the student's current achievement level.

For the pretest and mid-year test, the general goal areas or strands within a state's standards become the goals measured. In the Math Goals Survey 6+ CA V2, Number Sense, Algebra and Functions, Measurement and Geometry, Statistics and Probability, and Mathematical Reasoning, are the strands aligned to California's standards. In the Math Goals Survey 6+ CA V1, Number Sense, Estimation Computation, Algebra and Functions, Geometry, Measurement, and Statistics, Data, and Probability, are the strands aligned to California's standards.

For the posttest, the general goal areas or strands within a state's standards become the goals measured. In the Algebra I End of Course, Linear Equations, Quadratic Equations, Algebraic Operations, and Problem Solving are the strands aligned to California's standards. In the Geometry End of Course, Spatial Relationships, Measurement, Geometric Relationships, and Problem Solving are

This test is scored on a Rasch unIT (RIT) scale, a measurement scale developed to simplify the interpretation of test scores. This scale is used to measure student achievement and student growth on an equal-interval scale so that a change of one unit indicates the same change in growth, regardless of the actual numerical values. RIT scores range from about 150 to 300 and indicate a student's current achievement level along a curriculum scale for a particular subject. In addition to the overall RIT score, the measures also provide goal score ranges to aid in the identification of a student's instructional levels.

Testing Schedule

Pretest

The CST-math results are readily available for the majority of students. However, because of the make-up of the participating Algebra and Geometry classes some students were in the 8th grade while others were in the 9th, 10th, and 11th grades when they took the test. Consequently, scores obtained were on three different tests: General Math, Algebra I, and Geometry. Since the tests covered different topics in mathematics and were not aligned, the scores were not useable as a pretest. A more detailed explanation of the items assessed by the three CST-math is available in the California Blueprints Preface (California Department of Education, 2006).

School districts administered ALT and MAP versions of the NWEA Math Goals Survey 6+ CA. To minimize the loss of instructional time, teachers typically administered the test during one class period with one day of make-up testing in their own classrooms or school computer lab. Since the ALT version is a paper-pencil test, the logistics behind delivering and recovering the tests proved difficult. In some cases, copies of the tests were misdirected or unintentionally received at a different school. Teachers also had problems completing and returning the answer sheets. In some cases, the tests were not administered. In other cases, answer sheets were packaged to be returned but became delayed, misplaced or lost entirely. Upon receiving the answer sheets, researchers noted that some answer sheets were incomplete or improperly labeled. The above factors contributed to data loss on the ALT version of the test.

For the MAP version of the test, teachers typically administered the test at the school's computer lab. Some schools that tested with the MAP version pulled students from their PE classes, while other schools tested students in small groups on classroom computers or small computer labs. Given the different testing scenarios, some students were over-looked and not tested at all. Students absent during the testing, students new to the school, and students who recently left the school were not tested. Students guessed on too many items yielded invalid scores. These factors contributed to data loss on the MAP version of the test.

Mid-Year NWEA-MAP

At mid-year, at the beginning of the 2nd academic semester (February) we asked the district to once again test the participating classes using the NWEA-MAP Math-Goals Survey 6+. Although we supervised the testing more closely, there continued to be problems. At some schools access to the computer labs was problematic because the labs were routinely used to teach scheduled classes and the required testing schedule interfered with on-going instruction. At still others, technical issues continued to plague the reporting of test scores.

End of Year Posttest

In the spring of 2006 (May – June), the CST Algebra and Geometry tests were administered to all of the students in the district. These test scores became available in September and were used in the outcome analyses.

NWEA MAP test were also administered during the same timeframe. The appropriate end of course Algebra and Geometry tests were given to students. At the end of academic year, we supervised the district NWEA-MAP testing more closely, and did not assume that the district routines were effective in generating the test results we needed. There continued to be problems in scheduling, the use of the software, and the reporting of results. We were much more successful at obtaining results because the district point of contact secured access to the IT structures more directly and we began enlisting the help of the various principals at the schools. The final counts of the data are provided in the following section.

Test Administration

Table 7 shows the kinds of tests that were administered at each location. After the intervention, two tests were administered – the NWEA test and the CST test. There were two versions of the NWEA test – a computerized adaptive version and a paper-and-pencil version. Both the CST and the NWEA tests tested either geometry or algebra (depending on the courses that students were enrolled in.) The NWEA test of basic math skills that was administered in the Fall served as the pretest for all analyses.

Table 7. Test Administration

		ES	SD
NWEA	**Algebra**	Computerized adaptive	Paper-pencil Except for 2 teachers who use computerized adaptive (no pretest for control)
	Geometry	Computerized adaptive	Paper-pencil Except for 2 teachers who use computerized adaptive
CST	**Algebra**	Standard test	Standard test (no pretest for control)
	Geometry	Standard test	Standard test

Note. In all cases the NWEA general test of math skills was used as the pretest.

Several problems were encountered in testing. They are described below:

1. The NWEA tests of Geometry and Algebra tested distinct constructs in math and the results could not be combined onto one scale. Despite initial assurances from NWEA that the tests were scaled together, we discovered that they were not. This required us to run separate analyses for Algebra and Geometry, thereby reducing the sample size.
2. The lack of compatible scaling described in (1) prevented us from doing a repeated-measures design using a longer time series. We collected data on three occasions and were planning to do an HLM-type repeated measures analysis; however, the test of general math knowledge that was administered in the Fall and Winter was not scaled with tests given in the Spring. An exception to this was in San Jose where four teachers accidentally re-administered the test of general math ability in the Spring, which allows us to conduct a limited repeated measures analysis as planned.

3. The use of both paper and pencil as well as computerized adaptive version of tests in San Diego produced initial results that were distributed in peculiar ways. On closer examination it became apparent that the two kinds of tests were not scaled together and the paper and pencil test led to a considerable floor effect (there were two teachers in San Diego who administered the computerized adaptive test.) In the San Diego report we eliminated cases below the floor (i.e., the results for students of the two control teachers who administered the computerized adaptive test) and performed analyses only on students who took the paper-and-pencil test, with the caveat that the analysis applied only to higher-performing students, since only those above the floor were included in the analysis. We then operated under the assumption that above the floor the computerized adaptive and paper and pencil tests were equated.
4. Pretest scores were unavailable for algebra students in San Diego. The use of the pretest score as a covariate in the combined analysis for algebra would lead to the deletion of almost all cases in Sand Diego (assuming we drop cases that are missing data.) We proceeded to do the combined analysis without using the pretest – this is a tradeoff in precision (we increase the sample size but eliminate the pretest covariate; the benefit is that the result is potentially more relevant to both sites because it uses cases from both sites.)
5. CST pretest measures could not be used in the analysis. The CST pretest results were reported on several incompatible scales that would have required us to subset the data for different students almost immediately. All students took the same NWEA pretest of general math, so we used this as the covariate instead.

The problems described above led us to consider NWEA outcomes only in a limited way in the combined report. The NWEA pretest served a useful purpose as a pretest covariate given the unavailability of CST pretest scores.

Statistical Analysis and Reporting

The basic question for the statistical analysis was whether, following the intervention, students in the *GC* classrooms had higher math scores than those in the control classrooms. The mean impact is estimated using multi-level models that account for the clustering of students in classes, which provides a more accurate, and often more conservative, assessment of the confidence we should have in the findings. We use SAS PROC MIXED (from SAS Institute Inc.) as the primary tool for this work. To increase the precision of our estimate, we include students' pretest scores in the analysis. In our experience, these are good predictors of achievement; including them as covariates in the impact analysis reduces the error variance, which makes it easier to discern the treatment impact.

In addition to the basic analysis of the mean impact, the plan for the study identifies the teacher- and student-level covariates that we expect (through theory or prior research) to make a difference in the effectiveness of the program being tested. The analysis tests for the interactions between these covariates and the experimental condition.

In addition to examining impacts and interactions where we anticipate effects, to better understand unexpected results, we use other demographics, teacher characteristics, and supplementary observational data in exploratory analyses to generate additional hypotheses about which factors potentially moderate the treatment impact.

Our analyses produce several results: among them are the coefficients for fixed effects, effect sizes, and *p* values. These are found in all the tables where we report the results of the statistical models.

Estimates. The estimate can be thought of as a prediction of the size of an effect. Specifically, it is how much we would predict the outcome to change for a one-unit increase in the corresponding variable. We are often most interested in the estimate associated with the experimental conditions, which is the expected change in outcome in going from control to treatment, holding other variables constant.

Effect sizes. We also translate the difference between treatment and control into a standardized effect size by dividing the difference by the amount of variability in the outcome (also called the standard deviation). This allows us to compare the results we find with results from other studies that use different measurement scales. In studies involving student achievement, effect sizes as small as 0.1 (one tenth of a standard deviation) are sometimes found to be important educationally. When possible we also report the effect size of the difference after adjusting for pretest, since that provides a more precise estimate of the effect (i.e. in theory, with many replications, we would expect the adjusted effect size on average to be closer to the true value).

***p* values.** The *p* value is very important because it gives us a gauge of how confident we can be that the result we are seeing is not due simply to chance. Specifically, it tells us what the probability is that we would get a result with a value as large as – or larger than – the absolute value of the one observed when in fact there is no effect. Roughly speaking, it tells us the risk of concluding that the treatment has had an effect when in fact it hasn't. Thus a *p* value of .1 gives us a 10% probability of that happening. We can also think of it as the level of confidence, or the level of belief we have that the outcome we observe is not simply due to chance. While ultimately depending on the risk tolerance of the user of the research, we suggest the following guidelines for interpreting *p* values:

1. We have a high level of confidence when $p \leq .05$. (This is the level of confidence conventionally referred to as "statistical significance.")
2. We have some confidence when $.05 < p \leq .15$.
3. We have limited confidence when $.15 < p \leq .20$.
4. We have no confidence when $p > .20$.

East Side Results

Our report of the ESUHSD results begins with an analysis of the composition of the experimental groups that were formed through the randomization of teachers prior to training and student testing. We also provide an accounting of the loss of teachers and students through mobility during the year and, importantly in this case, through difficulties in collecting the pretest data. In the second part of this section, we report on the implementation of the program including what we learned about the professional development, use of the technology and the issues raised by the participants.

The third section contains the quantitative results. We examined the impact of Graphing Calculators (*GC*) separately for Algebra and Geometry. For each of these strands we used two-outcome measures: NWEA tests of Algebra or Geometry and the California Standards Test (CST) of Algebra or Geometry. First, we report the impact of *GC* on Algebra learning, and then we report the impact of *GC* on Geometry learning. We use the NWEA test of general math achievement from the fall assessment as a covariate in the analysis.

The basic question for the statistical analysis was whether, following the intervention, students in *GC* classrooms had higher scores than those in control classrooms. In addition to looking at the main effect of *GC* and doing related analyses of covariance (ANCOVAs), we estimated the interactions of condition (*GC* versus control) with the pretest and English proficiency. In particular, we were interested in whether *GC* is more effective with English proficient students compared to non-English proficient students and whether *GC* was differentially effective for low- and high-performing students. These two moderators were identified before the intervention started and we report the results of both moderators in all the analyses. If the moderator effects are significant, we also graph the results.

Our plan initially also was to examine teacher experience as a moderator systematically throughout the analyses. However, we find that the participants in this study were, with one exception, relatively experienced math teachers so those analyses were not included here.

The impact of the TI-Navigator system that was introduced in the second semester will be addressed in the context of the combined data.

Formation and Attrition of the Experimental Groups

Groups as Initially Randomized and Pretest Attrition

The randomizing process does not guarantee that an experiment's groups will be perfectly matched. It simply guarantees that there is no selection bias. It is important to inspect the two groups to determine whether there is an imbalance between the *GC* and control groups on important factors that may affect the results. This section addresses the nature of the groups in each of the content areas.

First, we trace the process by which cases were subset, leading to the samples used in the analyses. We focus on student counts, with the recognition that teacher and class numbers are also determined by the process of subsetting cases. We give a table of student, class and teacher counts later in the discussion.

Altogether, we collected data on 4984 students across both East Side and San Diego. At East Side, there were 3126 students and the remaining 1858 students are at San Diego. (The results for the San Diego arm of this experiment are described in separate reports.) From here forward we consider only the students at East Side.

To begin with, each student can be characterized in terms of whether he or she: 1) is taking a Geometry or Algebra class; 2) is assigned to treatment or control (it is possible for a student to switch conditions or have no assignment status in one semester or the other or both); and 3) whether he or she has test scores in the Fall or Spring. In the analyses to follow, we remove students who switched conditions or were not initially assigned a condition. We also remove students who did not have an NWEA pretest score or NWEA or CST posttest score.

Counts of Cases in the Attrition and Formation Section

In the sections that follow, where we explore attrition and formation, *we consider only cases that have been assigned to treatment or control in the Fall* (shaded rows above). The total number of cases is 1525. These can be broken down further:

There were 667 cases enrolled in Geometry in the Fall. Of these, 586 have non-missing treatment assignment in the Fall and Spring (81 are missing treatment assignment for the Spring.) Out of the 586 cases, 348 have a Fall score, and of these, 317 have both a Fall and a Spring score.

There were 858 cases enrolled in Algebra in the Fall. Of these, 628 have non-missing treatment assignment in the Fall and Spring (230 are missing treatment assignment for the Spring.) Of the 628 cases, 318 have a Fall score, and of these 310 have both a Fall and a Spring score.

Table 8 through Table 32 report the number of teachers, classes, and students available for analysis given availability of CST posttest scores. We observe that between the initial teacher randomization and class assignment and the later gathering of pretest data, there was substantial attrition. Issues in obtaining pretest data through the computerized NWEA system led to a loss of 4 teachers, 14 classes and approximately half the students enrolled in the classes of teachers initially randomized. It is clear also from inspection of the table that there was substantially more attrition from the *GC* group than from the control. There were no other sources of attrition prior to the pretest. Fortunately, the portion of teachers lost at this point was much smaller. In other words, many teachers remained in the experiment but with fewer classes with pretests or with classes with fewer students with pretests.

Subsequent attrition was far less severe. No teachers were lost. One class was lost. Nineteen students who had pretests did not have posttests. Because the students who were not tested initially were still in the classes, we find that the number of students with posttests and who had been enrolled in the fall was quite large.

We do not believe that there were any *GC*-related reasons for the differential attrition and therefore it does not represent an obvious source of bias. However, whenever a substantial loss is encountered, a level of uncertainty enters into the interpretation of results.

Table 8. Teachers in *GC* and Control Groups

		Initially randomized	Completed program	Teachers with at least one student having an NWEA pretest	Teachers with at least one student having a CST posttest	Teachers with at least one student having an NWEA pretest and a CST posttest
Algebra	***GC***	8	8	6	8	6
	Control	8	8	8	8	8
Geometry	***GC***	5	5	3	5	3
	Control	3	3	3	3	3
Totals		**24**[a]	**24**	**20**	**24**	**20**

Notes. A total of 22 teachers were initially randomized. Two teachers taught both Algebra and Geometry subjects.

Table 9. Classes in *GC* and Control Groups

		Initially randomized	Completed program	Classes with at least one student with an NWEA pretest	Classes with with at least one student with a CST posttest	Classes with at least one student with an NWEA pretest and a CST posttest
Algebra	***GC***	19	18	13	17	12
	Control	21	20	19	20	19
Geometry	***GC***	14	14	7	14	7
	Control	8	8	8	8	8
Totals		**62**	**60**	**47**	**59**	**46**

Table 10. Students in *GC* and Control Groups

		Initially randomized at the start of academic year	Students with an NWEA pretest[a]	Completed program	Students with an NWEA pretest[b]	Students with a CST posttest[b]	Students with both an NWEA pretest and a CST posttest[b]
Algebra	***GC***	401	110	266	96	243	93
	Control	457	253	362	222	349	217
	Totals	**858**	**363**	**628**	**318**	**592**	**310**
Geometry	***GC***	396	140	339	133	300	131
	Control	271	208	247	190	211	171
	Totals	**667**	**348**	**586**	**323**	**511**	**302**
Totals		**1525**	**711**	**1214**	**641**	**1103**	**612**

[a] Students who were initially randomized

[b] Students who completed the program

Post Randomization Composition of the Experimental Groups

We were able to examine the English proficiency of the original sample of students assigned to classes of the teachers originally randomized. This was done separately for students enrolled in Algebra and Geometry classes. We see in Table 11 and Table 12 that English proficiency was not distributed evenly between the conditions in spite of randomization. For the Algebra classes, there are proportionally more non-proficient students in the control group than in the *GC* group. For the Geometry classes, there are proportionally more non-proficient students in the *GC* group compared to the control group. Chi-square tests indicate that despite randomization, English proficiency was not balanced between conditions. The imbalance may lead the estimate of the impact to depart from its true value.

Table 11. Comparison of English Proficiency between *GC* and Control Group Algebra Students

Condition	English Proficiency: Not proficient	English Proficiency: English proficient	Totals
GC	134	266	**400**
Control	191	266	**457**
Totals	**325**	**532**	**858**[a]
Statistics	**DF**	**Value**	***p* value**
Chi-square	1	6.23	.01

[a] This count is for Algebra students assigned to a treatment condition in the Fall.

Table 12. Comparison of English Proficiency between *GC* and Control Group Geometry Students

Condition	English Proficiency: Not proficient	English Proficiency: English proficient	Totals
GC	95	301	**396**
Control	6	265	**271**
Totals	**101**	**566**	**667**[a]
Statistics	**DF**	**Value**	***p* value**
Chi-square	1	59.38	<.01

[a] This count is for Geometry students assigned to a treatment condition in the Fall.

Characteristics of the Experimental Groups Defined by Pretest

From this point in the analysis forward, we treat the teachers, classes and students as the sample for analysis.

With randomization, we expect the pretest scores to be equally distributed between *GC* and control groups, but in any single randomization there may be discrepancies between the distributions due to chance. In the following tables we use the effect size measure as a way to consider the size of the initial differences (this is a measure of the extent of imbalance on the pretest in spite of randomization and is not the result of a cause.) For the Algebra students, as shown in Table 13, the groups were reasonably well matched on the pretest scores for the fall assessment. For the Geometry students *GC* and control groups had slightly different average pretest scores as shown in Table 14. However, when we accounted for the fact that outcomes for students of the same teacher tend to be dependent by modeling these dependencies, the discrepancy became less discernable (for both the Algebra and Geometry the *p value*s rose to levels above .20). In the analyses that follow, we add the pretest covariate in order to increase the precision of the impact estimate.

Table 13. Differences in Pretest Scores for Algebra Students in *GC* and Control Groups

Descriptive statistics	Raw group means	Standard deviation	Number of students	Standard error	Effect size
GC	225.89	16.08	110	1.53	0.13
Control	224.79	14.50	253	0.91	
***t* test for difference between independent means**	**Difference**		**DF**	***t* value**	***p* value**
Condition (*GC* - control)	1.93		361	1.13	.26
Note. Pretest scores for Fall NWEA test					

Table 14. Differences in Pretest Scores for Geometry Students in *GC* and Control Groups

Descriptive statistics	Raw group means	Standard deviation	Number of students	Standard error	Effect size
GC	238.54	12.57	140	1.06	.26
Control	235.25	13.03	208	0.90	
***t* test for difference between independent means**	**Difference**		**DF**	***t* value**	***p* value**
Condition (*GC* - control)	3.28		346	-2.34	.02
Note. Pretest scores for Fall NWEA test					

Attrition After the Pretest

From among the students who had pretests, we identified those who had posttests for the CST test. We see in Table 15 that 91% of the Algebra students who had a pretest also had a CST posttest Algebra score. A Chi-square test indicates no difference between the two conditions in the proportion of students missing the posttest score.

Table 15. Availability of Pre- and Posttest Scores for Algebra Students

Condition	Having both pre- and posttest scores	Having pretest scores only	Totals
GC	100	10	**110**
Control	230	23	**253**
Totals	**330**	**33**	**363**
Statistics	**DF**	**Value**	***p* value**
Chi-square	1	0.00	1.00
Note. Pretest refers to Fall NWEA test; posttest refers to Spring CST			

Table 16 shows that students with a posttest may have scored a bit higher on the pretest compared to students without a posttest. The effect could easily be due to chance. The lack of differential attrition and the close performance shown in Table 15 rules out the possibility that bias was introduced by retaining different students in the two conditions (where the difference is measured in terms of how well students perform at the outset.)

Table 16. Difference in Pretest Scores for Algebra Students with Pre- and Posttest Vs. Posttest Only

Descriptive statistics	Raw group means	Standard deviation	Number of students	Standard error	Effect size
Having pretest scores only	220.97	15.43	33	2.69	0.23
Having both pre- and posttest scores	224.52	14.95	330	0.82	
***t* test for difference between independent means**	**Difference**		**DF**	***t* value**	***p* value**
(Having pretest scores only) – (Having both pre- and posttest scores)	-3.56		361	-1.30	.19
Note. Pretest refers to Fall NWEA test; posttest refers to Spring CST					

We see in Table 17 that 91% of the Geometry students who had a pretest also had a CST posttest Geometry score. However, a Chi-square test indicates that despite the small rate of dropout overall, a greater proportion of students dropped out of the control condition than out of the *GC* condition. The *p* value is <.01 which indicates that this difference in attrition rates is not likely to happen by chance.

Table 17. Availability of Pre- and Posttest Scores for Geometry Students

Condition	Having both pre- and posttest scores	Having pretest scores only	Totals
GC	134	6	**140**
Control	183	25	**208**
Totals	**317**	**31**	**348**
Statistics	**DF**	**Value**	***p* value**
Chi-square	1	6.17	.01
Note. Pretest refers to Fall NWEA test; posttest refers to Spring CST			

Table 18 shows that, as with the Algebra classes, students with no score for the posttest scored slightly lower on the pretest. We have some confidence that this effect is not due to chance. While involving a very small number of students, there is the possibility that this pattern of attrition biases the *GC* effect down somewhat.

Table 18. Difference in Pretest Scores for Geometry Students with Pre- and Posttest versus Posttest Only

Descriptive statistics	Raw group means	Standard deviation	No. of students	Standard error	Effect size
Having pretest scores only	233.42	16.09	31	2.89	-0.24
Having both pre- and posttest scores	236.88	12.56	317	0.71	
***t* test for difference between independent means**	**Difference**		**DF**	***t* value**	***p* value**
(Having pretest scores only) – (Having both pre- and posttest scores)	-3.46		346	-1.43	.15

Note. Pretest refers to Fall NWEA test; posttest refers to Spring CST

Implementation Results

In this section we describe more fully the aspects of the implementation that characterize this intervention. We used the following questions to guide our descriptions and analysis: What resources are needed to manifest the graphing calculator and the TI-Navigator condition? Are there differences in the extent, quality, and type of implementation of the materials? Our perspective takes into account three levels of resources needed to implement the intervention: those resources provided by TI, those provided by the school district and individual schools, and those provided by the teacher. We discuss each level separately. We also studied the features of the implementation to identify possible variables related to the outcome measures.

In year 1, TI did not communicate an explicit implementation model to the *GC* teachers for either the graphing calculator or for the TI-Navigator based graphing calculator system. TI did communicate an implicit *ideal* implementation framework based on the guidelines of the National Council of Teachers of Mathematics (NCTM) *Principles and Standards for School Mathematics* as outlined in the resource materials. During the training, the implicit implementation framework was rarely modeled by the trainer. Teachers were not directed to select particular activities nor were they directed to adopt the style of inquiry mathematics implied by the exploratory framework of the activities.

It is difficult to characterize deviations from the recommended practices given the implicit nature of the implementation framework. Nevertheless, given that there is an implicit structure following NCTM (2000) guidelines and since all of the teachers in the study are additionally provided structure by the California mathematics teaching standards, we can outline certain characteristics that mark low, medium and high implementation environments. We also recognize that this type of field-based implementation has integrity of its own and that certain elements must be present before it can be distinguished from the control condition. The characteristics and activities which identify an *ideal GC* condition are described in the Issues in Rating the Level of Implementation section.

Teacher Background

Teachers' background, attitudes, and preparation are commonly thought to impact the level of implementation (Fullan & Profret, 1977; Mukti, 2000; Supovitz & Turner 2000; Thompson, 2005). During the randomization process teachers identified themselves according to years of teaching experience and the initial teacher pair was formed correspondingly so that the bias due to teaching experience would be distributed among the groups evenly. As part of our data collection we asked the teachers to provide us with information regarding their backgrounds. The following tables summarize the background characteristics of the teachers in the study.

In general, most of the teachers in the study were highly experienced professionals and held college degrees with either a mathematics major or minor and in one case a master's degree.

Table 19. Total Number of Years Teaching Experience[a]

		Early career (0-3 Years)	Emerging professional (4-6 Years)	Mid-career professional (7-15 Years)	Highly experienced professional (15+ Years)
Condition	**Number of teachers**	**No.**	**No.**	**No.**	**No.**
GC	10	2	0	2	6
Control	11	2	0	3	6

[a] Refers to any teaching experience regardless of location or time period

Table 20. Total Number of Years Teaching in Grade Level[a]

		0-3 Years	4-6 Years	7-15 Years	15+ Years
Condition	**Number of teachers**	**No.**	**No.**	**No.**	**No.**
GC	9	2	0	4	3
Control	9	1	0	3	5

[a] Not necessarily consecutive

Note. One GC teacher and 2 control teachers did not report this information.

Table 21. Extent of Math Preparation in College

		Some	Minor	Major	Masters
Condition	**Number of teachers**	**No.**	**No.**	**No.**	**No.**
GC	9	0	0	9	0
Control	11	1	1	8	1

Note. One *GC* teacher did not report this information.

Table 22. Recent Professional Development

Condition	Number of teachers	Recent participation No.	No participation No.
GC	10	7	3
Control	11	9	2

Previous training in graphing calculators

Two of the *GC* teachers indicated that they had been trained extensively in the use of earlier versions of TI graphing calculators. One of these teachers further specified that she was a TI "mentor" teacher and had created workshops for other teachers in the use of calculator-based laboratory equipment. Three teachers have experience teaching with a graphing calculator for more than 10 years. One teacher reported attending trainings sponsored by TI and the school district. Another teacher reported attending TI-Leadership training and several 1 to 2 day workshops.

A teacher with some, but not extensive experience teaching with a graphing calculator, reported attending a variety of calculator workshops sponsored by TI, Casio, and NCTM. These workshops consisted of summer workshops and year long monthly classes. This teacher also reported attending one to two hour trainings at California Mathematics Council conferences. Another teacher with beginning experience teaching with a graphing calculator reported attending a workshop on teaching Algebra II with a graphing calculator.

There no major differences in previous training with graphing calculators between the *GC* and the *GC+Nav* groups.

Implementation of the Graphing Calculator (*GC*) Program

Training Observations

During a five-day interactive workshop prior to the start of the academic year (August 8th – 12, 2005), the teachers randomly assigned to the *GC* group were introduced to the TI-84 graphing calculator system and its capabilities. The week-long workshop was conducted by a TI consultant, a former California Certified teacher. During the workshop participating teachers used the graphing calculator to solve a variety of problems similar to those that the students would encounter during the normal course of instruction. This workshop described by TI as a typical five-day T^3 Summer Institute was held at the East Side District Offices. On average each training day began at 8:00 AM and ended around 2:45 PM, with a 30 minute break for lunch and two 15 minutes breaks (one each morning and afternoon. The first two days of the training were held in a small conference room where teachers practiced connecting and using the presentation tools, concentrating on the use of the emulator software recently developed by TI called, Smart View®. The conference table arrangement allowed participation of all the members fairly seamlessly. The participants verbalized frequently, talked out loud as they performed keystrokes, called for help without identifying anyone in particular and anyone that was able, helped. The instructor allowed a loose environment using the information on the handouts to guide the hands-on activities. The last three days of the training were held in a much larger room with separate tables and included practice in a computer lab. The computer lab gave teachers experience logging onto the online support system and creating electronic flash cards (StudyCards™).Of the 10 teachers randomly assigned to the treatment group, 9 participated for the entire week. One teacher was out of the country and was not able to attend. This teacher was provided with DVDs containing video of the entire training workshop (San Diego version).

Teachers were directed to use the calculator technologies in the normal course of instruction, using the materials, calculators, and presentation technologies as appropriate instructional opportunities arise. Since there is no set curriculum that accompanies the use of the graphing calculator, teachers are expected to select from a variety of materials including those provided by TI to integrate calculator-based instructional activities into their lessons. Note that calculators are used heavily throughout mathematics and science instruction to facilitate numerical computation and so are present in the typical learning environment as computational tools. We are however more interested in aspects of calculator use during graphing and various other analytical modalities rather than computational activities.

The training covered several topics focused on building skill and navigating the various functions of the *GC*. These topics were covered within the context of various activities outlined in the Explorations series of workbooks (TI materials provided to all *GC* teachers) geared towards helping the teacher and students use the calculator for exploring concepts in Algebra and Geometry. According to TI, the activities in the workbooks are tailored to specific student and teacher needs and are aligned to content standards. A typical training day covered approximately 14 different activities, each day focusing on developing increasing flexibility and knowledge of the calculator functions and interfaces. Topics concerning pedagogical choices and classroom management issues were rarely addressed. There was some discussion and time devoted to planning and adapting the activities for use in the individual teacher's classroom, but this was limited to approximately 2 hours of the week-long session and came about as a consequence of discussions among instructor and researchers.

A participant self-assessment focusing on calculator skills was administered at the beginning and at the end of the week-long session. Questions regarding basic four function computation and standard graphing functions ([Y=] [WINDOWN] [GRAPH] [TRACE] [ZOOM]) indicted that teachers had relatively high level of skills prior to the training and little changed. Teachers responded to a question regarding the advanced feature set using Polar and Parametric graphing indicated that they evaluated themselves with low level skills and did not feel differently after the training. This feature set was not addressed during the training. Teachers reported having gained the most in the following areas: using CATALOG and CATALOG help; linking calculators and computers to send and receive data and files; using the DRAW menu in coordinate Geometry; and memory management (archiving, grouping and deleting files).

Materials

After the training, teachers were sent a digital projector, a teacher's edition graphing calculator, a calculator for each student in an eligible class, a calculator based laboratory CBL/CBRs sufficient for use with a class size of 40, connectors for each calculator, resource workbooks, activity books, the SmartView® emulator software, various application software including the Cabri® Jr. Dynamic Geometry application, and a TI online account to access support and additional materials. See Appendix C for a complete list of materials.

The teachers were provided storage cases for the student edition calculators. All of the equipment was delivered to the teachers at their schools.

At the start of the experiment TI thought that the *GC* teachers would have access to computers in their classrooms. Although this is true in general, after the initial deployment of the materials we identified several issues as problematic. The issues can be summarized into three categories. Although three areas were identified, the issues are intertwined.

- **Availability** – does the teacher have a desktop or laptop computer to use with the SmartView® software emulator?
- **Compatibility** – If the teacher has a desktop or laptop computer, is the operating system, speed, memory, etc. compatible with and able to adequately drive the emulator software?
- **Supportability** – What is the level of on-going technical support of the available desktop or laptop computer?

Availability

Most teachers do have either a desktop or laptop computer available to them for use in the classroom. Desktops computers are typically found either on the teacher's desk or off to the side of the room, neither location makes it easy to use with a projector given the constraints imposed by power and connectivity requirements. A few teachers had laptops available to them. The laptops do not have the same physical constraints as the desktop computers do, but suffer from the same operating system compatibility problems.

Another related concern is security. At least one teacher in the study is sharing a room with another teacher that is not part of the study. The equipment (digital projector, demonstration calculators, laptop, etc.) cannot be secured easily. The *GC* teacher in this circumstance requested from the school, a roll-around cart where the equipment can be secured and removed from the room to be placed in a storage closet when not in use. This took several months to arrive and so this teacher did not implement all of the components of the system in the same timeframe as other teachers in the study.

Compatibility

More often than not, the desktop computer is an Apple rather than a PC and is usually an older version operating system or in the case where the operating system is current the latest service pack has not been installed. The original emulator software disk did not contain an Apple version of the software consequently none of the teachers with Apple computers were able to test the emulator software. A disk with the Apple version of the software was sent to any teacher requesting it.

Besides the operating system compatibility issues, the hardware concerns are more difficult to resolve. The likelihood that more memory can be added to either a desktop or laptop is very slim. School districts do not typically upgrade hardware piecemeal, they usually upgrade in large batches, upgrading a grade level or a school. If the problem with the PC is insufficient memory or slow speed, there is no school-level remedy available to the teacher.

Supportability

Most of the schools in this district have an Information Technology (IT) person that reports to a central district IT group. Requests for upgrades and troubleshooting go to central IT then routed back to the more local IT person. These requests can take upwards of three weeks to resolve. In some cases, the issues cannot be resolved because they require expenditures of funds that the school district has not allocated.

After reviewing all of the issues, we gave TI a list of recommendations to ensure that equipment issues were not a barrier to the use of graphing calculators in the classroom. TI responded by providing all *GC* group teachers with laptops with the emulator software already installed and tested. Additionally, they agreed to provide laptops to newly inducted *GC* group teachers in year two of the study during their initial professional development week. We continued to monitor equipment support issues through periodic check-in with the teachers.

Other Equipment Issues

TI envisioned that the teachers could assign each student a calculator for their own personal use rather than just provide a class set where use would be restricted to class time. Teachers however were reluctant to distribute the calculators because of the possibility that students would lose or break them. Teachers took several precautions to avoid losing calculators. Teachers received the school-bus yellow calculators that are easily identifiable as student versions and can only be purchased by school districts, but teachers thought it insufficient identification to protect them from loss. Some teachers used the school library's book check-out system to further tag and track the calculators. Those using this system etched each calculator with a number and recorded the number given to each student. Other teachers recorded the unique serial number noted in the memory section of the calculator before distributing to the students. Still others used the class set system, but also provided the students the opportunity to check out calculators overnight or over the weekend.

Perhaps the most ubiquitous and problematic issues became the operating system variation among the various pieces of hardware (graphing calculators and laptops) and the memory management requirements. Eight of the teachers tried downloading StudyCards™ Applications and removing from the graphing calculator some of the language and game applications via the linking cables. Of these teachers, all reported that several of the student calculators would not respond to the commands issued through the linking cables. After contacting TI via their online support, the teachers became aware of the different operating systems and the need to ensure that all calculators have the same operating system in order to take advantage of the calculator communication ability. Teachers reported that upgrading the operating systems on the student calculators was extremely time consuming and many of the teachers abandoned the use of the communication abilities early in the study. Only two teachers reported using the communication ability after mid-year (source: email exchanges and classroom observations).

Memory management is an ongoing issue with the calculators. Calculators have limited memory available and require regular maintenance to ensure access to software applications such as Algebra I (ALG1CH1 etc.) and Cabri® Jr. At the beginning of the workbook, Topics in Algebra (2004), that all of the *GC* teachers received as part of their training materials, TI clearly states that a TI Connectivity Cable and TI Connect™ software is required in order to run these applications. TI also advises that if an Archive Full error message appears while installing the Algebra I application, the graphing calculator does not have sufficient memory and that Applications and/or archived variables must be deleted to make room.

Many of the teachers complained that they neither had the time nor the desire to upgrade and memory manage the student calculators. This is a serious barrier to the implementation of the more robust applications that are available on the graphing calculators. Willing teachers turn into reluctant teachers very quickly faced with the daunting task of managing 70+ calculators.

Classroom Settings

The classroom setting was observed during normal classroom instruction the last week of April and the first week of May. Teachers were not asked to prepare specific lessons for observation, but teachers could determine when researchers came to observe within the two week observation window. All *GC* and TI-Navigator group teachers were formally observed at least once and a sample of control teachers were formally observed. In addition to the formal observations, researchers frequently observed classrooms from both groups during school visits to coordinate testing, drop-off materials, or to discuss implementation problems.

In the *GC* classrooms, we observed the graphing calculators used with graphing lessons that explored a “family of curves”, data collection and analysis, and a short lesson using the Cabri® Jr. software. In all, teachers reported using the calculator as a computational device more often than a device for exploration and investigation, but they understood the goal of trying to incorporate the technology more fully. Neither of the beginning teachers attempted any activities beyond the computational functionality of the calculators. Typical classroom routines included taking attendance followed by checking homework, moving to a new concept, practicing the new concept and assigning homework. We also observed students taking quizzes. Summarized from survey responses, we provide a breakdown of the time spent using calculators and the topics covered in Appendix E, Calculator System Usage.

Physical Settings

Most teachers in both groups had traditional classroom layouts consisting of individual student desks arranged in rows and facing towards a whiteboard, the designated “front” of the classroom. Two classrooms were observed in modified traditional set up of individual student desks paired next to each other, arranged in rows and facing front.

All teachers were observed to have a whiteboard and conventional overhead projector at the front of the room. Often the overhead projector was located on a rolling cart to ease storage and to accommodate different room arrangements. These resources were available and typically used during a lesson. Most teachers had a television in the classroom mounted from the ceiling or on a rolling cart, but no teacher observed used the television during a lesson. Teachers also had a computer located either in the front corner or rear corner of the classroom designated for teacher use. These computers were used for reporting attendance and creating lessons, but were not used during any of the observed lessons. A few teachers also had designated computer labs, but none of the observed lessons featured this technology, although a couple of the teachers commented that they did use Geometer's SketchPad with their Geometry students.

Most notably, teachers did not have enough power outlets in the classroom to accommodate more devices than were already present. Most teachers powered their overhead and television through one outlet connected to a power strip. We report this feature in particular because the TI-Navigator system requires access to a power source for at least the charging cycle of the hubs and the laptop-projector components. Because of the limited power availability, a TI-Navigator system setup may not be possible in some classrooms.

Most teachers and students in the *GC* and TI-Navigator classrooms were observed to have the TI-84 graphing calculator or other graphing calculators available for use during class assignments. Few control group teachers were observed to have calculators available during a lesson. Those *GC* and control classrooms with calculators present appeared to be using the calculators as a computation device more often than not.

During the weeks set aside for CST review and preparation students are not allowed to use calculators at all for any reason because the assessment is conducted without the aid of calculators. Consequently, during this time all calculator use was suspended.

Instructional groups

There are four distinct types of student groupings in the Algebra classroom: entering 9th graders with varying degrees of prior instruction in Algebra, but not considered to have *failed* Algebra I, these are designated as "first-time freshmen"; Accelerated or Honors Algebra students, these students are expected to do well and consequently follow an aggressive schedule and typically cover more material than first-time freshmen; Sheltered Algebra students are those identified as needing more English language support during instruction; and those students in 10th, 11th, or 12th grade that have not succeeded in passing Algebra I in a previous attempt, these students are designated as "repeaters".

Geometry classes are similarly grouped. There are fewer "repeater" classrooms because Geometry credit is not needed to graduate high school.

Implementation of the Graphing Calculator with TI-Navigator™ (*GC*+Nav) Program

Training Observations

In January, from among the *GC* group teachers, five teachers were randomly selected to use the TI-Navigator based system. Together with a group selected from among the San Diego *GC* group, they attended a three-day training at the East Side Union district offices. The training was conducted by a university professor that is part of a school of education engaged in teacher preparation and education. Training concentrated on system set-up, managing of applications, and various activity skills. There was direct attention paid to issues of how students could be organized to engage in the activities and how to structure the classroom for optimum use. Even so, we noted that it was difficult for teachers to take in the new aspects of the system. As the implementation of the system proceeded to the classroom setting, we confirmed our suspicions as three of the teachers that received the training did not set-up the TI-Navigator system at all. Seemingly, from teacher feedback, mid-year implementations are hard to negotiate, because it requires planning and new routines to be incorporated in the classroom. Both of these activities take time and can potentially impact an already hectic schedule.

Materials

At the end of the training session, these teachers received the TI-Navigator system software and hardware sufficient for connecting a classroom set of calculators. Texas Instruments provided teachers with on-going support for the all of the equipment replacing any malfunctioning equipment or software. Teachers commented that all of the packages looked intimidating because they were large and full of “gadgets”.

Classroom Setting

While TI provided all *GC* and *GC*+Nav teachers with a notebook computer and data projector, few *GC* and *GC*+Nav teachers were observed using the technology. The principal differences in classroom setting were noted in the two classrooms that were using the TI-Navigator technology. In these classrooms, the systems were up and connected, student desks were clustered around the hubs (groups of 4 students to one hub) and the students and teachers had developed routines where they systematically logged-in to the system at the start of class. Students wasted no time in getting their calculator networked to the distribution hub connected to the teacher’s laptop and working on the warm-up problems. The teacher noted attendance using the laptop as the students were acknowledged by the log-in process. Both the level of activity and the discussion about mathematics was high among the students. Summarized from survey responses we provide a breakdown of the time spent using calculators and the topics covered in Appendix E, Calculator System Usage.

The instructional groups were the same as noted for the *GC* classrooms.

Issues in Rating the Level of Implementation

We begin our discussion of implementation issues with a summary of the perspectives through which we viewed implementation. For our purposes, we define implementation as a specified set of activities designed to put into practice a program of know dimensions. Accordingly, implementation processes are purposeful and are described in sufficient detail so that observers can detect the presence and strength of the “specific set of activities” named as the implementation program. In this case, we are interested in the implementation of TI-84 graphing calculators and TI-Navigator into two types of instructional settings, Algebra I and Geometry classrooms.

In addition to implementation processes, we also note that there are *intervention* processes and outcomes. These we define as the processes required to set the implementation processes in motion. Intervention processes in this case include, a description of the implementation program, the training supplied by TI to the teachers, the equipment and materials deployed, the settings and resources (materials, time, and other personnel) supplied by the schools, and the on-going support provided by TI.

The first issue we confront is the non-specified nature of the implementation program. As noted before, the integration of graphing calculators and/or TI-Navigator into classroom instruction is ill-defined as an intervention program. There are no specific principals that guide the use of this technology. At best, we have some general guidelines that address emerging patterns of fruitful use. The training provided some pedagogical strategies, but these were hidden within the form of activities and never explicitly discussed. For this study, what, when, and how the materials were to be used were left to the individual teacher’s judgment and discretion. The number of activities to be used for the study was not specified. Once we realized that the intervention was so loosely defined, we turned our attention to developing a framework whereby we might characterize practices that use the technology in productive ways.

Framework

In developing the observation framework we considered four sources: recent research on learning with technology, National Council of Teachers of Mathematics (NCTM) Principles and Standards for School Mathematics, California Standards for the Teaching Profession (1997), National Board for Professional Teaching Standards, and the International Society for Technology in Education (ISTE) Technology Standards for Teachers.

How People Learn (2000) explains how technology might be used to:

- introduce real-world problems into the classroom

- provide thinking scaffolds and tools for problem-solving
- provide multiple ways to represent data
- use networking capabilities to stimulate communication for presentation, feedback, reflection, and revision
- cause change in the teacher's role in the classroom

While recent research such as How People Learn (2000) provides guidance, researchers also indicate we have just begun to understand the potential of these technologies as change agents for social interactions to promote learning.

According to the NCTM Principles and Standards for School Mathematics (2000) the most fruitful uses of technology during instruction takes place in an environment of student-centered practices that allow students to explore complex problems and mathematical ideas and to present their solutions. This type of environment promotes discussions that challenge and extend students' mathematical ideas thereby helping students to develop deeper mathematical understandings. From the California and National Board teaching standards, we noted areas that articulate elements present in robust student learning environments and coupled them with the ideas expressed in the and ISTE principles. These we categorized into three main concepts:

- *classroom routines* that presume the use of the *GC* and TI-Navigator technologies not as an add-on, but as part of the flow of working with and learning mathematics
- *classroom interactions* that are marked by shared responsibilities among teacher and students
- *lesson structures* that focus on learning, where learning outcomes are clearly articulated, frequent monitoring of learning is present and discussed by both teacher and students, and students question, offer conjectures, and allow for alternative approaches and solutions

These ideas led to the construction of an observation protocol that can be found in Appendix B. We began using this observation protocol to more closely identify and profile those emergent practices that distinguish the major attributes of the *GC* condition. We used focused interviews with a select group of *GC* teachers to further explore how they conceptualized the use of the technology in the classroom. These findings are summarized in the case profiles section.

Definition of the intervention program

The second issue we confront is the intervention process. Since there are multiple processes these became a set of issues. The intervention process in this case consists of series of stages: definition of the intervention program, teacher training, distribution of materials, classroom set-up, teacher planning, teacher use, and student use. We have, in part, addressed the definition issue in the framework, but several aspects remain undefined.

Although teachers appear to be very comfortable with the basic aspects (common computation functions, graphing of y =) of the calculator, the extended feature set requires initial training and practice before the teacher can use them with ease. To what extent should the training address the feature set and how many activities should the teachers be asked to implement during a semester or academic year? The research is silent on the "quantity" issue which is only one aspect of the classic dosage question, how much, how often? What we do know is that classroom routines must be established in order to manage the usage of technology in the classroom.

Should the training be designed to address different aspects of the calculator or TI-Navigator system according to the pedagogical strategy rather than "Topics in Algebra"? Current training focuses strongly on introducing and practicing the functions of the calculator with an orientation towards "student activities", but does not actually discuss the reasoning and pedagogical strategies that are required. We have many more questions regarding the relationship between the training and the implementation program than we are prepared to answer. Our instruments and protocols did not specifically focus on trying to answer these questions directly, but we were able to understand the general nature of what is required for teachers to manifest a level of implementation distinguishable from the control condition. In general, we noted several indicators that must be present for the teachers and students to enact graphing calculator supported Algebra and Geometry lessons. Teachers must be reasonably familiar:

- and proficient with the graphing calculator feature set, software, networking capabilities, and have general troubleshooting skills
- with the curriculum and standards and how the activities outlined in the training can be incorporated into the flow of instruction
- with methods for establishing new routines that incorporate the systems and gradually scaffold students into becoming proficient with the systems' functionality set so that they can access the more powerful features
- with methods for creating classroom organizations that take advantage of the new capabilities

Student learning of the functionality of the calculator is not addressed in the training directly. The functionality of the graphing calculator system is complex enough so that developing proficiency on the part of the student must be incorporated into the dynamic of the lessons. Teachers can not take for granted that students are familiar with graphing calculators and their capabilities beyond the four computation functions. Students must be reasonably familiar:

- with the calculator and software so that access to the applications and other analyses and data functions become routine and not the focus of the lessons
- with group work dynamics so that they can rely on each other for support regarding the feature set and release the teacher to concentrate on helping them develop mathematical understandings

Those teachers that were successful in using the graphing calculator and TI-Navigator system all displayed the above indicators.

Time constraints

A third issue is the time available for learning, planning, experimentation, and lesson instruction. Many teachers expressed that they did not have any time to play and plan with the graphing calculator system. This factor potentially has the ability to derail any desired implementation. As can be seen from the case profile data, even those teachers that were successful implementers still lamented the lack of time available. We noted this in this study and made recommendations to TI to modify the training in year 2 of the study to include daily planning time so that teachers could discuss which activities might work given the pacing and the curricular demands, how they might organize the classroom and to practice the activities as if they were going to teach it to their students.

Teacher Case Profiles

Three teachers in the ESUHSD teachers were interviewed on successes, barriers, and implementation of the *GC* and TI-Navigator technology.

Most teachers used activities they had done successfully in previous years. Teacher tended not to venture beyond what they felt comfortable doing. While a variety of activities were presented during the trainings, it was the familiar activities that were implemented.

Teachers also expressed concern that calculators did not provide students with the opportunity to learn the arithmetic. Teacher and researchers observed students primarily using the calculators as computational devices without fully exploiting the calculator's additional capabilities.

Teachers also expressed time constraints as a significant problem. In an already busy day, teachers did not believe they had time to effectively learn and introduce the technology. Many teachers believed with more time they would be able to better implement the technology.

_GC_ (Teacher ID #93)

Teacher ID #93 has had extensive experience teaching with the graphing calculator. The teacher served as a calculator mentor teacher and taught many in-service trainings. The teacher feels very comfortable with the technology and knows which functions students grasp easily and are most fruitful for grounding student understanding.

To implement a successful graphing calculator lesson, the teacher believes first in setting specific ground rules and establishing routines prior to distributing calculators. Teachers must be very familiar with the equipment and lesson being taught. Teachers must read the equipment documentation; understand how the equipment functions and its set up, and practice the lesson before using it with students. While a specific activity may be new to the teacher, familiarity with the equipment and its basic functions facilitates the implementation.

When asked about a successful lesson, the teacher listed several that worked well, such as using the CBL/CBR motion detectors and plotting real data. While the CBL/CBR activity was new, the teacher had experience using the CBL/CBR sensors previously and taught lessons with the equipment. The teacher likes the CBL/CBR sensors because it involves the collaborative participation of the whole class.

The teacher found the initial *GC* training helpful, but did not use the Cabri® Jr. application because it was unfamiliar and the teacher does not teach Geometry classes.

Problems with activities mainly stemmed from low achieving students relying on the calculator for basic computations. For this reason, the teacher does not allow low achieving students to access the calculators on a daily basis.

TI-Navigator (Teacher ID #88)

TI-Navigator teacher #88 has had a lot of training using the graphing calculator, specifically with Algebra II and Trigonometry.

The teacher views the technology as a tool to better understand the material and does not want the calculator to be perceived as a shortcut and circumvent students' mathematical learning. The teacher observes that many students perceive the calculator as a "magical" device that always provides the correct answer without really understanding the concept. The teacher believes the TI-Navigator implementation primarily benefited the middle achieving students and the low achieving students needed more help.

When planning a lesson, the teacher is mainly interested in teaching the concept. The teacher asks "how can I use the calculator in terms of the time that I have?" and "what do I want to get across, concept-wise, what is the best way to get that concept across?" The teacher believes the calculator to be a great tool, however some concepts require more scaffolding.

When asked to describe a successful lesson, teacher highlights the Navigator's™ instant feedback through Learning Checks and Quick Polls. Those activities were used in class to review and discuss California state assessment practice problems. The teacher also believes the graphing calculators are very helpful for analyzing graphs and coordinate Geometry activities. Yet the teacher emphasizes students needed to be able to do the graphs by hand. Ultimately with regards to lessons, the teacher believes in taking a leap of faith with an activity and being flexible. Due to timing constraints, some lessons had to be extended to the next day, but the teacher always had a willingness to try anything if it was believed to further student understanding.

The teacher is less inclined to use the calculators when students use them as a computational device without understanding the arithmetic. The teacher wants students to understand the calculator is a supplemental tool and that the material must still be learned. As a result the teacher requires students to manually show their work.

The time required to teach with the calculator was significant. The teacher used the Cabri® Jr. application and found more time spent on students constructing the figures than on developing the concept. When planning time allowed, the teacher constructed the figure and distributed it to the students through TI-Navigator, but the day did not always allow that. The teacher also found that the Puzzle Pack application and student developed games on the calculator to be distracting for the students. The teacher was able to remove unwanted applications with the TI-Navigator, but it became a burden to constantly maintain the calculators. The teacher also supplemented the TI-Navigator and SmartView setup with a TI overhead Viewscreen because SmartView would lose connection and worked too slowly with the Cabri® Jr. With the two projections devices, the teacher found it difficult to determine which device would work best. Toward the end of the year, students began to forget their calculators and the TI-Navigator could not be used with the entire class.

The teacher notes the TI-Navigator system was not ready for use directly out of the box. Implementing the system required time and adaptation to fit into the curriculum. The teacher preferred better alignment to the curriculum, but the teacher believed the TI-Navigator training was very helpful. There were lots of logistics with configuring the classroom correctly and managing the wires between calculators and hubs. All set up issues were successfully solved through persistence and conversations with the TI-Navigator trainer and TI. Once the TI-Navigator was setup successfully in the classroom, it appeared to function correctly with few interruptions.

The teacher believed the training was sufficient but may not have been for other teachers and more training always helps. The teacher was disappointed that the TI-Navigator was not ready to use out of the box and wanted more instruction on how to adapt lessons to the curriculum like the TI-Navigator supplement available for the Prentice Hall textbook.

TI-Navigator (Teacher ID No.80)

Teacher #80 taught a class of “repeater” Algebra students without the standard textbook. The teacher has seven years experience teaching with the graphing calculator for basic graphing functions. The teacher learned to use regressions with Algebra II students five years ago and in the past two years adapted regression lessons for freshman classes.

The teacher’s biggest concern with adopting new technology is if it will yield a significant payoff. The teacher asks, “how can I be sure this will work and if it’s better than what I’m already doing?” The teacher does not claim to be very technology savvy reporting to be the last person to the get a CD player or cell phone. The teacher was intimidated by the large box of TI-Navigator components and connections. After a few attempts, the teacher was unable to initially setup the TI-Navigator consistently and was unable to continue implementing the equipment.

The teacher sees the calculator only as a tool that ties things together. Students have to understand that graphical representations on the calculator are only approximations. Students need the basic understanding first and then introduce the technology. The teacher primarily uses the calculator to explore prior knowledge or illustrate real applications of concepts with data collection. But sometimes the teacher will introduce a concept with a calculator through an initial discovery. The teacher’s style varies depending on the concept being taught.

To plan the lesson, the teacher believes teachers must understand how students learn to deliver the lesson effectively, determine what concept is being taught and how to best deliver the concept. If applicable the teacher will determine the best place to insert the calculator in the lesson to further expand student understanding. However teachers must understand the equipment well before introducing it in a lesson

When asked about a successful lesson, the teacher reports using a calculator lesson where students plotted class data and determined the line of best fit. This lesson the teacher used previously by hand, but adapted it for use with the calculator. The teacher believes that "I stuck to old things and did not bring a lot of new things."

Primarily the teacher was unable to get the TI-Navigator functioning consistently and as a result was unable to implement it. The teacher also believes students need to learn the manual arithmetic and learn what happens "behind the scenes" before relying on the calculator. The teacher found it difficult to manage the supplies and keep track of the equipment. The biggest barrier to the teacher was finding the time to learn and use the technology. Because the equipment was not ready to be used out of the box, the teacher found it difficult to make time during an already busy day to adapt the technology to the classroom. It was difficult to learn and use the technology while still teaching. The teacher felt overwhelmed by the number of components and asks "so much could be done with the TI-Navigator, where do I begin?"

The teacher believed the training was sufficient, but wanted more specific things for the curriculum. The teacher taught a low achieving student Algebra class with a variety of materials, but believed TI-Navigator could be more easily applied to teachers using a standard textbook.

Summary of Implementation Results

Common themes that arose during the overall study are reinforced by the focused interviews we conducted with the case profile teachers. We noted that the teachers in the *GC* and *GC+Nav* group that implemented the technologies shared the following themes:

- Previous experiences using calculator-based technologies in the classroom, either as a teacher or as a student
- Are well versed in both functionality and troubleshooting of the failure modes
- Have established routines for students to follow which include attention to providing students with activities that facilitate learning keystrokes and functionality
- Use the technologies to focus on the value-added type activities; capitalize on the multiple representations, visualization capabilities, and allow students to probe deeper into a concept having already established the concept in more traditional ways

Most teachers used activities they had done successfully in previous years. Teacher tended not to venture beyond what they felt comfortable doing. While a variety of activities were presented during the trainings, it was the familiar activities that were implemented or activities teachers self-created and tailored to a particular lesson. Teachers expressed interest in having ready-made activities that could be easily dropped into a lesson.

Teachers also expressed concern that calculators did not provide students with the opportunity to learn the arithmetic. Teachers wanted to see students use the calculator as a tool for learning and not a computational device. While researchers observed students primarily using the calculators for computation without fully exploiting the calculator's additional capabilities, teachers that implemented the *GC* and *GC+Nav* report a focus on exploratory activities.

Teachers expressed time constraints as a significant problem. In an already busy day, teachers did not believe they had time to effectively learn and introduce the technology. Many teachers believed with more time they would be able to better implement the technology. Additionally, teachers report that the current training was insufficient to provide them with the tools to expand beyond already familiar activities.

In summary, the issues we faced in measuring the level of implementation prevented us from formally rating the implementation integrity of the *GC* and *GC+Nav* classrooms and employ anything other than a dichotomous measure (using vs. non-using) to characterize implementation. Additionally we note that the actual length of implementation for the *GC* group was approximately 19 weeks and approximately 10 weeks for the *GC+Nav* group out of a possible 30 week instructional year. The shorten length of time for the implementation is due to two main contributions, loss due to class roster settling (impacts the *GC* group only) and loss due to CST preparation and administration. Teachers are reluctant to invest the time to teach the students how to use the technology when class rosters are still fluctuating for the approximately 6 weeks at the beginning of the school year. Mid-way through April, teachers turn their attention to preparing the students for CSTs and consequently suspend "normal" instruction. All of the quantitative results should be viewed with regard to these shorten times of implementation. Recall also that *GC+ Nav* implementation was begun mid-year..

For year 2, we plan to work more closely with the teachers to help them implement the technology and to help us formulate reasonable measures. By using our modified observation tool we hope to capture more information regarding both the nature of the implementation and additional indicators so that we can begin to address the issue of "quality" of implementation and provide some guidance as to emerging accomplished practices.

Overview of Quantitative Results

In all cases, our analysis of the quantitative results takes the same form. We first provide the descriptive statistics in the raw form, that is, with no statistical adjustments. These descriptive statistics include the number of students, classes and teachers. We then compute the average impact with and without adjusting for the prior score. These results are in standard deviation units (so the results are labeled an 'effect size') and indicate whether there is an overall difference between the treatment and control groups. We then present the results of statistical models where we estimate whether the impact of the intervention depends on the level of certain moderator variables. Below the descriptive for instance, we show the results of a model that tests whether there is a differential impact across the prior score scale. That is, we test for the interaction of treatment with the prior score. The fixed factor part of the table provides estimates of the factors of interest, in particular, whether being in a *GC* or a control class makes a difference for the average student. At the bottom of the table we give results for technical review – these often consist of random effects estimates which are added to the analysis to account for the fact that the individual results that come from a common upper-level unit (e.g., class or teacher) tend to be similar (i.e., the observations are dependent.) In some cases, to account for these dependencies, we model fixed rather than random effects but don't present the individual fixed effects estimates. Modeling the dependencies results in a more conservative estimate of the treatment impact. We note that the number of cases used to compute the effect size will often be larger than the number used in the mixed model analysis because to be included in the latter analysis a student has to have both a pretest and a posttest score.

We address the impact on Algebra and then the impact on Geometry outcomes. Within each content area we consider the results for the NWEA outcome measure and then the California Standards Test (CST) outcome. For each outcome measure we provide a statistical analysis of the impact of *GC* controlling for pretest and examine the interaction of *GC* with pretest, that is, we examine whether students initially scoring higher or lower on the pretest differentially benefited from *GC*. We then examine the influence of English proficiency as a potential moderator of the impact of *GC*.

Impact of *GC* on Algebra

NWEA Outcomes

Figure 2 summarizes counts of students we use in this first analysis of Algebra NWEA Outcomes.

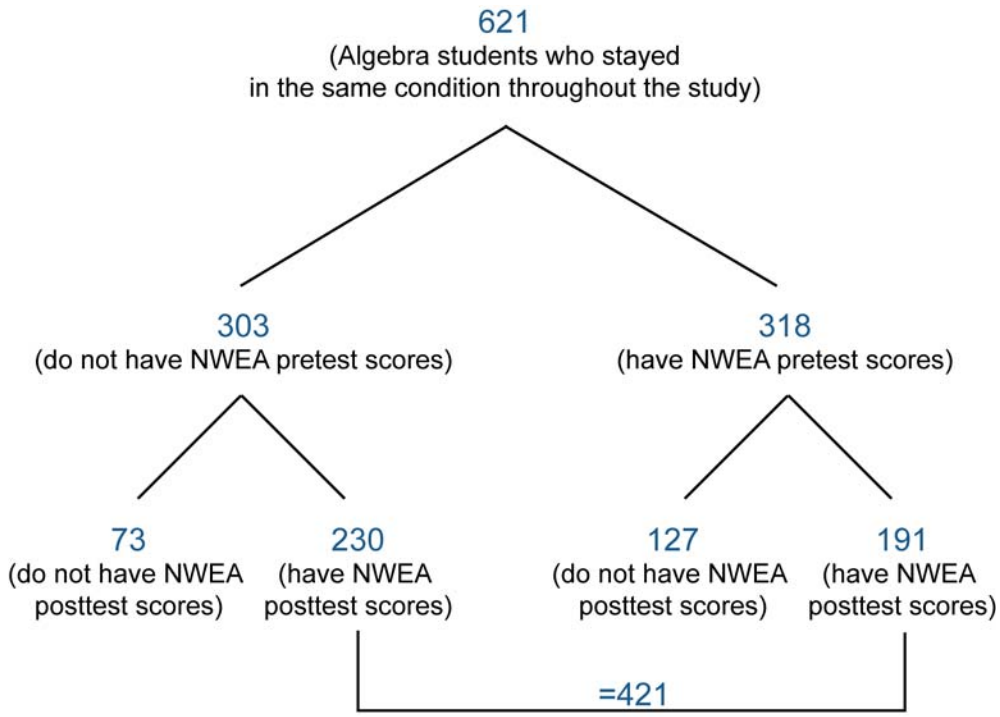

Figure 2. Counts of Algebra Students Having Both NWEA Pre- and Posttests

Analysis Including Pretest

Table 23. Impact of *GC* on NWEA Algebra Outcomes

Descriptive statistics	Raw group means	Standard deviation	No. of students	No. of classes	No. of teachers	Unadjusted effect size
GC	235.43	13.22	190	16	7	0.20
Control	232.23	18.61	231	15	7	
Mixed model: Fixed factors	**Estimate of coefficient**		**Standard error**	**DF**	***t* value**	***p* value**
Intercept	236.39		1.94	10	122.16	<.01
Pretest score (centered at the mean)	0.73		0.06	174	11.31	<.01
Condition (*GC* = 1; control = 0)	-0.94		2.96	10	-0.32	.76
Pretest score by condition interaction	-0.09		0.11	174	-0.89	.37
Mixed model: Technical details for random components	**Estimate of variance component**		**Standard error**		***z* value**	***p* value**
Teacher mean achievement	16.02		10.20		1.57	.06
Within teacher variation	96.87		10.31		9.40	<.01

Note. 621 students were included in the analysis of the Algebra outcome. Of these 421 had a posttest and were included in the calculation of the unadjusted effect size. Of the original sample, 191 students had pretest and posttest scores. Of these we removed 3 because they were influential points (where the determination of influential points was based on the model above.) The remaining 188 students were used to compute the adjusted effect size and the results presented in the table above. The number of teachers for the reduced sample of 188 students is 12 (6 control and 6 *GC*). The number of classes is 26 (13 controls and 13 *GC*.) Teachers are modeled as a random factor.

Table 23 shows the estimated impact of *GC* on students' performance on the NWEA test of Algebra. The unadjusted effect size is 0.20 calculated on the basis of all the students who took the posttest. However, the effect size when adjusted for the pretest is -0.05. This much smaller impact is based on students with both pre and post tests and then adjusted for the pretest. For a student with an average score on the pretest, there is roughly a 0.94-point disadvantage to being in the *GC* group. The *p* value for this effect is .76 from which we conclude that we have no confidence that this effect is due to something other than chance. The value for the interaction between *GC* and the prior score is -0.09 with a *p* value of .37. Again we have no confidence that this effect is due to something other than chance.

As a visual representation of the results described in Table 23, we present a scatterplot in Figure 3, which shows student performance at the end of the year in Algebra, as measured by the NWEA test, against their performance on the NWEA general math test in the fall. These graphs show where each student fell in terms of his or her starting point (horizontal x-axis) and his or her outcome score (vertical y-axis). Each dot plots one student's post-intervention score against his or her pre-intervention score. The darker dots represent *GC* students; the lighter dots, control students. The shaded area in the lower right of the graph is the area of negative change (i.e., where students lost ground).

The two lines are the predicted values on the posttest for students in the *GC* and control conditions as determined using the estimated fixed effects in the model.

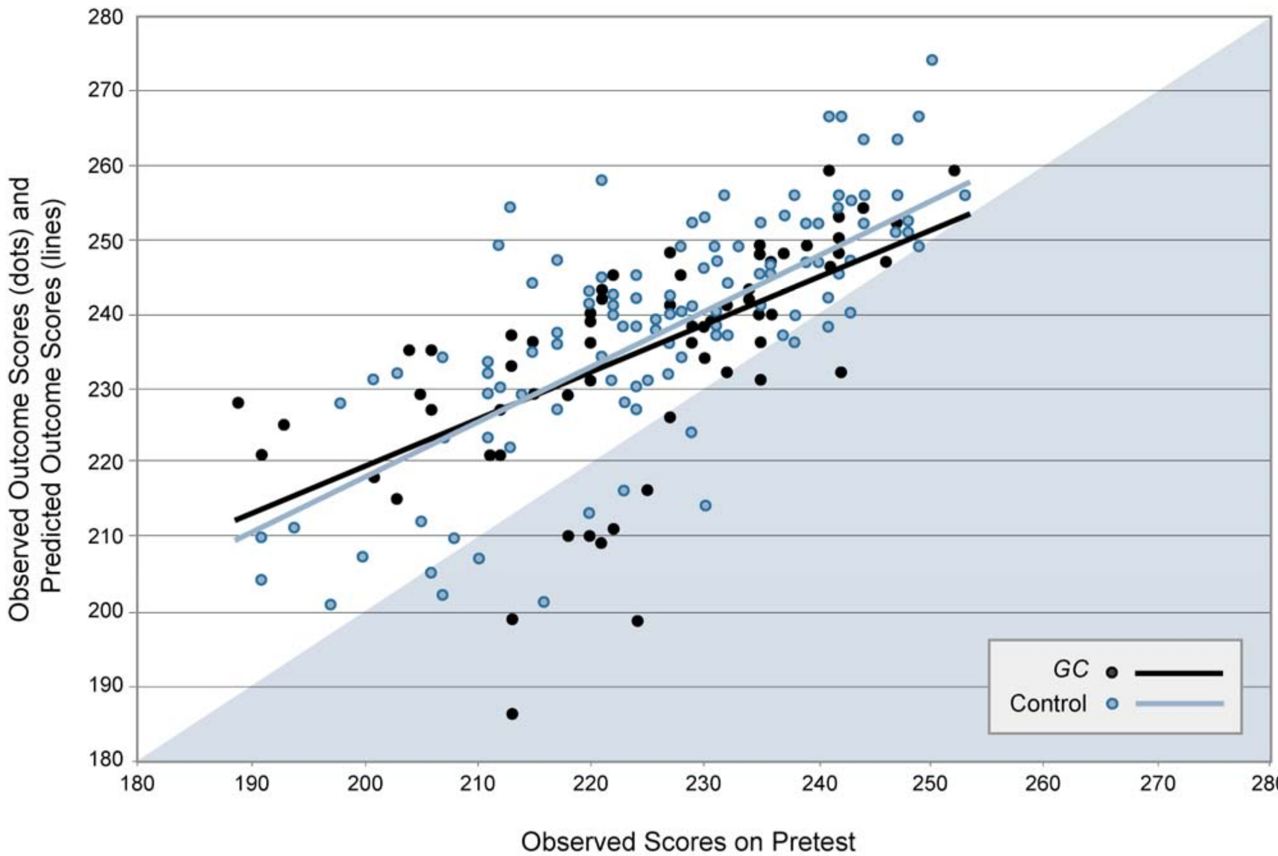

Figure 3. Comparison of Predicted and Actual NWEA Algebra Outcomes for *GC* and Control Group Students

The lack of appreciable separation between the lines represents the negligible impact of *GC* and the only slight difference in slope represents the low value for the interaction.

Analysis Including English Proficiency as a Moderator

Table 24 shows the estimated moderating effect of English proficiency on the Algebra outcome as measured by the NWEA test. Here we were interested in whether *GC* was more or less effective for students not proficient in English compared to English proficient students. Although we did not detect an impact in the previous analysis, we pursued the question of whether the net impact was weak masking the fact that differences for subgroups were in opposite directions, thereby effectively cancelling out the overall *GC* effect.

The coefficient associated with the English proficiency by condition interaction is -4.56 with a *p* value of .28, which gives us no confidence that there is a moderating effect of being an English-language-learner.

Table 24. Moderating Effect of English Proficiency on NWEA Algebra Outcomes

Mixed model: Fixed factors	Estimate of coefficient	Standard error	DF	*t* value	*p* value
Intercept	234.68	2.45	10	95.82	<.01
Pretest score (centered at the mean)	0.66	0.05	173	13.70	<.01
Condition (*GC* = 1; control = 0)	2.154	4.22	10	0.51	.62
English proficiency status (proficient = 1; not proficient= 0)	3.37	2.34	173	1.44	.15
English proficiency status by condition interaction	-4.56	4.21	173	-1.08	.28
Mixed model: Technical details for random components	**Estimate of variance component**	**Standard error**		***z* value**	***p* value**
Teacher mean achievement	17.02	10.73		1.59	.06
Within teacher mean variation	89.82	9.59		9.37	<.01

Note. 621 students were included in the analysis of the Algebra outcome. Of these 421 had a posttest and were included in the calculation of the unadjusted effect size. Of the original sample, 191 students had pretest and posttest scores. Of these we removed 3 because they were influential points (where the determination of influential points was based on the model above.) The remaining 188 students were used to compute the results presented in the table above. The number of teachers for the reduced sample of 188 students is 12 (6 control and 6 *GC*). The number of classes is 26 (13 control and 13 *GC*.) Teachers are modeled as a random factor.

CST Outcomes

Figure 4 summarizes the counts of students we use in the analysis of Algebra CST Outcomes.

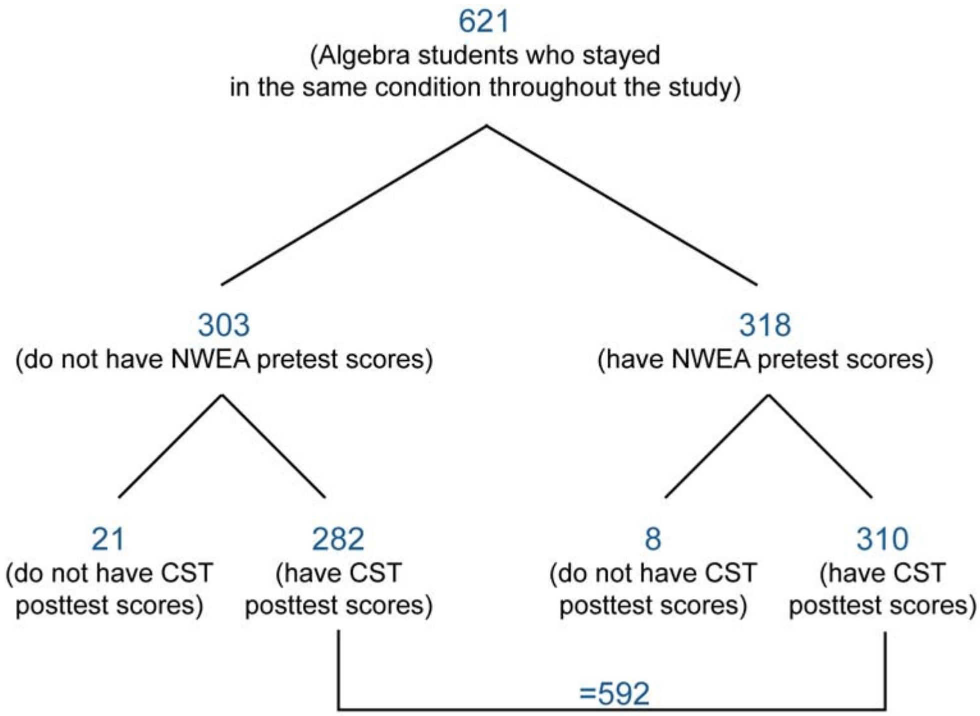

Figure 4. Counts of Algebra Students Having Both NWEA Pretests and CST Posttests

Analysis Including Pretest

Table 25 shows the estimated impact of *GC* on students' performance in the CST test of Algebra. The effect size is .01.

Table 25. Estimated Impact of *GC* on CST Algebra Outcomes

Descriptive statistics	Raw group means	Standard deviation	No. of students	No. of classes	No. of teachers	Unadjusted effect size
GC	306.27	45.61	243	16	8	0.01
Control	305.66	51.66	349	20	8	
Mixed model: Fixed factors	**Estimate of coefficient**		**Standard error**	**DF**	***t* value**	***p* value**
Intercept	310.33		3.68	12	84.34	<.01
Pretest score (centered at the mean)	2.42		0.20	290	12.29	<.01
Condition (*GC* = 1; control = 0)	0.06		6.49	12	0.01	.99
Pretest score by condition interaction	-0.49		0.33	290	-1.49	.14
Mixed model: Technical details for random components	**Estimate of variance component**		**Standard error**		***z* value**	***p* value**
Teacher mean achievement	48.75		64.17		0.76	.22
Within teacher variation	1297.42		108.42		11.97	<.01

Note. 621 students were included in the analysis of the Algebra outcome. Of these 592 had a posttest and were included in the calculation of the unadjusted effect size. Of the original sample, 310 students had pretest and posttest scores. Of these we removed 4 because they were influential points (where the determination of influential points was based on the model above.) The remaining 306 students were used to compute the adjusted effect size and the results presented in the table above. The number of teachers for the reduced sample of 306 students is 14 (8 control and 6 *GC*). The number of classes is 31 (19 control and 12 *GC*.) Teachers are modeled as a random factor.

The effect size when adjusted for the pretest is a negligible <0.01. (We calculated this using the impact estimate from a simple model that adjusted only for the pretest.) The *p* value for this effect is .99 leading us to conclude that we have no confidence that this effect is not simply due to chance.

The value for the interaction between *GC* and the prior score is -0.49 with a *p* value of .14. We have some confidence that this effect is not simply due to chance.

As a visual representation of the results described in Table 25, we present a scatterplot in Figure 5, which shows student performance at the end of the year in Algebra, as measured by the CST test, against their performance on the NWEA general math test in the fall. As with Figure 3 the two lines are the predicted values on the posttest for students in the *GC* and control conditions as determined using the estimated fixed effects in the model. However, in this case the difference in the slope of the two lines is more distinct with the line representing the *GC* group higher for the students at the lower end of the pretest distribution.

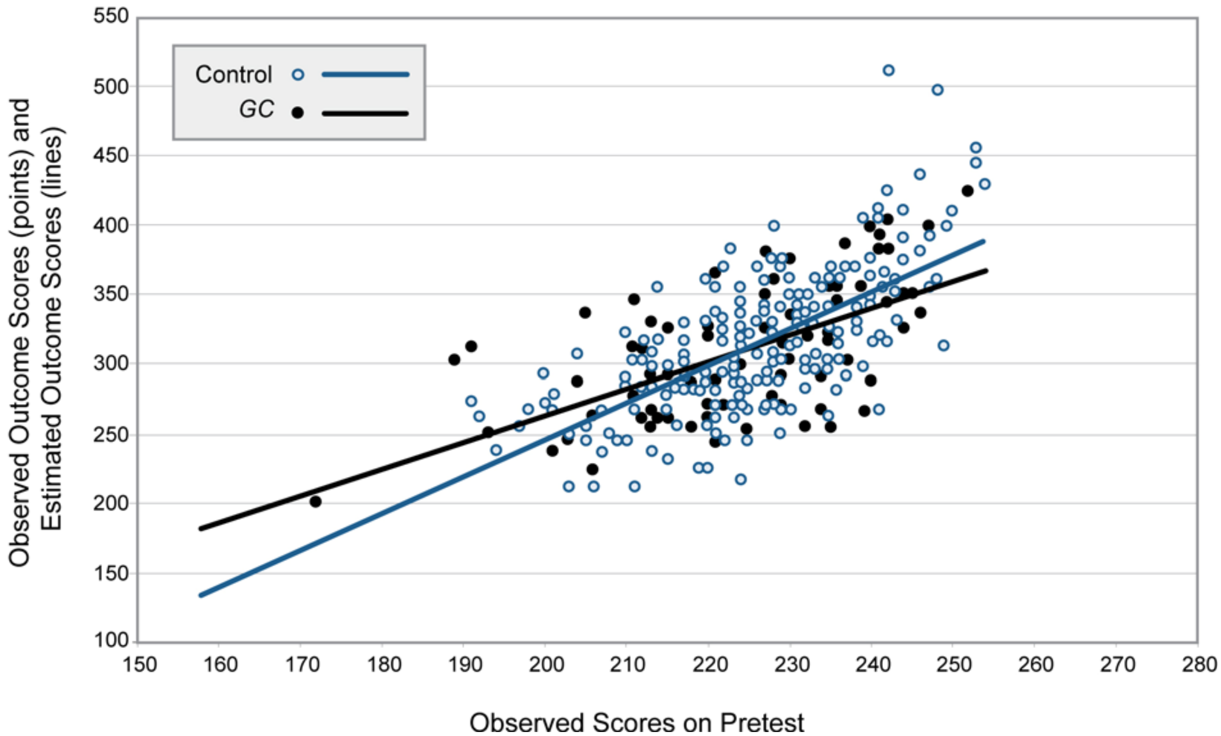

Figure 5. Comparison of Predicted and Actual CST Algebra Outcomes for *GC* and Control Group Students

Figure 6 shows the predicted difference between the *GC* and control groups for different points along the prior score scale. In this graph the estimated difference between *GC* and control groups is expressed as the straight line in the middle of the shaded bands – it is the predicted outcome for a *GC* student minus the predicted outcome for a control student. Around the difference line, we provide gradated bands representing confidence intervals. These confidence intervals are an alternative way of expressing uncertainty in the result. The band with the darkest shading surrounding the dark line is the "50-50" area, where the difference is considered equally likely to lie within the band as not. The region within the outermost shaded boundary is the 95% confidence interval—we are 95% sure that the true difference lies within these extremes. Between the 50% and 95% confidence intervals we also show the 80% and 90% confidence intervals. Consistent with the results in Table 32, there is evidence of a differential impact of the intervention across the prior score scale as measured by the NWEA test. Considering the points representing the median student in the bottom and top quartiles, it appears that *GC* has more benefit for the lower scoring students. However, neither point is sufficiently far from zero to be confident that it warrants a firm conclusion.

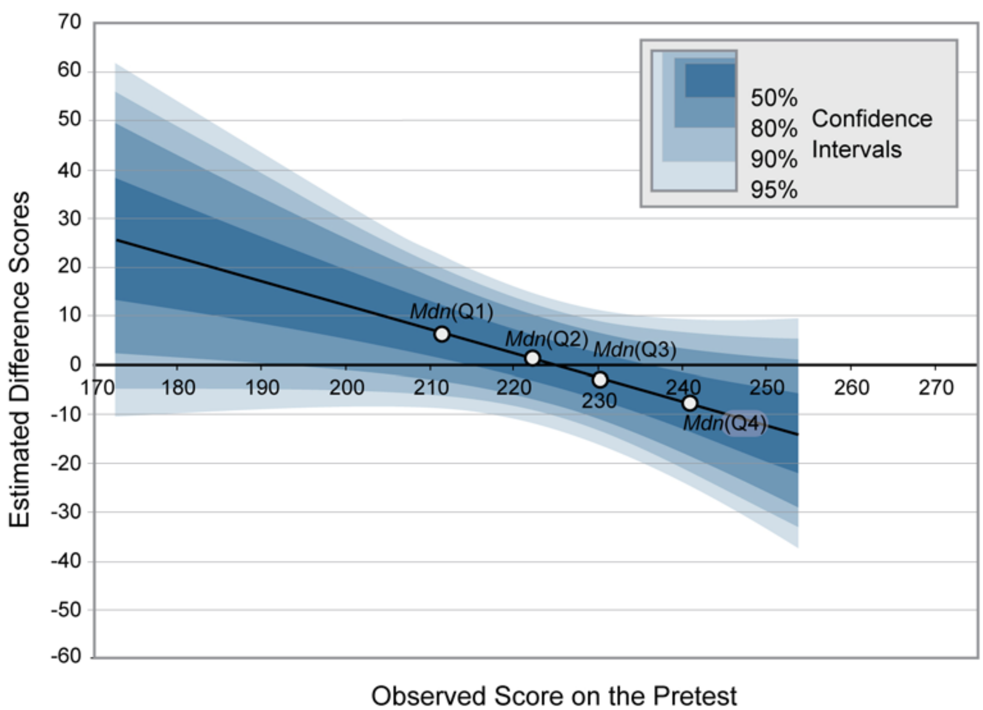

Figure 6. Differences between *GC* and Control Group CST Algebra Outcomes: Median Pretest Scores for Four Quartiles Shown

Figure 7 presents the same information represented in Figure 6 but this time in the form of a bar graph showing the predicted difference between *GC* and control conditions for students at the medians of the first and fourth quartiles of the pretest measure. The bar graph includes the 80% confidence interval as a marker at the top of the bars. These markers are an alternative representation of the 80% band in Figure 6 and are meant to be interpreted as: for either *GC*-control comparison, we are 80% sure that the true difference between conditions would place the tops of the bars simultaneously within the confidence interval markers. We see that for a student at the median of the first quartile there is little difference in the predicted outcomes in the two conditions and there is a substantial amount of overlap in the confidence intervals. The same applies to a student at the median of the fourth quartile.

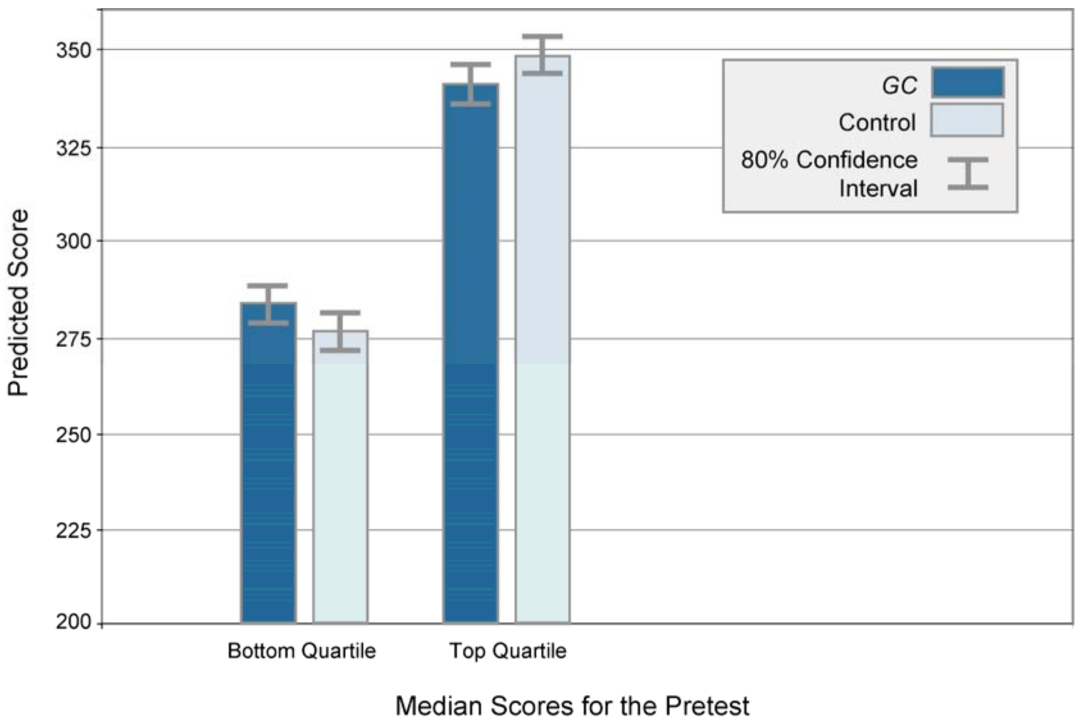

Figure 7. Difference Between *GC* and Control Group CST Algebra Outcomes: Median Students in Top and Bottom Quartiles

The overlap of the confidence intervals shows that the contrast between *GC* and control for the lower and high scoring students can easily be a matter of chance. Even though the differences between *GC* and control for these students at the median of the bottom and top quartile are small (and the confidence markers overlap), we can see that the direction of the difference changes for the two pairs of bars. The information in Table 25 gives us some confidence that this reversal is not due to chance.

Analysis Including English Proficiency as a Moderator

Table 26 reports the statistical analysis of the moderating effect of English proficiency on the Algebra outcome as measured by the CST test of Algebra. Here we were interested in whether *GC* was more or less effective for students not proficient in English compared to English proficient students.

Table 26. Moderating Effect of English Proficiency on CST Algebra Outcomes

Mixed model: Fixed factors	Estimate of coefficient	Standard error	DF	*t* value	*p* value
Intercept	307.06	5.00	12	61.35	<.01
Pretest score (centered at the mean)	2.37	0.17	290	14.00	<.01
Condition (*GC* = 1; control = 0)	10.97	12.55	12	0.87	.40
English proficiency (proficient = 1; not proficient = 0)	8.41	5.62	290	1.50	.14
English proficiency by condition interaction	-18.97	13.63	290	-1.39	.17
Mixed model: Technical details for random components	**Estimate of variance component**	**Standard error**		***z* value**	***p* value**
Teacher mean achievement	47.51	63.75		0.75	.23
Within teacher mean variation	1418.66	118.09		12.01	<.01

Note. Of these 621 students were included in the analysis of the Algebra outcome. Of these 592 had a posttest and were included in the calculation of the unadjusted effect size. Of the original sample, 310 students had pretest and posttest scores. Of these we removed 4 because they were influential points (where the determination of influential points was based on the model above.) The remaining 306 students were used to compute the results presented in the table above. The number of teachers for the reduced sample of 306 students is 14 (8 control and 6 *GC*). The number of classes is 31 (19 control and 12 *GC*.) Teachers are modeled as a random factor.

The coefficient for English proficiency is 8.41 indicating that on average the students not proficient in English scored somewhat lower (the *p* value of .14 indicates we have some confidence that this difference is due to the status rather than chance). The coefficient associated with the English proficiency by condition interaction is -18.97. This is an estimate of the difference between English-proficient and non-proficient students in impact of *GC*. The negative moderating effect of being an English-language-learner has a *p* value of .17. While this gives us limited confidence that this effect is due to something other than chance, it warrants consideration, which we can do in terms of the following figure.

Figure 8 displays the predicted values on the posttest, for a student who has an average score on the pretest. This is done separately for an English proficient student and a non-English proficient student. The bars in the graph are presented with their 80% confidence intervals. (As before, the levels of uncertainty are based on a joint distribution within each category of English-proficiency. A simultaneous 80% confidence interval for both levels of English proficiency would yield a more conservative result.) The result reflects the information in the table: English proficiency moderates the *GC* impact. Specifically, there is some evidence that *GC* is beneficial for low-English-proficient students, while the opposite is true for high-English-proficient students. We point out however that there is considerable overlap in the confidence intervals which limits our ability to draw strong conclusions about the moderating impact of English proficiency.

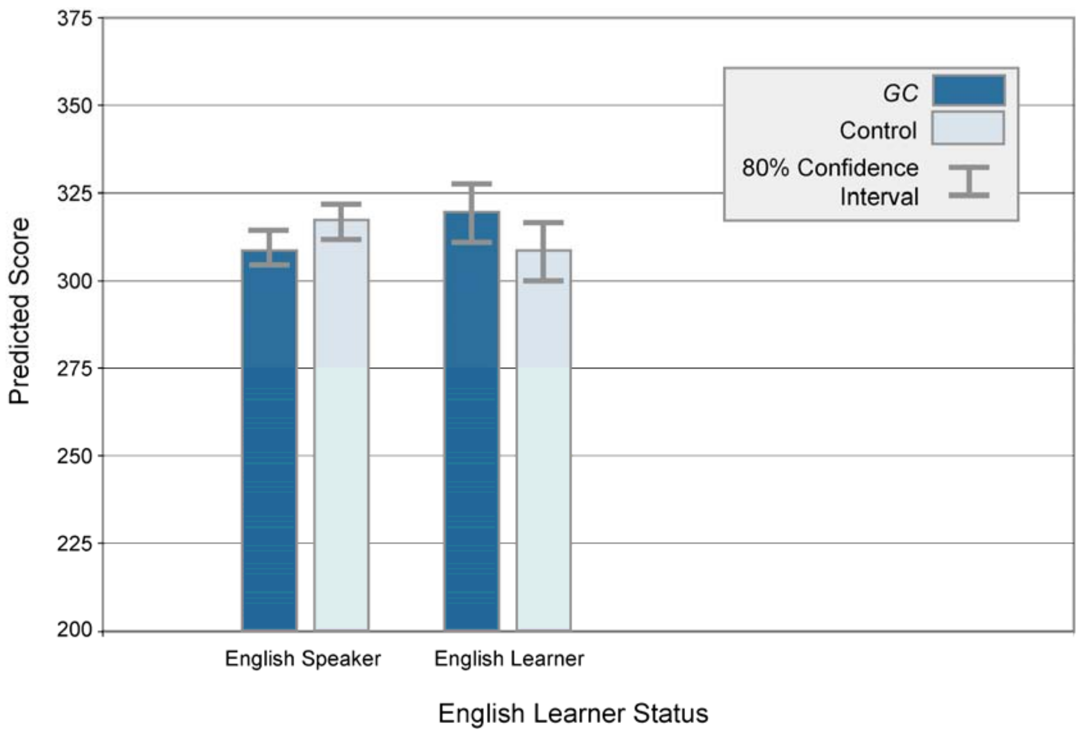

Figure 8. Difference Between *GC* and Control Group CST Algebra Scores: Median Students in Top and Bottom Quartiles

Impact of *GC* on Geometry

We also examined the impact of *GC* on student performance in Geometry. We assessed student performance in Geometry using the NWEA test of Geometry and the CST. The statistical analyses take the same form as those for Algebra.

NWEA Outcomes

Figure 9 summarizes the counts of students we use in the analysis of Geometry NWEA Outcomes.

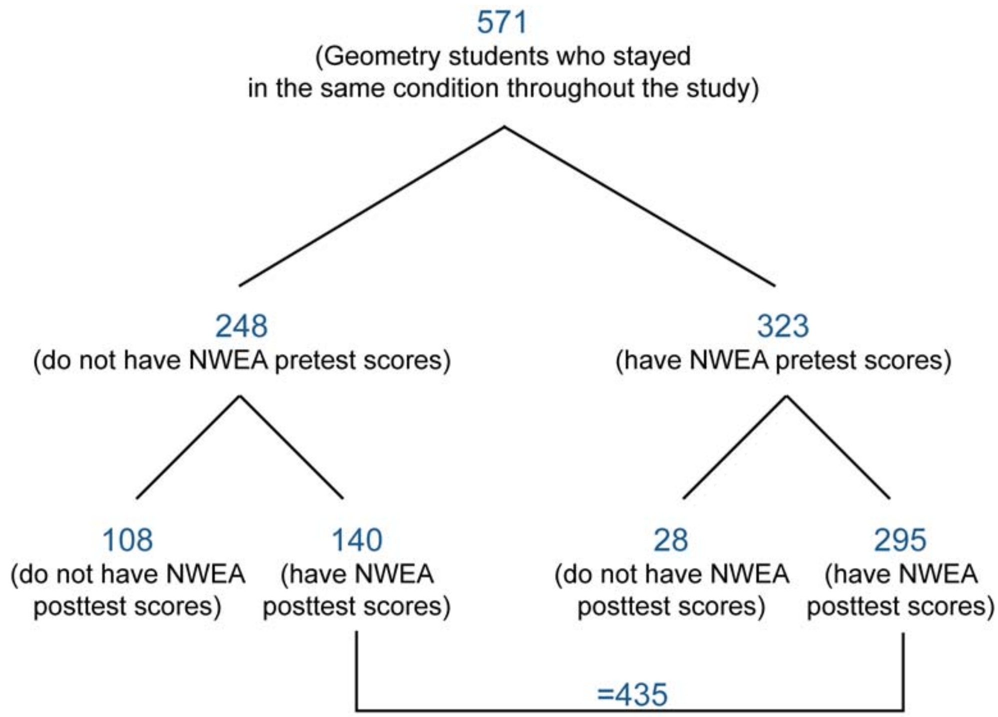

Figure 9. Counts of Geometry Students Having Both NWEA Pre- and Posttests

Analysis Including Pretest

Table 27. Estimated Impact of *GC* on NWEA Geometry Outcomes

Descriptive statistics	Raw group means	Standard deviation	No. of students	No. of classes	No. of teachers	Unadjusted effect size
GC	240.09	17.80	232	12	6	-0.17
Control	242.87	15.03	203	6	3	
Mixed model: Fixed factors	**Estimate of coefficient**		**Standard error**	**DF**	***t* value**	***p* value**
Intercept	243.44		2.86	4	85.18	<.01
Pretest score (centered at the mean)	0.75		0.06	284	12.59	<.01
Condition (*GC* = 1; control = 0)	3.30		4.06	4	0.81	.46
Pretest score by condition interaction	-0.14		0.09	284	-1.65	.10
Mixed model: Technical details for random components	**Estimate of variance component**		**Standard error**		***z* value**	***p* value**
Teacher mean achievement	23.06		17.40		1.33	.09
Within teacher variation	75.37		6.32		11.92	<.01

Note. 571 students were included in the analysis of the Geometry outcome. Of these 435 had a posttest and were included in the calculation of the unadjusted effect size. Of the original sample, 295 students had pretest and posttest scores. Of these we removed 3 because they were influential points (where the determination of influential points was based on the model above.) The remaining 292 students were used to compute the unadjusted effect size and the results presented in the table above. The number of teachers for the reduced sample of 292 students is 6 (3 control and 3 *GC*). The number of classes is 15 (8 control and 7 *GC*.) The teacher factor is modeled as fixed.

Table 27 shows the estimated impact of *GC* on students' performance on the NWEA Geometry test. For a student with an average score on the pretest, there is roughly a 3.30-point advantage to being in the *GC* group; that is, we predict that an average student would score 3.30 points higher on the outcome measure if he or she is in a *GC* class instead of a control class. The unadjusted effect size is -0.17, a negative value. The effect size when adjusted for the pretest is positive 0.20. (We calculated this using the impact estimate from a simple model that adjusted only for the pretest.) The *p* value for this effect, however, is .46 giving us no confidence that this effect is due to something other than chance. The value for the interaction between *GC* and the prior score is -0.14 with a *p* value of .10. We have some confidence that this effect is due to something other than chance.

As a visual representation of the results described in Table 27, we present a scatterplot in Figure 10, which shows student end-of-year performance in Geometry as measured by the NWEA test, against their performance on the NWEA general math test in the Fall. The two lines are the predicted values on the posttest for students in the *GC* and control conditions as determined using the estimated fixed effects in the model. (In the graph it appears that the variance in the outcome is not constant across the prior score scale. We adjusted for heteroscedasticity in the model.)

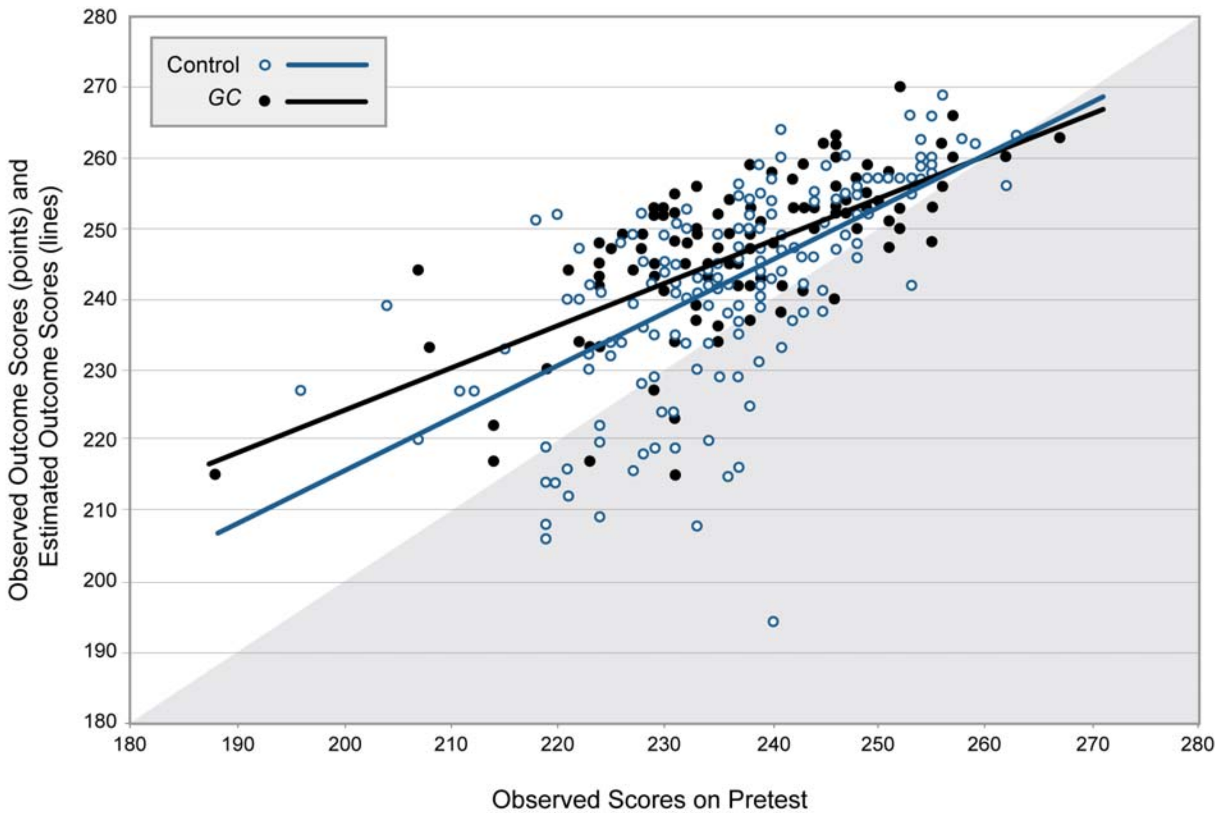

Figure 10. Comparison of Predicted and Actual NWEA Geometry Outcomes for *GC* and Control Group Students

As a visual representation of this result, Figure 11 shows the predicted difference between the *GC* and control groups for different points along the prior score scale. This graph takes the same form and has the same interpretation as Figure 6. Consistent with the results in Table 27, there is evidence of a differential impact of *GC* on Geometry performance as measured by the NWEA test. Specifically, along the prior score scale, we see that *GC* has a positive impact for students performing at the lower end of the prior score scale and no effect for the high-performing students.

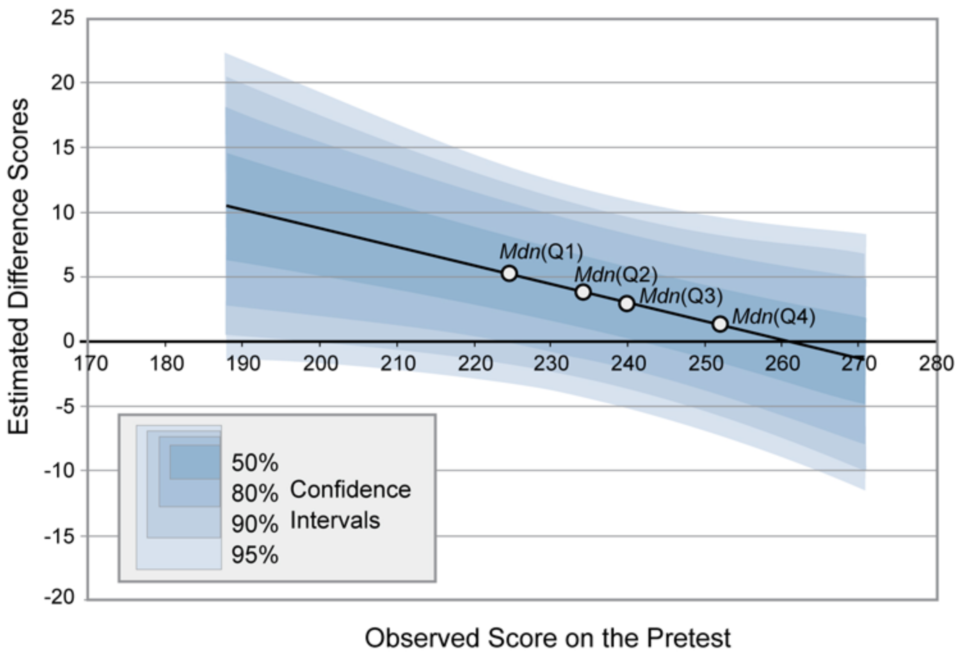

Figure 11. Differences Between *GC* and Control Group NWEA Geometry Outcomes: Median Students for Four Quartiles Shown

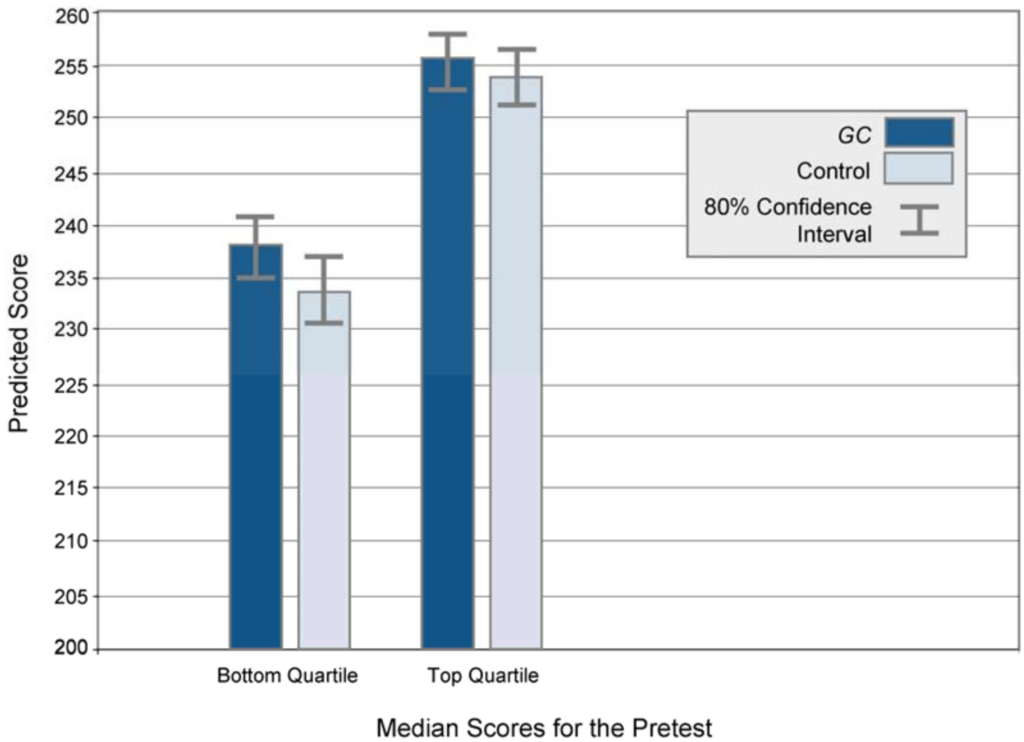

Figure 12. Difference between *GC* and Control Group NWEA Geometry Outcomes: Median Students in Top and Bottom Quartiles

Figure 12 presents the same information represented in Figure 11 but this time in the form of a bar graph showing the predicted difference between *GC* and control conditions for students at the medians of the first and fourth quartiles of the pretest measure. For either *GC*-control comparison, we are 80% sure that the true difference between conditions would place the tops of the bars simultaneously within the confidence interval markers. We see that for a student at the median of the first quartile there is an advantage to being in the *GC* condition. There is no overlap in the confidence intervals which suggests that we should have confidence in the estimated impact at this point on the scale. The advantage of being in *GC* for a student at the median of the fourth quartile is substantially less, and there is a high degree of overlap in the confidence intervals suggesting that *GC* has no impact for high-performing students.

Analysis Including English Proficiency as a Moderator

Table 28 shows the estimated moderating effect of English proficiency as measured by the NWEA test of Geometry. Here we were interested in whether *GC* was more or less effective for students not proficient in English compared to English proficient students.

In Table 12 we showed that there were only 6 non-proficient students in the control condition. This greatly reduces the precision of our estimates of the impact and moderator effects. Few of the effects of interest approach levels of statistical significance that allow us to have confidence that the results are for reasons other than chance. The lack of significance is probably driven by the small sample size.

Table 28. Moderating Effect of English Proficiency on NWEA Geometry Outcomes

Descriptive statistics	Raw group means	Standard deviation	No. of students	No. of classes	No. of teachers	Unadjusted effect Size
GC	240.09	17.80	232	12	6	-0.17
Control	242.87	15.03	203	6	3	
Mixed model: Fixed factors	**Estimate of coefficient**		**Standard error**	**DF**	***t* value**	***p* value**
Intercept	247.26		6.07	4	40.74	<.01
Pretest score (centered at the mean)	0.69		0.05	285	15.18	<.01
Condition (*GC* = 1; control = 0)	-3.21		7.27	4	-0.44	.68
English proficiency (proficient = 1; not proficient = 0)	-4.01		5.38	285	-0.75	.46
English proficiency by condition interaction	6.92		6.12	285	1.13	.26
Mixed model: Technical details for random components	**Estimate of variance component**		**Standard error**		***z* value**	***p* value**
Teacher mean achievement	25.48		19.23		1.32	.09
Within teacher mean variation	82.99		6.95		11.94	<.01

Note. 571 students were included in the analysis of the Geometry outcome. Of these 435 had a posttest and were included in the calculation of the unadjusted effect size. Of the original sample, 295 students had pretest and posttest scores. Of these we removed 1 because it was an influential point (where the determination of influential points was based on the model above.) The remaining 294 students were used to compute the results presented in the table above. The number of teachers for the reduced sample of 294 students is 6 (3 controls and 3 *GC*). The number of classes is 15 (8 controls and 7 *GC*.) The teacher factor is modeled as random.

CST Outcomes

Figure 13 summarizes the counts of students we use in the analysis of Geometry CST Outcomes.

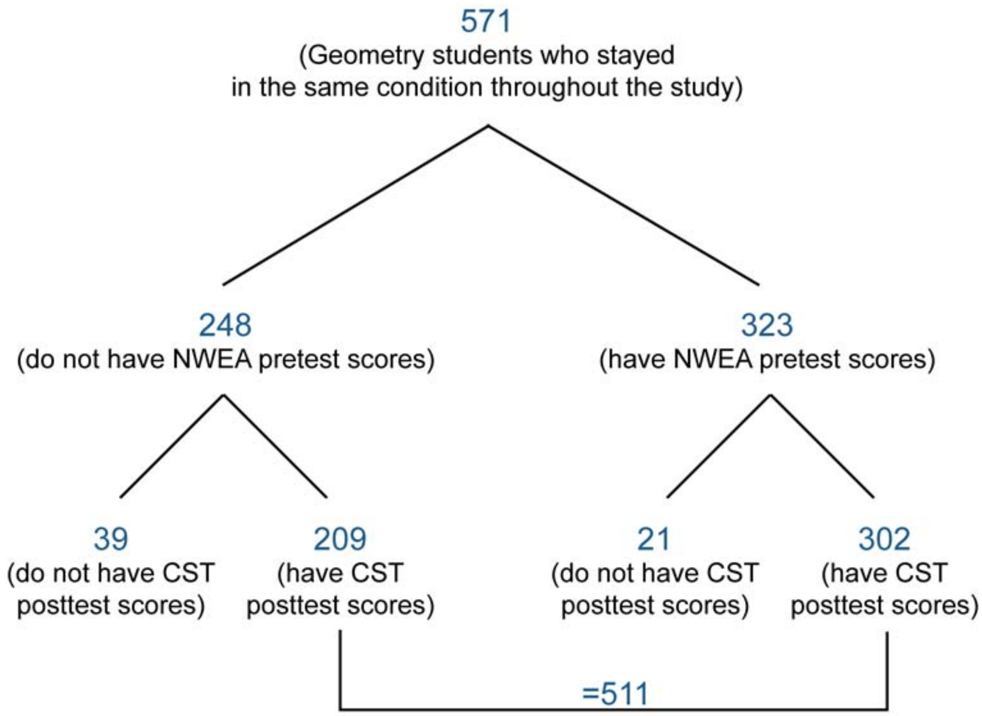

Figure 13. Counts of Geometry Students Having Both NWEA Pretest and CST Posttest

Analysis Including Pretest

Table 29. Estimated Impact of *GC* on CST Geometry Outcomes

Descriptive statistics	Raw group means	Standard deviation	No. of students	No. of classes	No. of teachers	Unadjusted effect size
GC	300.02	51.64	300	14	5	-0.23
Control	313.10	62.93	211	8	3	
Mixed model: Fixed factors	**Estimate of coefficient**		**Standard error**	**DF**	***t* value**	***p* value**
Intercept	312.07		9.18	4	33.98	<.01
Pretest score (centered at the mean)	3.56		0.25	289	14.50	<.01
Condition (*GC* = 1; control = 0)	4.04		13.10	4	0.31	.77
Pretest score by condition interaction	-0.57		0.38	289	-1.49	.14
Mixed model: Technical details for random components	**Estimate of variance component**		**Standard error**		***z* value**	***p* value**
Teacher mean achievement	228.85		183.00		1.25	.11
Within teacher variation	1281.74		106.63		12.02	<.01

Note. 571 students were included in the CST analysis of the Geometry outcome. Of these 511 had a posttest and were included in the calculation of the unadjusted effect size. Of the original sample, 302 students had pretest and posttest scores. Of these we removed 5 because they were influential points (where the determination of influential points was based on the model above.) The remaining 297 students were used to compute the unadjusted effect size and the results presented in the table above. The number of teachers for the reduced sample of 292 students is 6 (3 control and 3 *GC*). The number of classes is 15 (8 control and 7 *GC*.) The teacher factor is modeled as random.

Table 29 shows the estimated impact of *GC* on students' performance on the CST test of Geometry. The effect size is -0.23. The effect size when adjusted for the pretest is 0.06. The estimated effect for a student who scores at the average level on the pretest is 4.04 with a p-value of .77. We have no confidence that this effect is due to something other than chance. The value for the interaction between *GC* and the prior score is -0.57 with a *p* value of .14. We have some confidence that this effect is due to something other than chance.

As a visual representation of the results described in Table 29, we present a scatterplot in Figure 14, which shows student end-of-year performance in Geometry, as measured by the CST test, against their performance on the NWEA general math test in the fall. The cross over of the lines for *GC* and control is a representation of the interaction.

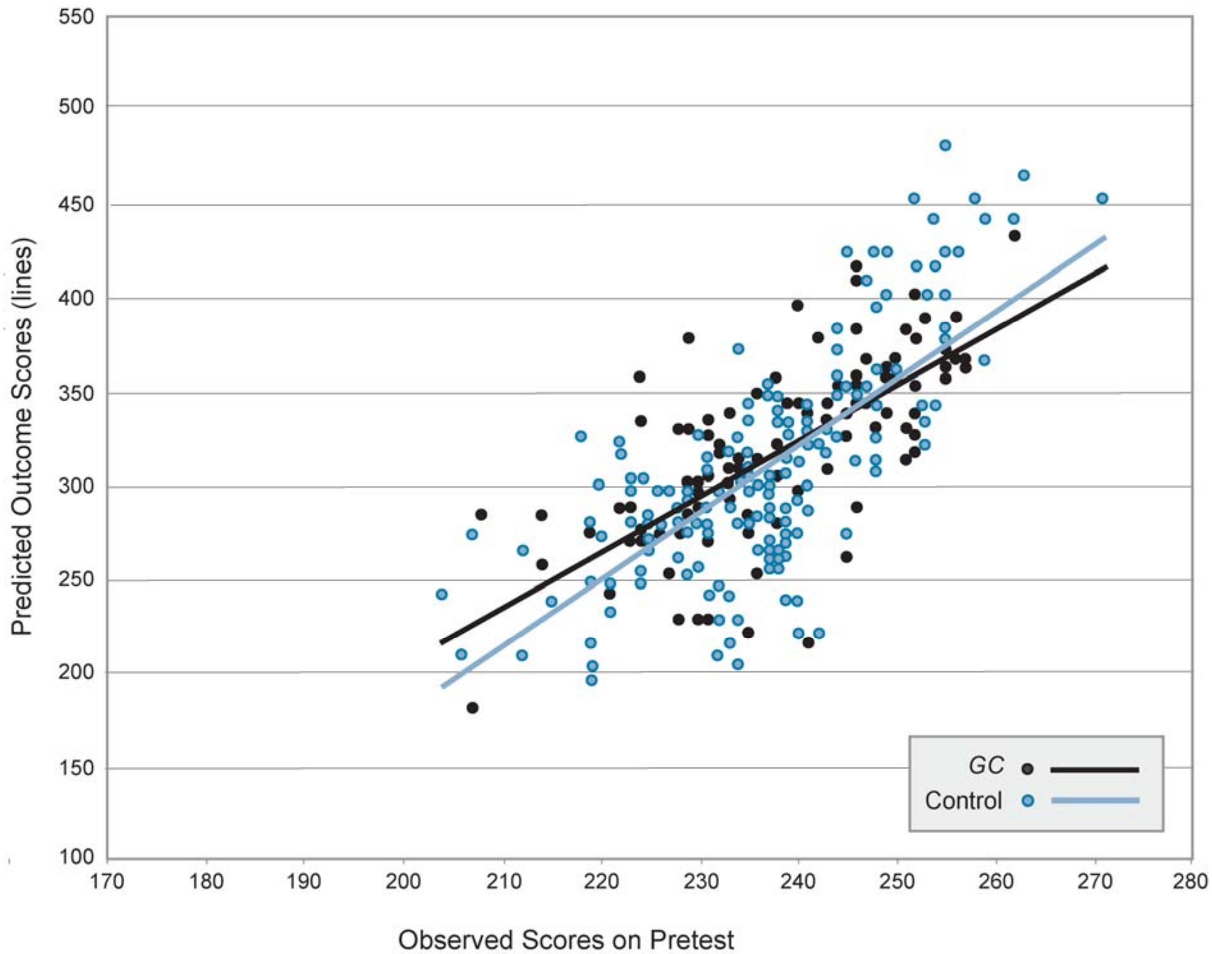

Figure 14. Comparison of Predicted and Actual CST Geometry Outcomes for *GC* and Control Group Students

Another way to visualize this interaction is provided in Figure 15, which shows the predicted difference between the *GC* and control groups for different points along the prior score scale. This graph takes the same form and has the same interpretation as Figure 6. Consistent with the results in Table 29, there is evidence of a differential impact of *GC* on Geometry performance as measured by the NWEA test. Specifically, along the prior score scale, we see that *GC* has a positive impact for students performing at the lower end of the prior score scale and no effect for the high-performing students.

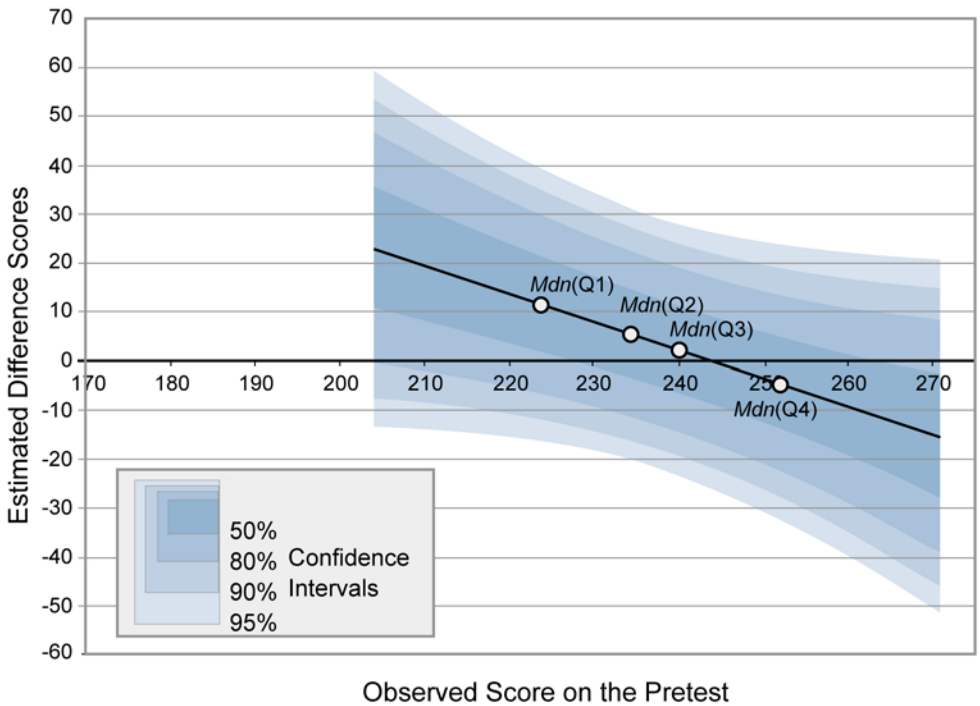

Figure 15. Differences between *GC* and Control Group CST Geometry Outcomes: Median Students for Four Quartiles Shown

Figure 16 presents the same information represented in Figure 15 but this time in the form of a bar graph showing the predicted difference between *GC* and control conditions for students at the medians of the first and fourth quartiles of the pretest measure. For either *GC*-control comparison, we are 80% sure that the true difference between conditions would place the tops of the bars simultaneously within the confidence interval markers. We see that for a student at the median of the first quartile there is an advantage to being in the *GC* condition. There is limited overlap in the confidence intervals which suggests that we should have some confidence in the estimated impact at this point on the scale. The advantage of being in *GC* for a student at the median of the fourth quartile is substantially less, and there is a high degree of overlap in the confidence intervals suggesting that *GC* has no impact for high-performing students.

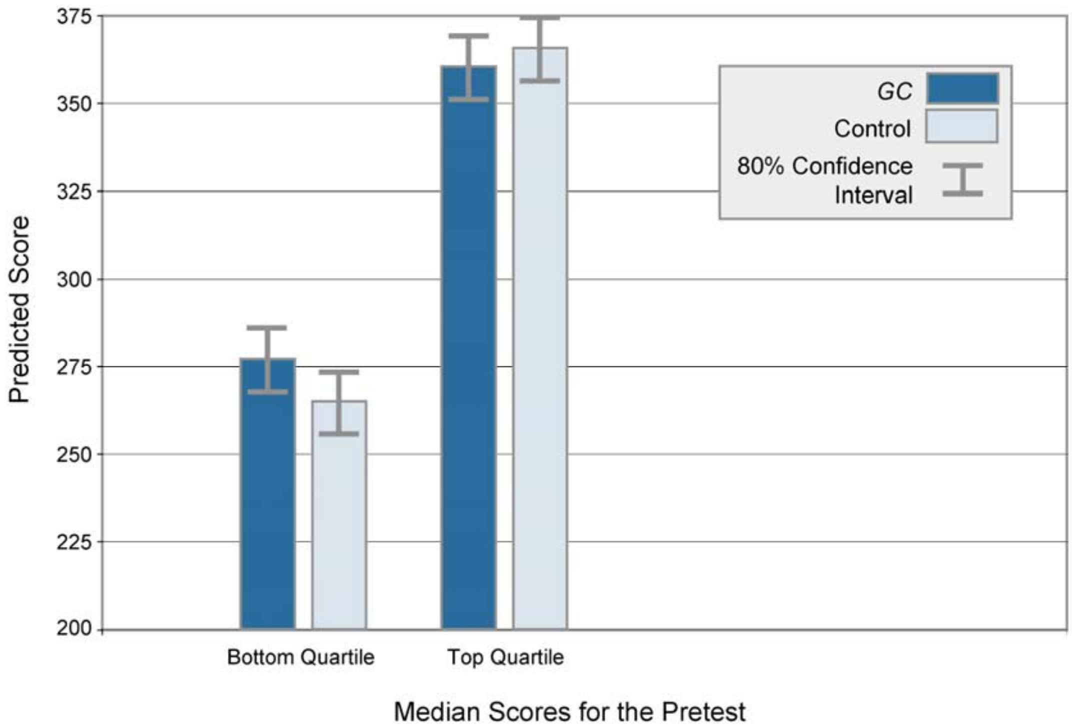

Figure 16. Differences Between *GC* and Control Group CST Geometry Outcomes: Median Pretest Scores in Top and Bottom Quartiles

Analysis Including English Proficiency as a Moderator

There were only two controls in the category of students not proficient in English who had CST scores. We therefore did not perform an analysis of the moderating effect of English proficiency on the *GC* effect as measured by the CST test of Geometry.

Repeated Measure Analysis Using General Math

Many students were given an intermediate administration of the NWEA general math test approximately at semester break. As mentioned previously, we considered the possibility of a repeated measure analysis using the data from fall, winter and spring. However the outcome measure was a different test from the first two tests—either End of Course Algebra I or End of Course Geometry. We concluded that the constructs tested in the final test were not comparable to the first two. An exception to this occurred with four teachers who mistakenly administered the NWEA general math test in the spring as well. While these teachers and their classes counted as attrition in the analyses presented so far, they did offer an opportunity to conduct a repeated measure analysis. For these teachers, and some of their 107 students, we did have data on three measurement occasions on the same scale. The advantage to an analysis using three points is that it allows estimation of change over time using an additional intermediate measurement of performance. We performed a repeated-measures analyses using HLM for these cases. We do not report the result because the sample size of teachers was too small to achieve adequate power, and because the *GC* condition could be easily confounded with teacher characteristics. Not surprisingly, there was no main effect of *GC*, but this simply reflects lack of statistical power.

Discussion of East Side Results

Our randomized experiment in East Side Union High School District provides some evidence of the impact of a program for using graphing calculators for Algebra and Geometry in the context of the particular implementation in this district. The research was weakened by difficulties in obtaining test results, especially for the pretest. However, in the absence of any reason to believe that the testing difficulties were related to the experimental conditions, we proceeded with an analysis of the outcomes.

The results of our experiment were modest. For the Algebra classes, we did not find an impact of the graphing calculator program in comparison to the control classes for either the NWEA or the CST outcome measures. We did find some evidence that the program favored students with lower prior achievement scores and favored students not proficient in English. While this was not evident in the analysis of the NWEA results, we did find for the California state test of Algebra an interaction between the experimental condition and the pretest as well as between the experimental condition and the student's English proficiency in which we have some degree of confidence. For the Geometry classes, the results based on the NWEA test results showed a relatively strong impact of the *GC* program (an effect size of .41) in which we have limited confidence. We can also conclude with some confidence that the effect was stronger for students who started out at a lower level in general math achievement. The available sample of Geometry students did not provide enough to test whether there was a differential impact for the students not proficient in English.

As we discussed in the Methods section, our analyses which involve subgroups of teachers have a much smaller sample size. Under these conditions only relatively large effects can be detected with any degree of confidence. In our report of the combined analysis from this and another district, we will be able to pool samples from both locations.

We did not investigate whether the impact was different depending on teacher years of experience because almost all the teachers were veterans and highly qualified to teach their subjects. We have also deferred the analysis of implementation differences as well as of the impact of TI-Navigator to our report on the analysis that combines the two sites. The larger number of teachers in the combined analysis allows us to detect the impacts of variables associated with teachers and the associated student clusters.

The implementation of the program met with several challenges including technical and pedagogical. These problems are typical of field settings where training and support face realistic constraints. We found considerable evidence that the *GC* condition was differentiated from the control condition so to some extent challenges were addressed. Our observations of implementation weaknesses serve both to inform the improvement of the Texas Instruments' technologies and services and to caution districts that may be considering implementing a program like this that technical support and professional development aimed at usage that addresses local standards and existing resources are critical.

It is important, finally, to emphasize that we are testing an implementation of a specific program in one particular school district. Our experiment cannot be used as definitive evidence of the value of graphing calculators but, as a randomized experiment, it does provide a valuable point of reference as researchers accumulate more evidence. As a local experiment to inform decisions in ESUHSD, we believe it supports the continued implementation and improvement of the program as is occurring in the second phase of the research now in operation.

San Diego Results

As with ESUHSD, our report of the SDCS results begins with an analysis of the composition of the experimental groups that were formed through the randomization of teachers prior to training and student testing. We also provide an accounting of the loss of teachers and students through mobility during the year and, importantly in this case, through difficulties in collecting the pretest data. In the second part of this section, we report on the implementation of the program including what we learned about the professional development, use of the technology and the issues raised by the participants.

The third section contains the quantitative results. We examined the impact of Graphing Calculators (*GC*) separately for Algebra I and Geometry. For each of these strands we used two outcome measures: NWEA tests of Algebra or Geometry and the California Standards Test (CST) of Algebra or Geometry. First, we report the impact of *GC* on Algebra I learning, and then we report the impact of *GC* on Geometry learning. We use the NWEA test of general math achievement from the fall assessment as a covariate in the analysis.

The basic question for the statistical analysis was whether, following the intervention, students in *GC* classrooms had higher scores than those in control classrooms. The study design also allows us to make comparisons between the *GC+Nav*, *GC*, and control groups. We report the results of that analysis in the combined report that encompasses both sites.

In all of the reports, in addition to looking at the main effect of *GC* and doing related analyses of covariance (ANCOVAs), we estimated the interactions of condition (*GC* versus control) with the pretest and English proficiency. In particular, we were interested in whether *GC* is more effective with English proficient students compared to non-English proficient students and whether *GC* was differentially effective for low- and high-performing students. These two moderators were identified before the intervention started and we report the results for both moderators in all the analyses. If the moderator effects are significant, we also graph the results.

Formation and Attrition of the Experimental Groups

Groups as Initially Randomized and Pretest Attrition

Of the original 22 teachers who agreed to participate, four teachers had to be excused from the study after the randomization and before the pretest measures. Around October 2005, two control teachers and two treatment teachers left the study. One control teacher was reassigned to a new school and no longer taught eligible classes. The second control teacher had to leave due to personal reasons unrelated to the study. One treatment teacher was reassigned to a different school district and no longer taught eligible classes. The second treatment teacher asked to leave the study because the perceived responsibilities of implementing the treatment would be overly demanding.

The randomizing process guarantees that there is no intentional or unintentional bias in the selection of teachers and students into the treatment or the control condition. It does not, however, guarantee that the groups will be perfectly matched. It is important to inspect the two groups to determine whether, in spite of randomization, there are any significant differences on factors that affect the outcome[1]. The following tables address the nature of the groups in each of the content areas. They also show that between the initial teacher randomization and class assignment and the

[1] In technical terms, randomization ensures a lack of bias due to selection, but we are interested in knowing whether the particular estimate resulting from this randomization may be far from the true value as a result of chance imbalances on factors that affect the outcome.

later gathering of pretest data, there was substantial attrition. Issues in obtaining pretest data through the computerized NWEA system and excluding teachers who switched their assignment condition during intervention led to a loss of 7 teachers, 15 classes and 875 students enrolled in the classes of teachers initially randomized. It is clear also from inspection of the table that there was substantially more attrition from the *GC* group than from the control. Fortunately, the portion of teachers lost at this point was much smaller. In other words, many teachers remained in the experiment but with fewer classes with pretests or with classes with fewer students with pretests.

We do not believe that there were any *GC*-related reasons for the differential attrition and therefore it does not represent an obvious source of bias. However, whenever a substantial loss is encountered, a level of uncertainty enters into the interpretation of results.

The following tables also show that subsequent attrition was far less severe. They report the number of teachers, classes, and students available for analysis given availability of CST posttest scores. No teachers were lost. One class was lost. Nineteen students who had pretests did not have posttests. Because the students who were not tested initially were still in the classes, we find that the number of students with posttests and who had been enrolled in the fall was quite large.

Table 30. Teachers in *GC* and Control Groups

		Sample at the start of the academic year	Completed study year 1	Teachers with at least one student having NWEA pretest with/without CST posttest	Teachers with at least one student having CST posttest with/without NWEA pretest	Teachers with at least one student having NWEA pretest and CST posttest
Algebra I	***GC***	6	6	4	6	4
	Control	4	3	1	2	1
Geometry	***GC***	6	5	4	4	4
	Control	6	6	4	5	4
Totals		**22**	**20**[a]	**13**	**17**	**13**

[a] One control teacher left for reasons related to the study and one treatment teacher left the study due to personal reasons; hence these teachers have no classes that qualify for the experiment.

Table 31. Classes in GC and Control Groups

		Sample at the start of the academic year	Completed study year 1	Classes with Students who have NWEA pretest with/without CST posttest	Classes with students who have CST posttest with/without NWEA pretest	Classes with students who have NWEA pretest and CST posttest
Algebra I	***GC***	17	17	12	16	12
	Control	5	3	1	3	1
Geometry	***GC***	13	9	8	8	8
	Control	17	17	13	16	13
Totals		**62**	**60**[a]	**47**	**59**	**46**

[a] One control teacher left for reasons related to the study and one treatment teacher left the study due to personal reasons; hence these teachers have no classes that qualify for the experiment.

Table 32. Students in *GC* and Control Groups

		Sample at the start of the academic year	Students with an NWEA pretest[a]	Completed program	Students with an NWEA pretest with/ without a CST posttest[b]	Students with a CST posttest with/ without an NWEA pretest[b]	Students with both an NWEA pretest and a CST posttest[b]
Algebra I	***GC***	466	282	328	220[c]	305	214
	Control	101	22	60	1[c]	51	1
	Totals	567	304	388	221	356	215
Geometry	***GC***	415	208	197	154	184	146
	Control	508	313	368	240	343	230
	Totals	923	521	565	394	527	376
Totals		**1490**	**825**	**953**	**615**	**883**	**591**

[a] Students who were initially randomized

[b] Students who completed the program

[c] Due to the large difference in number of students in GC and control group in Algebra , we will not consider this test outcome in further analyses.

Post Randomization Composition of the Experimental Groups

In checking for balance in the composition of the experimental groups, we examine student English proficiency and student pretest outcomes. This was done separately for active students enrolled in Algebra I and Geometry classes. These analyses are based on the full sample size of 1490 students who were initially randomized.

Student Variables

English Proficiency

We were able to examine English proficiency levels of the original sample of students assigned to classes of the teachers originally randomized. This was done separately for students enrolled in Algebra I and Geometry classes. The high p-value in Table 33 indicates that English proficiency was distributed evenly between conditions in the Algebra I classes. The same does not hold true for the Geometry classes where there are proportionately more non-English-proficient students in the control condition in spite of randomization. The randomization assures that the experiment remains unbiased although the imbalance must be noted.

Table 33. Comparison of English Proficiency between *GC* and Control Group Algebra I Students

	English Proficiency		
Condition	Not proficient	English proficient	Totals
GC	116	350	**466**
Control	29	72	**101**
Totals	**145**	**422**	**567**
Statistics	**Value**	**DF**	***p* value**
Chi-square	0.64	1	.43

Table 34. Comparison of English Proficiency between *GC* and Control Group Geometry Students

	English Proficiency		
Condition	Not proficient	English proficient	Totals
GC	41	374	**415**
Control	90	418	**508**
Totals	**131**	**792**	**923**
Statistics	**Value**	**DF**	***p* value**
Chi-square	11.52	1	<.01

Characteristics of the Experimental Groups Defined by Pretest

With randomization, we expect the pretest scores to be equally distributed between *GC* and control groups, but in any single randomization there may be discrepancies between the distributions due to chance. In the following tables we use the effect size measure as a way to consider the size of the initial differences (this is a measure of the extent of imbalance on the pretest in spite of randomization and is not the result of a cause.) For the Algebra I students, shown in Table 35, we note that there were only 22 control students and we see a significant difference on average test scores between the *GC* and control group. Those 304 students were students who are initially randomized and have an NWEA pretest score; however, when we consider only the students who completed the program, the number of students who have an NWEA pretest reduced to 221(220 treatment students, 1 control students), as shown in table 7. In the analyses that follow, we will not add NWEA pretest as a covariate for the impact estimate for Algebra I.

Table 35. Differences in Pretest Scores for Algebra I Students in GC and Control Groups

Descriptive statistics	Raw group means	Standard deviation	Number of students	Standard error	Effect size
GC	221.37	13.76	282	0.82	-0.30
Control	226.73	15.68	22	3.34	
***t* test for difference between independent means**	**Difference**		**DF**	***t* value**	***p* value**
Condition (*GC* - control)	-5.36		302	-1.74	.08

Note. Pretest scores for Fall NWEA test

For the Geometry students, *GC* and control groups had slightly different average pretest scores as shown in Table 36.

Table 36. Differences in Pretest Scores for Geometry Students in *GC* and Control Groups

Descriptive statistics	Raw group means	Standard deviation	Number of students	Standard error	Effect size
GC	234.39	12.28	208	0.85	0.41
Control	228.02	13.66	313	0.77	
***t* test for difference between independent means**	**Difference**		**DF**	***t* value**	***p* value**
Condition (*GC* - control)	6.37		519	5.43	<.01

Note. Pretest scores for Fall NWEA test

Attrition after the Pretest

From among the students who had pretests and initially randomized, we identified those who had posttests for the CST test. When examining attrition rates, we compare the two conditions in terms of the proportion of cases without a pretest. We then do a statistical test of the equality of these proportions. We see in Table 39 that 89.5% of the Algebra I students who had a pretest also had a CST posttest score. A Chi-square test indicates that roughly an equal proportion of students dropped out of each of the conditions.

Algebra I

Table 37. Availability of Pre- and Posttest Scores for Algebra I Students

Condition	Having both pre- and posttest scores	Having pretest scores only	Totals
GC	252	30	**282**
Control	20	2	**22**
Totals	**272**	**32**	**304**[a]
Statistics		***p* value**	
Fisher's exact test[b]		.29	

[a] Pretest scores refer to Fall NWEA test; posttest refers to CST.

[b] Since some of the cells have expected counts less than 10, a Chi-square may not be a valid test. Therefore, Fisher's exact test is used here.

Table 38 shows that students with a posttest scored at about the same level on the pretest as students without a posttest.

Table 38. Difference in Pretest Scores for Algebra I Students with Pre- and Posttest Vs. Posttest Only

Descriptive statistics	Raw group means	Standard deviation	Number of students	Standard error	Effect size
Having pretest scores only	222.04	14.09	32	2.24	0.20
Having both pre- and posttest scores	219.34	12.67	272	0.85	
***t* test for difference between independent means**	**Difference**		**DF**	***t* value**	***p* value**
(Having pretest scores only) – (Having both pre- and posttest scores)	2.69		302	1.03	.30

Note. Pretest scores refer to Fall NWEA test; posttest refers to CST.

Geometry

We see in Table 39 that 87.5% of the geometry students who had a pretest also had a CST posttest score. A Chi-square test indicates that roughly an equal proportion of students dropped out of each of the conditions.

Table 39. Availability of Pre- and Posttest Scores for Geometry Students

Condition	Having both pre- and posttest scores	Having pretest scores only	Totals
GC	183	25	**208**
Control	273	40	**313**
Totals	**456**	**65**	**521**[a]
Statistics	**DF**	**Value**	***p* value**
Chi-square	1	0.06	.80

[a] Pretest scores refer to Fall NWEA test; posttest refers to CST.

Table 40 shows that students with no score for the posttest scored slightly lower on the pretest. Since roughly the same proportion of lower scoring students dropped out of each of the conditions we assume that this does not introduce bias.

Table 40. Difference in Pretest Scores for Geometry Students with Pre- and Posttest Vs. Posttest Only

Descriptive statistics	Raw group means	Standard deviation	No. of students	Standard error	Unadjusted effect size
Having pretest scores only	224.75	13.99	65	1.73	-0.49
Having both pre- and posttest scores	231.39	13.22	456	0.62	
***t* test for difference between independent means**	**Difference**		**DF**	***t* value**	***p* value**
(Having pretest scores only) – (Having both pre- and posttest scores)	-6.64		519	3.76	<.01

Note. Pretest scores refer to Fall NWEA test; posttest refers to CST.

Implementation Results

In this section we describe more fully the aspects of the implementation that characterize this intervention. We used the following questions to guide our descriptions and analysis: What resources are needed to manifest the graphing calculator and the TI-Navigator condition? Are there differences in the extent, quality, and type of implementation of the materials? Our perspective takes into account three levels of resources needed to implement the intervention: those resources provided by TI, those provided by the school district and individual schools, and those provided by the teacher. We discuss each level separately. We also studied the features of the implementation to identify possible variables related to the outcome measures.

In year 1, TI did not communicate an explicit implementation model to the *GC* teachers for either the graphing calculator or for the TI-Navigator based graphing calculator system. TI did communicate an implicit *ideal* implementation framework based on the guidelines of the National Council of Teachers of Mathematics (NCTM) *Principles and Standards for School Mathematics* as outlined in the resource materials. During the training, the implicit implementation framework was rarely modeled by the trainer. Teachers were not directed to select particular activities nor were they directed to adopt the style of inquiry mathematics implied by the exploratory framework of the activities.

It is difficult to characterize deviations from the recommended practices given the implicit nature of the implementation framework. Nevertheless, given that there is an implicit structure following NCTM (2000) guidelines and since all of the teachers in the study are additionally provided structure by the California mathematics teaching standards, we can outline certain characteristics that mark low, medium and high implementation environments. We also recognize that this type of field-based implementation has integrity of its own and that certain elements must be present before it can be distinguished from the control condition. The characteristics and activities which identify an *ideal GC* condition are described in the Issues in Rating the Level of Implementation section.

Teacher Background

Teachers' background, attitudes, and preparation are commonly thought to impact the level of implementation (Fullan & Profret, 1977; Mukti, 2000; Supovitz & Turner 2000; Thompson, 2005). During the randomization process teachers identified themselves according to years of teaching experience and the initial teacher pair was formed correspondingly so that the bias due to teaching experience would be distributed among the groups evenly. As part of our data collection we asked the teachers to provide us with information regarding their backgrounds. The following tables summarize the background characteristics of the teachers in the study.

In general, most of the teachers in the study were mid-career professionals and most held college degrees with a mathematics major or at least some coursework in mathematics.

Table 41. Total Number of Years Teaching Experience

		Early career (0-3 Years)	Emerging professional (4-6 Years)	Mid-career professional (7-15 Years)	Highly experienced professional (15+ Years)
Condition	Number of teachers	No.	No.	No.	No.
GC	8	1	2	2	3
Control	8	1	0	3	4

Note. Refers to any teaching experience regardless of location or time period.

Table 42. Total Number of Years Teaching in Grade Level

		0-3 Years	4-6 Years	7-15 Years	15+ Years
Condition	Number of teachers	No.	No.	No.	No.
GC	8	1	3	3	1
Control	8	2	1	2	3
Note. Not necessarily consecutive					

Table 43. Extent of Math Preparation in College

		Some	Minor	Major	Masters
Condition	Number of teachers	No.	No.	No.	No.
GC	8	1	1	5	0
Control	8	2	1	6	0

Table 44. Recent Professional Development

		Recent participation	No participation
Condition	Number of teachers	No.	No.
GC	8	7	1
Control	8	7	1

Previous training in graphing calculators

Control teachers had limited experience with graphing calculators. Most control teachers had been exposed to the graphing calculator at some point in their careers and two teachers reported having had recent graphing calculator training. Trainings included summer T^3 trainings and workshops in the 1990's with the TI-82's and TI-83's.

The *GC* teachers tended to have slightly more experience using graphing calculators. One teacher reported not having attended a graphing calculator training in the past two years, but reports over 10 years of using graphing calculators. Trainings included TI-83 training specifically targeting calculus and statistics, TI summer institutes, and district sponsored training targeting the integration of the calculator with the textbook.

There no major differences in previous training with graphing calculators between the *GC* and the *GC+Nav* groups.

Implementation of the Graphing Calculator (*GC*) Program

Training Observations

During a five-day interactive workshop prior to the start of the academic year (August 1 – 5 2005), the teachers randomly assigned to the *GC* group were introduced to the TI-84 graphing calculator system and its capabilities. The week-long workshop was conducted by a TI consultant, a former California Certified teacher. During the workshop participating teachers used the graphing calculator to solve a variety of problems similar to those that the students would encounter during the normal course of instruction. This workshop described by TI as a typical five-day T^3 Summer Institute was held at Point Loma Nazarene University, Mission Valley Campus. On average each training day began at 8:00 AM and ended around 2:45 PM, with a 30 minute break for lunch and two 15 minutes breaks (one each morning and afternoon). All five days of the training were held in a large classroom where teachers practiced connecting and using the presentation tools, concentrating on the use of the emulator software recently developed by TI called, Smart View®. Teachers sat in small clusters of two or three. Most of these teachers knew each other fairly well because they had attended other math training workshops together or taught at the same school. Only one teacher was new to teaching and new to the district. The instructor allowed a loose environment using the information on the handouts to guide the hands-on activities. Of the 10 teachers randomly assigned to the treatment group, 7 participated for the entire week. Two teachers had summer school assignments. Of these teachers one came every day at 11:30, and the other only attended the last half of the first day. One teacher did not attend any of the training dates at all and dropped from participation in the study. The teacher that only participated in the training briefly was provided with DVDs containing video of the entire training workshop.

Teachers were directed to use the calculator technologies in the normal course of instruction, using the materials, calculators, and presentation technologies as appropriate instructional opportunities arise. Since there is no set curriculum that accompanies the use of the graphing calculator, teachers are expected to select from a variety of materials including those provided by TI to integrate calculator-based instructional activities into their lessons. Note that calculators are used heavily throughout mathematics and science instruction to facilitate numerical computation and so are present in the typical learning environment as computational tools. We are however more interested in aspects of calculator use during graphing and various other analytical modalities rather than computational activities.

The training covered several topics focused on building skill and navigating the various functions of the *GC*. These topics were covered within the context of various activities outlined in the Explorations series of workbooks (TI materials provided to all *GC* teachers) geared towards helping the teacher and students use the calculator for exploring concepts in Algebra I and Geometry. According to TI, the activities in the workbooks are tailored to specific student and teacher needs and are aligned to content standards. A typical training day covered approximately 14 different activities, each day focusing on developing increasing flexibility and knowledge of the calculator functions and interfaces. Topics concerning pedagogical choices and classroom management issues were rarely addressed. There was some discussion and time devoted to planning and adapting the activities for use in the individual teacher's classroom, but this was limited to approximately 1 hour of the week-long session and came about as a consequence of discussions among instructor and researchers.

A participant self-assessment focusing on calculator skills was administered at the end of the week-long session. Questions regarding basic four function computation and standard graphing functions ([Y=] [WINDOWN] [GRAPH] [TRACE] [ZOOM]) indicted that teachers self-reported that they had relatively high level of skills. Teachers responses to questions regarding creating and executing simple programs, creating and operating with matrices, the advanced feature set using Polar and Parametric graphing, using the replay features—RCL, ENTRY, and Enter, and using the DRAW menu in coordinate Geometry indicated that they evaluated themselves with low level skills. Most of these features were not addressed during the training.

We conducted a short focus group with the teachers on the last day of the training. We asked them to identify the benefits and challenges of the T^3 Summer Institute for teachers planning on using the graphing calculator in the classroom. While the general feeling was that the training was great for establishing technical fluency with the hardware, teachers did not feel prepared or comfortable to use the technology in their classroom with their students. Among the challenges that teachers identified were:

- Don't know how long kids will take to do these applications

- Don't know the "advisable" amount of time to spend on calculators each week
- Don't know why each application would be used in the classroom

We additionally asked the teachers what advice they would give TI regarding the redesign of the T^3 Summer Institute:

- Plan for lesson plan integration with teachers
- Create geometry/algebra groups to focus on core curriculum
- Use jigsaw method to deeply study then present different programs
- Have all the necessary equipment for the classroom in the training, so teachers can feel comfortable and safe about first day introduction to students
- Create list of best methods to use with certain math problems
- Give specific guidance about how to use the technology in the classroom

Materials

After the training, teachers were sent a digital projector, a teacher's edition graphing calculator, a calculator for each student in an eligible class, a calculator-based laboratory CBL/CBRs sufficient for use with a class size of 40 students, connectors for each calculator, resource workbooks, activity books, the SmartView® emulator software, various application software including the Cabri® Jr. Dynamic Geometry application, and a TI online account to access support and additional materials. See Appendix C for a complete list of materials.

The teachers were provided storage cases for the student edition calculators. All of the equipment was delivered to the teachers at their schools.

At the start of the experiment TI thought that the *GC* teachers would have access to computers in their classrooms. Although this is true in general, after the initial deployment of the materials we identified several issues as problematic. The issues can be summarized into three categories. Although three areas were identified, the issues are intertwined.

- **Availability** – Does the teacher have a desktop or laptop computer to use with the SmartView® software emulator?
- **Compatibility** – If the teacher has a desktop or laptop computer, is the operating system, speed, memory, etc. compatible with and able to adequately drive the emulator software?
- **Supportability** – What is the level of on-going technical support of the available desktop or laptop computer?

Availability

Most teachers do have either a desktop or laptop computer available to them for use in the classroom. Desktops computers are typically found either on the teacher's desk or off to the side of the room, neither location makes it easy to use with a projector given the constraints imposed by power and connectivity requirements. A few teachers had laptops available to them. The laptops do not have the same physical constraints as the desktop computers do, but suffer from the same operating system compatibility problems.

Compatibility

More often than not, the desktop computer is an Apple rather than a PC and is usually an older version operating system or in the case where the operating system is current the latest service pack has not been installed. The original emulator software disk did not contain an Apple version of the software consequently none of the teachers with Apple laptops were able to test the emulator software. A disk with the Apple version of the software was sent to any teacher requesting it.

Besides the operating system compatibility issues, the hardware concerns are more difficult to resolve. The likelihood that more memory can be added to either a desktop or laptop is very slim. School districts do not typically upgrade hardware piecemeal, they usually upgrade in large batches, upgrading a grade level or a school. If the problem with the PC is insufficient memory or slow speed, there is no school-level remedy available to the teacher.

Supportability

Most of the schools in this district have an Information Technology (IT) person that reports to a central district IT group. Requests for upgrades and troubleshooting go to central IT then routed back to the more local IT person. These requests can take upwards of three weeks to resolve. In some cases, the issues cannot be resolved because they require expenditures of funds that the school district has not allocated.

After reviewing all of the issues, we gave TI a list of recommendations to ensure that equipment issues were not a barrier to the use of graphing calculators in the classroom. TI responded by providing all *GC* group teachers with laptops with the emulator software already installed and tested. Additionally, they agreed to provide laptops to newly inducted *GC* group teachers in year two of the study during their initial professional development week. We continued to monitor equipment support issues through periodic check-in with the teachers.

Other Equipment Issues

Other issues arose as teachers began to use the different components of the system. One teacher requested that his digital projector be mounted on the ceiling so that it would be out of the way and he would not need to secure it on a daily basis. He had discovered that if the projector was bumped, the handshake connection to the laptop was interrupted causing the SmartView® software to crash, halting the lesson for several minutes. This teacher finally abandoned the use of the digital projector and emulator software favoring the overhead screen projector and ViewScreen™ panel because this combination of equipment provided him the necessary demonstration capabilities without the potential downtime.

As noted in the Introduction, part of the implementation strategy included giving students unrestricted access to calculators. Teachers however were reluctant to distribute the calculators because of the possibility that students would lose or break them. Teachers took several precautions to avoid losing calculators. Teachers received the school-bus yellow calculators that are easily identifiable as student versions and can only be purchased by school districts, but teachers thought it insufficient identification to protect them from loss. Some teachers used the school library's book check-out system to further tag and track the calculators. Those using this system etched each calculator with a number and recorded the number given to each student. Other teachers recorded the unique serial number noted in the memory section of the calculator before distributing to the students. Still others used the class set system, but also provided the students the opportunity to check out calculators overnight or over the weekend on an individual basis as requested by the student.

Perhaps the most ubiquitous and problematic issues were due to the operating system variation among the various pieces of hardware (graphing calculators and laptops) and the memory management requirements. Four of the teachers tried downloading StudyCards™ Applications and removing from the graphing calculator some of the language and game applications via the linking cables. Of these teachers, all reported that several of the student calculators would not respond to the commands issued through the linking cables. After contacting TI via their online support, the teachers became aware of the different operating systems and the need to ensure that all calculators have the same operating system in order to take advantage of the calculator communication ability. Teachers reported that upgrading the operating systems on the student calculators was extremely time consuming and many of the teachers abandoned the use of the communication abilities early in the study.

Memory management is an ongoing issue with the calculators. Calculators have limited memory available and require regular maintenance to ensure access to software applications such as Algebra I (ALG1CH1 etc.) and Cabri® Jr. At the beginning of the workbook, Topics in Algebra

(2004), that all of the *GC* teachers received as part of their training materials, TI clearly states that a TI Connectivity Cable and TI Connect™ software is required in order to run these applications. TI also advises that if an Archive Full error message appears while installing the Algebra I application, the graphing calculator does not have sufficient memory and that Applications and/or archived variables must be deleted to make room.

Many of the teachers complained that they neither had the time nor the desire to upgrade and memory manage the student calculators. This is a serious barrier to the implementation of the more robust applications that are available on the graphing calculators. Willing teachers turn into reluctant teachers very quickly faced with the daunting task of managing 70+ calculators.

Classroom Settings

The classroom setting was observed during normal classroom instruction from April 12th to 14th. Teachers were not asked to prepare specific lessons for observation, but teachers could determine when researchers came to observe within the observation window. All *GC* and TI-Navigator group teachers were formally observed at least once and a sample of control teachers were formally observed. In addition to the formal observations, researchers frequently observed classrooms from both groups during school visits to coordinate testing, drop-off materials, or to discuss implementation problems.

In the *GC* classrooms, we observed the graphing calculators used with graphing lessons that explored a “family of curves”, data collection and analysis, and a short lesson using the Cabri® Jr. software. In all, teachers reported using the calculator as a computational device more often than a device for exploration and investigation, but they understood the goal of trying to incorporate the technology more fully. Neither of the beginning teachers attempted any activities beyond the computational functionality of the calculators. Typical classroom routines included taking attendance followed by checking homework, moving to a new concept, practicing the new concept and assigning homework. We also observed students taking quizzes. Summarized from survey responses, we provide a breakdown of the time spent using calculators and the topics covered in Appendix E, Calculator System Usage.

Physical Settings

Most teachers in both groups had traditional classroom layouts consisting of individual student desks arranged in rows and facing towards a whiteboard, the designated “front” of the classroom. Two classrooms were observed in modified traditional set up of individual student desks paired next to each other, arranged in rows and facing front.

All teachers were observed to have a whiteboard and conventional overhead projector at the front of the room. Often the overhead projector was located on a rolling cart to ease storage and to accommodate different room arrangements. These resources were available and typically used during a lesson. Most teachers had a television in the classroom mounted from the ceiling or on a rolling cart, but no teacher observed used the television during a lesson. Teachers also had a computer located either in the front corner or rear corner of the classroom designated for teacher use. These computers were used for reporting attendance and creating lessons, but were not used during any of the observed lessons. A few teachers also had designated computer labs, but none of the observed lessons featured this technology, although a couple of the teachers commented that they did use Geometer’s SketchPad with their Geometry students.

Most notably, teachers did not have enough power outlets in the classroom to accommodate more devices than were already present. Most teachers powered their overhead and television through one outlet connected to a power strip. We report this feature in particular because the TI-Navigator system requires access to a power source for at least the charging cycle of the hubs and the laptop-projector components. Because of the limited power availability, a TI-Navigator system setup may not be possible in some classrooms.

Most teachers and students in the *GC* and TI-Navigator classrooms were observed to have the TI-84 graphing calculator or other graphing calculators available for use during class assignments. Few control group teachers were observed to have calculators available during a lesson. Those *GC* and control classrooms with calculators present appeared to be using the calculators as a computation device more often than not.

During the weeks set aside for CST review and preparation students are not allowed to use calculators at all for any reason because the assessment is conducted without the aid of calculators. Consequently, during this time all calculator use was suspended.

Instructional groups

There are four distinct types of student groupings in the Algebra classroom: students with varying degrees of prior instruction in Algebra, but not considered to have *failed* Algebra I, these are designated as "first-time freshmen"; Accelerated or Honors Algebra students, these students are expected to do well and consequently follow an aggressive schedule and typically cover more material than first-time freshmen; Sheltered Algebra students are those identified as needing more English language support during instruction; and those students in 10th, 11th, or 12th grade that have not successfully passed Algebra I in a previous attempt, these students are designated as "repeaters".

Geometry classes are similarly grouped. There are fewer "repeater" classrooms because Geometry credit is not needed to graduate high school.

Implementation of the Graphing Calculator with TI-Navigator (*GC*+Nav) Program

Training Observations

In January, from among the *GC* group teachers, five teachers were randomly selected to use the TI-Navigator based system. Together with a group selected from among the East Side Union High School District *GC* group, they attended a three-day training at the East Side Union district offices. The training was conducted by a university professor that is part of a school of education engaged in teacher preparation and education. Training concentrated on system set-up, managing of applications, and various activity skills. There was direct attention paid to issues of how students could be organized to engage in the activities and how to structure the classroom for optimum use. Even so, we noted that it was difficult for teachers to take in the new aspects of the system. Three teachers of the five selected did not attend the January training. Of these teachers only one was able to set-up the TI-Navigator system successfully. The other two did not set-up the TI-Navigator system at all. From teacher feedback, mid-year implementations are hard to negotiate, because it requires planning and new routines to be incorporated in the classroom. Planning and establishing new routines take time and can potentially impact an already hectic schedule.

Materials

At the end of the training session, these teachers received the TI-Navigator system software and hardware sufficient for connecting a classroom set of calculators. Texas Instruments provided teachers with on-going support for the all of the equipment replacing any malfunctioning equipment or software.

Classroom Setting

While TI provided all *GC* and *GC*+Nav teachers with a notebook computer and data projector, few *GC* and *GC*+Nav teachers were observed using the technology. The principal differences in classroom setting were noted in the two classrooms that were using the TI-Navigator technology. In these classrooms, the systems were up and connected, student desks were clustered around the hubs (groups of 4 students to one hub) and the students and teachers had developed routines where they systematically logged-in to the system at the start of class. Students wasted no time in getting their calculator networked to the distribution hub connected to the teacher's laptop and working on the warm-up problems. The teacher noted attendance using the laptop as the students were acknowledged by the log-in process. Both the level of activity and the discussion about mathematics was high among the students. Summarized from survey responses we provide a breakdown of the time spent using calculators and the topics covered in Appendix E, Calculator System Usage.

The instructional groups were the same as noted for the *GC* classrooms.

Issues in Rating the Level of Implementation

We begin our discussion of implementation issues with a summary of the perspectives through which we viewed implementation. For our purposes, we define implementation as a specified set of activities designed to put into practice a program of know dimensions. Accordingly, implementation processes are purposeful and are described in sufficient detail so that observers can detect the presence and strength of the "specific set of activities" named as the implementation program. In this case, we are interested in the implementation of TI-84 graphing calculators and TI-Navigator into two types of instructional settings, Algebra I and Geometry classrooms.

In addition to implementation processes, we also note that there are *intervention* processes and outcomes. These we define as the processes required to set the implementation processes in motion. Intervention processes in this case include, a description of the implementation program, the training supplied by TI to the teachers, the equipment and materials deployed, the settings and resources (materials, time, and other personnel) supplied by the schools, and the on-going support provided by TI.

The first issue we confront is the non-specified nature of the implementation program. As noted before, the integration of graphing calculators and/or TI-Navigator into classroom instruction is ill-defined as an intervention program. There are no specific principals that guide the use of this technology. At best, we have some general guidelines that address emerging patterns of fruitful use. The training provided some pedagogical strategies, but these were hidden within the form of activities and never explicitly discussed. For this study, what, when, and how the materials were to be used were left to the individual teacher's judgment and discretion. The number of activities to be used for the study was not specified. Once we realized that the intervention was so loosely defined, we turned our attention to developing a framework whereby we might characterize practices that use the technology in productive ways.

Framework

In developing the observation framework we considered four sources: recent research on learning with technology, National Council of Teachers of Mathematics (NCTM) Principles and Standards for School Mathematics, California Standards for the Teaching Profession (1997), National Board for Professional Teaching Standards, and the International Society for Technology in Education (ISTE) Technology Standards for Teachers.

How People Learn (2000) explains how technology might be used to:

- introduce real-world problems into the classroom
- provide thinking scaffolds and tools for problem-solving
- provide multiple ways to represent data
- use networking capabilities to stimulate communication for presentation, feedback, reflection, and revision
- cause change in the teacher's role in the classroom

While recent research such as How People Learn (2000) provides guidance, researchers also indicate we have just begun to understand the potential of these technologies as change agents for social interactions to promote learning.

According to the NCTM Principles and Standards for School Mathematics (2000) the most fruitful uses of technology during instruction takes place in an environment of student-centered practices that allow students to explore complex problems and mathematical ideas and to present their solutions. This type of environment promotes discussions that challenge and extend students' mathematical ideas thereby helping students to develop deeper mathematical understandings. From the California and National Board teaching standards, we noted areas that articulate elements present in robust student learning environments and coupled them with the ideas expressed in the and ISTE principles. These we categorized into three main concepts:

- *classroom routines* that presume the use of the *GC* and TI-Navigator technologies not as an add-on, but as part of the flow of working with and learning mathematics
- *classroom interactions* that are marked by shared responsibilities among teacher and students
- *lesson structures* that focus on learning, where learning outcomes are clearly articulated, frequent monitoring of learning is present and discussed by both teacher and students, and students question, offer conjectures, and allow for alternative approaches and solutions

These ideas led to the construction of an observation protocol that can be found in Appendix B. We began using this observation protocol to more closely identify and profile those emergent practices that distinguish the major attributes of the *GC* and the *GC+Nav* conditions. We used focused interviews with a select group of teachers to further explore how they conceptualized the use of calculator-based technologies in the classroom. These findings are summarized in the Teacher Case Profiles section.

Definition of the intervention program

The second issue we confront is the intervention process. Since there are multiple processes these became a separate set of issues. The intervention process in this case consists of series of stages: definition of the intervention program, teacher training, distribution of materials, classroom set-up, teacher planning, teacher use, and student use. We have, in part, addressed the definition issue in the framework, but several aspects remain undefined.

Although teachers appear to be very comfortable with the basic aspects (common computation functions, graphing of y =) of the calculator, the extended feature set requires initial training and practice before the teacher can use them with ease. To what extent should the training address the feature set and how many activities should the teachers be asked to implement during a semester or academic year? The research is silent on the "quantity" issue which is only one aspect of the classic dosage question, how much, how often? What we do know is that classroom routines must be established in order to manage the usage of technology in the classroom. In this study the routines established by some of the teachers in the *GC* and *GC+Nav* group point to the complexity of the issues these teachers are trying to facilitate.

Should the training be designed to address different aspects of the calculator or TI-Navigator system according to the pedagogical strategy rather than "Topics in Algebra"? Current training focuses strongly on introducing and practicing the functions of the calculator with an orientation towards "student activities", but does not actually discuss the reasoning and pedagogical strategies that are required. We have many more questions regarding the relationship between the training and the implementation program than we are prepared to answer. Our instruments and protocols did not specifically focus on trying to answer these questions directly, but we were able to understand the general nature of what is required for teachers to manifest a level of implementation distinguishable from the control condition. In general, we noted several indicators that must be present for the teachers and students to enact graphing calculator supported Algebra I and Geometry lessons. Teachers must be reasonably familiar:

- and proficient with the graphing calculator feature set, software, networking capabilities, and have general troubleshooting skills
- with the curriculum and standards and how the activities outlined in the training can be incorporated into the flow of instruction
- with methods for establishing new routines that incorporate the systems and gradually scaffold students into becoming proficient with the systems' functionality set so that they can access the more powerful features

- with methods for creating classroom organizations that take advantage of the new capabilities

Student learning of the functionality of the calculator is not addressed in the training directly. The functionality of the graphing calculator system is complex enough so that developing proficiency on the part of the student must be incorporated into the dynamic of the lessons. Teachers can not take for granted that students are familiar with graphing calculators and their capabilities beyond the four computation functions. Students must be reasonably familiar:

- with the calculator and software so that access to the applications and other analyses and data functions become routine and not the focus of the lessons
- with group work dynamics so that they can rely on each other for support regarding the feature set and release the teacher to concentrate on helping them develop mathematical understandings

Those teachers that were successful in using the graphing calculator and TI-Navigator system all displayed the above indicators.

Time constraints

A third issue is the time available for learning, planning, experimentation, and lesson instruction. Many teachers expressed that they did not have any time to play and plan with the graphing calculator system. This factor potentially has the ability to derail any desired implementation. As can be seen from the case profile data, even those teachers that were successful implementers still lamented the lack of time available. We noted this in this study and made recommendations to TI to modify the training in year 2 of the study to include daily planning time so that teachers could discuss which activities might work given the pacing and the curricular demands, how they might organize the classroom and to practice the activities as if they were going to teach it to their students.

Teacher Case Profiles

Four teachers were interviewed on the successes, barriers, and implementation of the *GC* and TI-Navigator technology.

TI-Navigator (Teacher ID #65)

Teacher ID #65 has over 20 years experience using graphing calculators and has taught with the graphing calculators for over eight years. She has attended trainings by TI that focused on applying the TI-83 to calculus and AP classes, and although not a TI instructor, she has given her own presentations of fractals using the TI-83. She believes she has more experience than most teachers and is current on practices using the calculator for calculus and pre-calculus instruction. She feels confident using the lists functions, drawing graphs, shading under curves, and creating animations. Given her extensive familiarity with the TI-82, 83 and 91, she felt transferring to the TI-84 was easy. She also finds the TI support phone number to be very helpful and used the support to walk her through setting up the TI-Navigator step by step.

This teacher uses TI-Navigator extensively in her classes. She likes the screen capture feature of the TI-Navigator system because she can monitor students closely. In the past, she would demonstrate with the calculator and expect students to follow along without really knowing if in fact they were following. The screen capture feature allows her to check on students at various times during the process. She also uses the TI-Navigator system to replace activities originally done on paper, such as vocabulary review. She differentiates instruction by giving independent exploration activities to students who are more skilled with the calculator.

To implement the TI-Navigator successfully, this teacher has established routines. She numbers and assigns each student a specific calculator that no other student will use. This ensures the student the same calculator every time. She also permits students to take their specific calculator home on an individual basis. Typically, the teacher will set up and troubleshoot the TI-Navigator system while students are performing a warm-up activity. She introduces a topic, teaches a lesson that integrates the calculator, and ends the lesson with a quiz and screen capture for her later review.

The teacher suggests that TI-Navigator teachers be well versed in its functions before implementing it in the classroom. She recommends that a teacher come in early to make sure the TI-Navigator is set up and working properly. In the beginning, she believes teachers should spend at least the first five days instructing students on how to use the calculator and TI-Navigator. The TI-Navigator must be used every day for some time so that it becomes a habit and is considered part of the ongoing lesson flow rather than a periodic addition.

In terms of limitations, the teacher believes time to be the biggest constraint. She would like to use the TI-Navigator extensively every day, but there is not a lot of time in her schedule. Time is wasted when the TI-Navigator malfunctions because lesson flow is interrupted and quick instructional alternatives are not always ready to hand. Malfunctions would usually happen when the teacher dismantled the TI-Navigator for practice at home. Additionally, she would experience compatibility issues when trying to use files created on a home computer rather than on the provided laptop. Another limitation is the lack of access to the internet. The teacher does not have reliable internet access in her room. She has been able to access the TI support sites infrequently, but reports she has not found anything applicable to her classes. The teacher would prefer to have activities that are already designed and ready to be used. Student mobility is a separate issue, but not unrelated as it can have an impact on accessibility. Due to the high mobility rate at the school, several calculators have not been returned and as a consequence having sufficient calculators could become an issue.

The teacher reports that using the calculators alone takes up a lot of time at first, but the students look forward to using them. There is a tension in using the calculators all of the time. There is a tendency for students to use them as computational crutches and become indiscriminately reliant on them. As an isolated issue, this is not problematic but because the CST requires that students take the math exam without the aid of a calculator, overdependence can be an opportunity for failure.

She has tried the SmartView software to demonstrate calculator keystrokes, but believes it is too slow to be of much use. She prefers the overhead ViewScreen for presentation functions, such as demonstrating CabrJr. constructions.

In describing a successful lesson, the teacher recalls geometry explorations where she blended the technology usage with traditional paper methods. She has students draw and measure angles and line segments on the calculator. She has the students repeat the same procedure using paper. However she notes the accuracy of the calculator was difficult to replicate on the paper and would change her lessons to start with paper and move to calculator constructions.

For the students to be successful, the teacher believes students need to know the locations of the keys. She has found the TI poster to be very helpful. She also believes the students need to know how to maintain the calculators, such as keeping the OS current and transfer files using the linking cables.

GC **(Teacher ID#58)**

Teacher ID#58 has had previous training with the TI-83 calculators when she started using the Algebra Explorations textbook. The training consisted of eight meetings once a month that focused on using the TI-83 and CBRs with the textbook. At the end of the meetings, the teacher felt confident on how the textbook authors expected teachers to use the TI-83s. Since then, she has also attended several summer trainings on the TI-83 sponsored by district. She reports having taken nearly every training the district offered that focused on algebraic thinking, geometry, and Geometer SketchPad. The teacher has also taken training on using the Carnegie Learning Cognitive Tutor program in the classroom that focused on using the TI-82.

When planning a lesson, the teacher plans to attract the students' attentions, add to their knowledge, expand their knowledge, and translate the knowledge into real world situations. She believes the novelty of using the calculator with the digital projector would draw the students' attentions. Her primary goal is to have the students see a pattern and make conjectures about the pattern. When working with the calculator, she wants students to use the calculator to do something they could not do without the calculator. She sees that the calculator could do something else to analyze the data. In her opinion, the technology package is a good way of showing the students what more can be done.

From a teacher perspective the teacher believes the students are very tech savvy and can easily learn to use the device. She believes the calculator should be introduced right away as a valuable tool that can be used from the beginning for special activities. She does not want the students to misuse the calculator and rely on it as a crutch. She also believes expectations must be set when the calculators are introduced that it is not a computational crutch, that calculator games are not permitted during class, and that these contracts should be formalized in English and Spanish.

To be successful with the calculators, teachers need to be good at troubleshooting the problems that may arise. She observed that all the male students would change the settings on the calculator to make it do something "crazy". With that, the teacher needs to be well versed in setting the calculator window, modes, and floating decimal points. She believes that many of these small issues can ruin an activity and that teachers should be able to easily identify and correct these.

For the students, the teacher believes that students can easily conquer the calculator if they have the confidence to explore. Students need to understand that they cannot "break" the calculator and that there is a logical explanation for why something did not work out. The teacher encourages students to do more exploring and structures activities to explore more. She will send home assignments that allow students to explore, such as manipulating slopes, and drawing conclusions about why it happened.

The teacher describes a successful activity demonstrating compound interest to the students. When the calculator is used in conjunction with the data projector and Microsoft Excel, students could visually see the rapid growth due to compounded interest when compared to simple interest. The class applied the concept to a window washing business and a comparison of cell phone plans. The teacher also found the SmartView key history to be very useful when presenting to the students.

The teacher could not think of a calculator activity that was not successful with the students. If an activity did not seem to be working well, the teacher would suspend the activity for another time or make modifications for the next periods. She believed that giving the students a paper and pencil in addition to the calculator gave students an added sense of security when doing any activity.

The teacher did not believe the training was sufficient. She likes using the Cabri Jr. activities with the students, but did not think the training covered those topics adequately. She also would have liked to learn more about using the CBR because that is an area she felt she needed more development. She tried to search the TI website for further instruction, but reports she could not find anything that was applicable.

***GC* (Teacher ID #70)**

Teacher ID #70 is a relatively new teacher (fifth year) with little experience teaching with a graphing calculator. He reports having taught with a graphing calculator before, but never having attended a formal training that focused on teaching with the graphing calculator. His experience primarily stems from using a graphing calculator as a math major in college and presenting graphs on an overhead projector in the classroom. He does not think he is a superior calculator implementer but believes he does a strong job with what he has available.

When planning a lesson, the teacher focuses on time planning, the ability to explain concepts in a way the students will understand, and the ease of use for students. He uses the calculator in lessons only where the calculator is applicable and does not force the calculator. The calculator lesson needs to have a point and produce a successful method of discovery that can be clearly explained to the students. He does not want to spend time using the calculator when the same thing can be done more easily through manual graphing on the board. He sees the calculator as a tool to introduce or reinforce a concept. He likes using the calculator to help students understand a concept and further prove points in geometry. Once the lesson is planned, he runs through it quickly to see if it works and where modifications need to be made. But he recommends spontaneity with the calculator. When a calculator lesson was not successful, he would scrap it and move on, but at other times he would use the calculator impromptu if he saw the opportunity.

The teacher believes anyone who teaches with graphing calculators must understand the calculator is a tool and not a “solve-all” technology. He believes teachers must have a basic comfort with the technology and be willing to take risks – that is push buttons and figure out what happens. He also believes students can find out much more about the technology if they are given the opportunity to explore and teachers need to be willing to learn things about the calculator from students. He believes students want to show their skills and teach the teacher. It pleases him to see students genuinely interested in using the technology.

The teacher believes that students must not see the calculator as a short-cut. The unsuccessful students were the ones who did not think they had to learn the arithmetic because of the calculator. Students still have to challenge themselves to learn what is being taught and he teaches the students that the calculators do not make things simple. He constantly reinforces the lessons and makes things deeper and more difficult for students. Operationally, he believes students have to learn the basic procedures of the calculator and recommends having a quick find sheet to help students and teachers.

When asked to describe a successful lesson, the teacher describes a geometry activity where students constructed a triangle and measured the angles. By manipulating the triangle, students were able to draw conclusions about the interior and exterior angles. His students also manipulated a triangle between parallel lines and were able to conclude the height remained constant. The teacher noted that the training mainly focused on the algebra activities with very little emphasis on geometry. He created most of the geometry activities on his own through experimentation.

Activities involving the SmartView presentation software did not work well. The teacher complains the calculator-computer connection was much too slow to be of use. He prefers to use the overhead ViewScreen instead. He also finds it difficult to draw with the Cabri Jr. program, but believes the students could learn to be more adept at the drawing. He would like a type of game that will encourage students to draw on the calculator and explore the Cabri Jr. application.

The teacher did not believe the training was sufficient. The training did not have enough activities that pertained to geometry and the algebra activities presented seemed too complex for teachers who were techno-phobic. He did not think the activities presented in the books directly pertained to the classroom and wanted a training that applied directly to the classroom. He would have liked more training on how to create and teach students the creation of programs on the calculator. Although he had developed good activities for the classroom, he was unable to use all of them and would have liked support in implementation and would have preferred more guidance on integrating activities in the classroom.

Control (Teacher ID #69)

Teacher ID #69 is a control teacher, but uses TI-83 graphing calculators regularly in his classroom. He has about two years of experience teaching with a graphing calculator and reports an understanding of the calculator's capabilities from his college math and engineering classes. He uses the CD-ROMs and literature that come with the TI-83s to figure functions out as needed for lessons, but has not had formal training. He has attended district sponsored professional developments that focus on specific features, such as using lists and recursive sequences. He views the calculators as a great tool to bridge into higher level activities, but only uses the calculators as the material dictates.

The teacher does not plan a lesson with the calculator specifically in mind. Use of the calculator largely depends on the topic. He sees the calculator as a higher order tool and will use the calculator to reinforce a concept, or as a secondary approach, after the goal of the lesson has already been established. He tries to blend the calculator in as often as possible, but finds it difficult to use the calculator for certain topics or conceptual lessons. However when planning a lesson, the teacher believes timing and pacing is critical to breaking down a lesson for students. He also tries to use partner sharing and cooperative learning as much as possible.

Operationally, the teacher believes that teachers should know all the calculator functions and features. The teacher also has to keep in mind that different features can present in different ways for a given lesson. He also believes teacher should use more of the display options that come with calculators. This allows teachers to see what options are available to the teacher. For implementing the technology, teachers must be able to make quick changes and troubleshoot error messages. He recommends knowing what features are possible so he can present and enhance lessons as needed.

To be successful with the calculator, the teacher believes that the students must understand the basic use of the calculator as a tool, such as inputting data. Many of his ELL students have yet to be exposed to the language of mathematics and would need to be taught symbol meaning and usage prior to using a calculator, such as the difference between a negative and minus sign. He also believes students could benefit from cross-collaboration with students in other classes.

When asked to describe a successful lesson, the teacher cites a lesson where paired students graphed on a calculator and dry erase board simultaneously. The students manipulated and analyzed different families of graphs on the calculator and dry-erase board. Students also graphed systems of equations and used the window and trace function on the calculators to further visualize the solutions and understand more advanced concepts.

When asked about a lesson that did not work well with the students, the teacher cites activities in geometry and the lack of scaffolds for the calculator portion of the activity. Many of his students did not know how to correctly perform a calculator function or the correct way to input data into a calculator.

The teacher does not believe he has had enough training to teach with the graphing calculators. He recognizes that he is unfamiliar with and does not use a number of calculator features. His district trainings were not specific to lessons and he does not believe the district is the best option for trainings.

Summary of Implementation Results

Common themes that arose during the overall study are reinforced by the focused interviews we conducted with the case profile teachers. We noted that the teachers in the *GC* and *GC+Nav* group that implemented the technologies shared the following common themes:

- Previous experiences using calculator-based technologies in the classroom, either as a teacher or as a student

- Are well versed in both functionality and troubleshooting of the failure modes
- Have established routines for students to follow which include attention to providing students with activities that facilitate learning keystrokes and functionality
- Use the technologies to focus on the value-added type activities; capitalize on the multiple representations, visualization capabilities, and allow students to probe deeper into a concept having already established the concept in more traditional ways

Most teachers used activities they had done successfully in previous years. Teacher tended not to venture beyond what they felt comfortable doing. While a variety of activities were presented during the trainings, it was the familiar activities that were implemented or activities teachers self-created and tailored to a particular lesson. Teachers expressed interest in having ready-made activities that could be easily dropped into a lesson.

Teachers also expressed concern that calculators did not provide students with the opportunity to learn the arithmetic. Teachers wanted to see students use the calculator as a tool for learning and not a computational device. While researchers observed students primarily using the calculators for computation without fully exploiting the calculator's robust capabilities, teachers report a focus on exploratory activities.

Teachers expressed time constraints as a significant problem. In an already busy day, teachers did not believe they had time to effectively learn and introduce the technology. Many teachers believed with more time they would be able to better implement the technology. Additionally, teachers report that the current training was insufficient to provide them with the tools to expand beyond already familiar activities.

In summary, the issues we faced in measuring the level of implementation prevented us from formally rating the implementation integrity of the *GC* and *GC+Nav* classrooms and employ anything other than a dichotomous measure (using vs. non-using) to characterize implementation. Additionally we note that the actual length of implementation for the *GC* group was approximately 19 weeks and approximately 10 weeks for the *GC+Nav*group out of a possible 30 week instructional year. The shorten length of time for the implementation is due to two main contributions, loss due to class roster settling (impacts the *GC* group only) and loss due to CST preparation and administration. Teachers are reluctant to invest the time to teach the students how to use the technology when class rosters are still fluctuating for the approximately 6 weeks at the beginning of the school year. Mid-way through April, teachers turn their attention to preparing the students for CSTs and consequently suspend "normal" instruction. All of the quantitative results should be viewed with regard to these shorten times of implementation. Recall also that *GC+ Nav* implementation was begun mid-year.

For year 2, we plan to work more closely with the teachers to help them implement the technology and to help us formulate reasonable measures. By using our modified observation tool we hope to capture more information regarding both the nature of the implementation and additional indicators so that we can begin to address the issue of "quality" of implementation and provide some guidance as to emerging accomplished practices.

Overview of Quantitative Results

In all cases, our analysis of the quantitative results takes the same form. We first provide the descriptive statistics in the raw form, that is, with no statistical adjustments. These descriptive statistics include the number of students, classes and teachers. We then compute the average impact with and without adjusting for the prior score. These results are in standard deviation units (so the results are labeled an 'effect size') and indicate whether there is an overall difference between the treatment and control groups. We then present the results of statistical models where we estimate whether the impact of the intervention depends on the level of certain moderator variables. Below the descriptives for instance, we show the results of a model that tests whether there is a differential impact across the prior score scale. That is, we test for the interaction of treatment with the prior score. The fixed factor part of the table provides estimates of the factors of interest, in particular, whether being in a *GC* or a control class makes a difference for the average student. At the bottom of the table we give results for technical review – these often consist of random effects estimates which are added to the analysis to account for the fact that the individual results that come from a common upper-level unit (e.g., class or teacher) tend to be similar (i.e., the observations are dependent.) In some cases, to account for these dependencies, we model fixed rather than random effects but don't present the individual fixed effects estimates. Modeling the dependencies results in a more conservative estimate of the treatment impact. We note that the number of cases used to compute the effect size will often be larger than the number used in the mixed model analysis because to be included in the latter analysis a student has to have both a pretest and a posttest score.

Unlike in the attrition section, where we were interested in students who remained in or left the experiment, our analyses are based on students who completed the intervention. We removed those who switched conditions or who dropped out of the intervention. As shown in table 6, there is a total of 953 students (out of 1490) who remained in the experiment through the end. Out of those, 615 have an NWEA pretest score and 883 have a CST posttest score. We break this sample down into Algebra and Geometry subgroups in the following analyses.

We address the impact on Algebra I and then the impact on Geometry outcomes. Within each content area we consider the results for the NWEA outcome measure and then the California Standards Test (CST) outcome. For each outcome measure, when pretest scores are available, we provide a statistical analysis of the impact of *GC* controlling for pretest and examine the interaction of *GC* with pretest, that is, we examine whether students initially scoring higher or lower on the pretest differentially benefited from *GC*. We then examine the influence of English proficiency as a potential moderator of the impact of *GC*.

Impact of *GC* on Algebra I

NWEA Outcomes

Loss of cases led to the availability of scores for only a single teacher for NWEA outcomes in Algebra. With only one teacher in the control group, we are unable to estimate separately the teacher and treatment effects. We therefore do not present these results.

CST Outcomes

Figure 17 summarizes the counts of students we use in the analysis of Algebra CST Outcomes.

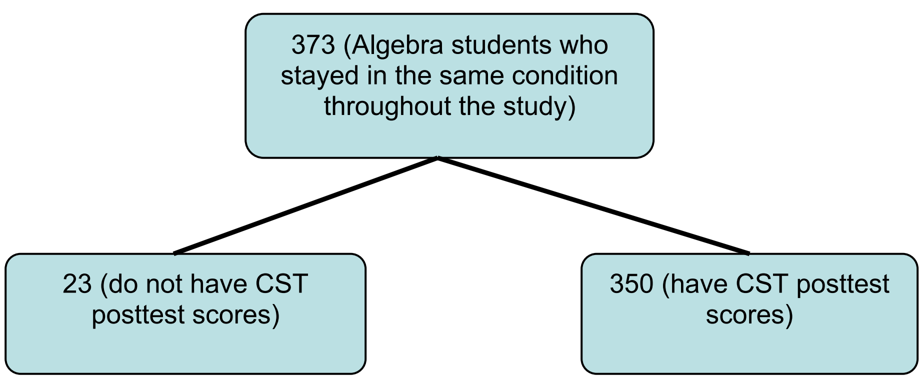

Figure 17. Counts of Algebra I Students Having CST Posttests

Analysis

Table 45 shows the estimated impact of *GC* on students' performance on the CST test of Algebra. The unadjusted effect size is -0.05 and is calculated using scores for students who took the posttest**.** We do not report the adjusted effect size since there are no pretest scores available for the control group. We see that students in the treatment group on average perform 60.78 points higher than in the control group. For this model we treated teachers as a fixed factor, so the generalizability of this result to other samples of teachers depends on non-statistical arguments about the similarity of cases.

Table 45. Estimated Impact of *GC* on CST Algebra Outcomes

Descriptive statistics	Raw group means	Standard deviation	No. of students	No. of classes	No. of teachers	Unadjusted effect size
GC	288.33	48.50	301	3	2	-0.05
Control	290.37	36.98	49	16	6	
Fixed factors	**Estimate of coefficient**		**Standard error**	**DF**	***t* value**	***p* value**
Intercept	237		45.67	342	5.19	<.01
Condition (*GC* = 1; control = 0)	60.78		45.82	342	1.33	.19

349 students were used to compute the estimates of the Fixed factors in the table above (one case was an influential point and was removed.) The number of teachers for the reduced sample of 349 students is 8 (2 control and 6 *GC*). The number of classes is 19 (3 control and 16 *GC*.) Teachers are modeled as a fixed factor.

Analysis Including English Proficiency as a Moderator

Table 46 reports the statistical analysis of the moderating effect of English proficiency on the Algebra outcome as measured by the CST test of Algebra. Here we were interested in whether *GC* was more or less effective for English proficient compared to non-proficient students.

Table 46. Moderating Effect of English Proficiency on CST Algebra Outcomes

Mixed model: Fixed factors	Estimate of coefficient	Standard error	DF	*t* value	*p* value
Intercept	5.47	0.14	339	38.30	<.01
Condition (*GC* = 1; control = 0)	0.14	0.14	339	0.96	<.34
English proficiency (proficient = 1; not proficient = 0)	0.10	0.05	339	2.14	<.03
English proficiency by condition interaction	<-0.01	0.05	339	-0.09	.93

349 students were used to compute the results presented in the table above (one case was an influential point and was removed.) The number of teachers for the reduced sample of 349 students is 8 (2 control and 6 *GC*). The number of classes is 19 (3 control and 16 *GC*.) Teachers are modeled as a fixed factor.

The coefficient associated with the English proficiency by condition interaction is <-.01 with a *p* value of .93. We have no confidence that English proficiency moderates the treatment impact.

Impact of *GC* on Geometry

We also examined the impact of *GC* on student performance in Geometry. We assessed student performance in Geometry using the NWEA test of Geometry and the CST. The statistical analyses take the same form as those for Algebra I.

NWEA Outcomes

Figure 18 summarizes the counts of students we use in the analysis of Geometry NWEA Outcomes.

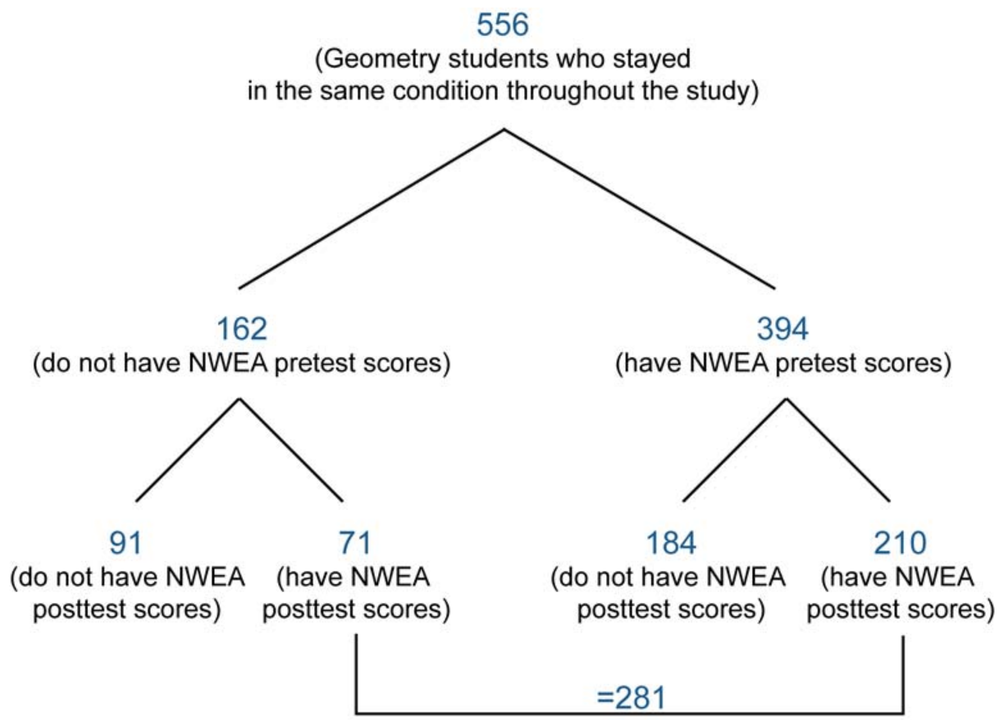

Figure 18. Counts of Geometry Students Having Both NWEA Pre- and Posttests

Analysis Including Pretest

The NWEA outcomes in Geometry at San Diego were problematic for several reasons. First, all except two teachers administered the paper-and-pencil form of the NWEA test. The remaining two teachers administered the computerized adaptive version of the test. The paper-and-pencil version of the test and the computerized adaptive version have different properties, with the paper-and-pencil version exhibiting a noticeable floor effect. When we included the results for all students in one analysis, the score range for students taking the computerized adaptive test extended beyond the floor cut point for the rest of the students. Essentially we had two score distributions superimposed on each other. Putting them into one analysis greatly distorted the results.

We eliminated from analysis two teachers, both in the control condition, who administered the computerized adaptive version of the test. We recommend that the results of analyses involving the NWEA posttest be interpreted with caution because we have selectively removed cases from one of the conditions to deal with the problem described above.

For the control cases that were removed, scores were distributed along the range of the prior score scale. Removing them leads to a distribution of controls that is more highly clustered in the lower range of the pretest scale than the distribution for treatment - that is, it exacerbates the chance imbalance between conditions on the pretest. As a results of this, we limit our inferences to the upper-end of the pretest scale (i.e., to the higher-performing students) where we have greater overlap of treatment and control cases.

Table 47. Estimated Impact of *GC* on NWEA Geometry Outcomes

Descriptive statistics	Raw group means	Standard deviation	No. of students	No. of classes	No. of teachers	Unadjusted effect size
GC	253.46	7.48	150	8	4	0.77
Control	248.24	5.94	131	9	4	
Mixed model: Fixed factors	**Estimate of coefficient**		**Standard error**	**DF**	***T* value**	***p* value**
Intercept	250.02		2.01	4	124.50	<.01
Pretest score (centered at the mean)	0.23		0.04	201	5.78	<.01
Condition (*GC* = 1; control = 0)	1.87		2.72	4	0.69	.53
Pretest score by condition interaction	0.19		0.06	201	3.21	<.01
Mixed model: Technical details for random components	**Estimate of variance component**		**Standard error**		**Z value**	**p value**
Teacher mean achievement	8.28		7.07		1.17	.12
Within teacher variation	24.66		2.45		10.05	<.01

Note 281 students are included in the calculation of the unadjusted effect size. 210 students had both pretest and posttest scores. We removed 1 because he/she was an influential point or outlier (where the determination of influential points was based on the model above.) The remaining 209 students (86 in the control group, and 123 in the treatment group) were used to compute the adjusted effect size and the results presented in the table above. The number of teachers for the reduced sample of 209 students is 6 (2 control and 4 *GC*). The number of classes is 13 (5 control and 8 *GC*.)

Table 47 shows the estimated impact of *GC* on students' performance on the NWEA geometry test. The unadjusted effect size is 0.77. The effect size when adjusted for the pretest is 0.16 with a *p* value of .96. We have no confidence that the observed average difference is due to something other than chance. (However, we remind the reader that with the small teacher sample, the power to detect small but potentially meaningful effects is limited.)

For a student with an average score on the pretest, there is roughly a 1.87-point advantage to being in the *GC* group; that is, we predict that an average student would score 1.87 points higher on the outcome measure if he or she is in a *GC* class instead of a control class. The p-value for this estimate also gives us no confidence that there is a real difference between conditions for a student at this level of the pretest.

The value for the interaction between *GC* and the prior score is 0.19 with a *p* value of <.01.

As a visual representation of the results described in Table 47, we present a scatterplot in Figure 19, which shows student end-of-year performance in Geometry, as measured by the NWEA test, against their performance on the NWEA general math test in the Fall. The two lines are the predicted values on the posttest for students in the *GC* and control conditions as determined using the estimated fixed effects in the model.

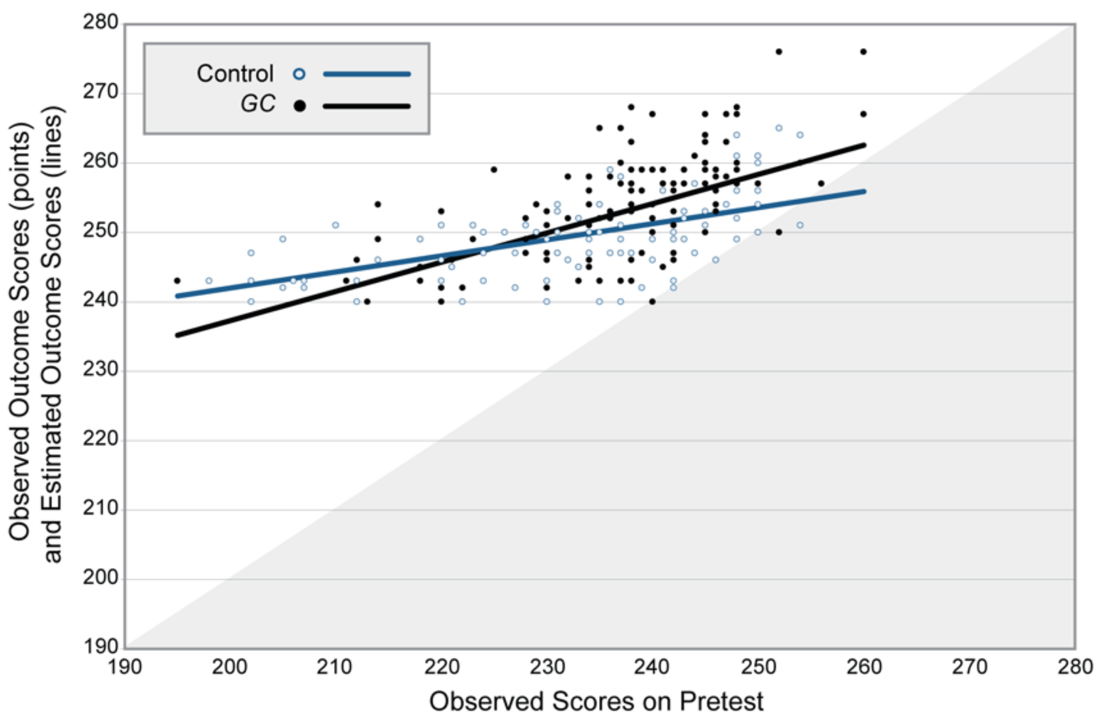

Figure 19. Comparison of Predicted and Actual NWEA Geometry Outcomes for *GC* and Control Group Students

As a visual representation of this result, Figure 20 shows the predicted difference between the *GC* and control groups for different points along the prior score scale. Consistent with the results in Table 47, there is evidence of a differential impact of *GC* on geometry performance as measured by the NWEA test. Specifically, along the prior score scale, we see that *GC* has a positive impact for students performing at the high end of the prior score scale. (Given that there are few treatment cases at the lower end of the prior score scale, we cannot draw inferences about what happens in that range.) (A limitation of the current model is that we do not figure in the natural variation in slopes across teachers (i.e., the variation we would see without the treatment effect.) If this variation is high, then with a small sample of teachers, chance may easily lead to an imbalance in slopes between the conditions. In the current model, the standard error for the interaction assumes no such variation.)

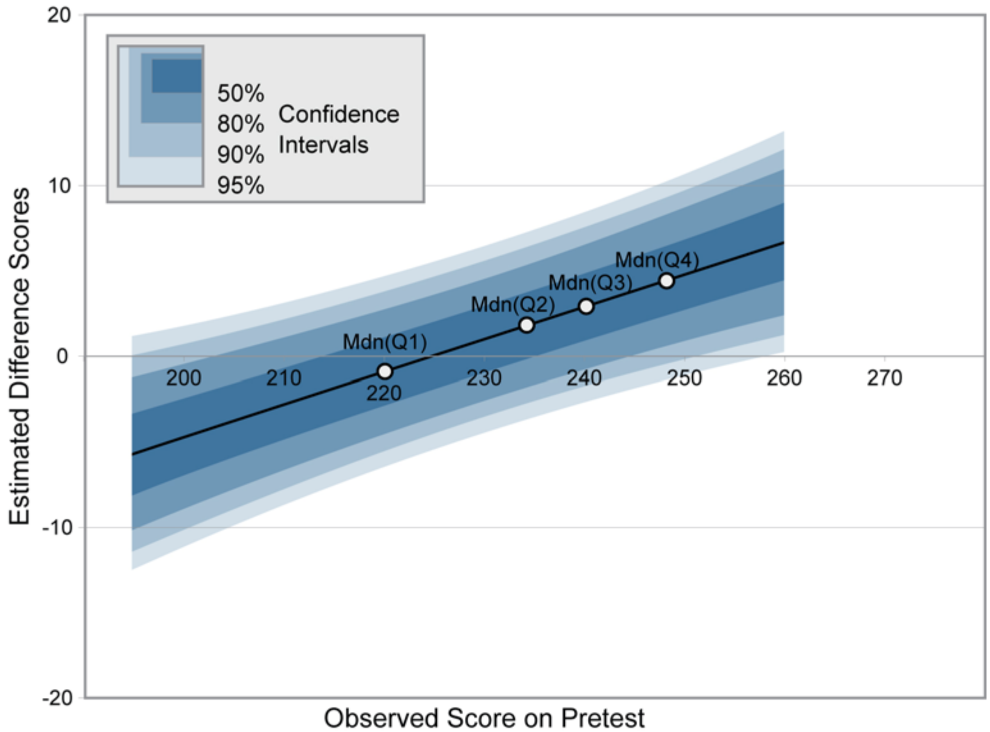

Figure 20. Differences Between *GC* and Control Group NWEA Geometry Outcomes: Median Students for Four Quartiles Shown

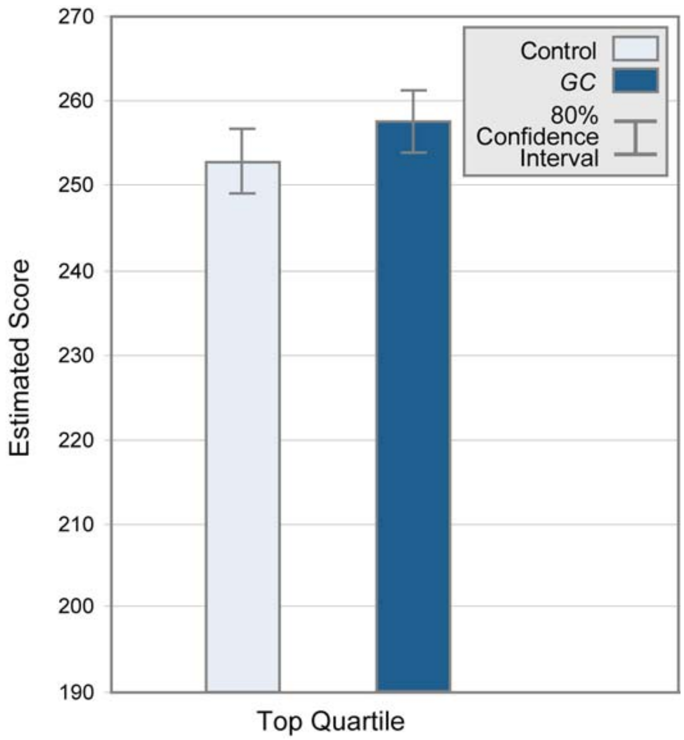

Figure 21. Difference Between *GC* and Control Group NWEA Geometry Outcomes: Median Students in Top Quartile

Figure 21 presents the same information represented in Figure 20 but this time in the form of a bar graph showing the predicted difference between GC and control conditions for students at the median of the fourth quartile of the pretest measure. We are 80% sure that the true difference between conditions would place the tops of the bars simultaneously within the confidence interval markers. There is no overlap in the 80% confidence intervals which is consistent with the observation that treatment benefits higher-performing students, but this result should be interpreted keeping in mind the caveat in the previous paragraph.

Analysis Including English Proficiency as a Moderator

Table 48 shows the estimated moderating effect of English proficiency as measured by the NWEA test of Geometry. Here we were interested in whether *GC* was more or less effective for English proficient compared to non-proficient students.

The coefficient associated with the English proficiency by condition interaction is 1.76 with a *p* value of .60. We have no confidence that English proficiency moderates the treatment impact.

Table 48. Moderating Effect of English Proficiency on NWEA Geometry Outcomes

Mixed model: Fixed factors	Estimate of coefficient	Standard error	DF	*t* value	*p* value
Intercept	252.85	2.63	4	96.02	<.01
Pretest score (centered at the mean)	0.33	0.03	197	10.71	<.01
Condition (GC = 1; control = 0)	-0.45	4.02	4	-0.11	0.92
English proficiency status (proficient = 1; not proficient = 0)	-2.53	2.09	197	-1.21	0.23
English proficiency status by condition interaction	1.76	3.36	197	0.52	.60
Mixed model: Technical details for random components	**Estimate of variance component**	**Standard error**		***z* value**	***p* value**
Teacher mean achievement	10.67	10.11		1.06	.15
Within teacher mean variation	25.95	2.60		9.97	<.01

The sample used for this analysis is identical to the one reported in Table 47.

CST Outcomes

Figure 22 summarizes the counts of students we use in the analysis of Geometry CST Outcomes.

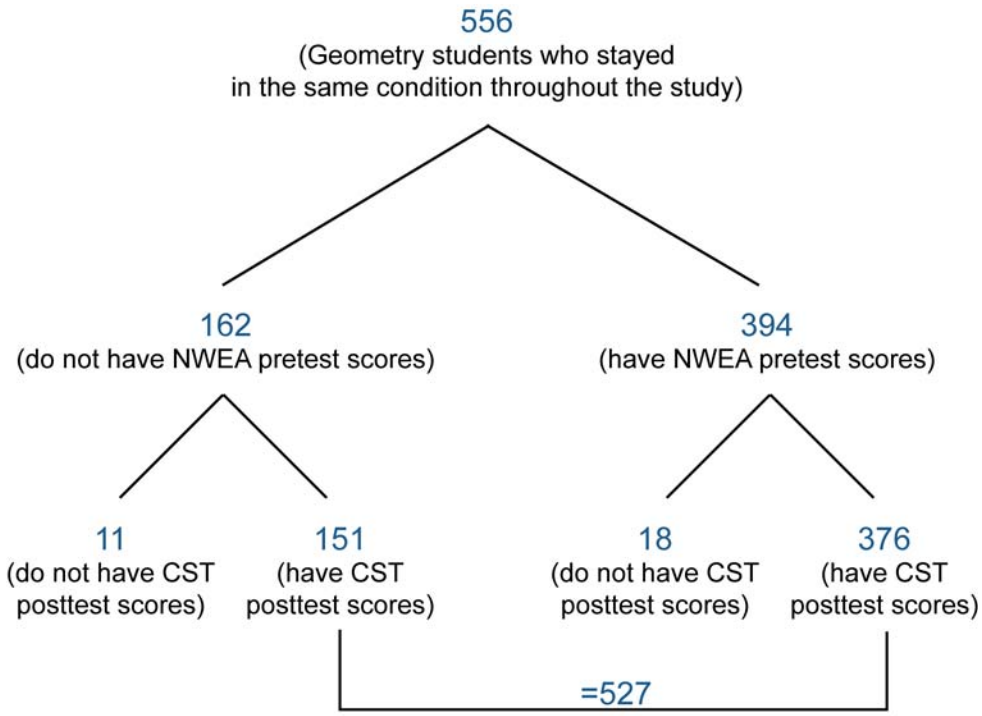

Figure 22. Counts of Geometry Students Having Both NWEA Pretests and CST Posttests

Analysis Including Pretest

Table 49. Estimated Impact of *GC* on CST Geometry Outcomes

Descriptive statistics	Raw group means	Standard deviation	No. of students	No. of classes	No. of teachers	Unadjusted effect size
GC	336.17	67.69	184	8	4	0.86
Control	286.17	46.55	343	16	5	
Mixed model: Fixed factors	**Estimate of coefficient**		**Standard error**	**DF**	***t* value**	***p* value**
Intercept	298.59		10.25	6	29.13	<.01
Pretest score (centered at the mean)	2.45		0.20	365	12.09	<.01
Condition (*GC* = 1; control = 0)	0.59		16.33	6	0.04	.97
Pretest score by condition interaction	1.63		0.37	365	4.35	<.01
Mixed model: Technical details for random components	**Estimate of variance component**		**Standard error**		***Z* value**	***p* value**
Teacher mean achievement	390.97		272.44		1.44	.08
Within teacher variation	1494.41		110.43		13.53	<.01

Note: 527 students had a posttest and were included in the calculation of the unadjusted effect size. 376 students had pretest and posttest scores. Of these we removed 1 because he/she was an influential point or outlier (where the determination of influential points was based on the model above.) The remaining 375 students are used in the analysis (230 controls, 145 treatment cases.) The number of classes is 24 (16 control and 8 treatment.) The number of teachers is 9 (5 control and 4 treatment.)

Table 49 shows the estimated average impact of *GC* on students' performance on the CST test of geometry. The effect size is 0.86. The effect size when adjusted for the pretest is 0.52. The *p* value for this effect is .98, meaning that we have no confidence that the estimated average difference is due to something other than chance.

The estimated impact for a student with an average pretest score is 0.59 with a *p* value of .97. We have no confidence that the effect is different from zero at this point along the pretest scale.

The value for the interaction between *GC* and the prior score is 1.63 with a *p* value of <.01.

As a visual representation of the results described in Table 49, we present a scatterplot in Figure 23, which shows student end-of-year performance in geometry, as measured by the CST test, against their performance on the NWEA general math test in the fall. The cross over of the lines for *GC* and control is a representation of the interaction.

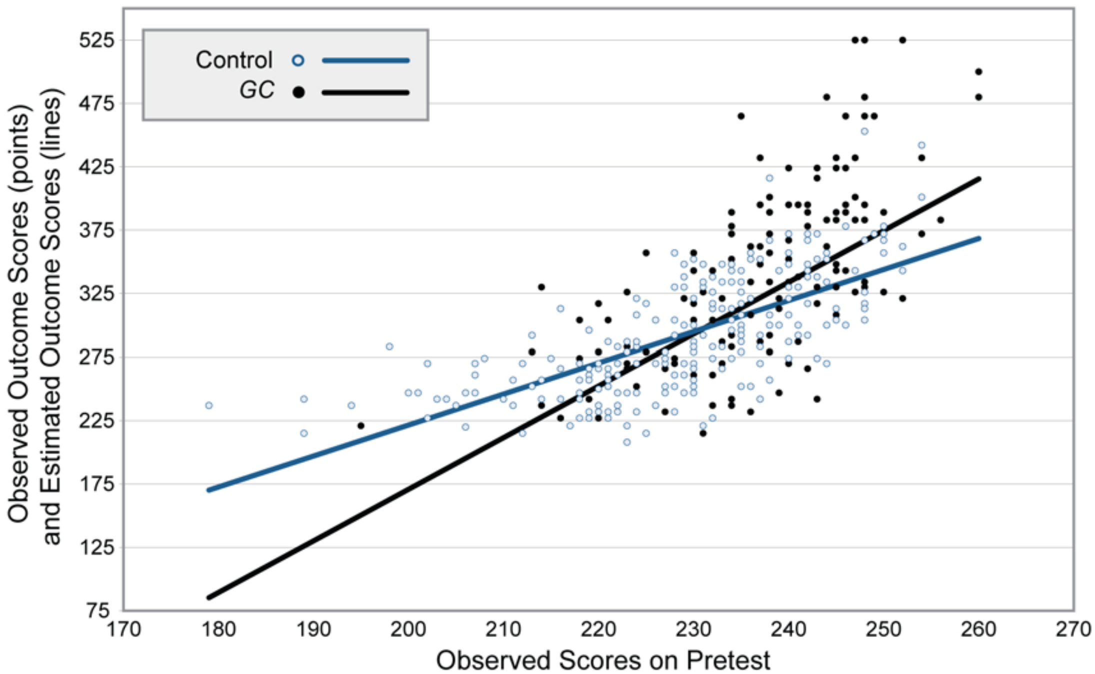

Figure 23. Comparison of Predicted and Actual CST Geometry Outcomes for *GC* and Control Group Students

Another way to visualize this interaction is provided in Figure 24, which shows the predicted difference between the *GC* and control groups for different points along the prior score scale. Consistent with the results in Table 49, there is evidence of a differential impact of TI on geometry performance as measured by the CST test. Specifically, along the prior score scale, we see that TI has a positive impact for students performing at the high end of the prior score scale. (Given that there are few treatment cases at the lower end of the prior score scale, we cannot draw inferences about what happens in that range; also, we remind the reader that we do not model the ambient variation in slopes among teachers, therefore, with only nine teachers, there may be a chance imbalance in slopes between the conditions.)

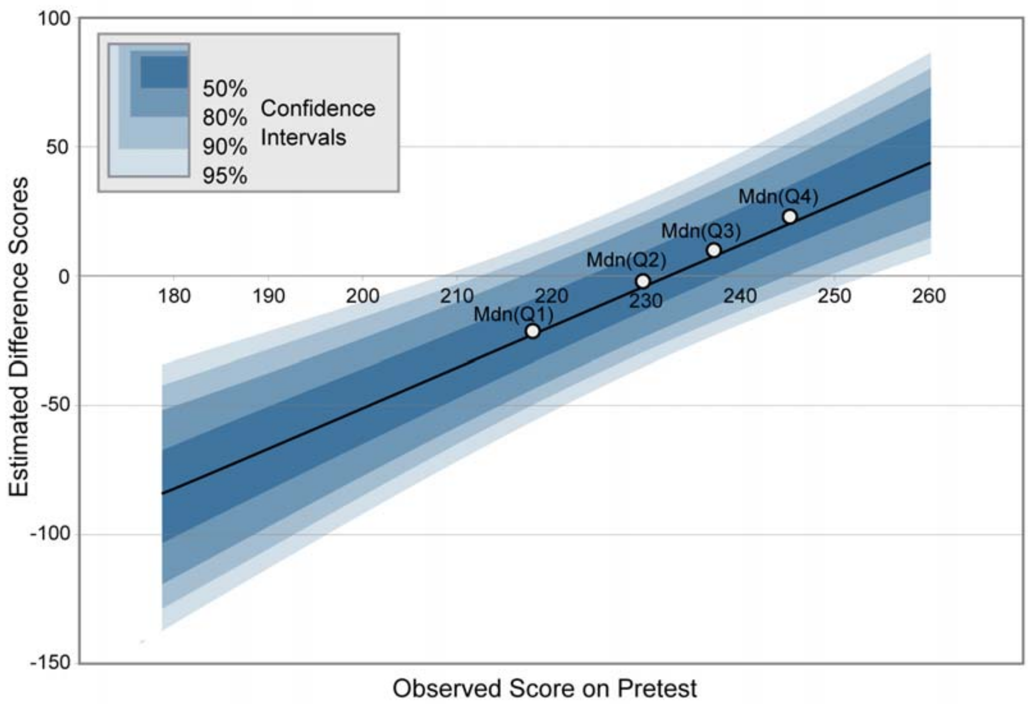

Figure 24. Differences between *GC* and Control Group CST Geometry Outcomes: Median Students for Four Quartiles Shown

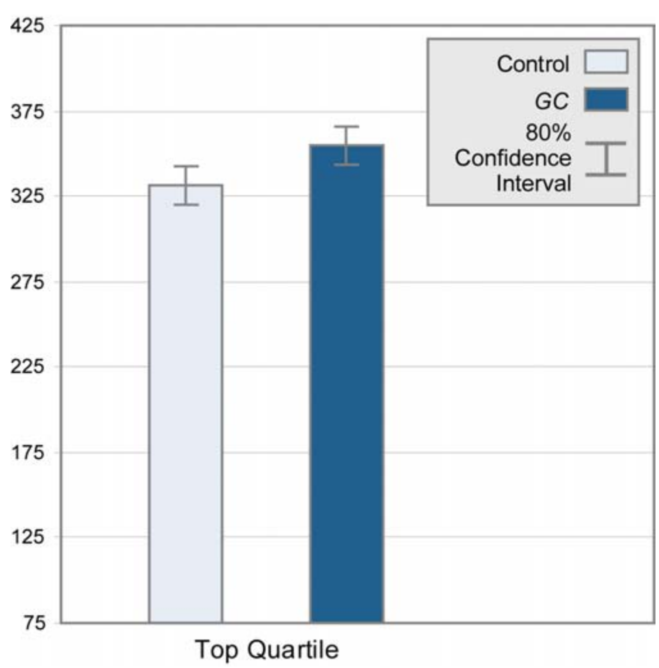

Figure 25 presents the same information represented in Figure 24 but this time in the form of a bar graph showing the predicted difference between *GC* and control conditions for students at the median of the fourth quartile of the pretest measure. We are 80% sure that the true difference between conditions would place the tops of the bars simultaneously within the confidence interval markers. There is limited overlap in the confidence intervals meaning that there is evidence of a differential impact at the high end of the prior score scale.

Figure 25. Differences between *GC* and Control Group CST Geometry Outcomes: Median Pretest Scores in Top Quartile

Analysis Including English Proficiency as a Moderator

Table 50 shows the estimated moderating effect of English proficiency as measured by the CST test of Geometry. Here we were interested in whether treatment was more or less effective for English proficient compared to non-proficient students.

The coefficient associated with the English proficiency by condition interaction is 37.29. This effect estimates the difference in impact between English-learners and non-English-learners. The positive moderating effect of being an English-language-learner has a *p* value of .01. This means that we expect a result as large or larger than the absolute value of the one observed 1% of the time when in fact there is no moderating effect. By our standards, we are very confident that this effect is systematic, assuming variation in this effect due to teacher effects is minimal. (As before, we caution that the result may be at least in part due to a chance imbalance on teacher effects.)

Table 50. Moderating Effect of English Proficiency on CST Geometry Outcomes

Mixed model: Fixed factors	Estimate of coefficient	Standard error	DF	*t* value	*p* value
Intercept	309.07	12.56	6	24.60	<.01
Pretest score (centered at the mean)	2.96	0.18	364	16.26	<.01
Condition (GC = 1; control = 0)	-30.20	21.48	6	-1.41	.21
English proficiency status (proficient = 1; not proficient = 0)	-11.32	7.55	364	-1.50	.13
English proficiency status by condition interaction	37.29	15.09	364	2.47	.01
Mixed model: Technical details for random components	**Estimate of variance component**	**Standard error**		***z* value**	***p* value**
Teacher mean achievement	450.70	323.14		1.39	.08
Within teacher mean variation	1546.60	114.51		13.51	<.01

Note. The sample used for this analysis is identical to the one reported in Table 49.

The bar graph in Figure 26 shows how to interpret the interaction: The low English proficient students do better in the control condition than in the treatment condition, whereas there is virtually no difference in between treatment and control students who are English-fluent.

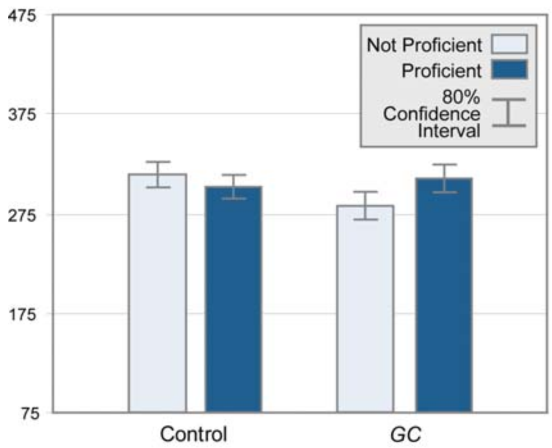

Figure 26. Moderating Effect of English Proficiency on CST Geometry Outcomes

Discussion of San Diego Results

Our randomized experiment in San Diego City Schools provides some evidence of the impact of a program using graphing calculators and the TI-Navigator classroom networking system for Algebra I and Geometry in the context of the implementation in this district. The research was weakened by difficulties in obtaining test results, especially for the pretest. However, in the absence of any reason to believe that the testing difficulties were related to the experimental conditions, we proceeded with an analysis of the outcomes for the *GC* condition only. We have deferred the impact of TI-Navigator (the *GC+Nav* condition) to our report on the analysis where we combine outcomes from two sites.

The results of our experiment were modest. For the Algebra I classes, we did not find a significant average impact of the intervention. Also we did not find that the program was more- or less-beneficial for English language learners as compared to English proficient students. For Algebra I students we were limited to using the CST outcome.

For the Geometry classes, the results based on the NWEA test outcomes showed no average impact of *GC,* but for students at the upper end of the pretest scale, those in the *GC* condition outperformed their counterparts in the control condition. There was no moderating effect of English proficiency status on the treatment effect. The validity of inferences concerning this outcome measure was limited by the selective removal of two control teachers. For the CST outcome, the pretest also moderated the treatment effect in the same direction as when we analyzed the outcomes using the NWEA test as the posttest measure. Different from the NWEA outcome, the interaction between English proficiency and treatment was statistically significant, mainly due to the fact that Low English Proficient students performed worse in the treatment condition than in the control condition.

All of these analyses and results are limited by the small sample of teachers. Both the average effect of treatment and the interaction effects involving treatment may be highly sensitive to the sample used and could change significantly upon re-sampling. Under these conditions only relatively large effects can be detected with any degree of confidence. In our report of the combined analysis from this and another district, we will be able to pool samples from both locations.

We did not investigate whether the impact was different depending on teacher years of experience because almost all the teachers were veterans with certifications designating them as highly qualified to teach their subjects. We have also deferred the analysis of implementation differences to our report on the analysis that combines the two sites. The larger number of teachers in the combined analysis allows us to detect the impacts of variables associated with teachers and their respective student clusters.

The implementation of the program met with several challenges both technical and pedagogical. These problems are typical of field settings where training and support activities face realistic constraints. We found considerable evidence, detailed in Appendixes A and B, that the *GC* condition was differentiated from the control condition so to some extent challenges were addressed. Our observations of implementation weaknesses serve both to inform the improvement of the Texas Instruments' technologies and services and to caution districts that may be considering implementing a program like this that existing resources are critical to successfully scale-up and impact usage.

It is important, finally, to emphasize that we are testing an implementation of a specific program in one particular school district. Our experiment cannot be used as definitive evidence of the value of graphing calculators but, as a randomized experiment, it does provide a valuable point of reference as researchers accumulate more evidence. As a local experiment to inform decisions in SDCS, we believe it supports the continued implementation and improvement of the program as is occurring in the second phase of the research now in operation.

Combined Results

Formation of the Experimental Groups

Groups as Initially Randomized

The randomization process guarantees that there is no intentional or unintentional bias in the selection of teachers and students into the treatment or the control condition. It does not, however, guarantee that the groups will be perfectly matched. It is important to inspect the two groups to determine whether, in spite of randomization, there are any significant differences on factors that affect the outcome[2]. Table 51 through Table 53 address the nature of the experimental groups; they show the distribution of teachers, classes, grades, and students between *GC* and control conditions. This is the complete number of students in the experiment at the time that the experiment began in the fall of 2005.

Table 51. Students in *GC* and Control Groups

		Initially randomized	Students with an NWEA pretest with/ without a post test score[a]	Completed program	Students with an NWEA pretest with/ without a CST posttest[b]	Students with a CST posttest with/ without an NWEA pretest[b]	Students with both an NWEA pretest and a CST posttest[b]
Algebra	***GC***	867	392	594	314	544	305
	Control	558	275	422	223	398	218
	Totals	1425	667	1016	537	942	523
Geometry	***GC***	811	348	536	287	484	277
	Control	779	521	615	430	554	401
	Totals	1590	869	1151	717	1038	678
Totals		**3015**	**1536**	**2167**	**1254**	**1980**	**1201**

[a] Students who were initially randomized

[b] Students who completed the program

[2] In technical terms, randomization ensures lack of bias, but we are interested in knowing whether the particular estimate resulting from this randomization may be far from the true value as a result of chance imbalances on factors that affect the outcome

Table 52. Teachers in *GC* and Control Groups

		Initially randomized	Completed program	Teachers with at least one student with NWEA pretest with/without a CST posttest	Teachers with at least one student with CST posttest with/without an NWEA pretest	Teachers with at least one student with NWEA pretest and CST posttest
Algebra	***GC***	14	14	10	14	10
	Control	11	10	9	10	9
Geometry	***GC***	10	9	7	9	7
	Control	9	9	7	8	7
Totals		**44**[a]	**42**	**33**	**41**	**33**

[a] Two teachers taught both Algebra and Geometry subjects.

Table 53. Classes in *GC* and Control Groups

		Initially randomized	Completed program	Classes with at least one student with NWEA pretest with/without a CST posttest	Classes with at least one student with CST posttest with/without an NWEA pretest	Classes with at least one student with NWEA pretest and CST posttest
Algebra	***GC***	36	35	25	33	24
	Control	26	23	20	23	20
Geometry	***GC***	27	23	15	22	15
	Control	25	25	21	24	21
Totals		**114**	**106**	**81**	**102**	**80**

Post Randomization Composition of the Experimental Groups

In checking for balance in the composition of the experimental groups, we examine teacher experience first, followed by student level variable English proficiency and pretest outcomes.

Teaching Experience

During the randomization process teachers identified themselves according to years of teaching experience. The initial teacher pair was formed correspondingly so that teaching experience would be balanced between the groups. Table 54 summarizes this information. The statistical test is meant to confirm whether we were successful in pairing according to this criterion.

Table 54. Years of Teaching Experience

	Number of teachers		
Condition	**0 to 3 years**	**4 or more years**	**Totals**
GC	3	15	18
Control	3	16	19
Totals	6	31	37
Statistics	**DF**	**Value**	***p* value**
Fisher's exact test	1	0.01	.94

Note. Fisher's exact test is reported because 50% of the cells have expected counts less than 10. We are missing information about years of experience from 5 teachers.

Randomization resulted in years of teaching experience being evenly balanced between *GC* and control teachers. The large *p* value of .94 is consistent with this assertion.

Student Variables

From Table 51, we see that 3015 students were enrolled in the study. Out of those, 1425 students are from Algebra classes and 1590 students are from Geometry classes.

English Proficiency Distribution

Table 55 and Table 56 show the distribution of English proficient/non-proficient students in each group.

Algebra

Table 55. English Proficiency for Algebra *GC* and Control Groups

	English proficiency		
Condition	**No**	**Yes**	**Totals**
GC	250	616	**866**
Control	220	338	**558**
Totals	**470**	**954**	**1424**
Statistics	**DF**	**Value**	***p* value**
Chi-Square	1	17.11	<.01

Note. Information about English proficiency is missing for 1 student.

We see that the number of students per grade was not distributed evenly between the conditions in spite of randomization. There are proportionally less students who are proficient in the control group than in the *GC* group. Chi-square tests confirm that this characteristic was not balanced between conditions. The imbalance may lead the estimate of the impact to depart from its true value.

Geometry

We also see that the number of students in English proficient/non-proficient was not distributed evenly between the conditions in spite of randomization. There are proportionally less students who are non English proficient in the control group than in the *GC* group. Chi-square tests confirm that this characteristic was not balanced between conditions. The imbalance may lead the estimate of the impact to depart from its true value.

Table 56. English Proficiency for Geometry *GC* and Control Groups

	English proficiency		
Condition	**No**	**Yes**	**Totals**
GC	136	675	**811**
Control	96	683	**779**
Totals	**232**	**1358**	**1590**
Statistics	**DF**	**Value**	***p* value**
Chi-Square Test	1	6.3	.01

Characteristics of the Experimental Groups as Defined by Pretest

We also checked whether randomization resulted in a balance on pretest scores, a variable that we include in most of our analyses to increase the precision of our estimates. As shown in Table 51, the total number of students with pretest scores was much less than the anticipated 3015. Only 1536 students received a NWEA pre test score. The difficulties in implementation of the pretests are detailed in the individual reports. The following is a test of difference between *GC* and control group on the subject of Algebra and Geometry.

Algebra

For Algebra, the attrition rate is 53.2%. We account for the attrition in pretest scores as reported by NWEA (scoring agency). We are left with a sample size of 667 students with pretest scores in Algebra. The following analysis tests for balance in pretest scores and is based on these sample sizes.

The *GC* and control groups had slightly different average pretest scores on Algebra, as shown in Table 57. However, when we accounted for the fact that outcomes for students of the same teacher tend to be related by factoring these dependencies in the model, the *p* value increased to .33, indicating that the difference we are seeing is very likely due to chance.

Table 57. Difference in Algebra Pretest Scores between Students in the *GC* and Control Groups

Descriptive statistics: Pretest outcomes	Raw group means	Standard deviation	Number of students	Standard error	Effect size[a]
GC	221.78	14.45	392	0.73	-0.22
Control	224.94	14.58	275	0.88	
***t* test for difference between independent means**	**Difference**		**DF**	***t* value**	***p* value**
Condition (– control)	-3.16		665	2.77	.01

[a] The difference we are measuring is not an effect of treatment (the usual sense of effect size) but a result of chance differences in the randomization.

Geometry

The attrition rate for Geometry is 45.35%. We account for the attrition in pretest scores as reported by NWEA (scoring agency). We are left with a sample size of 869 students with pretest scores in Geometry. The following analysis tests for balance in pretest scores and is based on these sample sizes.

As with NWEA Algebra, the *GC* and control groups had slightly different average pretest scores in Geometry. Again, when we accounted for the fact that outcomes for students of the same teacher tend to be related by modeling these dependencies, the *p* value increased to 0.21, indicating that this difference is likely due to chance.

Table 58. Difference in Geometry Pretest Scores between Students in the *GC* and Control Groups

Descriptive statistics: Pretest outcomes	Raw group means	Standard deviation	Number of students	Standard error	Effect size[a]
GC	236.06	12.55	348	0.67	0.39
Control	230.91	13.86	521	0.61	
***t* test for difference between independent means**	**Difference**		**DF**	***t* value**	***p* value**
Condition (– control)	5.15		867	-5.57	<.01

[a] The difference we are measuring is not an effect of treatment (the usual sense of effect size) but a result of chance differences in the randomization.

Attrition after the Pretest

Based on the information above there are 1536 who has science pretest scores. However, we did not receive science posttest scores for all of these students. We will be looking at them by subject.

Algebra

Table 59 shows the attrition of enrolled students that occurred after taking the pretest. Chi-square tests confirm that this attrition of students was balanced between conditions.

We observe that 65 students (or 9.7%) are missing posttest scores due to a variety of reasons including being absent during testing or not being able to complete the test. Table 60 shows that students with no score for the posttest (having pretest scores only) scored lower on the pretests. The low *p* value confirms that high performers are overrepresented in our study sample. Thus, we can be less confident of the applicability of findings for lower scoring students.

Table 59. Availability of Pre- and Posttest Scores for Algebra Students

Condition	Having both pre- and posttest scores	Having pretest scores only	Totals
GC	352	40	**392**
Control	250	25	**275**
Totals	**602**	**65**	**667**
Statistics	**DF**	**Value**	***p* value**
Chi-square	1	0.23	.63

Table 60. Difference in Pretest Scores for Students Having Pre- and Posttest Scores Vs. Posttest Only

Descriptive statistics	Raw group means	Standard deviation	Number of students	Standard error	Effect size[a]
Having pretest scores only	220.17	14.05	65	1.74	-0.22
Having both pre- and posttest scores	223.4	14.61	602	0.6	
***t* test for difference between independent means**	**Difference**		**DF**	***t* value**	***p* value**
(Having pretest scores only) – (Having both pre- and posttest scores)	-3.23		665	1.7	.09

[a] The difference we are measuring is not an effect of treatment (the usual sense of effect size) but a result of chance differences in the randomization.

Geometry

Table 61 shows us that there are proportionally more students in the control group who were originally enrolled but did not take the CST posttest, as compared to *GC* group. The low *p* value on the Chi-square test confirms that this attrition was some balanced between conditions. The small imbalance may lead the estimate of the impact to depart from its true value.

We observe that 96 students (or 11.05%) are missing posttest scores. Table 62 shows that students with no score for the posttest (having pretest scores only) scored lower on the pretests. The low *p* value confirms that high performers are overrepresented in our study sample. Thus, we can be less confident of the applicability of findings for lower scoring students.

Table 61. Availability of Pre- and Posttest Scores for Geometry Students

Condition	Having both pre- and posttest scores	Having pretest scores only	Totals
GC	317	31	**348**
Control	456	65	**521**
Totals	**773**	**96**	**869**
Statistics	**DF**	**Value**	***p* value**
Chi-square	1	2.7	.10

Table 62. Difference in Pretest Scores for Students Having Pre- and Posttest Scores Vs. Posttest Only

Descriptive statistics	Raw group means	Standard deviation	Number of students	Standard error	Effect size[a]
Having pretest scores only	227.55	15.17	96	1.55	-0.45
Having both pre- and posttest scores	233.65	13.22	773	0.48	
***t* test for difference between independent means**	**Difference**		**DF**	***t* value**	***p* value**
(Having pretest scores only) – (Having both pre- and posttest scores)	-6.09		867	4.19	<.01

[a] The difference we are measuring is not an effect of treatment (the usual sense of effect size) but a result of chance differences in the randomization.

Implementation Results

The results of the implementation of the program in the two districts are detailed in the individual reports for the two sites.

Overview of Quantitative Results

The primary topic of our experiment was the impact of graphing calculators on student performance on the CST and NWEA tests. We will first address the impact on algebra achievement and then the impact on geometry achievement.

Unlike in the attrition section, where we were interested in students who remained in or left the experiment, our analyses are based on students who completed the intervention. We removed those who switched conditions or who dropped out of the intervention. As shown in Table 51, there is a total of 2167 students (out of 3015) who remained in the experiment through the end. Out of those, 1254 have an NWEA pretest score and 1980 have a CST posttest score. We break this sample down into Algebra and Geometry subgroups in the following analyses.

In the following sections, our analysis of the quantitative results takes the same form. Within each content area, we first estimate the average impact of the intervention on student performance. These results are presented in terms of effect sizes.

We then show the results of mixed model analyses where we estimate whether the impact of the intervention depends on the level of certain moderator variables. For instance, we show the results of a model that tests whether there is a differential impact across the prior score scale. We also model the potential moderating effects of gender and years of teaching experience. We provide a separate table of results for each of these moderator analyses. The fixed factor part of each table provides estimates of the factors of interest. For instance, in the case where we look at the moderating effect of a student's prior score, we show whether being in a *GC* or a control class makes a difference for the average student. We also show whether the impact of the intervention varies across the prior score scale. At the bottom of the table we give results for technical review – these often consist of random effects estimates which are added to the analysis to account for the fact that the individual results that come from a common upper-level unit (e.g., class or teacher) tend to be similar (i.e., the observations are dependent.) In some cases, to account for these dependencies, we model fixed rather than random effects but do not present the individual fixed effects estimates. Modeling the dependencies results in a more conservative estimate of the treatment impact.

We note that the number of cases used to compute the effect size often will be larger than the number used in the mixed model analysis because to be included in the latter analysis a student has to have both a pretest and a posttest score.

Impact of *GC* on Outcomes

Algebra CST Outcomes

Effect Size Analysis

Our next set of analyses addresses Algebra achievement as measured by CST Algebra. Table 63 provides a summary of the sample we used in the analyses and the results for the comparison of *GC* and control group performance for Algebra. This table gives information about all the students in the original sample for whom we have a posttest score. This shows the means and standard deviations as well as a count of the number of students, classes, and teachers in that group. The last column provides the effect size, which is the size of the difference between the means for *GC* and control in standard deviation units.

Table 63. Overview of Sample and Impact of *GC* on Algebra Achievement

Condition	Means	Standard deviations	No. of students	No. of classes	No. of teachers	Effect size
GC	296.34	48.03	544	33	14	-0.17
Control	304.78	52.15	398	23	10	

Analysis not Including Pretest as a Moderator

We described earlier the situation in San Diego where we did not have pretest scores for students in Algebra classes. This precluded testing any interaction involving prior score and location. Any other interactions involving pretest (such as prior score with treatment condition) would not apply to San Diego. We therefore took prior score out of this analysis.

In Table 64 we show the results of a model where we estimate separate effects for each location by including an interaction between location and treatment. From the result, we have no confidence that the true effect is different from zero at San Diego, and further, we have no confidence that the treatment effect is different between the two districts.

Table 64. Impact of *GC* on Algebra Achievement

Fixed effects	Estimate	Standard error	DF	*t* value	*p* value
Predicted value for a control student in SD	291.48	22.74	19	12.82	<.01
Impact of *GC* for a student in SD	-14.49	25.03	19	-0.58	.57
Location (1= ES; 0 = SD)	12.37	24.14	19	0.51	.61
Interaction of location and *GC*	17.07	27.73	19	0.62	.55
Random effects[a]	**Estimate**	**Standard error**		***z* value**	***p* value**
Teacher mean achievement	472.30	191.53		2.47	<.01
Within-teacher variation	2147.31	100.21		21.43	<.01

[a] Teachers were modeled as a random factor.

Geometry CST Outcomes

Effect Size Analysis

Next we consider outcomes for CST Geometry. Table 65 provides a summary of the sample we used in the analysis and the results for the comparison of CST Geometry scores for students in *GC* and control groups. This table gives information about all the students in the original sample for whom we have a posttest score. The interpretation of this table is the same as Table 63.

Table 65. Overview of Sample and Impact of *GC* on Geometry Achievement

Condition	Means	Standard deviations	No. of students	No. of classes	No. of teachers	Effect size
GC	313.76	60.79	484	22	9	0.30
Control	296.43	54.91	554	24	8	

Analysis Including NWEA General Math Pretest as a Moderator

In the following table, we show the result of the impact of *GC*, but as it is moderated through the effects of the NWEA General Math pretest and location. We chose to present this complex model because of the significant interaction effects. Taken together, they suggest that the impact of *GC* on Geometry CST outcomes depends on a student's incoming math proficiency and that this differential effectiveness may depend on the practices of a particular district.

Table 66. Impact of *GC* on Geometry Achievement

Fixed effects[a]	Estimate	Standard error	DF	*t* value	*p* value
Predicted value for a control student with an average pretest in SD	305.18	9.32	10	32.73	<.01
Impact of *GC* for a student with an average pretest in SD	6.00	14.91	10	.40	.70
Predicted change in control outcome for each unit increase on the pretest in SD	2.26	0.19	658	11.78	<.01
Location (1= ES; 0 = SD)	-1.94	14.20	10	-0.14	.89
Interaction of *GC* and location	-0.26	21.35	10	-0.01	.99
Interaction of *GC* and pretest	1.82	0.37	658	4.99	<.01
Interaction of pretest and location	1.00	0.31	658	3.19	<.01
Interaction of *GC*, pretest and location	-2.02	0.52	658	-3.87	<.01
Random effects[b]	**Estimate**	**Standard error**		***z* value**	***p* value**
Teacher mean achievement	315.63	166.28		1.90	.03
Within-teacher variation	1438.82	79.25		18.16	<.01

[a] Pairs of teachers used for random assignment are also modeled as a fixed factor but not included in this table.

[b] Teachers were modeled as a random factor.

As a visual representation of the results described in Table 66 we present two scatterplots in Figure 27 and Figure 28, which show student performance at the end of the year in Geometry, as measured by the CST, against their performance on NWEA General Math in the fall in East Side and San Diego. These graphs show where each student fell in terms of his or her starting point (horizontal x-axis) and his or her outcome score (vertical y-axis). Each point represents one student's post-intervention score against his or her pre-intervention score. The darker points represent *GC* students; the lighter points, control students. The shaded area in the lower right of the graph is the area of negative change (i.e., where students lost ground).

The two lines are the predicted values on the posttest for students in the *GC* and control conditions as determined using the estimated fixed effects in the model.

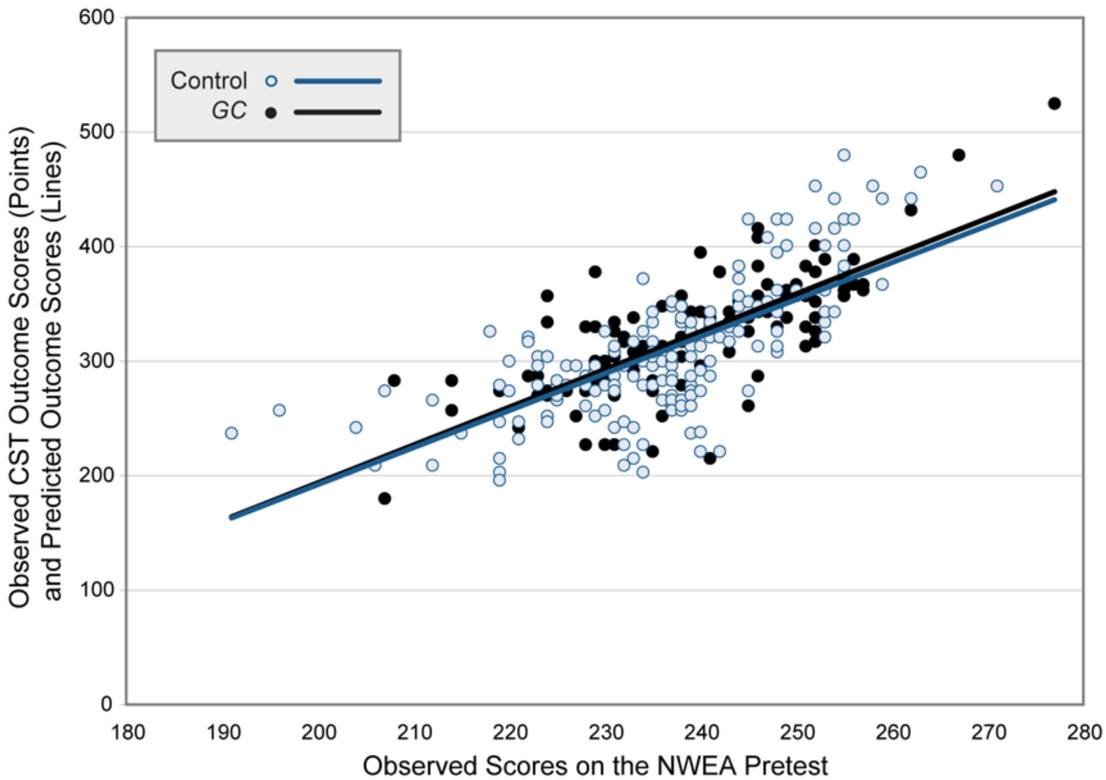

Figure 27. Comparison of Predicted and Actual Outcomes for *GC* and Control Group Students (Geometry Achievement in East Side Unified High School District)

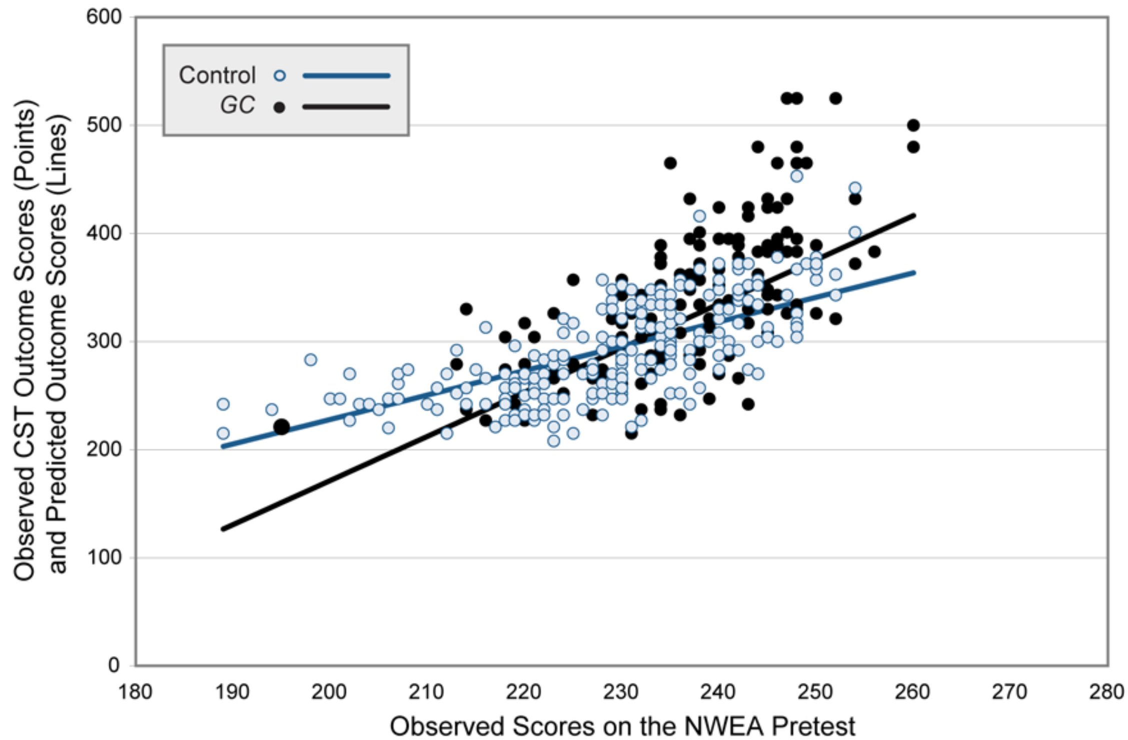

Figure 28. Comparison of Predicted and Actual Outcomes for *GC* and Control Group Students (Geometry Achievement in San Diego Unified School District)

The two scatterplots show distinctly different patterns – *GC* appears to be differentially effective for students along the prior score scale at San Diego, but not at East Side. Also, the relationship between the pre- and post-test performance appears to not be strictly linear at San Diego. This is not the case at East Side. Before running the analysis we checked whether the CST distribution was sufficiently skewed to warrant transforming the data and we determined that no such transformation was required. We also emphasize that very few students in the treatment condition are at the lower end of the pretest scale – a chance imbalance. However, we limit our inference to where the distributions of treatment and control cases overlap.

We find that in San Diego the *GC* students score higher on this sub-strand than the control group at the upper end of the pretest scale. For a student at the median of the third quartile, there is roughly an 11-point advantage to being in the *GC* group; that is, we predict that a student at the median of the third quartile would score 11 points higher on the outcome measure if he or she is in a *GC* class instead of a control class. For a student at the median of the fourth quartile, there is roughly a 27-point advantage to being in the *GC* group; that is, we predict that a student at the median of the fourth quartile would score 27 points higher on the outcome measure if he or she is in a *GC* class instead of a control class. At the low end of the scale control appears to outperform treatment, but due to lack of overlap in the distributions of the two groups in that section of the scale, we caution against drawing this conclusion. That is, there is a lack of common support at the low end of the scale which prevents us from drawing valid inferences concerning effects at that end.

We wish to emphasize that the high p value of .70 associated with the treatment effect in the previous table applies only to students at San Diego who are at the mean of the pretest score. This concurs with the graphed outcome where we don't see much difference in performance between treatment and control students in San Diego who are near the middle of the pretest scale. For East Side there does not appear to be any impact of the intervention on student performance across the prior score scale.

Impact of *GC+Nav* in Comparison to *GC* and Control Groups

In addition to looking at the main comparisons of interest which were covered in the previous sections, we were also interested in seeing whether students with different *GC* tools performed differently from each other. This meant that we split the *GC* group into two subgroups: those using the TI-Navigator with the calculator (*GC+Nav)*, and those using *GC* by itself. This produces three contrasts as shown in Figure 29 below. The three contrasts represent: 1 and 2) a treatment to control comparison (*GC*-only vs. control or *GC+Nav* vs. control) and 3) a treatment to treatment comparison (*GC+Nav* vs. *GC*-only). Breaking up the groups into smaller sub-sets sometimes makes it harder to detect variance or differences that are statistically significant. We therefore see this analysis as a preliminary outcome that will be combined with the results of the follow-on study from the second year in order to get a larger sample of teachers.

We also stress that teachers were assigned at random to all three conditions (i.e., there was a second stage randomization whereby some of the teacher who were in *GC*-only were re-randomized to remain in *GC*-only, while the rest were assigned to the *GC+Nav* condition.) This means that each contrast that is estimated is an unbiased estimate of the difference between two out of the three conditions. Finally, we have to be aware of the potential for inflated type-1error[3] which can occur when we consider multiple contrasts simultaneously.

---

[3] A type-1 error, also called a false-positive, is result that states that the intervention has had an effect when in fact it hasn't.

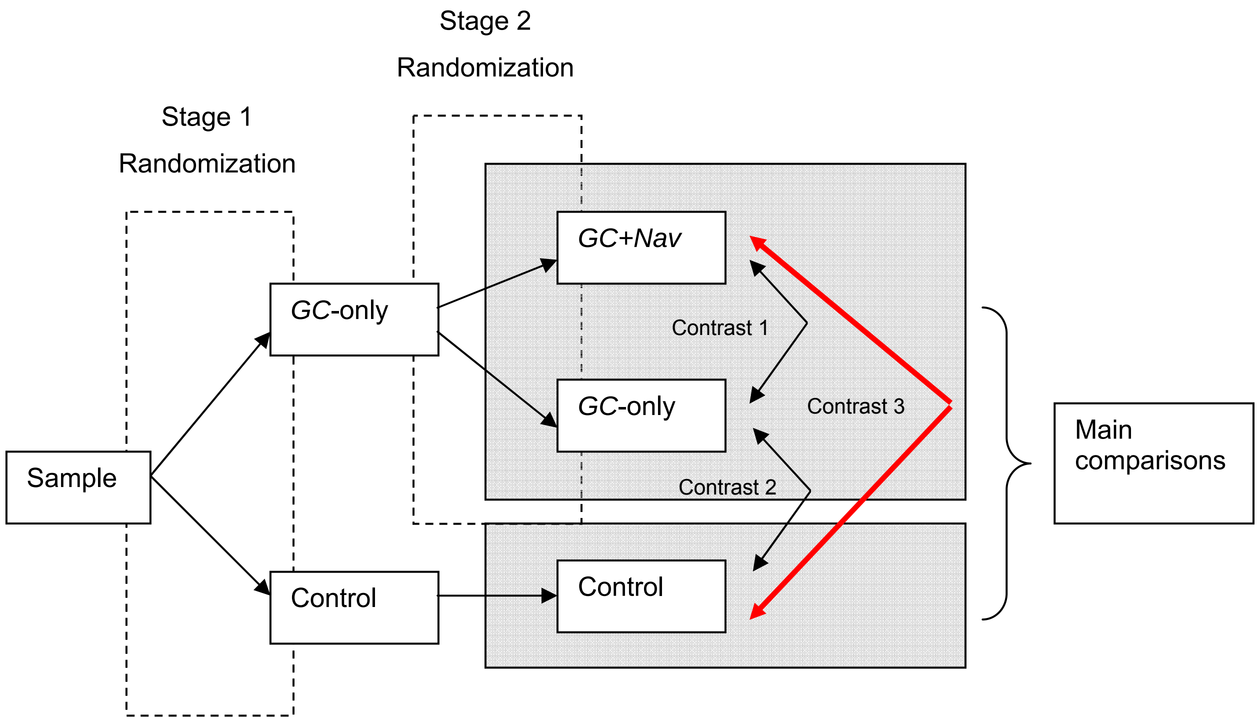

Figure 29. Description of Design

Algebra CST Outcomes

Descriptive Statistics

Table 67 provides a summary of the sample we used and the descriptive results for the analysis comparing *GC+Nav*, *GC*-only, and control group performance on CST Algebra in East Side and San Diego. The interpretation of this table is the same as for Table 63. We did not report an unadjusted effect size because it is not applicable to a three-group comparison.

Table 67. Overview of Sample and Impact of Treatment on Algebra Achievement (CST Outcomes)

	Condition	Means	Standard deviations	No. of students	No. of classes	No. of teachers
	GC+Nav	302.78	43.85	102	7	3
East Side	***GC*-only**	308.79	46.08	141	9	4
	Control	306.81	53.67	349	20	8
	GC+Nav	285.55	56.07	71	6	3
San Diego	***GC*-only**	289.19	45.77	230	10	3
	Control	290.37	36.98	49	3	2

There are several ways to examine whether there are differences among the three conditions in the estimated outcome. In the first approach we compared the performance of student in *GC*-only and *GC+Nav* to the performance of the controls. We included the NWEA General Math Fall test score and an indicator of site as covariates in the model. The results are displayed in Table 68. We see that the rows that describe the impacts of *GC+Nav* and *GC*-only have high *p* values which give us no confidence that the true difference in performance between either of these conditions and the control group is different from zero.

Analysis Including NWEA General Math Pretest as a Covariate

Table 68. Impact of *GC+Nav* and *GC* on Algebra Achievement

Fixed effects[a]	Estimate	Standard error	DF	*t* value	*p* value
Predicted value for a control student with an average pretest in SD	296.74	9.28	16	31.97	<.01
Predicted change in control outcome for each unit increase on the pretest	2.20	0.12	500	18.56	<.01
Impact of *GC+Nav*	-3.84	9.92	16	-0.39	.70
Impact of *GC*	1.41	7.56	16	0.19	.85
Location (1= ES; 0 = SD)	12.54	8.39	16	1.49	.15
Random effects[b]	**Estimate**	**Standard error**		***z* value**	***p* value**
Teacher mean achievement	78.00	78.68		0.99	.16
Within-teacher variation	1330.45	84.57		15.73	<.01

[a] We also modeled separate intercepts for matched pairs of teachers but estimates for these fixed effects are not included in this table. This implies that the intercept value applies to the reference matched pair.

[b] Teachers were modeled as a random factor.

The second approach used to examine if any of the contrasts are significant was to run pair-wise comparisons among the three conditions. This is simply an alternative rendering of the information just presented. We see in Table 69 estimates of the difference between each pair of conditions. The results substantiate the overall finding of no difference in performance among the three conditions. (We have not adjusted the *p* values to correct for the fact that we are looking at multiple comparisons; such an adjustment would only further raise the *p* values.)

Table 69. Pairwise Comparison of Three Groups of Treatment on Algebra Achievement

Comparison	Estimate	Standard error	DF	*t* value	*p* value
***GC* – (*GC+Nav*)**	5.25	9.09	16	0.58	<.57
***GC* – Control**	1.41	7.56	16	0.19	<.85
(*GC+Nav*) – Control	-3.84	9.92	16	-0.39	.70

As a third approach, we considered an overall test of the difference among conditions. That is, we considered whether there is evidence to reject the null hypothesis that the performance of students is the same under the three conditions. This test incorporated the pretest covariate, in order to increase precision. It also included a main effect for location. The type-3 test of fixed effects for the three conditions considered simultaneously has a p value of .84. We do not reject the null hypothesis of there being no difference in the performance of students across the three conditions.

Figure 30 displays the plot of each student's posttest against his/her pretest for the three conditions.

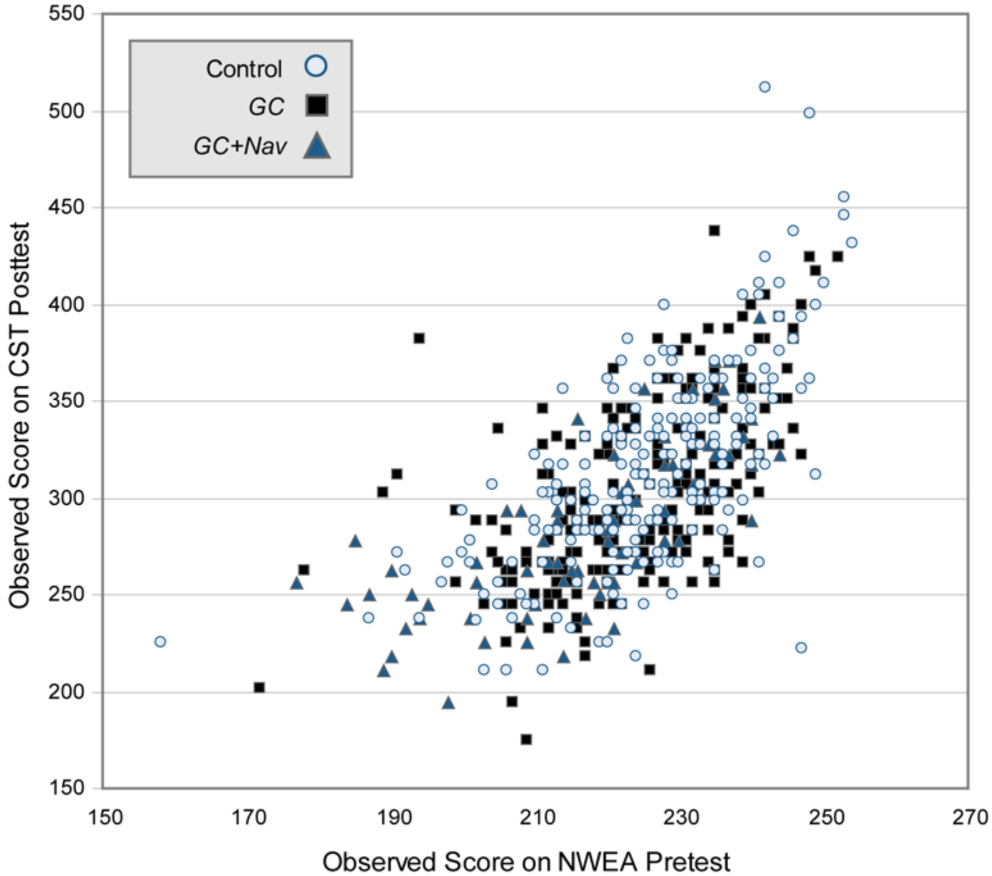

Figure 30. CST Algebra Outcomes for Control, *GC,* and *GC+Nav* Conditions

Geometry CST Outcomes

Descriptive Statistics

Next we considered the relationships among the three conditions and student performance on CST Geometry. Table 70 provides a summary of the sample we used and the descriptive results for the analysis comparing *GC+Nav*, *GC*-only and control group performance on CST Geometry in East Side and San Diego.

Table 70. Overview of Sample and Impact of Treatment on Geometry Achievement (CST Outcomes)

	Condition	Means	Standard deviations	No. of students	No. of classes	No. of teachers
ESUHSD	***GC+Nav***	306.92	50.55	154	7	3
	***GC*-only**	292.73	51.95	146	7	2
	Control	313.10	62.93	211	8	3
SDCS	***GC+Nav***	285.28	41.11	47	6	3
	***GC*-only**	353.64	66.27	137	4	1
	Control	286.17	46.55	343	16	5

Analysis Including NWEA General Math Pretest as a Covariate

Using the same approach as the one in the previous section, we start by comparing the performance of students in *GC*-only and *GC+Nav* to the performance of the controls. We included the NWEA General Math Fall test score and an indicator of site as covariates in the model. The results are displayed in Table 71. We see that the impact of *GC*-only has a *p* value of less than .01, which gives us a high-level of confidence that the observed effect is not simply due to chance. The estimate for the impact of *GC+Nav* has an associated *p* value that is very high, which gives us no confidence that the observed difference in performance between this condition and the control is due to something other than chance.

Table 71. Impact of *GC+Nav* and *GC* on Geometry Achievement

Fixed effects[a]	Estimate	Standard error	DF	*t* value	*p* value
Predicted value for a control student with an average pretest in SD	303.01	2.05	716	147.77	<.01
Predicted change in control outcome for each unit increase on the pretest	2.78	0.10	90	29.30	<.01
Impact of *GC+Nav*	0.63	3.68	716	0.17	.86
Impact of *GC*	16.04	3.42	716	4.07	<.01
Location (1= ES; 0 = SD)	-7.15	2.86	716	-2.50	.01
Random effects[b]	**Estimate**	**Standard error**		***z* value**	***p* value**
Teacher mean achievement	407.81	117.22		3.48	<.01
Within-teacher variation	961.77	117.04		8.22	<.01

[a] We also modeled separate intercepts for matched pairs of teachers but estimates for these fixed effects are not included in this table. This implies that the intercept value applies to the reference matched pair.

[b] Teachers were modeled as a random factor.

The second approach used to examine if any of the contrasts are significant was to run pair-wise comparisons among the three conditions. This is simply an alternative rendering of the information just presented. We observe in Table 72 estimates of the difference between each pair of conditions. The results substantiate the overall finding of an advantage of the *GC*-only condition over the control condition but no difference in performance between *GC+Nav* and the control condition. We see further that the third contrast, that between *GC*-only and *GC+Nav*, shows that students in *GC*-only outperform those in *GC+Nav*, and with a low *p* value for this contrast we have a high-level of confidence that the observed difference is not just the result of chance. (We have not adjusted the *p* values to correct for the fact that we are looking at multiple comparisons; such an adjustment would inflate the *p* values.)

Table 72. Pairwise Comparison of Three Groups of Treatment on Geometry Achievement

Comparison	Estimate	Standard error	DF	*t* value	*p* value
GC* - *GC+Nav	15.41	4.14	716	3.73	<.01
***GC* - Control**	16.05	3.42	716	4.70	<.01
***GC+Nav* - Control**	0.64	3.68	716	0.17	.86

As a third approach, we considered an overall test of the difference among conditions. That is, we considered whether there is evidence to reject the null hypothesis that the performance of students is the same under the three conditions. This test incorporated the pretest covariate, in order to increase precision. It also included a main effect for location. The type-3 test of fixed effects for the three conditions considered simultaneously has a *p* value smaller than .01. We reject the null hypothesis of there being no difference among students in their performance.

Figure 31 displays the plot of each student's posttest against his/her pretest for the three conditions.

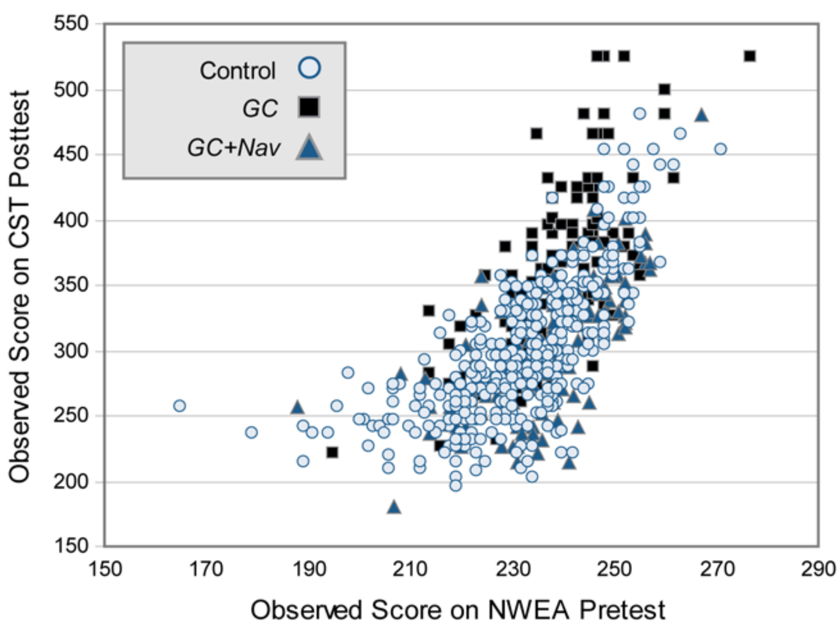

Figure 31. CST Geometry Outcomes for Control, *GC*, and *GC+Nav* Conditions

Exploratory Analysis

We also considered a number of measures from the classroom. These processes are potentially outcomes of *GC* as well as related to the student achievement outcome. As described in the implementation results section of this report, we measured the amount of time that teachers used *GC* and the amount of time students used *GC*. (Here, we return to considering the main comparison of interest, that between treatment and control, where members of the treatment group include all teachers and their students who are assigned to *GC* in the Fall, regardless of whether they are reassigned to *GC+Nav* in the Winter.)

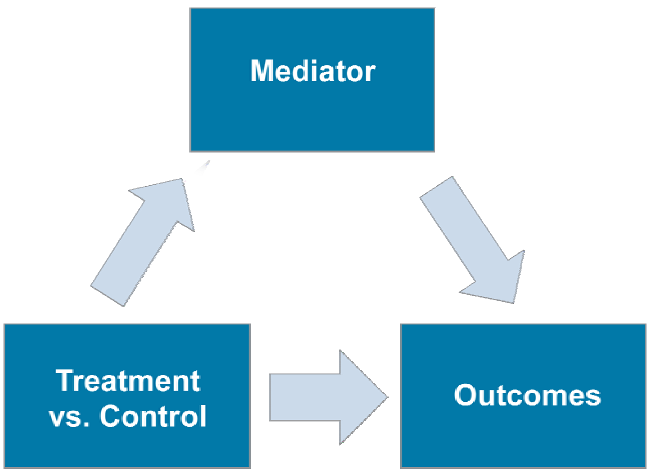

Figure 32. Relationships for Exploratory Analysis of Implementation Variables

When dealing with implementation variables, we can understand them as defining a distinct path or link between the intervention and student-level achievement, as illustrated in Figure 32. Part of the impact of *GC* on student outcomes may be mediated by the intermediate variables. *GC* can have a direct impact on both student outcomes and on instructional time, a teacher-level outcome. The link from instructional time to the student outcome is correlational but an important relationship to explore.

We may consider some events or conditions as outcomes, but it's more appropriate to analyze them as intermediate variables. Intermediate variables are important because they can often facilitate or block a pathway of the causal effect between randomization and student outcome. That is, these variables may mediate the effect of treatment on student outcomes. For instance, between the initial point of randomization and the final outcome measure, some event may occur that alters the program effect on the final outcome. Some examples include quality of implementation, teacher motivation, or certain classroom processes.

Thus in the next section we consider the potential mediating role of Teacher Usage.

Teacher Usage as an Outcome

Relationship between Condition and Teacher Usage

We look first at the impact of condition on amount of time teachers use *GC*. Usage is defined as minutes per class using *GC*. Our question here is whether *GC* resulted in a greater amount of time spent using graphing calculators. To compute the impact of *GC* on Teacher Usage, we regressed Teacher Usage on treatment status. Unlike the previous analyses, where the treatment indicator was at the teacher level and the outcome was at the student-level, in this analysis both the outcome and the assignment variable are at the teacher level[4].

[4] We used SAS PROC REG for the single-level analysis.

Table 73 shows that there was a significant difference between the two groups of teachers in the amount of time that they report using *GC*. Teachers in the treatment condition on average report using graphing calculators 9.65 minutes more per class than teachers in the control condition. With a low *p* value of <.01, we have high level of confidence that *GC* teachers were more likely to use *GC* than control teachers. Because teachers were randomized to conditions, differences between conditions in teacher-level outcome variables can also be regarded as caused by randomization.

Table 73. Impact of *GC* on Teacher Usage of *GC*

Fixed effects[a]	Estimate (minutes per class)	Standard error	DF	*t* value	*p* value
Intercept	5.10	2.05	1	2.49	.02
Impact of *GC*	9.65	2.94	1	3.29	<.01

Note. There are 37 teachers in total with 19 in control group and 18 in treatment group; we do not model separate effects for Algebra and Geometry teachers because some teachers taught both kinds of classes.

The results described in Table 73 are presented graphically in Figure 33. We see that Teacher Usage is higher in the treatment group compared to the control and the 80% confidence intervals do not overlap.

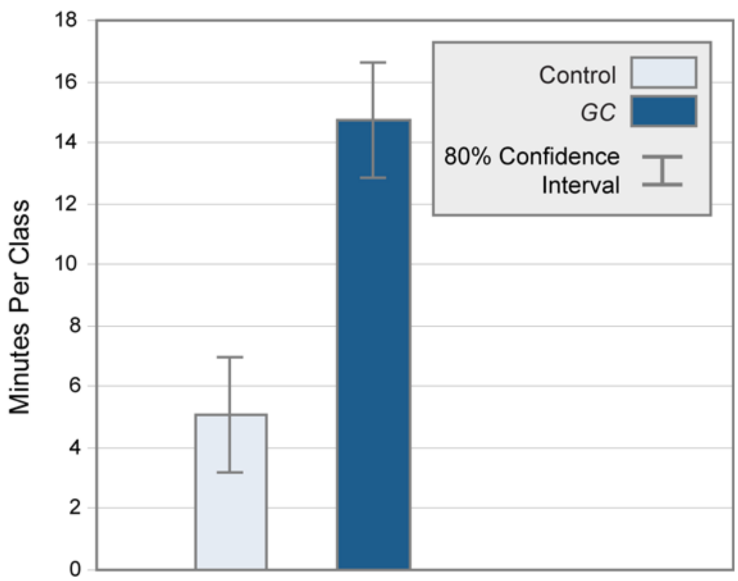

Figure 33. Difference in Teacher Usage between *GC* and Control Groups

Relationship between Teaching Experience, Condition and Teacher Usage

Before considering the potential mediating effect of Teacher Usage on student outcomes, we explored Teacher Usage as an outcome, in greater depth. Specifically we were also interested in the moderating effect of years of teaching on the relationship between treatment and usage. In other words, we wanted to know whether years of teaching influenced usage to different degrees in the treatment and control conditions. For instance, it could be the case that under the control condition experience matters less than under the treatment condition. This would happen for example if, absent the intervention, teachers use *GC* the same amount, more or less, regardless of years of teaching experience, but with the intervention, either the more- or less-experienced teachers become more motivated to use the technology. This would happen for instance if the less experienced teachers see the intervention as an opportunity to integrate the technology into their still-evolving teaching routines.

To answer this question we modeled the interaction of graphing calculator use with treatment status. We show the results of this analysis in Table 74. Using the guidelines for interpreting *p* values that we set forth at the beginning of this report, we have limited confidence that years of teaching experience moderates the treatment effect on teacher usage of *GC*. The point estimate for this effect is .38 which we interpret as: with each additional year of teaching experience the difference between treatment and control in usage of *GC* increases by .38 minutes per class. With a *p* value of .19, the 80% confidence interval for this effect is just short of crossing zero.

Table 74. Relationship of Teaching Experience and *GC* to Teacher Usage

Fixed effects[a]	Estimate	Standard error	DF	*t* value	*p* value
Predicted value for a teacher with an average number of years of teaching experience	8.31	3.87	1	2.15	.04
Impact of *GC* on teacher usage for a teacher with an average number of years of teaching experience	3.36	5.56	1	0.60	.55
Predicted change in control teacher usage for each additional year of teaching experience	-0.18	0.19	1	-0.98	.33
Interaction of *GC* and teaching experience	0.38	0.28	1	1.34	.19

Student CST Scores as Outcomes

Relationship between Teacher Usage and Student Outcomes

We wanted to explore the relationship between how much time was spent using *GC* and student achievement. The amount of time spent using *GC* is itself a causal outcome of the intervention. It stands between randomization and the student outcome, achievement.

To examine the mediating role of teacher usage, we consider three relationships, those between: 1) treatment status and student performance; 2) treatment status and the mediator (teacher usage); and 3) the mediator (teacher usage) and student performance.

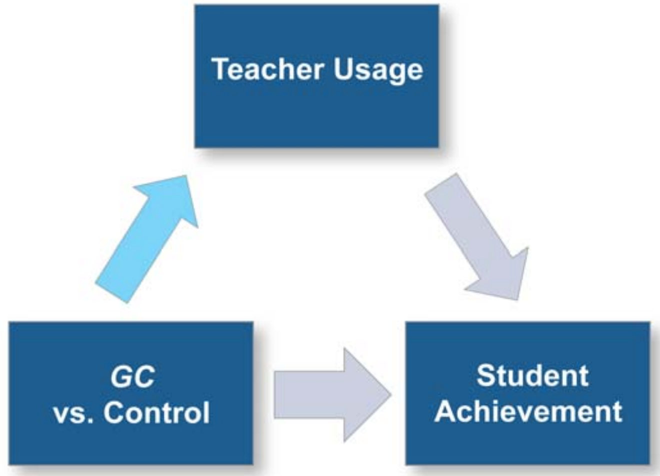

Figure 34. Arm 1: Causal Impact of *GC* on Teacher Usage

The first two of these are unbiased estimates of the impact of the intervention because units were randomly assigned to treatment or control. The third is correlational - we did not assign teachers to levels of usage so we cannot be sure whether Teacher Usage, or a factor confounded with usage, is the co-determinant of student performance. As with all our analyses, we model the dependency among observations among students who have the same teacher.

Common to all of these analyses is the link between the treatment condition and teacher usage (see Figure 34). We explored this relationship earlier and showed that there is a strong positive impact of *GC* on Teacher Usage.

To interpret the relationships represented by the other two arms in Figure 34, we have to consider the subject-specific outcomes. The causal impact represented by the base arrow was explored earlier, and the results will be reviewed here. The correlation between Teacher Usage and student performance (the third arm) will also be explored. (Note that subject-specific outcomes imply smaller teacher samples since most teachers taught one or the other subject.)

CST Algebra Outcome

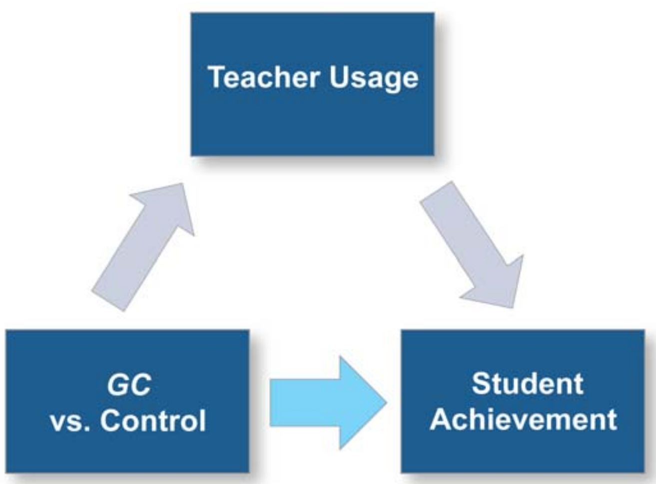

Figure 35. Arm 2: Causal Impact of *GC* on Student Achievement

Arm 2:

We consider the relationship between the treatment condition and student performance on CST Algebra. We showed in Table 68 that we have no confidence that students experiencing *GC* performed differently than students in the control condition in CST Algebra in the Spring.

Arm 3:

Even with no difference between the two groups, it is useful to explore whether there is a relationship between amount of teacher usage of *GC* and student achievement. The result of this analysis is purely correlational; we have not assigned teachers to levels of *GC* usage, therefore we cannot be sure whether it is *GC* teacher usage or some other variable correlated with instructional time with *GC* (e.g., teacher enthusiasm) that is the true cause of the student outcome[5].

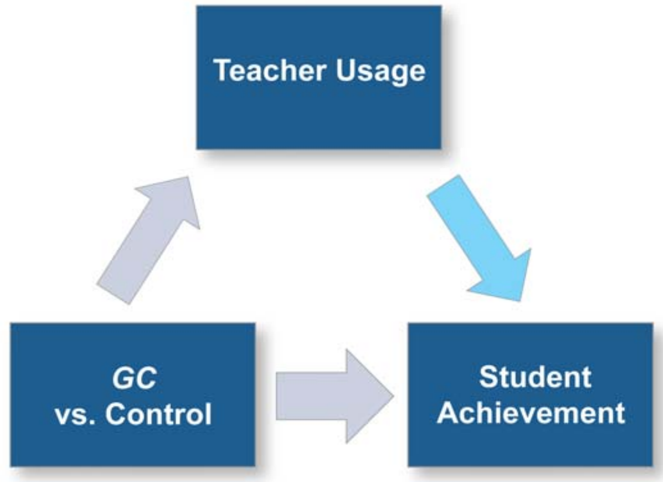

Figure 36. Arm 3: Causal Impact of Teacher Usage on Student Achievement

The result of the analysis of the relationship between Teacher Usage and student achievement on CST Algebra is given in Table 75. The pretest score was included as a covariate to increase the precision of our estimate. We see that there is a slight negative correlation between Teacher Usage and student performance on CST Algebra. Using our criteria for interpreting *p* values, we have some confidence that the observed difference is not simply due to chance.

[5] Since some potential mediator is bound to be different between treatment and control teachers, simply as a result of chance, it is important to not extensively identify such potential mediators post hoc. It would not be fair to continue testing other teacher-level outcomes, such as teacher enthusiasm, until we find some that are significant, in order to ‘explain’ the result.

Table 75. Association between *Teacher Usage* and Student CST Algebra Performance

Fixed effects[a]	Estimate	Standard error	DF	*t* value	*p* value
Intercept	309.90	3.27	316	94.84	<.01
Predicted change in outcome for each pretest point	2.31	0.12	204	19.19	<.01
The predicted change in student outcome for each minute increase in *Teacher Usage*	-0.44	0.26	316	-1.67	.10
Random effects	**Estimate**	**Standard error**		***z* value**	***p* value**
Teacher mean achievement	23.86	79.03		0.30	.38
Within-teacher variation	1473.71	119.03		12.38	<.01

[a] We also modeled separate intercepts for matched pairs of teachers but estimates for these fixed effects are not included in this table. This implies that the intercept value applies to the reference matched pair.

Considering all three arms together, we have a strong positive causal relationship between *GC* and Teacher Usage. We have no confidence that the intervention affects student performance. We have some confidence that Teacher Usage is negatively correlated with student achievement.

CST Geometry Outcome

Arm 2:

We consider the relationship between the treatment condition and student performance on CST Geometry. We showed in the section that described the impact of *GC* on CST Geometry that the result is complex. Integrating the outcomes from both sites, there does not appear to be an impact of the intervention at East Side. At San Diego, higher-performing students appear to benefit from the intervention and the results for lower-performing students are inconclusive.

Arm 3:

Next we consider whether there is a correlation between Teacher Usage and student achievement. The results of the analysis are described in Table 76. The pretest score was included as a covariate to increase the precision of our estimate. We see that there is a slight negative correlation between Teacher Usage and student performance on CST Geometry. Using our criteria for interpreting *p* values, we have some confidence that the observed difference is not simply due to chance.

Table 76. Association between *Teacher Usage* and Student CST Geometry Performance

Fixed effects[a]	Estimate	Standard error	DF	*t* value	*p* value
Intercept	283.65	2.87	425	98.96	<.01
Predicted change in outcome for each pretest point	3.12	0.13	250	24.38	<.01
The predicted change in student outcome for each minute increase in *Teacher Usage*	-0.23	0.15	425	-1.54	.12
Random effects	**Estimate**	**Standard error**		***z* value**	***p* value**
Teacher mean achievement	305.58	149.05		2.05	.02
Within-teacher variation	1572.38	143.25		10.98	<.01

[a] We also modeled separate intercepts for matched pairs of teachers but estimates for these fixed effects are not included in this table. This implies that the intercept value applies to the reference matched pair.

Considering all three arms together, we have a strong positive causal relationship between *GC* and Teacher Usage. In one location (San Diego), the intervention has a positive impact on student performance on CST Geometry for higher performing students. The impact for lower performers is null in East Side and inconclusive in San Diego. Finally, we have some confidence that Teacher Usage is negatively correlated with student achievement.

Discussion of Combined Results

The results from the experiments conducted in these two districts were inconclusive as to the primary question which was, did introducing graphing calculators along with professional development and support (collectively designated at *GC*) result in increased test scores in algebra and geometry. The reader should consult the individual reports for each district to get a finer grain detail on the implementation and the results. While both represented urban centers in California, the results were not identical as we saw in the findings for geometry. In one of the districts we found that GC benefited geometry students who initially scored higher on the math pretest but not students who initially scored on the low end. The inconsistency is intriguing but is not explained within the scope of these experiments.

We did find that the availability of the technology and the professional development increased the amount of calculator usage in the classes of those teachers. This is a rough indication that the treatment, considered as the classroom experience of the students, was on average implemented. The usage however, was not correlated with achievement in either algebra or geometry. In fact in both cases, calculator use was associated with lower student outcomes. The small number of teachers in each subgroup makes these analyses exploratory and of value mostly as pointers to factors that should be considered in subsequent experiments. For example, we detected that teacher years of experience was associated with both calculator usage and with experimental condition.

The experiment served also as a pilot for our subsequent work on the *TI-Navigator*, a wireless system for connecting calculators within the classroom. Since teachers were randomly assigned to using the *TI-Navigator* or continuing with the regular treatment condition for the second semester, we were able to look at the three way comparison of the control, calculator and calculator plus *TI-Navigator* conditions. For algebra outcomes we found no difference among these conditions. For geometry, the calculator alone condition outperformed both the control and the *TI-Navigator* conditions. The very small number of teachers involved in this phase of the experiment leaves the strong possibility that these differences had more to do with individual teachers than with the instructional tools available. Our second year experiment contrasting graphing calculators with and without *TI-Navigator* will shed additional light on these results.

The two individual site reports provide the districts with the local results both for the implementation and the achievement outcomes. As we noted, the results were not consistent between the districts so readers in each location should consider differences in population and implementation in interpreting the results. The primary value of this report combining data from both sites was to increase the number of teachers available for analysis of usage and of the *TI-Navigator* where the numbers were very small within each site. The results point especially to the potential for an impact of graphing calculators in geometry, an area with less existing research than found for algebra. We do not consider the outcome for the *TI-Navigator* to be representative of the potential impact both because of the short time it was in use and the small number of teachers—these factors made it a useful pilot for implementation but not a research result that can be used with any confidence.

References

California Department of Education. (2005, November). *Blueprints Preface.* Retrieved December 14, 2006 from: *http://www.cde.ca.gov/ta/tg/sr/blueprints.asp*

East Side Union High School. (2006, December). *Budget FAQs*. Retrieved December 14, 2006 from: *http://www.esuhsd.org/Business_Services/Finance/Budget_FAQs.html#Actual_&_Projected_R*

East Side Union High School. (2004, August). *News Release*. Retrieved December 14, 2006 from: *http://www.esuhsd.org/Administration/documents/News_8.26.04_2Prin.pdf*

East Side Union High School. (2005, January). *News Release*. Retrieved December 14, 2006 from: *http://www.esuhsd.org/Administration/documents/News_Release_2005/NR_20050119.pdf*

Fullan, M., & Promfret, A. (1977). Research on Curriculum and Instruction Implementation. *Review of Educational Research, 47*,(2), 335-397.

Heller, J., Curtis, D. Jaffe, R., and Verboncoeur, C. (2006). *Impact of Hand held Graphing Calculator Use on Student Achievement in Beginning Algebra*. Retrieved December 20, 2006 from: *http://education.ti.com/sites/US/downloads/pdf/heller_grcalcreport_2005.pdf*

Khoju, M., Jaciw, A. & Miller, G.I.(2006). *Effectiveness of Graphing Calculators in K-12 Mathematics Achievement: A Systematic Review.* Empirical Education Research Report. Palo Alto, CA: Empirical Education Inc.

Mukti, N. (2000). Computer Technology in Malaysia: Teacher's Background Characteristics, Attitudes, and Concerns. *The Electronic Journal on Information Systems in Developing Countries 3*,(8), 1-33.

Supovitz, J., & Turner, H. (2000). The Effects of Professional Development on Science Teaching Practices and Classroom Culture. *Journal of Research in Science Teaching, 37*,(9), 963-980.

Thompson, A. (2005). Scientifically Based Research: Establishing a Research Agenda for the Technology in Teacher Education Community. *Journal of Research on Technology in Education*, 37.

Appendix A1: East Side Surveys

Surveys were deployed via the web to both *GC* and Control group teachers beginning on January 16, 2006 and continued on a bi-weekly basis until mid-May. Response rates were calculated using a simple percentage calculation based on the ratio of actual received responses to the number of expected responses. There were 10 teachers in the *GC* group and 11 teachers in the Control group. All response rates were calculated based on these expectations. A total of nine surveys were deployed with an overall response rate of 89.95% for both groups, a 96.67% response rate for the *GC* teachers, and an 83.84% for the Control teachers.

In an effort to collect data equally from both groups, we sent the same survey to all of the teachers on all but three occasions. In these surveys, the topics were modified to allow for the differences between the learning environments across the two groups. Surveys 4 and 8 concentrated on questions regarding specific interactions with materials, and survey 9 focused on the teachers' overall experience with the various materials.

The survey topics were developed to account for the various aspects of teacher and student actions associated with instruction and learning. We surveyed each topic twice. In order to characterize the average time teachers and students spent using calculators, we used a repeated question strategy. On surveys 2 through 7 we asked two questions: a) During an average class, when you were teaching, about how many minutes do you normally use a calculator? b) During the last week, in an average class, about how many minutes did your students spend using a calculator? These questions, together with questions regarding the types of activities and types of calculators used, allow us to draw inferences about how much time was devoted to instruction using the calculator by teachers and by students in both the *GC* and Control groups. The typical survey took teachers no more than 15 minutes to answer with an average of 12 question stems. The last survey was slightly longer with 21 items.

Table A1-1. Survey Response Rates

Survey number	Date	Topic	*GC* response rate	Control response rate	Overall response rate
Survey 1	Jan. 16 - 20	Class structure	80%	45.45%	61.90%
Survey 2	Jan. 20 - Feb. 3	Resources and materials	100%	81.82%	90.48%
Survey 3	Feb. 13 - 17	Assessments	100%	100.00%	100.00%
Survey 4T*	Feb. 27 - Mar. 3	Interactions with TI materials	100%	N/A	100.00%
Survey 4C**	Feb. 27 - Mar. 3	Interactions with control materials	N/A	90.91%	90.91%
Survey 5	Mar. 10 - 17	Class structure	100%	81.82%	90.48%
Survey 6	Mar. 27 - 31	Resources and materials	100%	90.91%	95.24%
Survey 7	Apr. 10 - 14	Assessments	100%	90.91%	95.24%
Survey 8T*	Apr. 24 - 28	Interactions with TI materials	90%	N/A	90.00%
Survey 8C**	Apr. 24 - 28	Interactions with control materials	N/A	81.82%	81.82%
Survey 9T*	May 19	Final survey-overall experience with TI materials	100%	N/A	100.00%
Survey 9C**	May 19	Final survey-overall experience with control materials	N/A	90.91%	90.91%

*Asked only of *GC* teachers.

**Asked only of control teachers.

The following tables represent a subset of the final survey responses for the *GC* teachers only.

Table A1-2. During this year's study what percentage of math time did you spend using some part of your calculator package?

		0 %	1-25%	26-50%	51-75%	76-100%
	Number of teachers	No.	No.	No.	No.	No.
GC	10	1	3	5	1	0

Table A1-3. During this year's study what percentage of classroom time did you spend using Cabri® Jr.?

		0 %	1-25%	26-50%	51-75%	76-100%
	Number of teachers	No.	No.	No.	No.	No.
GC	5	2	1	2	0	0

Table A1-4. Please rate your experience using Cabri® Jr.

		Very Dissatisfied	Somewhat Dissatisfied	Neither Satisfied nor Dissatisfied	Somewhat Satisfied	Very Satisfied
	Number of teachers	No.	No.	No.	No.	No.
GC	5	0	0	2	1	2

Table A1-5. Would you recommend Cabri® Jr. to other teachers?

		Yes	No	Not Sure
	Number of teachers	No.	No.	No.
GC	5	4	0	1

Table A1-6 provides insight into how many teachers were actually using the TI-Navigator system by the end of the implementation period.

Table A1-6. During this years study, on average, what percentage of classroom time was spent using the TI-Navigator system?

		0 %	1-25%	26-50%	51-75%	76-100%
	Number of teachers	No.	No.	No.	No.	No.
GC+Nav	5	2	3	0	0	0

Table A1-7 shows the teacher that tried and failed to get the TI-Navigator system set-up and the teacher that was intimidated by the packaging.

Table A1-7. Please rate your experience using the TI-Navigator system.

		Very dissatisfied	Somewhat dissatisfied	Neither satisfied nor dissatisfied	Somewhat satisfied	Very satisfied
	Number of teachers	No.	No.	No.	No.	No.
GC+Nav	5	1	1	1	1	1

Table A1-8. Would you recommend the TI-Navigator system to other teachers?

		Yes	No	Not Sure
	Number of teachers	No.	No.	No.
GC+Nav	5	5	0	0

Appendix A2: San Diego Surveys

Surveys were deployed via the web to both *GC* and control group teachers beginning on January 16, 2006 and continued on a bi-weekly basis until mid-May. Response rates were calculated using a simple percentage calculation based on the ratio of actual received responses to the number of expected responses. There were 8 teachers in the *GC* group and 8 teachers in the control group. All response rates were calculated based on these expectations. A total of nine surveys were deployed with an overall response rate of 88.19% for both groups, an 84.72% response rate for the *GC* teachers, and a 91.67% for the control teachers.

In an effort to collect data equally from both groups, we sent the same survey to all of the teachers on all but three occasions. In these surveys, the topics were modified to allow for the differences between the learning environments across the two groups. Surveys 4 and 8 concentrated on questions regarding specific interactions with materials, and survey 9 focused on the teachers' overall experience with the various materials.

The survey topics were developed to account for the various aspects of teacher and student actions associated with instruction and learning. We surveyed each topic twice to characterize the natural variation in the classroom environment. In order to characterize the average time teachers and students spent using calculators, we used a repeated question strategy. On surveys 2 through 7 we asked two questions: a) During an average class, when you were teaching, about how many minutes do you normally use a calculator? b) During the last week, in an average class, about how many minutes did your students spend using a calculator? These questions, together with questions regarding the types of activities and types of calculators used, allow us to draw inferences about how much time was devoted to instruction using the calculator by teachers and by students in both the *GC* and control groups. The typical survey took teachers no more than 15 minutes to answer with an average of 12 question stems. The last survey was slightly longer with 21 items.

Table A2-1. Survey Response Rates

Survey number	Date	Topic	*GC* response rate	Control response rate	Overall response rate
Survey 1	Jan. 16 - 20	Class Structure	87.5%	75.0%	81.25%
Survey 2	Jan. 30 - Feb. 3	Resources and materials	100.0%	100.0%	100.00%
Survey 3	Feb. 13 - 17	Assessments	87.5%	100.0%	93.75%
Survey 4T*	Feb. 27 - Mar. 3	Interactions with TI materials	87.5%	N.A.	87.50%
Survey 4C**	Feb. 27 - Mar. 3	Interactions with Control materials	N.A.	87.5%	87.50%
Survey 5	Mar. 10 - 17	Class structure	62.5%	87.5%	75.00%
Survey 6	Mar. 27 – 31	Resources and materials	87.5%	100.0%	93.75%
Survey 7	Apr. 10 – 14	Assessments	87.5%	100.0%	93.75%
Survey 8T*	Apr. 24 - 28	Interactions with TI materials	100.0%	N.A.	100.00%
Survey 8C**	Apr. 24 - 28	Interactions with Control materials	N.A.	100.0%	100.00%
Survey 9T*	May 19	Final Survey- Overall experience with TI materials	62.5%	N.A.	62.50%
Survey 9C**	May 19	Final Survey- Overall experience with Control materials	N.A.	75.0%	75.00%

*Asked only of *GC* teachers

**Asked only of control teachers

The following tables represent a subset of the final survey responses for the *GC* teachers only. The questions represented were asked once at the end of the school year.

Table A2-2. During this year's study what percentage of math time did you spend using some part of your calculator package?

	0 %	1-25%	26-50%	51-75%	76-100%
GC	0%	80%	20%	0%	0%

Note. Three *GC* teachers did not provide responses.

Table A2-3. During this year's study what percentage of classroom time did you spend using Cabri Jr.?

	0 %	1-25%	26-50%	51-75%	76-100%
GC	40%	40%	20%	0%	0%
Note. Three *GC* teachers did not provide responses					

Table A2-4. Please rate your experience using Cabri Jr.

	Very Dissatisfied	Somewhat Dissatisfied	Neither Satisfied nor Dissatisfied	Somewhat Satisfied	Very Satisfied
GC	0%	0%	40%	60%	0%
Note. Three *GC* teachers did not provide responses					

Table A2-5. Would you recommend Cabri Jr. to other teachers?

	Yes	No	Not Sure
GC	60%	0%	40%
Note. Three *GC* teachers did not provide responses			

Table A2-6 provides insight into how many teachers were actually using the TI-Navigator system by the end of the implementation period.

Table A2-6. During this years study, on average, what percentage of classroom time was spent using the TI-Navigator system?

	0 %	1-25%	26-50%	51-75%	76-100%
GC+Nav	33.33%	66.67%	0%	0%	0%
Note. One *GC+Nav* teacher did not provide responses					

Table A2-7 shows the teacher that did not attempt to implement the TI-Navigator.

Table A2-7. Please rate your experience using the TI-Navigator system.

	Very dissatisfied	Somewhat dissatisfied	Neither satisfied nor dissatisfied	Somewhat satisfied	Very satisfied
GC+Nav	0%	0%	33.33%	66.67%	0%
Note. One *GC+Nav* teacher did not provide responses					

Table A2-8. Would you recommend the TI-Navigator system to other teachers?

	Yes	No	Not Sure
GC+Nav	66.67%	0%	33.33%
Note. One *GC+Nav* teacher did not provide responses			

Appendix B1: East Side Calculator System Use

This series of figures displays information derived from the surveys that contrasts several aspects of how the calculators and the TI-Navigator system were used during classroom instruction. Teachers were asked to check all that applied to their situation and were asked for descriptions of any use not already noted. Only teachers were surveyed and so the characterization of student usage is from the teachers' point of view.

GC and Control Group Usage

Figure B1-1 confirms what was seen in the classroom observations. Calculators are not used at all in many of the control classrooms, while the TI calculators in the *GC* group are present and being used.

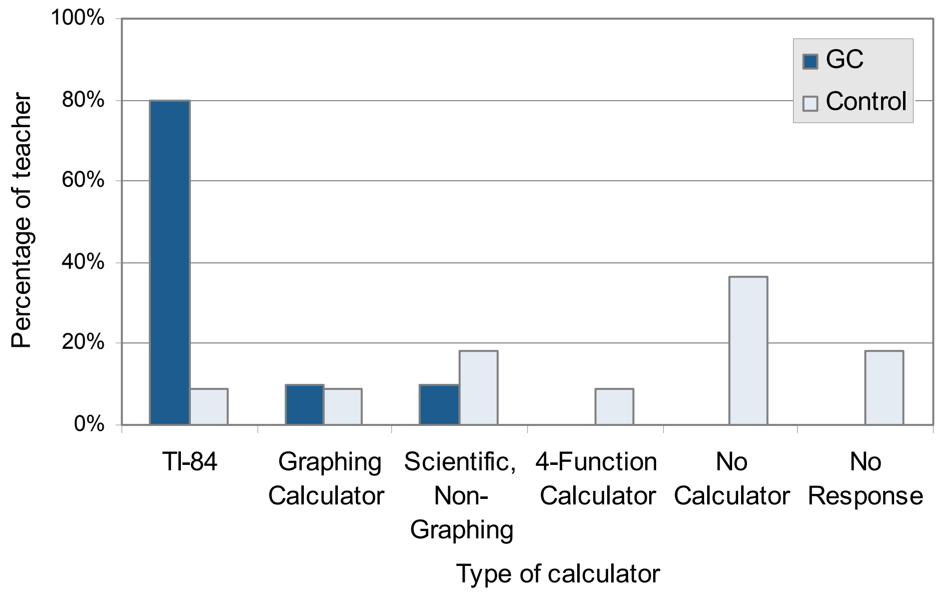

Figure B1-1. Types of Calculators Teachers Used While Teaching

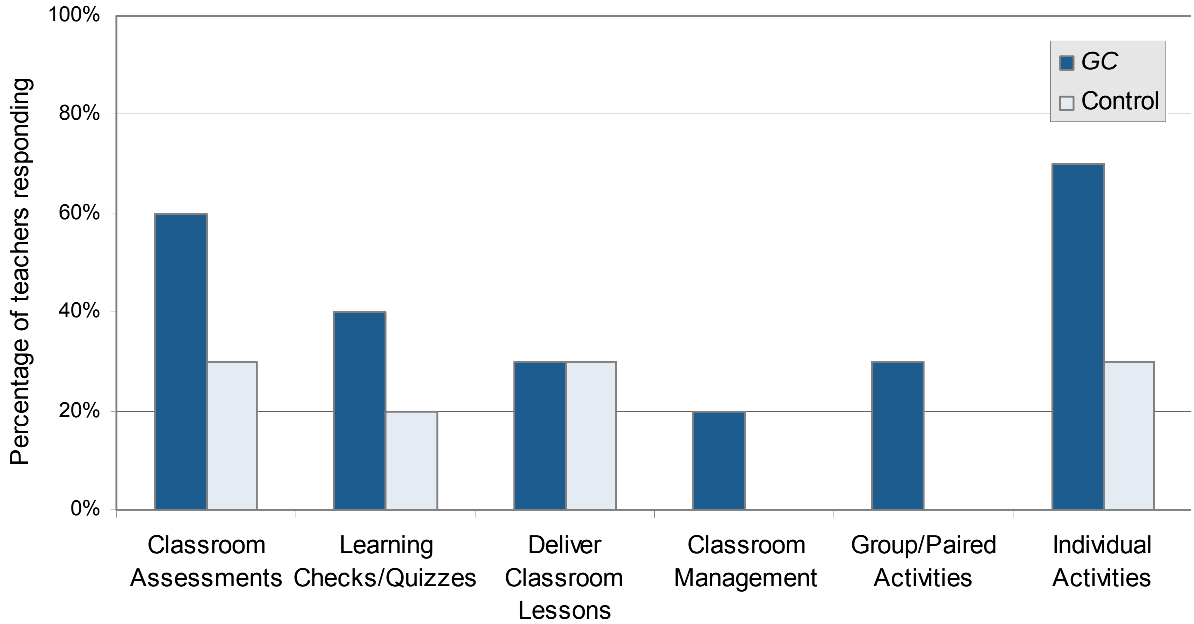

Figure B1-2. Teacher Tasks Performed with Graphing Calculator

Noteworthy in Figure B1-3 is that *GC* teachers' usage, even if it is just for computation, is higher than in the control condition.

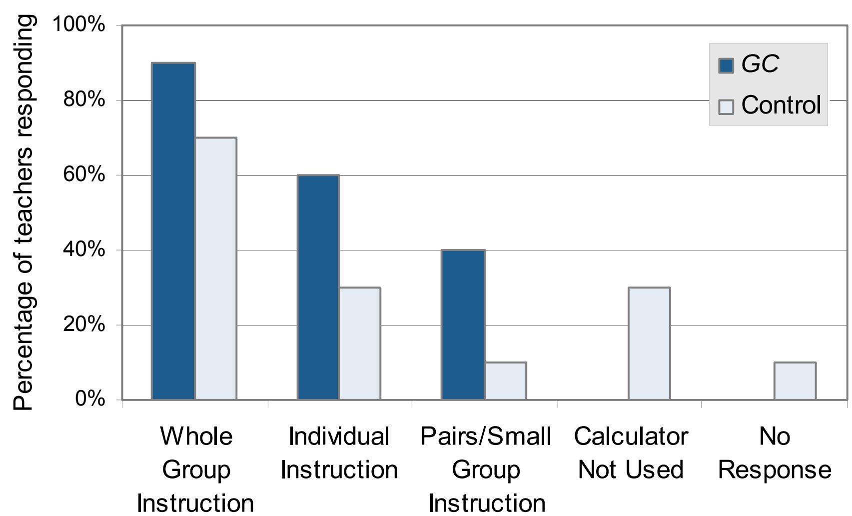

Figure B1-3. Classroom Organization While Teaching with a Calculator

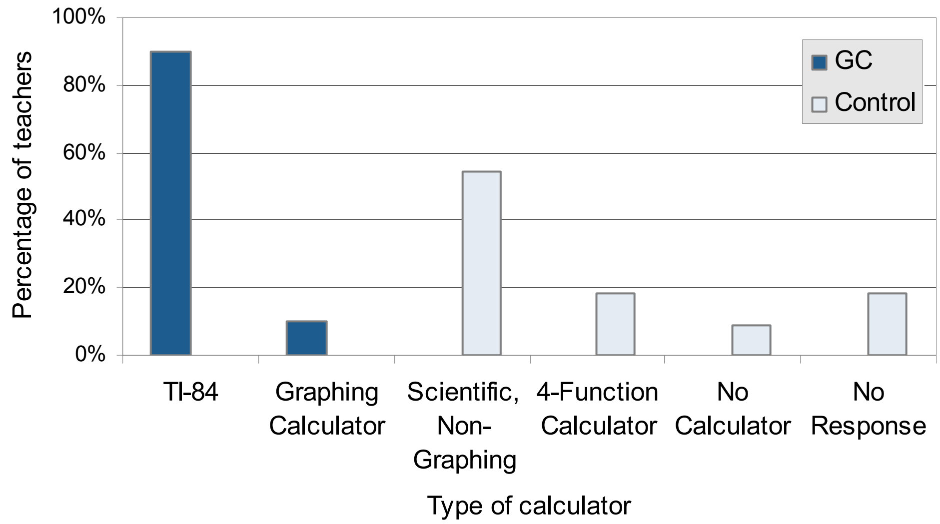

Figure B1-4. Types of Calculators Used by Students

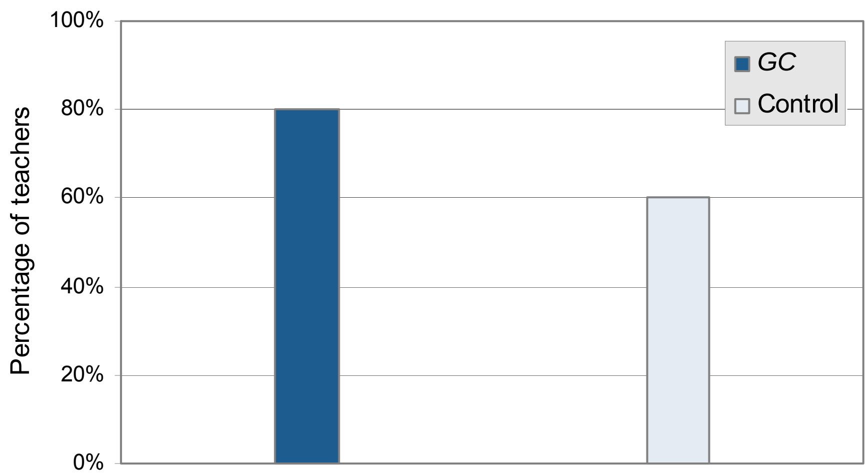

Figure B1-5. *GC* and Control Group Homework Assignments Using Calculators

Figure B1-6 shows that *GC* students used the calculator system to demonstrate problem solutions to the class.

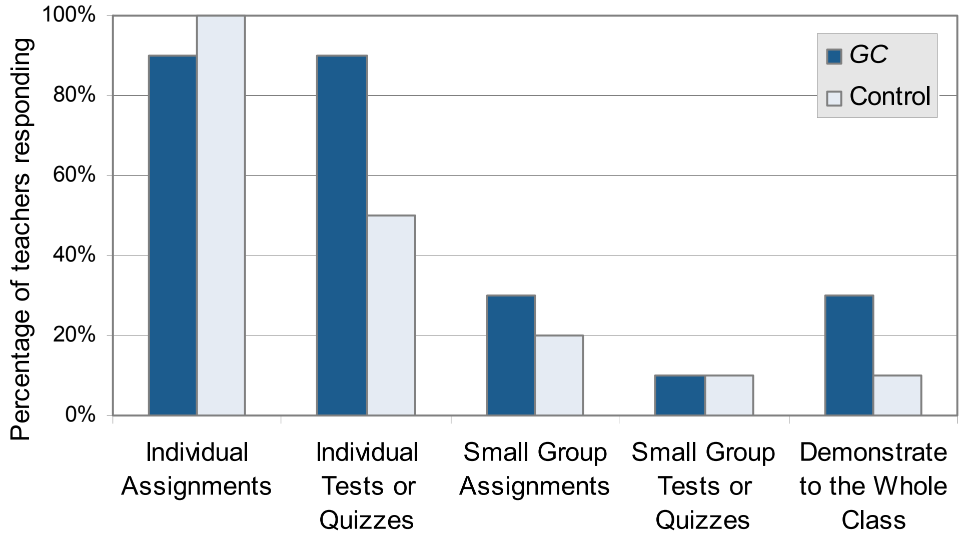

Figure B1-6. Student Calculator Use During Class

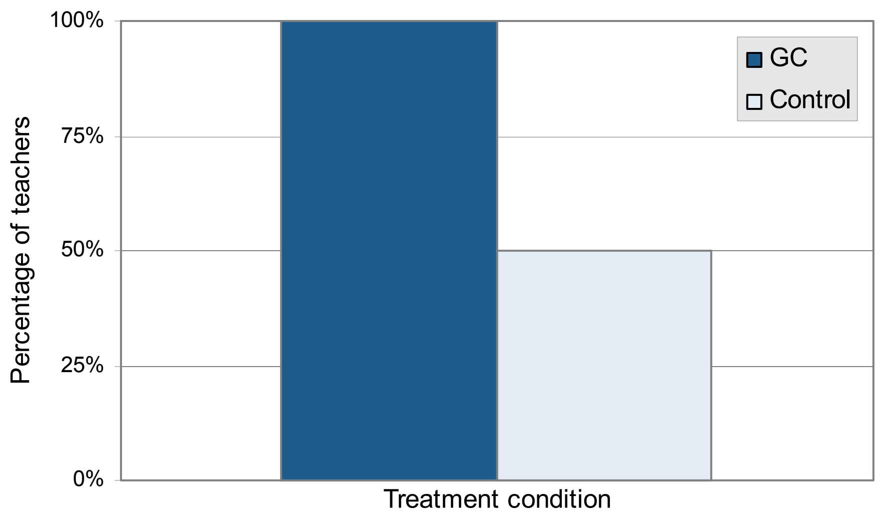

Figure B1-7. Availability of Calculators at Home

Comparison of Calculator System Use between *GC* and *GC+Nav* Groups

The next series of figures contrasts the *GC* group against the *GC+Nav* group in classroom organization, teacher tasks, and student tasks. Not surprisingly, the *GC+Nav* teachers indicated that they used their system for classroom management tasks such as taking attendance.

Classroom Organization

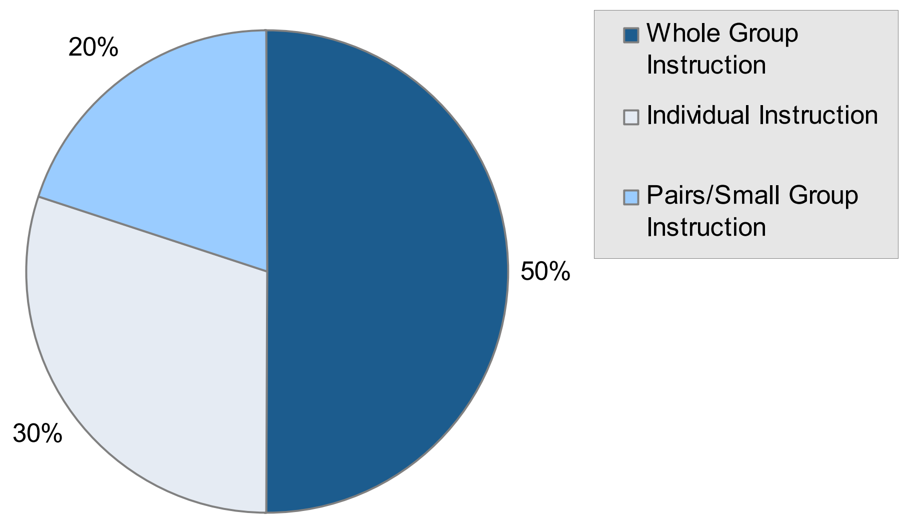

Figure B1-8. Classroom Organization by *GC* Group

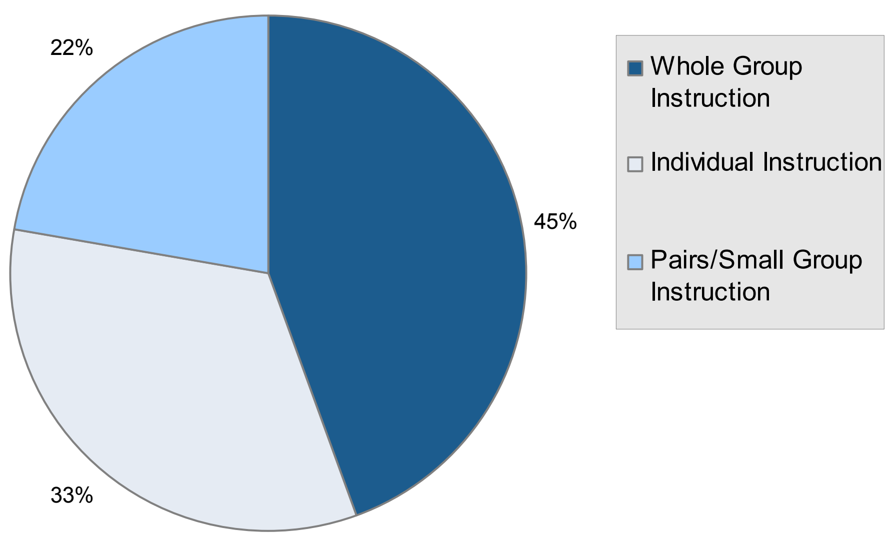

Figure B1-9. Classroom Organization by *GC+Nav* Group

Teacher Use of Calculator Systems

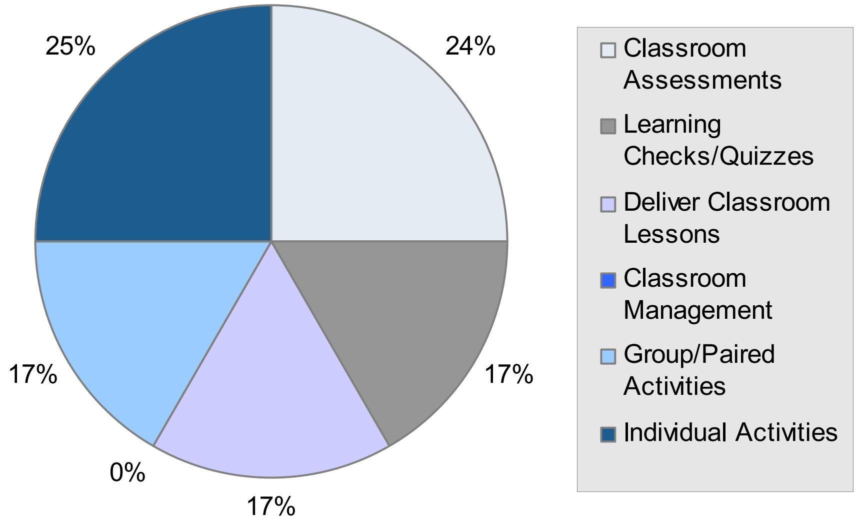

Figure B1-10. *GC* Teacher Use of Calculator Systems

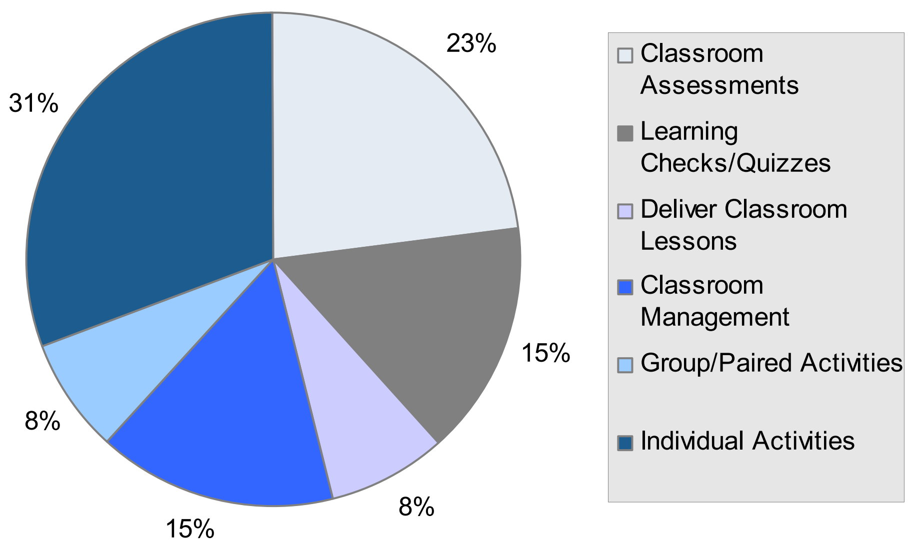

Figure B1-11. *GC+Nav* Teacher Use of Calculator Systems

Student Use of Calculator Systems

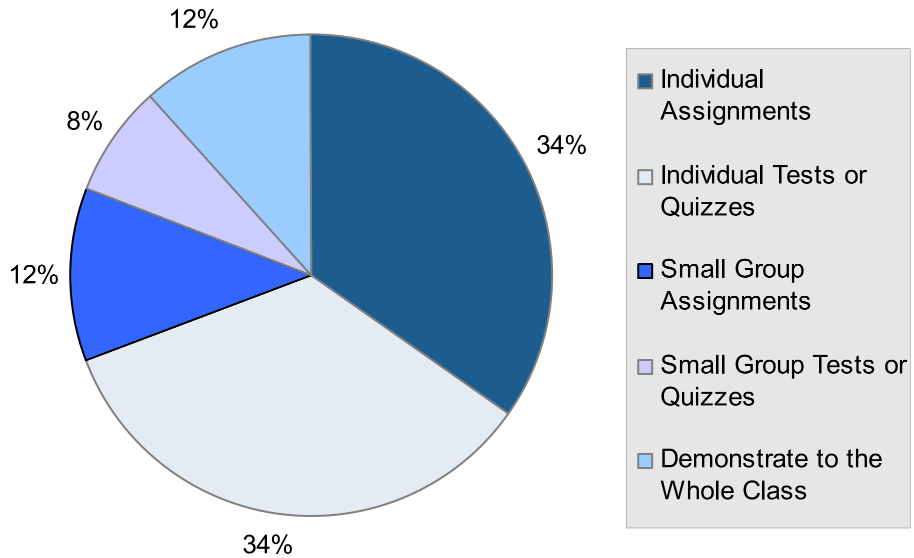

Figure B1-12. *GC* Student Use of Calculator Systems

In the following figure, note the slight increase in students demonstrating to the whole class.

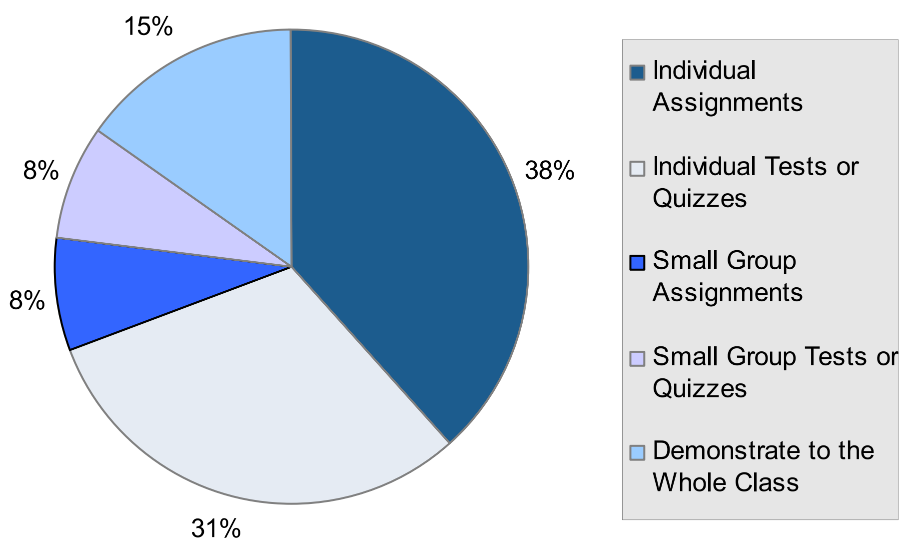

Figure B1-13. *GC+Nav* Student Use of Calculator Systems

Appendix B2: San Diego Calculator System Use

This series of figures displays information derived from the surveys that contrasts several aspects of how the calculators and the TI-Navigator system were used during classroom instruction. Teachers were asked to check all that applied to their situation and were asked for descriptions of any use not already noted. Only teachers were surveyed and so the characterization of student usage is from the teachers' point of view.

GC and Control Group Usage

Figure B2-1 and Figure B2-2 confirm what was seen in the classroom observations. The TI-84 graphing calculators are present and being used in most of the *GC* group classrooms, while calculators are not used at all by teachers or students in many of the control classrooms. The question was asked once at the beginning of February. We cannot characterize how the types of calculators may have changed after mid-semester.

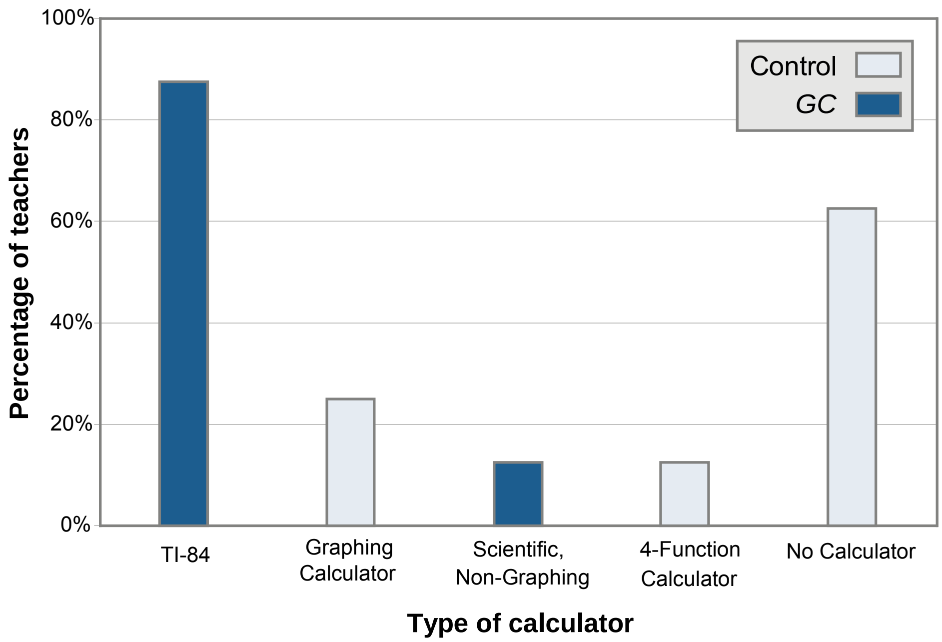

Figure B2-1. Types of Calculators Teachers Used While Teaching

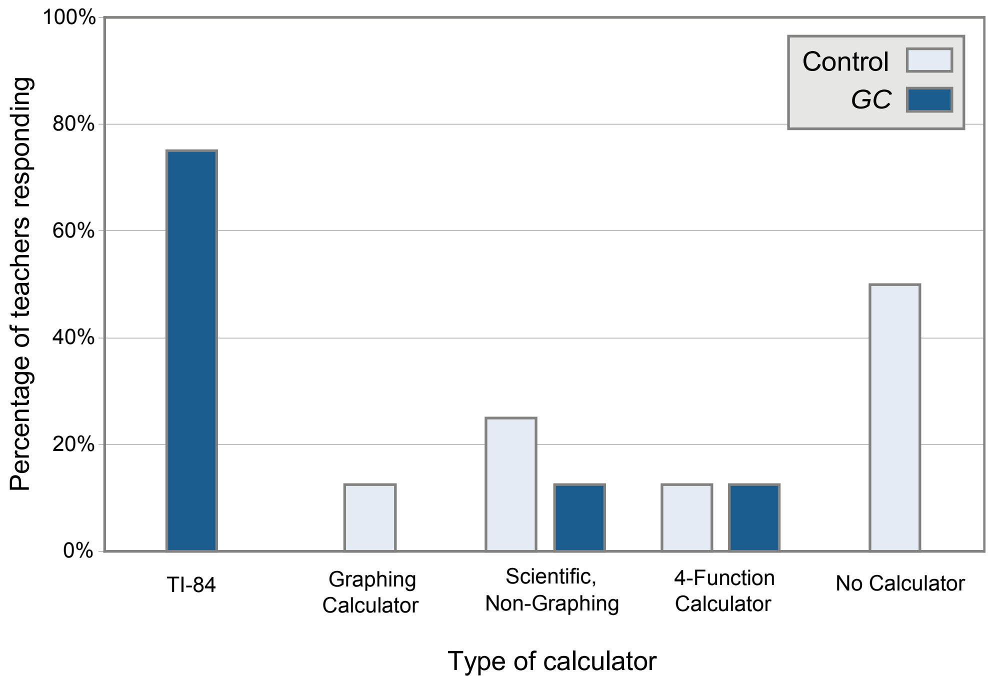

Figure B2-2. Types of Calculators Used by Students

Figure B2-3 shows the tasks for which teachers used their calculators. This question was asked at the end of February and end of April. The *GC* group responded consistently at both times, while the control group tended to report using the calculator for more tasks on the April survey.

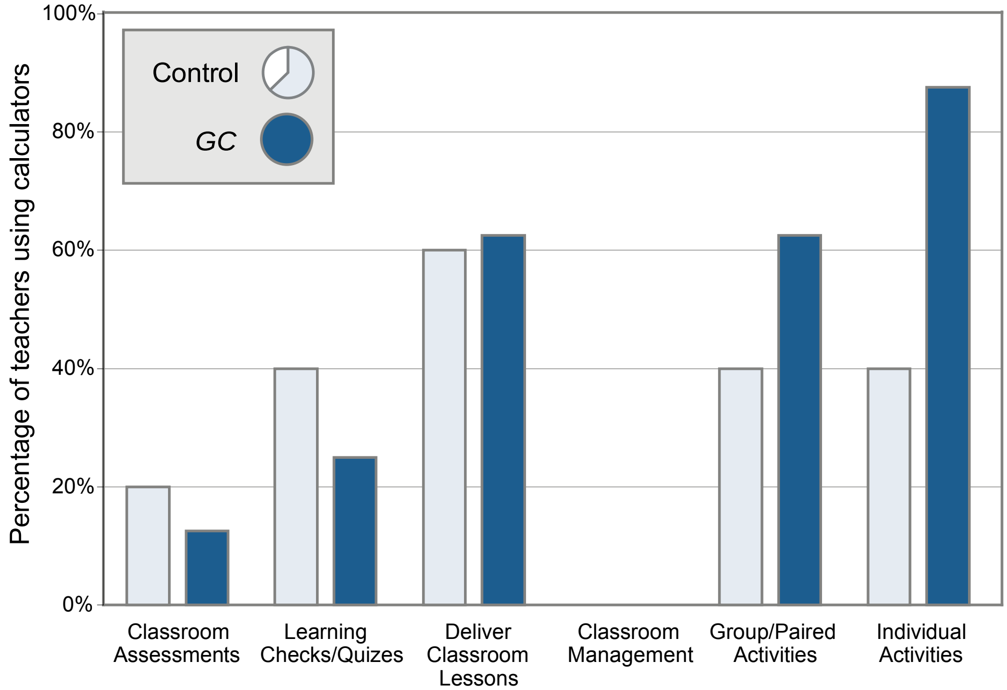

Figure B2-3. Teacher Tasks Performed with Graphing Calculator

Note. Legend pie charts reflect the percentage of teachers who responded to the survey question.

Figure B2-4 shows the classroom organization when teachers used their calculators while teaching. This question was asked at the end of February and end of April. All *GC* group teachers reported using each type of classroom organization on either the February or April survey. However *GC* teachers tended to report less individual and small group or paired instruction in April. The control teachers reported consistently across the two surveys; however more control teachers reported using whole group instruction in April. Noteworthy in Figure B2-3 is that *GC* teachers' usage, even if it is just for computation, is higher than in the control condition.

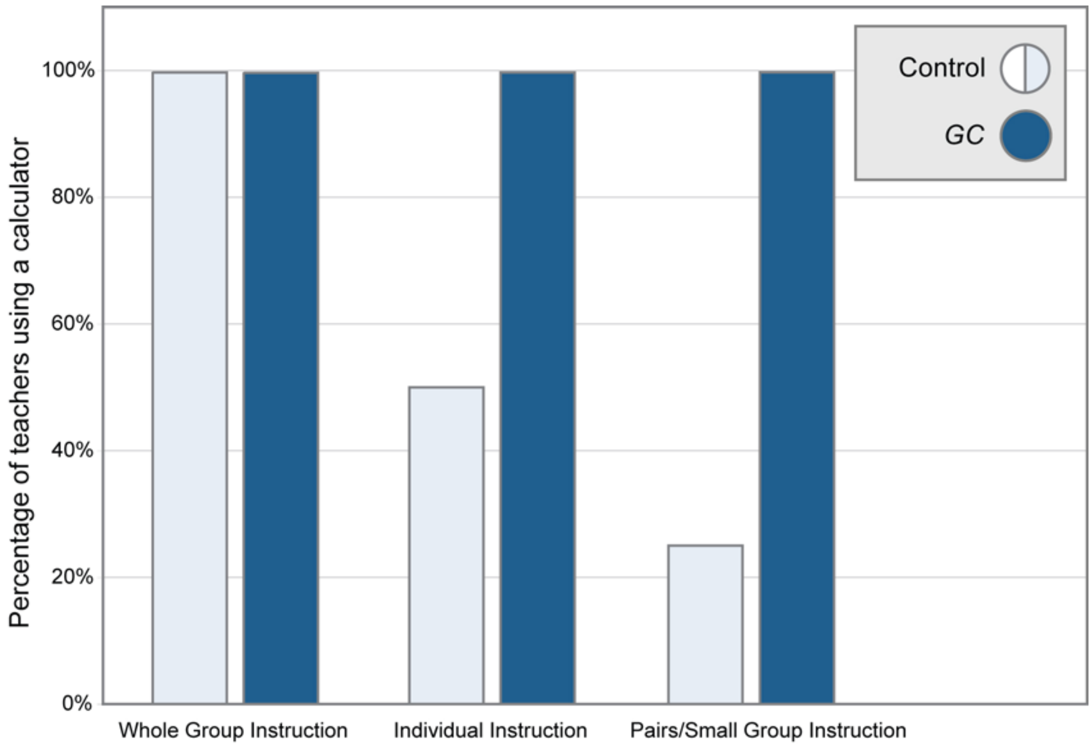

Figure B2-4. Classroom Organization While Teaching with a Calculator

Figure B2-5 shows the percentage of teachers who assigned homework that required the use of a calculator during the study. This survey question was asked in late February and again in late April. Both *GC* teachers and control teachers responded consistently over the two survey administrations.

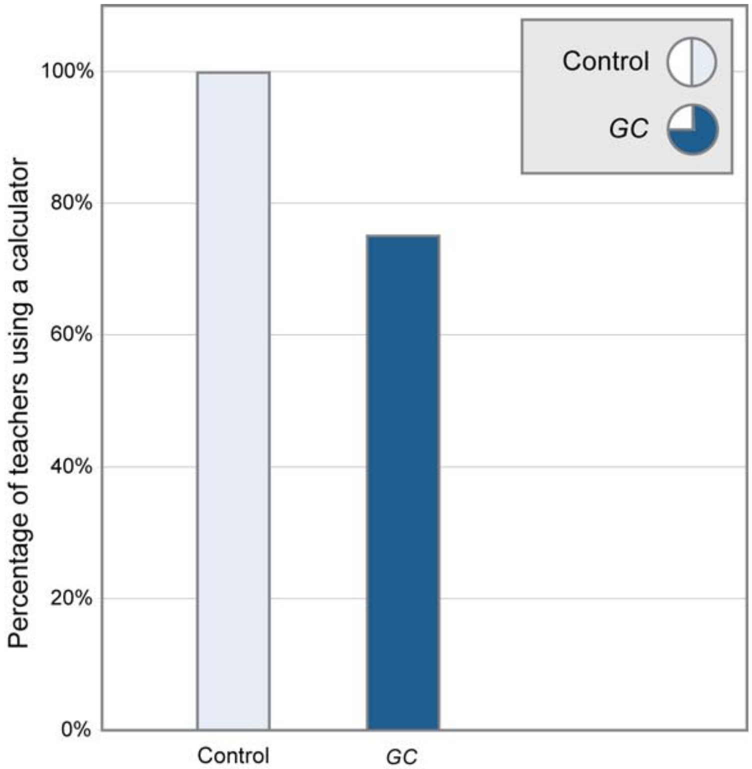

Figure B2-5. GC and Control Group Homework Assignments Using Calculators

Figure B2-6 shows how students used the calculators during class. In every *GC* classroom, students used the calculator system for individual and small group assignments during the study. Over the course of the study, *GC* teachers reported that students tended to use the calculator more than the control group for small group assignments and demonstration to the whole class. The survey question was asked in late February and again in late April. *GC* group teachers reported students using the calculator for more tasks in February than in April. Control group teachers reported consistently across both times the question was asked, but tended to report more instances of small group assignments and tests or quizzes in February.

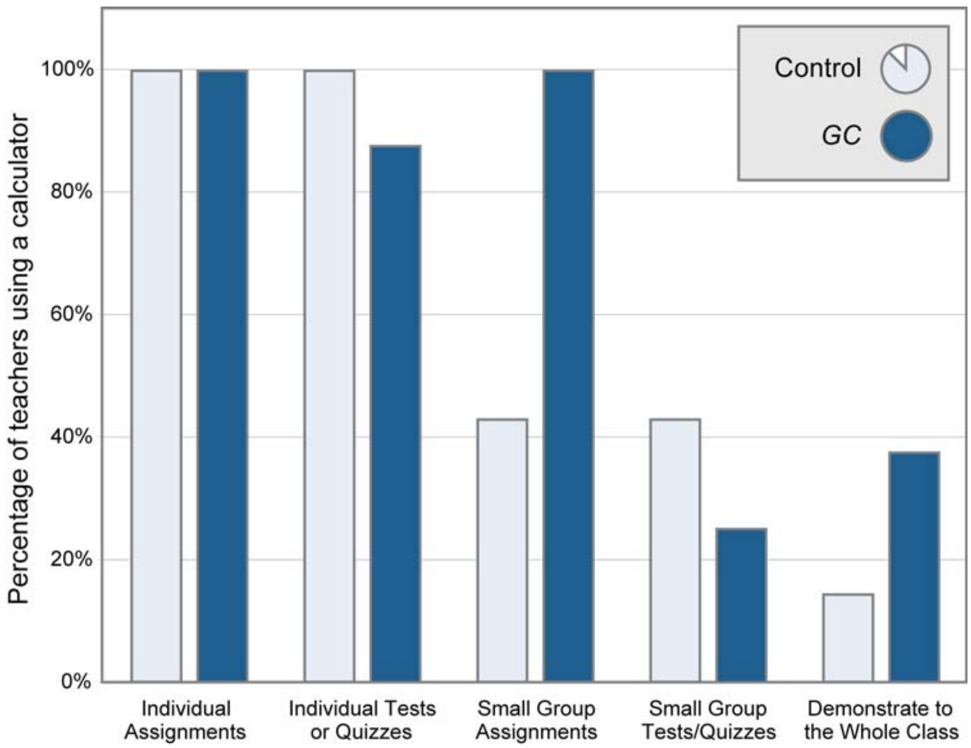

Figure B2-6. Student Calculator Use During Class

Figure B2-7 characterizes the calculator availability to the students. More than half of the *GC* teachers and all control teachers required their students to furnish their own calculators for use at home. All *GC* teachers furnished graphing calculators for use in class and three control teachers furnished at least a four function calculator for use in class. Four control teachers required students furnish their own calculator for class use and one control teacher did not respond. The question was asked once in February.

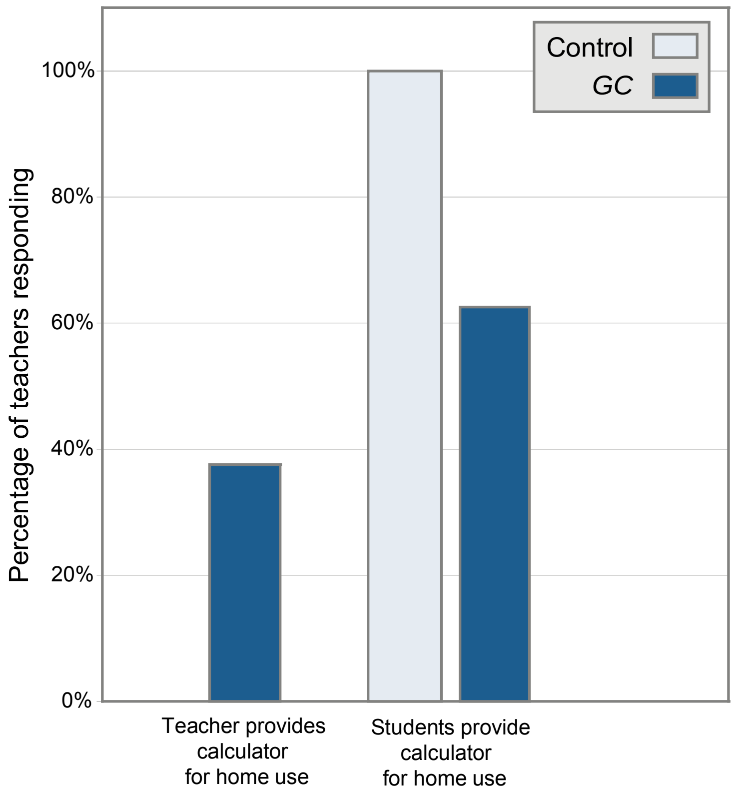

Figure B2-7. Availability of Calculators at Home

Note. One control teacher did not respond

Comparison of Calculator System Use between *GC* and *GC+Nav* Groups

The next series of figures contrasts the *GC* group against the *GC+Nav* group in classroom organization, teacher tasks, and student tasks. There was no difference between the *GC* group and the *GC+Nav* group in terms of classroom organization. Teachers in the *GC* group did not report any teacher use of the calculator for classroom assessments or student use of the calculator for small group tests or quizzes, while the *GC+Nav* group teachers reported some use in both of those categories.

Classroom Organization

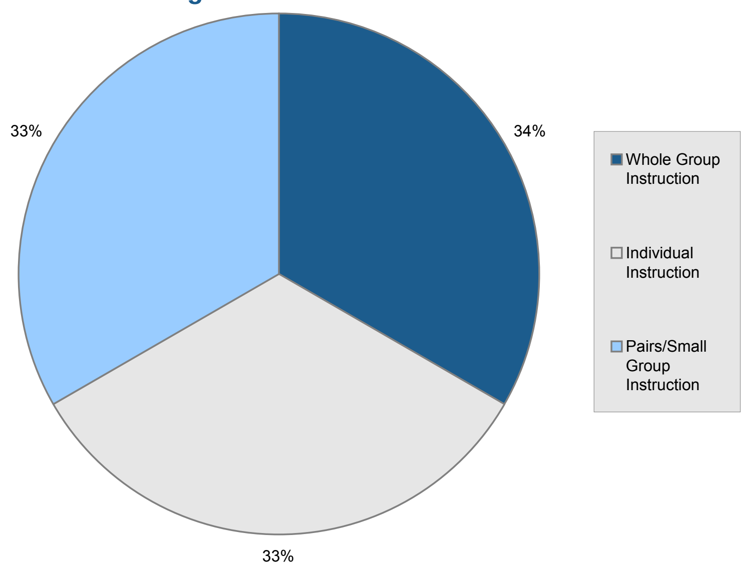

Figure B2-8. Classroom Organization by *GC* Group

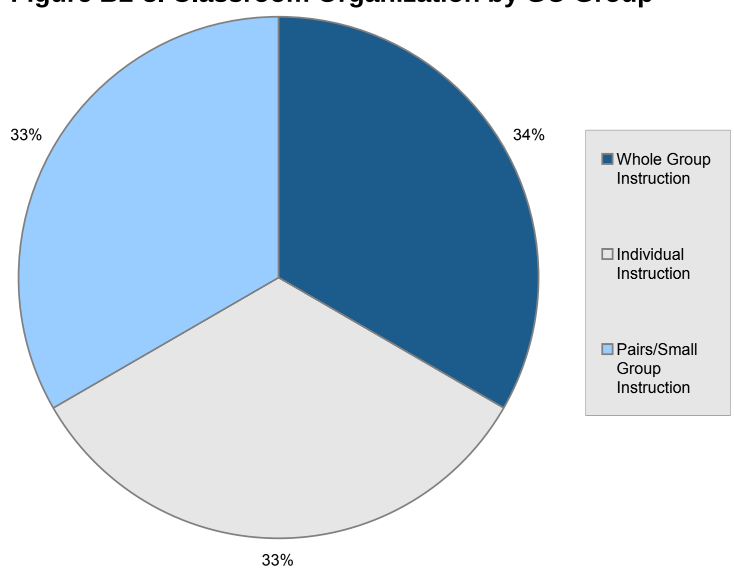

Figure B2-9. Classroom Organization by *GC*+Nav Group

Teacher Use of Calculator Systems

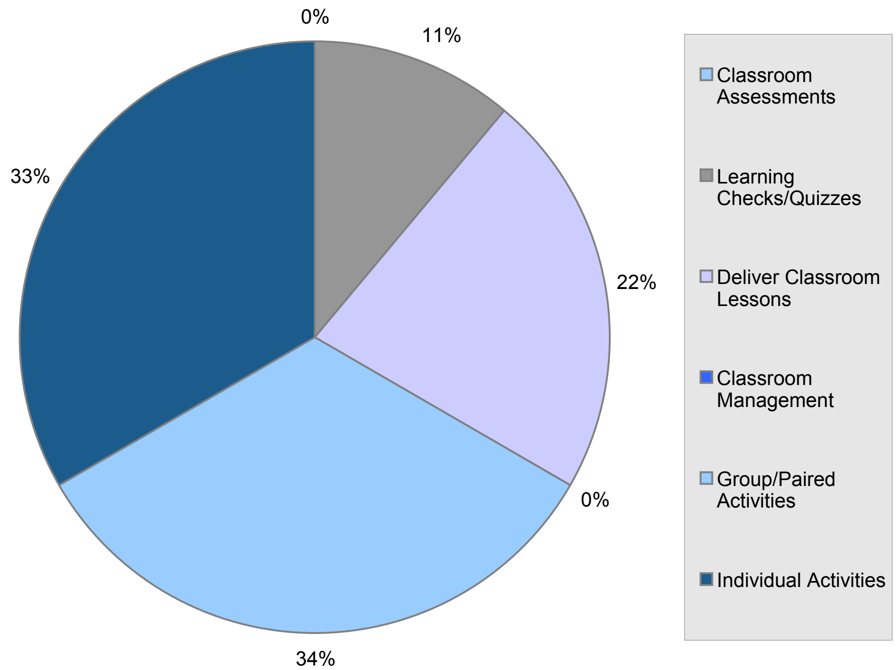

Figure B2-10. *GC* Teacher Use of Calculator Systems

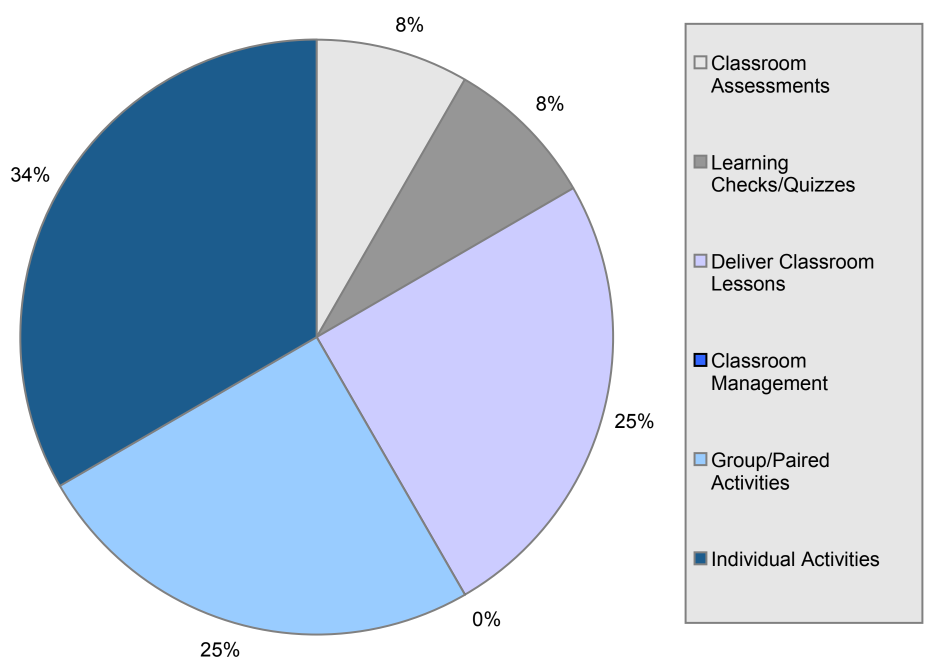

Figure B2-11. *GC*+Nav Teacher Use of Calculator Systems

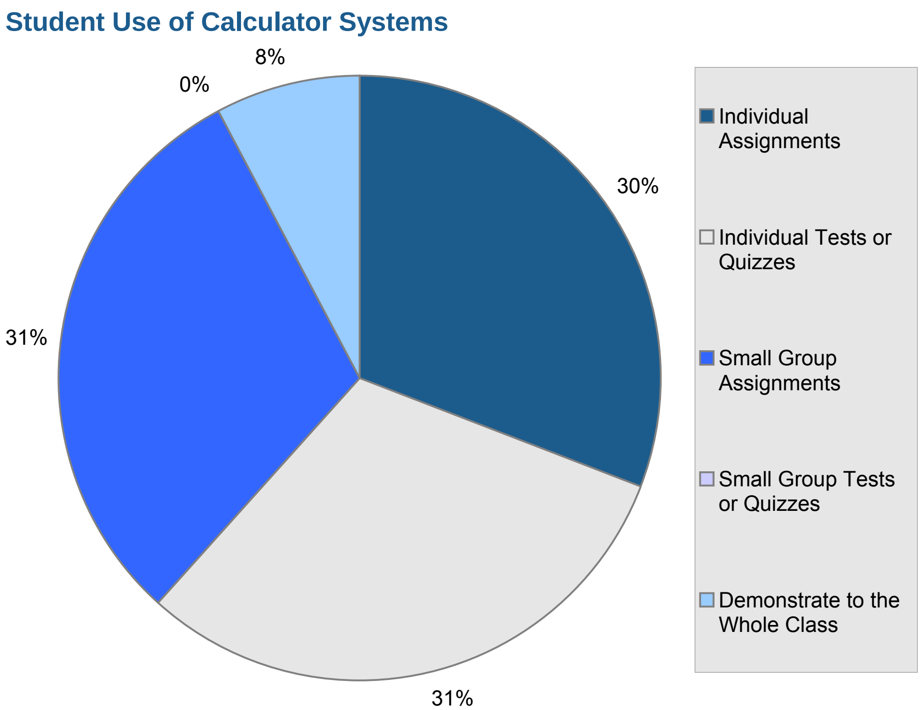

Figure B2-12. *GC* Student Use of Calculator Systems

In the following figure, note the slight increase in students demonstrating to the whole class.

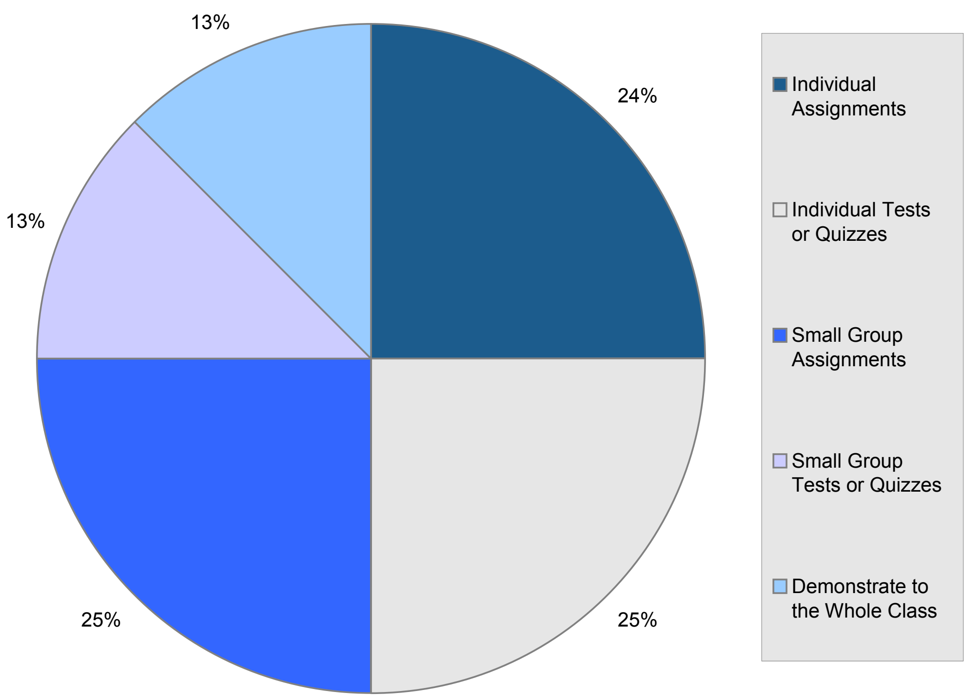

Figure B2-13. GC+Nav Student Use of Calculator Systems

Appendix C1: East Side Materials

Professional development training for the calculators was given to *GC* group teachers in August 2005. *GC* group teachers received study hardware consisting of one class set of TI-84+ Silver Edition graphing calculators for each of their Algebra I and Geometry classes, one notebook computer with TI SmartView™ software, and one InFocus® data projector in September 2005.

In December 2005, the *GC* group was randomized further into a TI-Navigator group and a *GC* stand alone group. Professional development training for the TI-Navigator system was given to the TI-Navigator group in January 2006. TI-Navigator group teachers received study hardware consisting of one TI-Navigator classroom system and one additional TI-Navigator network hub to be used with the hardware originally given to the *GC* group in September 2005.

Table C1-1. Materials Distributed

Date	Material	Number of teachers
August 2005	T3 Professional Development for Algebra and Geometry	10
August 2005	The activity resource workbooks provided to teachers during training: • Activities for Algebra with the TI-83 Plus • Cabri Jr.: Interactive Geometry Activities and Investigations • Exploring Mathematics with the Transformation Graphing Application • Exploring Mathematics with the Inequality Graphing Application • Exploring Mathematics with the Cabri Jr. Application • Exploring the Basics of Geometry with Cabri • Modeling Motion: High School Math Activities with the CBR • TI-84 Plus Graphing Calculator for Dummies • Topics in Algebra I	11
October 2005	TI-84+ Silver Edition Graphing Calculators Notebook computer with SmartView™ software InFocus® Data projector	11
January 2006	T3 Professional Development for High School Mathematics using the TI-Navigator™ TI-Navigator™ 32 User System Additional network hub for the TI-Navigator™ System The activity resource workbooks provided to teachers during training: • High School Mathematics using the TI-Navigator™ • Algebra using Real World Data: USA Today Activities for the TI-Navigator™ System	5

Appendix C2: San Diego Materials

Professional development training for the calculators was given to *GC* group teachers in August 2005. *GC* group teachers received study hardware consisting of one class set of TI-84+ Silver Edition graphing calculators for each of their Algebra I and Geometry classes, one notebook computer with TI SmartView™ software, and one InFocus® data projector in September 2005.

In December 2005, the *GC* group was randomized further into a *GC+Nav* group and a *GC* stand alone group. Professional development training for the TI-Navigator system was given to the *GC+Nav* group in January 2006. The *GC+Nav* group teachers received study hardware consisting of one TI-Navigator classroom system and one additional TI-Navigator network hub to be used with the hardware originally given to the *GC* group in September 2005.

Table C2-1. Materials Distributed

Date	Material	Number of teachers
August 2005	T^3 Professional Development for Algebra I and Geometry	9
August 2005	The activity resource workbooks provided to all teachers during training: • Activities for Algebra with the TI-83 Plus • *Cabri Jr*.: Interactive Geometry Activities and Investigations • Exploring Mathematics with the Transformation Graphing Application • Exploring Mathematics with the Inequality Graphing Application • Exploring Mathematics with the *Cabri Jr*. Application • Exploring the Basics of Geometry with *Cabri Jr.* • Modeling Motion: High School Math Activities with the CBR • TI-84 Plus Graphing Calculator for Dummies • Topics in Algebra I	11
October 2005	TI-84+ Silver Edition Graphing Calculators Notebook computer with *SmartView*™ software InFocus® Data projector	11
January 2006	T^3 Professional Development for High School Mathematics using the *TI-Navigator*™ *TI-Navigator* 32 User System Additional network hub for the *TI-Navigator* System The activity resource workbooks provided to teachers during training: • High School Mathematics using the *TI-Navigator*™ • Algebra using Real World Data: USA Today Activities for the *TI-Navigator*™ System	4

Appendix D1: East Side Observation Protocol

Research Program_TI-84______Site__________Date of Observation__________Observer__________

CLASSROOM ROUTINES USING:	Frequently	Sometimes	Rarely	Never
A. CALCULATORS				
Classroom routines are present that presume student use of calculator system (automaticity is displayed in settling down to work, logging-in, clean-up etc. with the technology, can happen at anytime within a lesson)				
Presentation by students is routine				
Teacher uses calculator to propose an *exploration* of new concept(s)				
Students use calculator to *explore* new concepts				
Teacher expected students to use a calculator on last night's homework				
Teacher uses the calculator to help the class correct the homework				
Students use the calculator to help themselves correct the homework				
B. TI-NAVIGATOR				
Teacher uses TI-Navigator to poll class				
Teacher uses TI-Navigator to collect student work				
Teacher uses TI-Navigator to share student work				
Teacher uses TI-Navigator to distribute assignments				
Teacher uses TI-Navigator activity center function				
Teacher uses the TI-Navigator screen capture to monitor students				

CLASSROOM INTERACTIONS MARKED BY SHARED RESPONSIBILITY	Frequently	Sometimes	Rarely	Never
Students ask teacher questions to further understanding of the material				
Students ask other students questions to further understanding of the material				
Teacher answers students' questions to further understanding of the material				
Teacher directs other students to answer the questions that further understanding of new material				
Students are encouraged to propose new directions for exploration				
Students take leadership roles				
Students choose from among options for exploration				

LESSON STRUCTURES THAT FOCUS ON LEARNING	Frequently	Sometimes	Rarely	Never
Teacher relates ideas to prior knowledge				
Teacher builds on student responses in introducing materials				
Teacher asks questions at different levels				
Teacher allows sufficient “wait time” for student responses				
Teacher models/enacts activity				
Students ask questions about content				
Students ask questions about skills				
Students are encouraged to answer each others’ question				
Students respond to teacher’s questions				
Teacher encourages student discussion				
Opportunities to practice are presented through hands-on activities				
Teacher explicitly states expected learning outcome				
Teacher encourages involvement of all students				
Teacher differentiates instruction				
Time is allocated for discussions about student assessment of knowledge				
Time is allocated for discussions about the range of student responses to an activity				
Teacher uses cohesive questioning strategy to elicit student higher order thinking skills				

Note. Guideline for response scale: Frequently (more than 5 occurrences); Sometimes (3-5 occurrences); Rarely (1-2 occurrences); Never (0)

Teacher denotes the leader of the instruction; Student(s) denotes the attendees at the class session.

Appendix D2: San Diego Observation Protocol

Research Program_TI-84______Site__________Date of Observation__________Observer_________

CLASSROOM ROUTINES USING:	Frequently	Sometimes	Rarely	Never
A. CALCULATORS				
Classroom routines are present that presume student use of calculator system (automaticity is displayed in settling down to work, logging-in, clean-up etc. with the technology, can happen at anytime within a lesson)				
Presentation by students is routine				
Teacher uses calculator to propose an *exploration* of new concept(s)				
Students use calculator to *explore* new concepts				
Teacher expected students to use a calculator on last night's homework				
Teacher uses the calculator to help the class correct the homework				
Students use the calculator to help themselves correct the homework				
B. TI-NAVIGATOR				
Teacher uses TI-Navigator to poll class				
Teacher uses TI-Navigator to collect student work				
Teacher uses TI-Navigator to share student work				
Teacher uses TI-Navigator to distribute assignments				
Teacher uses TI-Navigator activity center function				
Teacher uses the TI-Navigator screen capture to monitor students				

CLASSROOM INTERACTIONS MARKED BY SHARED RESPONSIBILITY	Frequently	Sometimes	Rarely	Never
Students ask teacher questions to further understanding of the material				
Students ask other students questions to further understanding of the material				
Teacher answers students' questions to further understanding of the material				
Teacher directs other students to answer the questions that further understanding of new material				
Students are encouraged to propose new directions for exploration				
Students take leadership roles				
Students choose from among options for exploration				

LESSON STRUCTURES THAT FOCUS ON LEARNING	Frequently	Sometimes	Rarely	Never
Teacher relates ideas to prior knowledge				
Teacher builds on student responses in introducing materials				
Teacher asks questions at different levels				
Teacher allows sufficient “wait time” for student responses				
Teacher models/enacts activity				
Students ask questions about content				
Students ask questions about skills				
Students are encouraged to answer each others’ question				
Students respond to teacher’s questions				
Teacher encourages student discussion				
Opportunities to practice are presented through hands-on activities				
Teacher explicitly states expected learning outcome				
Teacher encourages involvement of all students				
Teacher differentiates instruction				
Time is allocated for discussions about student assessment of knowledge				
Time is allocated for discussions about the range of student responses to an activity				
Teacher uses cohesive questioning strategy to elicit student higher order thinking skills				
Note. Guideline for response scale: Frequently (more than 5 occurrences); Sometimes (3-5 occurrences); Rarely (1-2 occurrences); Never (0) Note. Teacher denotes the leader of the instruction; Student(s) denotes the attendees at the class session.				

Appendix E: East Side and San Diego Interview Protocol

TI-84

Phone Interview Protocol

What kind of experience do you have with the graphing calculator?

What kinds of things do you consider when planning a lesson?

What kinds of things are helpful to a teacher to know, understand, or use the technology?

What does a teacher have to know in order to create the types of activities using the calculator?

Describe an activity that is very successful for the students:

Describe an activity that didn't work well with the students:

Do you think the training was sufficient?

What do students have to know about the calculators to be successful?