The Impact of Graphing Calculators on Student Performance in Beginning Algebra: An Exploratory Study

Gail Burrill
Michigan State University
Glenda Breaux
Michigan State University

Abstract

The pilot study, The Impact of Graphing Technology on Student Performance, focused on what students learned in algebra, how it was different for students with differential access to graphing calculators, the use of the technology on tasks of different cognitive demand, and whether the teachers' background and experience with graphing calculators might be related to student outcomes. The study considered two conditions: high quality professional development and high frequency calculator use on the part of the students and involved three different populations: 1) teachers who seldom or never used graphing calculators in their classrooms; 2) teachers who used graphing calculators in their classrooms but without a high degree of support and ongoing professional development; and 3) teachers with a high degree of support and ongoing professional development in the use of graphing calculators for instruction. Results indicate that access to and use of graphing calculators seems to increase achievement, achievement decreases for both users and nonusers of calculators as the cognitive demand of the tasks increases, and while the background and experience of the teachers seems to make a difference for the top 75 percent of the students, some students perform at very low levels with or without the technology.

The Impact of Graphing Calculators on Student Performance in Beginning Algebra: An Exploratory Study[1]

Since graphing calculators became available in 1985, they have been part of mathematics instruction at the secondary level in a variety of forms. Questions about whether this technology is an effective tool in teaching and learning mathematics have fueled discussion and debate ever since (Penglase & Arnold, 1996; Ruthven, 1996; Wu, 1997). Some studies found that the technology supports learning (Ellington, 2003, 2006; Graham & Thomas, 2000; Hollar & Norwood, 1999; Khoju, Jaciw, & Muller, 2005; Schwarz & Hershkowitz, 1999), while others raised concerns that the use of handheld graphing technology supplants learning or leads to dependency (Dancis, 2004; Hennessy, Fung & Scanlon, 2001; Simonson & Dick, 1997).

THE STUDY

Several issues that emerged when investigating syntheses of the literature on handheld graphing technology (Burrill et al, 2002; Ellington, 2003) laid the foundations for this pilot study, The Impact of Graphing Technology on Student Performance in Beginning Algebra. 1) Much of the research on the use of graphing technology in teaching and learning mathematics deals with mathematical content often taught in second year algebra, precalculus or calculus. Very few studies focused specifically on beginning algebra. 2) Some studies were based on short content units for which students were given the technology (Harskamp et al, 1998; Merriweather & Sharp, 1999). Van Streun and colleagues (2000), however, found that the technology made a difference for students only after a prolonged period of use but did not report the nature of student access. And while attention was paid to the attitudes and behaviors of teachers and students when using the technology, the preparation of the teachers to use the technology was rarely addressed. 3) In addition, questioned have been raised about the nature of student learning when using handheld technology (for example, Tucker, 1999). Studies often report how students perform on multi-step tasks or on tasks that require a higher level of understanding (Graham & Thomas, 2000; Thompson & Senk, 2001). Keller, Russell and Thompson (1999) reported that college students using the TI-92 with the technology were able to better solve more complex problems. But in all of these studies, what defined tasks as higher level or cognitively complex was not always clear.

As a consequence of these issues, the pilot study focused on what students learned in algebra, how it was different for students with differential access to graphing calculators, the use of the technology on tasks of different cognitive demand, and whether the teachers' background and experience with graphing calculators might be related to student outcomes. We looked at two factors in the context of the cognitive demand of tasks students were able to perform on pre-post tests. The first was the nature of student

access to the technology, including use on homework and ownership. The second was the background and experience of the teachers.

We identified three different populations: 1) teachers who seldom or never used graphing calculators in their classrooms and had little or no professional development related to their usage; 2) teachers who used graphing calculators in their classrooms but without a high degree of support and ongoing professional development; and 3) teachers with a high degree of support and ongoing professional development in the use of graphing calculators for instruction.

Teachers in the third group were randomly selected from members of Teachers Teaching with Technology (T^3), a group of teachers supported by Texas Instruments Education Technology, who attend at least one two-day professional development meeting a year, receive the most recent products related to graphing calculators and their use, and have access to training on how to use these products. These teachers, in turn, conduct workshops and seminars around the United States on using graphing calculators in the teaching and learning of secondary mathematics. Groups 1 and 2 were selected by recommendations from mathematics supervisors or department chairs in districts comparable to the teachers from the random sample of T^3 teachers. Those making the recommendations were asked to select teachers who were recognized as good teachers and who had assumed some form of leadership role at the local or state level. The final results for the pilot study involved nine teachers (three per group) from nine different school districts and 251 students. The original intent was to have at least four teachers from each group; unfortunately, district pressures about test taking and about research projects in general, teacher attrition during the study, and lack of consistent data over the study for the same group of students did not enable us to reach the desired goal. (In a full study, recruitment would have to be addressed as a critical issue in obtaining randomized matched samples).

Data were collected on student and teacher background, experience, attitudes and beliefs about mathematics and about technology, and about school environment, curriculum, and modes of instruction. Table 1 shows background data for teachers in each group.

Table 1
Teacher Background Data by Professional Development Group

PD Group*	School Location	School Size	Class/ Grade	Years experience	Text Publisher
Group 1 (non graphing calculator users)					
H	West/suburban	Over 1000	8	10	Addison Wesley
I	Midwest/suburban	500 -1000	8	2	Prentice Hall
J	West/suburban	Over 1000	8	11	McDougal Littell
Group 2 (graphing calculator users)					
E	West/urban	Over 1000	9	9	District Designed
F	Midwest/rural	Under 500	9	25	Glencoe
G	Midwest/rural	Under 500	8	17	Prentice Hall
Group 3 (T^3)					

A	Midwest/suburban	500-1000	8	29	Prentice Hall
B	Midwest/rural	Over 1000	8	17	Prentice Hall
C	South/urban	500-1000	9	30	No specific text

In Group 3, teachers' experience varied from 17 to 30 years, with an average of 25 years teaching experience; Group 2 teachers had an average of 17 years experience, and teachers in Group 1 an average of seven years. All but two of the teachers had given presentations and led workshops, in their own district or at a state or national meeting.

Information on how instruction was carried out was obtained through teacher interviews. Some teachers reported using whole group instruction as the primary mode of delivery; some reported using activities and investigations, and others reported involving students in small group work. Teacher goals for students in algebra were similarly varied: be able to pass the state test, learn by doing, describe the reasoning behind a particular mathematical outcome. Two teachers in Group 3 used the TI-Navigator, an interactive system that allows students and teachers to communicate and share information through the graphing calculator; some in Group 2 used interactive white boards and TI-Interactive. Six of the teachers taught beginning algebra to eighth graders in a middle school environment; three taught algebra to ninth graders in a high school setting. The course structures varied from 47 minutes per day to a longer block three days per week. All of these factors could potentially influence results.

To establish benchmark data for the students across each teacher, we constructed a pre-test that consisted of released items from the National Assessment of Educational Progress (NAEP) from grade 8 (1992, 1996, 2000), covering general mathematical knowledge, usually considered pre-requisite knowledge for beginning algebra. The post-test was constructed using items from published studies comparing the performance of students after calculator-based instruction versus instruction without an emphasis on calculators (see Appendix A for the pre- and posttests). All items on both tests were categorized in terms of their cognitive demand according to the mathematical task framework developed by Stein, Smith, Henningsen, and Silver (2000). A team of graduate students coded the items with disagreements mediated by the project leadership. The task framework describes lower-level tasks as those that involve memorization (e.g. committing definitions, rules, or procedures to memory) and using procedures without connections (PwoC) to meaning. Such tasks are those that:

- are algorithmic; use of the procedure is either specifically called for or is evident from prior instruction, experience, or placement of the task within the materials;
- have no connection to the concepts or meaning that underlie the procedure being used;
- focus on producing correct answers instead of on developing mathematical understanding;
- require no explanations or explanations that focus solely on describing the procedure that was used (see Figure 1).
-

Which of the following ordered pairs (x, y) is a solution to the equation $2x - 3y = 6$?

    A) (6, 3)   B) (3, 0)   C) (3, 2)   D) (2, 3)   E) (0, 3)

NAEP, 1996

Figure 1. Procedures without connections

Higher-level tasks involve procedures with connections to meaning and "doing mathematics". Procedures with connections (PwC) tasks are those that:

- explicitly or implicitly suggest pathways to follow that have close connections to underlying conceptual ideas (as opposed to narrow algorithms that are opaque with respect to underlying concepts);
- usually are represented in multiple ways, such as visual diagrams, symbols, and problem situations; making connections among multiple representations helps develop meaning;
- allow for general procedures to be followed but not mindlessly; students need to engage with conceptual ideas that underlie the procedures to complete the task successfully (See Figure 2).

A certain machine produces 300 nails per minute. At this rate, how long will it take the machine to produce enough nails to fill 5 boxes of nails if each box will contain 250 nails?

A) 4 min   B) 4 min 6 sec   C) 4 min 10 sec   D) 4 min 50 sec
E) 5 min

NAEP, 1996

Figure 2. Procedures with connections

Doing mathematics (DM) tasks are those that:

- require complex and non-algorithmic thinking; a predictable, well-rehearsed approach or pathway is not explicitly suggested by the task or task instructions;
- demand self-monitoring or self-regulation of one's own cognitive processes;
- require students to access relevant knowledge and experiences and make appropriate use of them in working through the task (see Figure 3).

The graph shows the speed (in meters per second) of a cyclist over a 10-minute period.

Is the distance the cyclist traveled greater in the first five minutes or the last?
Answer:
Reasoning:

Ruthven, 1996

Figure 3. Doing mathematics

The pre-test was given to check the initial comparability of the classes. Results indicated some significant performance differences among students of different teachers (See Figure 4). These differences tended to favor T^3 teachers (teachers 1, 2 and 3 in Group 3 from Figure 4) and disadvantage non-T^3 teachers who use calculators frequently (teachers 7, 8, and 9 in Group 2). These differences were adjusted for in the analyses related to teacher background and student performance. In addition one of the teachers in Group 2 (teacher 9) had students who performed significantly lower than students of all of the other instructors; analysis about the role of teachers was adjusted to account for this extreme difference.
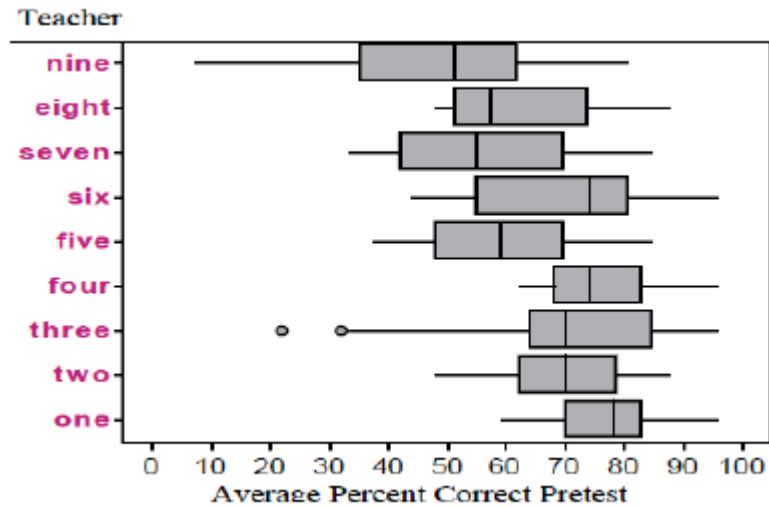

Figure 4. Percent correct on pretest by teacher[2]

In summary, pretest performance was heterogeneous among students sharing the same teacher, as well as across teachers in the same group. Within-teacher heterogeneity might have been partly due to placement policies with respect to who takes algebra at what grade level. Analyses and interpretation of posttest results take into account the initial performance differences.

RESULTS

The mean score across all of the classes on the posttest was 49 percent correct with a standard deviation of 18 percent and a range of 85 percentage points, from a low of 6 percent to a high of 91 percent correct. Figure 5 shows a scatterplot of the posttest results vs. the pretest results for all of the students. The square of the correlation, 0.487, indicates that approximately 48% of the variation in the posttest scores can be explained by knowing the pretest results, which leaves 52% of the variation in the scores due to other factors.

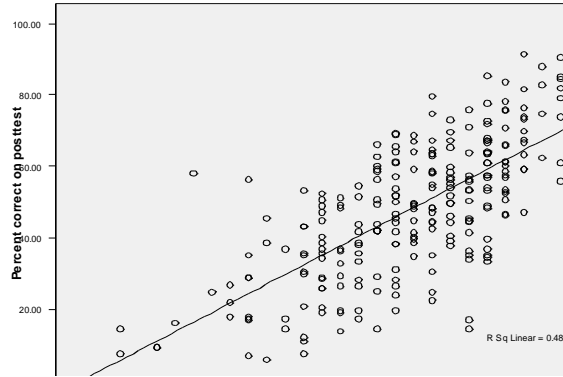[2] To preserve confidentiality, the teacher coding here is different than the coding in Table 1.

Figure 5. Scatterplot and correlation of pre-posttest results

The results of the analysis related to student experience with calculators, cognitive demand and classification of teacher with respect to professional development and training are described below.

Does student use of graphing calculators make a difference in their achievement?

The results related to this question on the posttest are analyzed using two lenses: student ownership and the use of calculators on homework (note that homework might be done during class using school provided calculators). A caution here is that student ownership might be associated with SES, which could not be factored out from the data collected. Figure 6 suggests a slight advantage to those owning a graphing calculator.
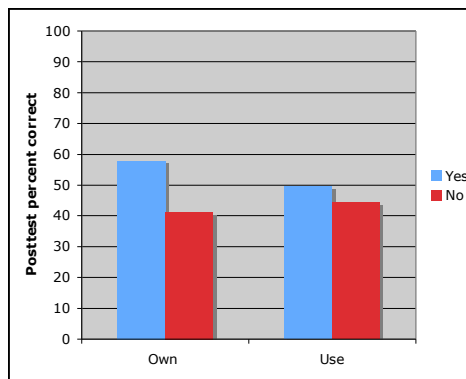

Figure 6. Posttest results related to student access to calculators

Tables 2 and 3 indicate that having access to calculators made a difference in achievement. Students who owned calculators (table 2) earned a significantly higher mean-percent correct on the posttest (57.8 percent) than those who did not (41.2 percent) with $p < 0.00$. Table 3 shows that using calculators on homework resulted in significantly higher performance on the posttest (44.4 percent correct to 53.8 percent

8

correct with $p < 0.00$.) The effect size[3] is 0.53 for calculator use on homework and 1.01 for calculator ownership. It is important to note here that use on homework and ownership are not independent, and ownership might be explained by socio-economic status. In a larger study, the design should tease out these factors.

Table 2
Posttest Mean Percent Correct by Calculator Ownership[4]

Do you own a graphing calculator		N	Mean % correct	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum Lower Bound	Maximum Upper Bound
Percent correct on posttest	no	119	41.18	16.78	1.54	38.13	41.18	5.76	91.38
	yes	117	57.84	15.99	1.48	54.92	57.84	16.97	90.52
	Total	236	49.44	18.37	1.20	47.09	49.44	5.76	91.38

Table 3
Posttest Mean Percent Correct by Calculator Use on Homework

Do you use a Calculator on Homework		N	Mean % correct	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum Lower Bound	Maximum Upper Bound
Percent correct on posttest	no	103	44.3876	17.60114	1.73429	40.9476	47.8275	5.76	91.38
	yes	127	53.8229	17.98345	1.59577	50.6649	56.9809	7.48	90.52
	Total	230	49.5975	18.38594	1.21233	47.2088	51.9863	5.76	91.38

Figures 7 and 8 show the spread related to experience with graphing calculators. Note that for the most part, some students did well with or without the calculator, which would be expected. The scores for the top half of those owning or using a graphing calculator for homework were higher than three fourths of the scores for those who did not. Some students performed at a very low level with or without a graphing calculator.

---

[3] Cohen's *d* with Hedges adjustment for sample size differences calculated using the *Effect Size Generator* program, Devilly, 2004
[4] Note that *n* represents only those students with consistent pre-posttest data for these questions.
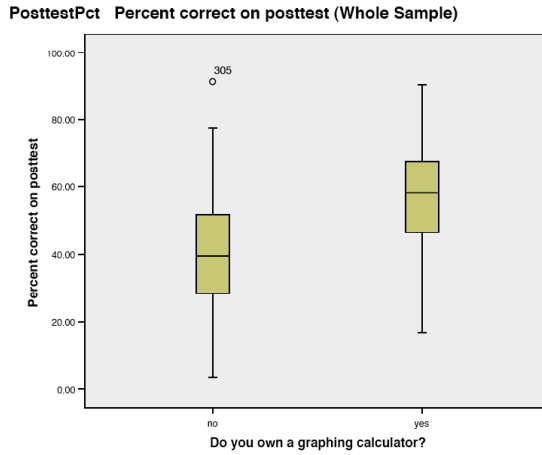
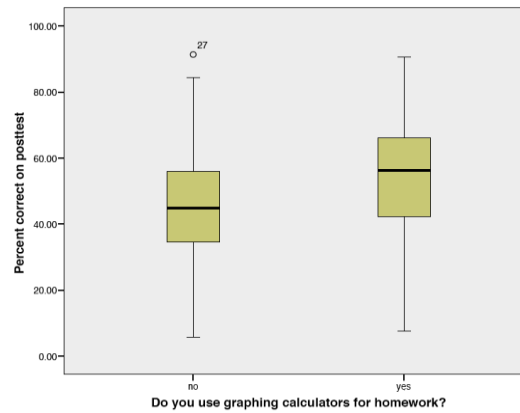Figure 7. Mean posttest score and calculator ownership

Figure 8. Mean posttest score and calculator use on homework

Is the use of calculators related to achievement levels and the cognitive demand of tasks?

Overall, as the cognitive demand increased, the achievement rate decreased (Figure 9), and the average percent correct was considerably lower for items classified as doing math. Table 4 shows individual scores varied a great deal, ranging from 0 to 100 percent correct for procedures without connections, from 0 to 98 percent correct for procedures with connections and from 0 to 82 percent correct for doing math.
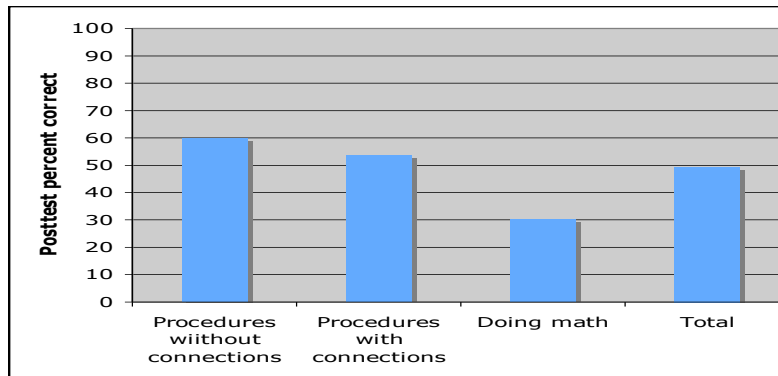

Figure 9. Mean posttest scores by cognitive demand

Table 4
Posttest Scores Related to Cognitive Demand

	N*	Mean % correct	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound	Lower Bound	Upper Bound
Procedures without connections	236	59.90	24.36	1.58	58.79	63.02	0.00	100.00

Procedures with connections	236	53.51	20.59	1.34	50.87	58.15	4.79	98.43
Doing math	236	30.07	19.00	1.24	27.63	32.51	0.00	82.14
Total	236	49.44	18.37	1.20	47.09	51.80	5.76	91.38

The box plots in Figures 10 and 11 show the variability and the shift in scores for those having access to graphing calculators. On items classified as procedures without connections three-fourths of students who owned graphing calculators scored above 52% on the posttest, while only half of those who did not own them scored above 52%; half of those who used graphing calculators for homework had a score above 63% correct, while only 25% of those who did not use calculators scored above 63%. On items classified as procedures with connections, the bottom quarter of the students who did not own a graphing calculator had a mean score of less than 32 percent while the bottom quartile scores ranged up to 42 percent for those who owned a calculator. Of those who used graphing calculators for homework, 25% more scored above 60% on the posttest than those who did not use the calculators. The differences on doing math were not as large, particularly for those using calculators on homework and generally reflect the overall low scores in this category for all students. Overall it appears that while the use of the technology generally seemed to raise the performance level of the top three fourths of the students, the performance of some students remained very low.
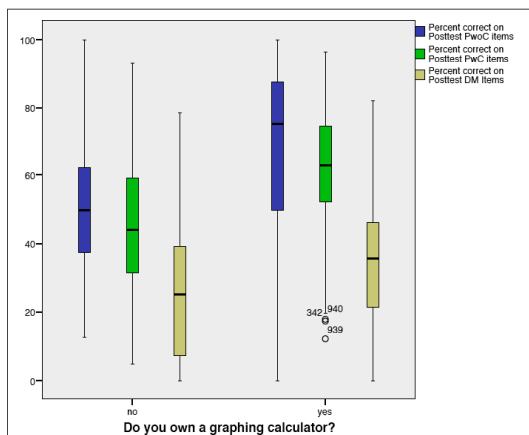

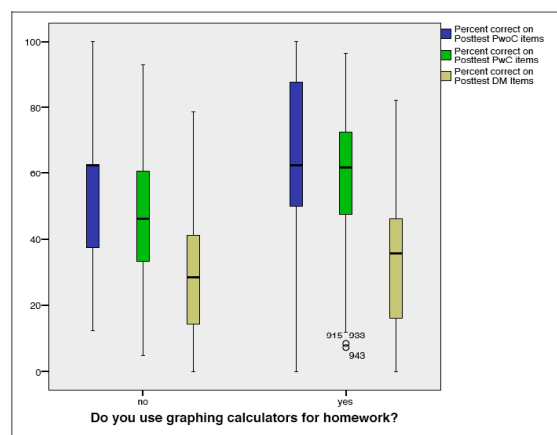Figure 10. Calculator ownership by cognitive demand

Figure 11. Calculator use on homework by cognitive demand

How was posttest performance related to professional development?

Some insight into this question can be found by investigating differences among students in the T[3] instructors' classes (PD group 3) versus students of instructors who used graphing calculators regularly in their instruction but were not T[3] instructors (PD group 2) and did not have the same opportunities for professional development and those

of instructors who made limited use of graphing calculators in their introductory algebra course (PD group 1).[5] The analysis for these data was done without respect to student calculator access. Because the pretest results indicated differences in the three groups at the onset of the study, the results on the post-test are examined in relation to pretest performance using the model $y_{ij} = \alpha_i + \beta x_{ij} + \varepsilon_{ij}$ where x_{ij} is the pretest score for student j of teacher i, y_{ij} is the posttest score for student j of teacher i, and ε_{ij} is random error (assumed independent), $j = 1, 2, \ldots, n_i$, $i = 1, 2, \ldots, 9$. The pretest score serves as a covariate and a null hypothesis of possible interest is H_0: $\alpha_{21} = \alpha_{41} = \alpha_{51} = \alpha_{61} = \alpha_{81} = \alpha_{91} = \alpha_{101} = \alpha_{111}$. The F-test rejects H_0 ($F = 528.7$, p-value < 0.0001). The raw posttest scores were adjusted by subtracting the estimated (linear) effect of the pretest score (the covariate effect) from the posttest percent score. The adjusted posttest percentage correct, y_{ij}^{adj}, is given by $y_{ij}^{adj} = y_{ij} - \hat{\beta}(x_{ij} - \bar{x})$, where $\bar{x}$ is the grand average of all pretest percent scores (excluding students of Teacher 9).

These analyses exclude the students of Teacher 9 because scores for that teacher's students were significantly lower than those for seven out of eight remaining teachers on both the pre and posttest (see Figure 4). As Teacher 9 was in the group using calculators on a regular basis (PD Group 2), those student scores had a negative effect on the entire group. This effect was magnified on the posttest where the students of Teacher 9 performed significantly worse than the students of all other teachers, and the performance gaps between this teacher's class and the others widened on the post test, while those for the other teachers contracted. The very small sample is particularly troubling in this case, and any observations have to be considered unstable.

The confidence intervals for the means of PD groups 1 and 2 overlap slightly after students of Teacher 9 are excluded. This had no effect on the relationship between the mean performance of students in PD groups 1 and 3, where the differences were not significant on the pretest but were significant on the posttest. Excluding Teacher 9 narrowed the pretest performance gap between PD group 2 and the others but did affect the significance of the differences among groups on the posttest. The change in effect size associated with membership in PD group 3 was 0.57 for the difference in the means of PD groups 3 and 1 and 1.04 for PD groups 3 and 2.

Table 5 shows the results of the adjusted scores for each group. The percent correct for the T^3 instructors high frequency use group is significantly different from the non-T^3 high frequency use group (p < 0.002) and from the low frequency users (p < 0.002). The effect size for mean differences in the posttest between PD groups 1 and 3 was 0.57, between groups 1 and 2 was 0.54, and between groups 2 and 3 was 1.04.

Table 5
Adjusted Posttest Mean Percent Correct by Professional Development Type (excluding students of Teacher 9)

PD/Usage Groups	N	Mean %	Std. Dev	Std. Error	95% Confidence Interval for Mean	Minimum	Maximum

PD/Usage Groups	N	correct			Lower Bound	Upper Bound		
1. non-T³, low frequency hgt use	62	50.31	10.62	1.35	47.61	53.01	29.99	76.09
2. non-T³, high frequency hgt use	33	48.20	11.68	2.03	44.06	52.34	12.47	64.47
3. T³, high frequency hgt use	116	55.98	12.62	1.17	53.66	58.30	21.93	80.40
Total	211	53.10	12.43	0.86	51.41	54.79	12.47	80.40

The box plots (Figure 12) suggest that while the top half of the students in the T^3 group scored higher than the top quarter of the students in the other two groups, the spread for the bottom quartile of the T^3 instructors' students was larger, and some students in that group had lower scores than students (with the exception of an outlier) in the other two groups.
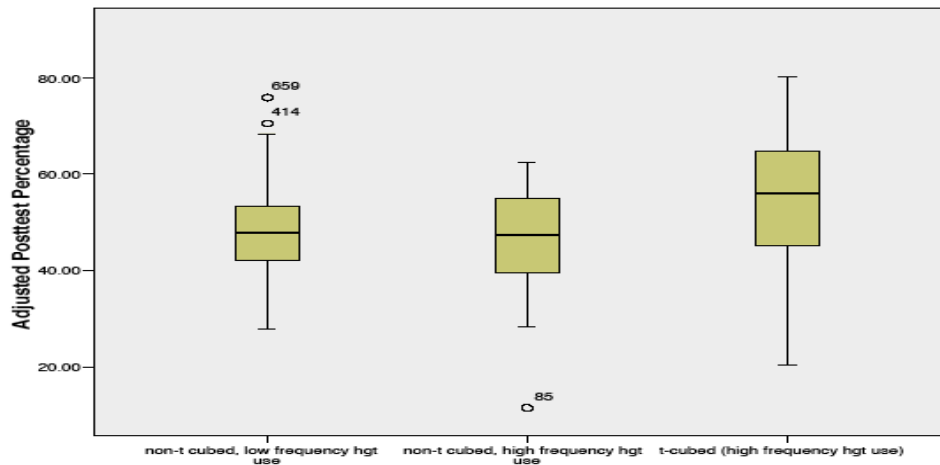

Figure 11. PD type and frequency of calculator use

The results relating professional development/calculator usage by teacher and the achievement by cognitive demand are similar to the overall results; as the cognitive demand increased, the performance level decreased. Using adjusted posttest scores (table 6), the mean percent correct for items classified as procedures without connections was significantly different for the students of T^3 instructors (65.86 percent) than the students of teachers in group 1 and group 2 (57.99 and 55.48 respectively). The mean percent correct for students of teachers in group 1, those using little technology, was greater than the mean for students of teachers in group 2, but the difference was not significant.

Table 6
Adjusted Posttest Mean Percent Correct by Professional Development Type/Cognitive Demand (excluding students of Teacher 9)

PD/Usage Groups	N	Mean % correct	Std. Dev	Std. Error	95% Confidence Interval for Mean	
					Lower	Upper

					Bound	Bound
Procedures without connections						
non-T^3, low frequency hgt use	62	57.99	17.10	0.86	53.65	62.33
non-T^3, high frequency hgt use	33	55.48	15.37	2.68	50.03	60.93
T^3. high frequency hgt use	116	65.86	21.61	2.01	61.88	69.83
Procedures with connections						
non-T^3, low frequency hgt use	62	53.51	14.23	1.81	49.89	57.12
non-T^3, high frequency hgt use	33	51.26	18.35	3.19	44.75	57.77
T^3, high frequency hgt use	116	58.57	14.92	1.07	53.83	58.05
Doing mathematics						
non-T^3, low frequency hgt use	62	41.05	8.38	1.06	38.92	43.18
non-, high frequency hgt use	33	44.39	10.09	1.76	40.81	47.97
T^3, high frequency hgt use	116	39.09	8.82	0.82	37.47	40.71

On items classified as procedures with connections, the mean percent correct for classes in the T^3 group (3), 58.57, was significantly higher than the means for the other two groups ($p < 0.01$ and $p < 0.02$). For doing mathematics items, the mean percent correct for all groups continued to decrease; students of non-T^3 teachers who used graphing calculators had a mean that was significantly higher than the other two groups ($p < .004$), and the mean percent correct for the T^3 instructors was the lowest of all three groups. Note that for this category, the standard deviations for all three groups were much smaller than for the other categories, around 9 percent, and uniformly the mean scores were low.
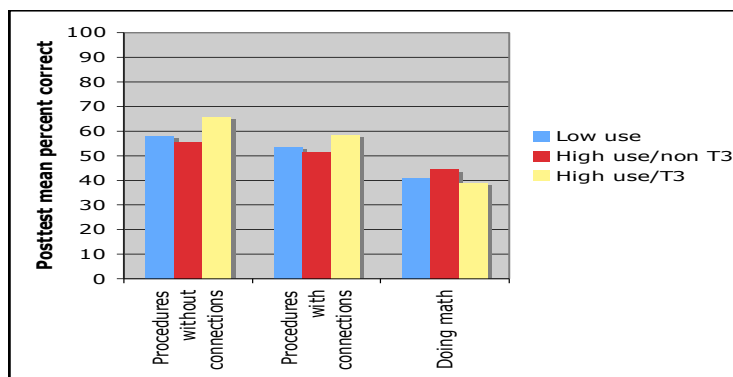

Figure 13. Adjusted posttest scores by professional development group and cognitive demand

The bar graph of these data (Figure 13) makes visible the decreasing achievement scores as the complexity of the questions increases. The graph also shows that the students of T^3 instructors had significantly higher mean percentages correct on problems categorized as procedures with and without connections.

RELEVANCE OF THE STUDY

Collecting reliable and robust information about what actually happens in classrooms is not easily done. This study highlights some of the critical issues that emerge when designing and conducting research into the use of technology in school classrooms. Specifically, these relate to obtaining and maintaining an intact and large enough sample when encountering obstacles such as districts' emphasis on only interventions directly related to high stakes assessments, teachers dropping out of the study, or reassignment of teachers to other classes or duties. Another area of difficulty is obtaining complete and reliable data from all of the students due to teacher fidelity to the study protocols, student (and teacher) absences and changing student populations. A third issue is the role that unexpected new influences such as the TI-Navigator or an interactive white board brought into the classroom during the study might play in the results. This pilot study, designed to look at teachers across a broad spectrum of schools, did not identify and control for factors that might offer alternate explanations for results, such as instructional materials, classroom and cultural norms, teacher experience or school policies. The focus and nature of a course called algebra can also vary in terms of content, expectations for students, rigor, and intent.

As a pilot with sample sizes too small to justify any real conclusions, the study does point to several areas that would benefit from further investigation and research. The initial findings related to cognitive demand demonstrate the importance of looking for such differences in studies about the use and effectiveness of handheld technology. Just using grand means obscures what may be happening, which suggests more attention should be given to item-level analysis. The performance on tasks rated by cognitive demand in the study highlights fairly convincingly the need to make more significant progress in engaging all students with higher-level mathematics tasks and to think more carefully about how the technology is being used in creating and implementing challenging tasks. The study provides some evidence that access and use of handheld graphing technology should routinely be part of the learning process if they are to be effective tools for learning, which suggests that frequency (and quality) of use of the technology needs to be taken into account and not just the presence of the technology. The study also suggests the role of professional development in helping teachers understand how to maximize the potential of graphing calculators in teaching beginning algebra might be important. Perhaps most importantly, however, looking only at significant gains in the differences between means hides indicators that these results do not seem to hold for students in the lower quartile, regardless of the professional development and experience of the teacher.

REFERENCES

Burrill, G., Breaux, G., Kastberg, S., Leatham, K., & Sanchez, W. (2002). Handheld graphing technology at the secondary level: Research findings and implications for classroom practice. Dallas, TX: Texas Instruments Corp. http://education.ti.com/research

Graham, A. T., & Thomas, M. O. J. (2000). Building a versatile understanding of algebraic variables with a graphic calculator. *Educational Studies in Mathematics, 41*(3), 265-282.

Dancis, J. (Feb. 24, 2004). Be wary of leading MD students down the California Path to Remedial Math. Presented at the MD State Board of Education meeting.

Devilly, G. J. (2004). The Effect Size Generator for Windows: Version 2.3 (computer programme). Centre for Neuropsychology, Swinburne University, Australia.) http://www.swin.edu.au/victims/resources/software/effectsize/effect_size_generator.html

Ellington, A. J. (2003). A meta-analysis of the effects of calculators on students' achievement and attitude levels in precollege mathematics classes. *Journal for Research in Mathematics Education*, 34(5), 433-463.

Ellington, A. J. (2006). The effects of non-CAS graphing calculators on student achievement and attitude levels in mathematics: A meta-analysis. School Science and Mathematics, 106(1), 16-27.

Harskamp, E. G., Suhre, C. J. M., & van Streun, A. (1998). The graphics calculator in mathematics education: An experiment in the Netherlands. *Hiroshima Journal of Mathematics Education, 6, 13-31.*

Hennessy, S., Fung, P., & Scanlon, E. (2001). "The role of the graphic calculator in mediating graphing activity". *International Journal of Mathematical Education in Science and Technology, 32, 267-290*

Hollar, J. C., & Norwood, K. (1999). The effects of a graphing-approach intermediate algebra curriculum on students' understanding of function. *Journal for Research in Mathematics Education,* 30(2), 220-226.

Keller, B. A., Russell, C., & Thompson, H. (1999). "A large-scale study clarifying the roles of the TI-92 and instructional format on students success in calculus" *International Journal of Computer Algebra in Mathematics Education, 6, 191-207*

Kohju, M., Jaciw, A., & Miller, G. (2005). *Effectiveness of graphing calculators in K–12 mathematics achievement: A systematic review*. Palo Alto CA: Empirical Education Inc.

Merriweather, M., & Tharp, M. L. (1999). "The effect of instruction with graphing calculators on how general mathematics students naturalistically solve algebraic problems" *Journal of Computers in Mathematics and Science Teaching, 18 (10), 7-22.*

National Assessment for Educational Progress Released Item. (1996). National Center for Educational Statistics.

Penglase, M., & Arnold, S. (1996). The graphic calculator in mathematics education: A critical review of recent research. *Mathematics Education Research Journal*, 8, 58-90.

Ruthven, K. (1996). Calculators in the mathematics curriculum: The scope of personal computational technology. In C. Laborde (Ed.), *International handbook of mathematics education,* (435-468). Dordrecht, The Netherlands: Kluwer.

Schwarz, B. B., & Hershkowitz, R. (1999). Prototypes: Brakes or levers in learning the function concept? The role of computer tools. *Journal for Research in Mathematics Education, 30*(4), 362-389.

Simonson, L. M., & Dick, T. P. (1997). "Teachers' perceptions of the impact of graphing calculators in the mathematics classroom". *Journal of Computers in Mathematics and Science Teaching, 16(2), 239-268*

Thompson, D. R., & Senk, S. L. (2001). "The effects of curriculum on achievement in Second-Year Algebra: The example of the University of Chicago School Mathematics Project" *Journal for Research in Mathematics Education, 32, 58-84*

Stein, M., Smith, M., Henningsen, M., Silver, E., (2000). *Implementing standards-based mathematics instruction: A casebook for professional development*. New York: Teachers College Press.

Tucker, *T*. W. (1999). *The American Mathematical Monthly*, 106(10), 910-914. http://www.jstor.org/stable/2589744

Van Streun, C., Harskamp, E., & Suhre, C. J. M. (2000). "The effect of the graphic calculator on students' solution approaches: A secondary analysis." *Hiroshima Journal of Mathematics Education*, 8, 27-39.

Wu, H. (1997). The mathematics education reform: Why you should be concerned and what you can do. *American Mathematical Monthly*, 104, 946-954.