The Effects of the Use of Handheld Devices on Student Achievement in Algebra.[1]

Irina Lyublinskaya[2] and Nelly Tournaki

Department of Education, College of Staten Island, City University of New York

Abstract

This study compares the achievement of students, enrolled in an integrated algebra course, taught with two different types of handhelds over a period of one year. One group was taught with TI-Nspire handhelds and was compared with another group taught with TI-84 graphing calculators. The teachers of both groups received on-going professional development. Student achievement was measured via a midyear school test, fall and spring semester grades, and New York State Regents exam scores. Results indicated that the group taught with TI-Nspire outperformed the other group in all assessments except the Regents exam. Further, analysis of scores indicated that girls outperformed boys in all assessments except the Regents exam, while there were no differences in achievement by race.

Keywords: handheld technology; gender differences; secondary education; algebra; student achievement.

---

[1] Submitted to Proceedings of the SITE 2009 conference

[2] Author's email: lyublinskaya@mail.csi.cuny.edu

The Effects of the Use of Handheld Devices on Student Achievement in Algebra

Just a few years ago, handhelds were widely considered to be organizers with limited memory and low-resolution grayscale screens. Today, educators are using handhelds for word-processing, Internet browsing, PowerPoint, grading, scheduling, attendance, lesson-planning, e-mailing, and data collection. There are literally thousands of software applications for handheld devices (Curtis, 2005).

Given their relatively low cost (e.g., $100 to $400), lower maintenance and support cost, security, portability, ease of use (minimum computer skills required), flexibility (ready availability), and battery life (typically 6 hours to more than 100 hours), handhelds are becoming an increasingly compelling choice of technology for K-12 classrooms because they enable a transition from the occasional, supplemental use of classroom computers and school computer labs to the frequent, integral use of portable computational devices (Soloway, Norris, Blumenfeld, Fishman, Krajcik, & Marx, 2001; Tinker & Krajcik, 2001). Research indicates that teachers and students respond favorably to handheld devices, which have the potential to affect learning positively across curricular topics and instructional activities. Teachers have indicated that students are more motivated, spend more time using technology, collaborate and communicate more, and in general benefit when they are given a portable and readily accessible tool (Vahey & Crawford, 2002). Students, in turn, have found handhelds easy to use, fun, and a useful for learning (van't Hooft, Diaz, & Swan, 2004). Further, researchers argue that classrooms with handheld computers differ fundamentally from those that have desktop computers in that users of handhelds interact with other computing devices as well as with each other at the same time (Cole & Stanton, 2003; Danesh, Inkpen, Lau, Shu, & Booth, 2001; Mandryk, Inkpen, Bilezkjian, Klemmer, & Landay, 2001; Roth, 2002). Overall, then handheld computers have the potential to support both personalized and collaborative learning. Roschelle and Pea (2002), for example, highlight three ways in which handheld devices have been used to increase learning collaboratively - classroom response systems, participatory simulations, and collaborative data gathering - and suggest there are many more such uses (Danesh et al., 2001; Mandryk et al., 2001; Roschelle, 2003). Finally, because of their small size, handheld computers support learning outside the classroom, any time or on any day of the week (Bannasch, 1999; Inkpen, 2001; Sharples, 2000; Soloway et al., 2001; Staudt & Hsi, 1999; Tinker, 1996).

The above mentioned advantages of handhelds – favorable responses from teachers and students, interaction with other computer devices, support for personalized and collaborative learning, support of learning outside the classroom – would do us no good if we didn't have evidence that handhelds increase student achievement. Given that in the past decades graphing calculators have been the most common form of handheld devices used in the classroom we will review the literature on their effectiveness. A meta-analysis by Ellington (2003) reviewed an inclusive set of 54 high-quality experimental studies. This meta-analysis shows a reliable positive effect of graphing calculator based interventions upon student achievement. In addition, the studies suggest that when graphing calculators are allowed on tests, gains extend from calculations and operations to conceptual understanding and problem solving performance. A second meta-analysis looked specifically at algebra. Khoju, Jaciw, & Miller (2005) screened available research using stringent quality-control criteria published by the U.S. Department of Education's What Works Clearinghouse. They selected eight high-quality studies examining the impact of graphing calculators on K-12 mathematics achievement. Four of these studies specifically assessed the impact on algebra learning. Across a wide variety of student populations and teaching conditions, use of graphing calculators with aligned instructional materials was shown to have a strong, positive effect on algebra achievement. Further, a study by Heller (2005) described and studied a model implementation, which included a new textbook, teacher professional development, and assessment—all aligned with the graphing technology by the theme of "Dynamic Algebra." This study establishes that the teachers and students who used graphing calculators most frequently learned the most. The National Center for Educational Statistics signature report, "The Nation's Report," (National Center for Education Statistics, 2001, p. 144) provides confidence that frequent use at the eighth grade level (but not at the fourth grade level) is associated with greater mathematics achievement, stating:

> Eighth-graders whose teachers reported that calculators were used almost every day scored highest. Weekly use was also associated with higher average scores frequent use. In addition, teachers who permitted unrestricted use of calculators and than less those who permitted calculator use on tests had eighth-graders with higher average scores than did teachers who did not indicate such use of calculators in their classrooms. The association between frequent graphing calculator

use and high achievement holds for both richer and poorer students, for both girls and boys, for varied students with varied race and ethnicity, and across states with varied policies and curricula.

Kastberg & Leatham (2005) warn teachers that the mere addition of calculator-based technology is not enough in order to develop deep conceptual understanding, and therefore achievement, of mathematics. Such technology needs to be used in an inquiry-based learning environment, which makes it possible for the students to get involved in collecting real time data, generating hypotheses, analyzing data, and drawing conclusions (Lyublinskaya, 2003a & b). Such a learning environment increases students' understanding of mathematical concepts and processes (Niess, 2001). Research indicates that it takes a great deal of education and experience to achieve a comfortable level of expertise in the use of technology as a teaching tool for helping students to learn (Fleener, 1995; Thomas & Cooper, 2000). Thus, when new technology is introduced to the teachers, there is a great need to address the pedagogical issues surrounding the use of this technology by providing teachers a forum to examine their pedagogical perspectives for using this technology in teaching and to explore when and how to use it in the classroom.

The scientific calculator found its place in mathematics education twenty-five years ago. But now the scientific calculator should be replaced with graphing and data collecting hand-held technology. Developmental/remedial mathematics teachers must accept the challenge of learning how to use the newest technology appropriately at the developmental level -- arithmetic through intermediate algebra (Laughbaum, 2003).

Up to this point research studies compare the effects of technology as they relate to absence of technology. That was necessary in order to provide evidence to the educational community that technology improves student achievement. Now that this has been established, we enter the era of having to look at the efficacy of new generations of technology as compared to old, and this is what the present study is examining. In our study, the student participants in the experimental group, received instruction with the new TI-Nspire handhelds and the control group received instruction through TI-84 graphing calculator – an existing technology. New generations of technology provide a variety of features that need to be tested empirically for their effectiveness.

In addition to the manipulation of technology, another component of our study was that all teacher participants were supported through a year-long professional development. The reason for the inclusion of the professional development component was that in recent years many schools that demonstrated improvement in student learning have used: a) technology as integral to their transformation and b) have shown that effective professional development is needed that helps teachers use technology to change how they teach core subject areas (Wolf, 2007).

The research questions

The following research questions guided the study: was TI-Nspire supported instruction more effective than the instruction with graphing calculators in improving high school students' algebra achievement in different level courses as measured by (1) mid-year assessment; (2) fall semester grades; (3) spring semester grades; and (4) scores on NYS Regents mathematics exam? Finally, (5) how are items (1)-(4) distributed by gender and race?

In the present study we measured student achievement through a variety of measures in order to detect whether different measures produce different results. More specifically, we compared teacher grades (fall and spring) to a school-made test, to a State-wide standardized test. We hypothesized that differences between our two groups will be more apparent when students are assessed by their teachers and less apparent when they are assessed through the state–wide test (i.e. NYS Regents). This is because large-scale assessments often face the difficulty of needing to be fair to multiple programs and educational jurisdictions – often with quite different philosophies and curricula. So, historically, they solved the problem by testing the intersection of various curricula, focusing on lowest-common-denominator, basic skills (Stancavage, Shepard, McLaughlin, Holtzman, Blankenship, Zhang, 2009).

Method

Student Participants

A High School in New York City was used as a site for this project. Its enrollment was 2,856, which is 160% of its planned size. The diverse population was roughly evenly divided among White (27%), Hispanic (30%), and African-American (37%) students, with a small Asian representation (6%). With 42% of students on free/reduced lunch, the school qualifies for substantial Title I aid. 7% of the students are ELL and 14% of the students are Special Education students. 58% of all students achieve passing grades in mathematics and achieve at least a 65 on NY State Regents mathematics exams (NYC DOE report card 2007-2008).

The study was conducted with 13 sections of Integrated Algebra (formerly NY Mathematics A) at three attainment levels:

- Level F: regular freshmen integrated algebra, students performed at the grade level (levels 3 and 4) on the 8th grade city mathematics exam – 1 experimental, 3 control sections
- Level R: reduced size freshmen integrated algebra for students who performed below grade level (levels 1 and 2) on the 8th grade city mathematics exam – 3 experimental, 2 control sections.
- Level A: repeater section of integrated algebra for students who did not pass Math A Regents exam as freshmen – 1 experimental, 3 control sections.

There were total of 294 students in experimental (TI-Nspire) and Control (TI-84) groups at the beginning of the school year with 131 students registered in the five sections in the experimental group and 163 students in the eight sections in the control group. The composition of experimental and control groups by course level, race, and gender are shown in Table 1.

Place Table 1 here

The school has high student mobility and dropout rate. By January the number of students dropped to 105 in experimental group and to 178 in control group. By June there were 107 students in experimental group and 171 students in control.

Teacher Participants

Experimental group: Four teachers volunteered to participate in the experimental group, knowing that they would receive training on a new technology, namely on TI-Nspire technology and professional development throughout the year. The descriptives for the teachers in the experimental group are shown in Table 2.

Place Table 2 here

Control group: Of all teachers assigned to teach Integrated Algebra three volunteered to participate in the control group, knowing that they will receive professional development on the existing technology used in the school. In addition, one teacher from experimental group also taught two sections of algebra in the control group (i.e. using TI-Nspire in one of the sections and TI-84 graphing calculators in two sections of the course). The descriptives for the four teachers in the control group are shown in Table 3.

Place Table 3 here

Further, multiple regression analysis was employed to determine whether there was a contribution of any teacher effects onto the study's dependent variables. More specifically, "teacher age" and "teacher experience" were put into the regression equation along with the independent variables of "control or experimental" and "course level". The R square of .24 for midyear assessment, .080 for fall semester grades, .081 for spring semester grades, and .032 for Regents indicates that only a small and not significant, amount of the variance is explained in students' scores by the regression models.

Study Design

The baseline was determined using an algebra readiness multiple-choice test developed by the school and administered in September 2007. An independent *t*-test was used to compare the scores of the experimental and control groups prior to the treatment. Weekly professional development workshops started in October 2007 for both groups of teachers and continued through mid-May 2008. TI-Nspire technology was introduced to the experimental classrooms in mid-October, while TI-84 graphing calculators were used with control group students from the beginning of the school year. Student surveys were used to determine frequency of technology use in the classroom. The survey results indicated that technology was used on average 1-2 times per week.

Algebra grades and scores were analyzed using a 2×3 factorial analysis of variance (ANOVA). The two independent variables were Technology Instruction (Experimental group using TI-Nspire vs. Control group using TI-84 graphing calculators) and Student Level (F/students on grade level vs. R/students who failed and repeat the same subject vs. A/students below grade level taking this course for the first time). The dependent variables were: (1) Midyear assessment scores in integrated algebra (a multiple choice test developed by the school's mathematics department) administered in January to assess students' progress towards NYS Regents Math A exam) (2) Fall semester grades of integrated algebra, given by teachers in February. (3) Spring Semester grades of integrated algebra, given by teachers in June. (4) New York State (NYS) Regents Math A exam scores from the exam administered at the end of June.

Students were assigned to different levels of integrated algebra courses by the school administration based on their performance on NYC 8th grade mathematics exam or NYS Mathematics Regents exam. Within each course level, the placement was made to maintain equal diversity by race, gender, and socio-economic status between the sections.

Materials

Technology Used in the Experimental and Control Groups.

The control group used classroom sets of TI-84 graphing calculators because that was the technology already adopted by the school. The experimental group used classroom sets of TI-Nspire™ handhelds. TI-Nspire was chosen for the experimental group because it represents not only a new generation of graphing calculator technology, but also an advance in the capabilities of a low-cost personal computing device that is reliable and easy-to-use, supporting a broad range of instructional models and advanced modes of assessment for teaching mathematics. TI-Nspire™ incorporates two new capabilities not available previously:

a. The ability to display multiple representations which are connected and in a single plane. That is, the multiple representation capability dynamically links graphical curves, axes, equations and tables in simultaneous displays, such that a change in one representation is transmitted to the others. This feature allows teachers to design new tasks for their students to address the NCTM standards focusing on connections between algebraic and geometric representations, on inquiry-based approaches to teaching and learning mathematics, etc.
b. The ability to document content. This document-based content system is an organized presentation of multiple screens of mathematics, which can be saved, shared, annotated, and revisited, giving teachers new ways of assessing their students' understanding of mathematics and technology.

In both groups students had access to handhelds during mathematics lessons. However, they were not available to them at home.

Instruments Used to Measure the Dependent Variables

Mid-year assessment was developed by the mathematics department as a multiple-choice test similar to NYS Regents exam that covered topics of the curriculum students have learned by January.

Fall and Spring semester grades were assigned to students by teachers based on their performance on in-class and out-of-class assignments, effort, and participation.

New York State (NYS) Regents Mathematics A exam scores were reported by the school. The Regents Examination is one of several large-scale tests developed to measure the extent to which schools throughout New York State achieve the content, concepts, and skills contained in the learning standards and core curricula.

Professional development model

In the present study teacher participants in both the control and experimental groups were supported through a year-long professional development.

Introductory Workshop to TI-Nspire Technology

In August 2007, the four high school mathematics teachers including the Mathematics Assistant Principal, volunteered to participate in a 3-day technology workshop given by one of the researchers. The workshop introduced them to the TI-Nspire technology and its applications for teaching high school mathematics. These four teachers formed an experimental group.

Control group teachers were selected on a volunteer basis at the beginning of the school year. Three of these teachers did not receive training on the new technology, since they were going to use graphing calculators that were already implemented into mathematics teaching in the school. One teacher from the experimental group also taught two sections in the control group, using only graphing calculators.

Professional Development Workshops

Both groups of teachers, experimental and control, met weekly for professional development workshops to develop inquiry based lessons. The experimental and control groups met separately. During the fall semester the experimental group met 12 times and the control group met 11 times for 2 hours each time. During spring semester both groups met 13 times for 2 hours each time. The purpose of the workshops was for the teachers to work cooperatively with their peers in planning and development of their own lesson activities while using either the new technology (TI-Nspire) or the existing technology (TI-84).

In both groups the workshops were guided by their facilitator. The professional development model cycled through the following stages of the content development with TI-Nspire or TI-84 technology: 1) to review the curriculum sequence for the two weeks following the workshop, to select the topic of the lesson, lesson's objectives, and to brainstorm lesson activities appropriate for the TI-Nspire or TI-84 environment (as a whole group, in pairs, or individually – varied by group and time of the year); 2) to meet together as a group for developing a lesson plan and activity documents for the lesson – putting together work developed by the teachers outside the meetings; 3) to present the lesson and activity at the workshop with demonstration of the TI-Nspire or TI-84 activity to the group, facilitator and/or principal investigator (PI), for peer review and critique – followed by necessary modification/revision of the lesson plan and documents; 4) to teach the lesson in class during the same week; teachers in both group were expected to teach the lessons they developed, but only experimental group teachers were observed by the facilitator and/or PI; 5) to reflect and discuss teaching and learning experience at the group meetings following the teaching of the lessons; and 6) to finalize curriculum materials based on teaching experience and peer feedback.

As a result of these meetings both groups of teachers developed series of lessons that have been peer reviewed and field tested. Lessons from experimental group were collected for further analysis. There were total of 13 lessons developed by the teachers in the experimental group, which included the following topics (in chronological order):

1. Writing Function Rule (written by whole group - October 2007)
2. Similar Figures (written by whole group - November 2007)
3. Solving Equations Graphically (written by whole group – December 2007)
4. Modeling with Real Function Rule (written by two teachers – February 2008)
5. Finding Rate of Change – (written by two teachers – February 2008)
6. Predicting Using Trend Lines – (written by two teachers – March 2008)
7. Using the Line of Best Fit to Make Predictions – (written by two teachers – March 2008)
8. Discovering Exponential Functions – (written by two teachers – March 2008)
9. Exploring Leading Coefficient of Quadratic Graphs – (written by two teachers – March 2008)
10. Discriminant (written by two teachers – April 2008)

11. Axis of Symmetry, Parabola (written by two teachers – April 2008)
12. What are trigonometry ratios (written by two teachers– May 2008)
13. Discovering Trigonometry Ratios (written by two teachers – May 2008)

Results

The number of cases varied from one assessment to another. On each assessment the number of scores was lower than initial enrollment. This could be explained by high mobility, low attendance rate, and high dropout rate of student body which are typical for the New York City public schools. The fact that number of scores collected on baseline assessment and midyear assessment is lower than the number of students in each group could also be explained by the fact that these assessments were scheduled on specific class days with one makeup day. Students who missed these two days were not assessed. Also, since school provides two types of high school diploma: local diploma for students who passed all courses but did not pass or take Regents exams, and Regents diploma for students who passed NYS Regents exams, students who are not interested in continuing their education usually do not take NYS Regents exams. More specific details on the number of collected data are shown in Table 4.

Place Table 4 here

Baseline Assessment

A multiple-choice algebra readiness test was developed and administered by the mathematics department of the school during the month of September. The scores of this test were used to compare the experimental and control groups before TI-Nspire technology was introduced in the classrooms. The mean for the experimental group was 50.91 (SD = 18.63 and N = 95) and the mean for the control group was 48.48 (SD = 23.60 and N = 134). A t-test for independent means indicated that the two groups were not significantly different in the beginning of the year t(227) = -2.428.

A 2 (Type of Technology: experimental vs. control) × 4 (Black vs. Hispanic vs. White vs. Asian) ANOVA was conducted to determine whether Race had any interaction with Type of Technology. The analysis did not reveal main effects or interaction, indicating that there were no significant differences in the performance among students of different races (see Table 5).

Place Table 5 here

Also, analysis on gender revealed no significant differences between males (M = 51.17, SD = 2.29, N = 90) and females (M = 49.48, SD = 2.09, N =112)

Analysis of Effects

Algebra scores were analyzed using a 2×3 factorial analysis of variance (ANOVA). The two independent variables were Type of Technology (Experimental/using TI-Nspire vs. Control/using TI-84 graphing calculators) and Student Level (F/students on grade level vs. R/students who failed before and were repeating the same subject vs. A/students below grade level). The dependent variables were: 1. Midyear assessment scores 2. Fall semester grades 3. Spring semester grades 4. New York State (NYS) Regents Mathematics A exam scores.

Midyear Assessment

A 2×3 ANOVA was conducted on the Midyear Assessment. Descriptive statistics are presented in Table 6.

Place Table 6 here

The analysis revealed a main effect for Type of Technology ($F(1, 1819.87) = 4.94$, $p < .05$), with the experimental group scoring significantly higher than the control group. Also, a main effect for Student Level ($F(2, 12103.15) = 32.847$ $p < .001$) was found. Post hoc Scheffe tests indicated that students on Level F scored significantly higher than both those of Level R ($p < .001$) and Level A ($p < .001$) and students on Level A scored significantly higher than Level R ($p < .05$). A two-way interaction (Type of Technology × Student Level) was not significant.

Supplemental analysis was performed on Race. 101 students in experimental group and 166 students in control group reported their races. A 2 (Type of Technology: experimental vs. control) × 4 (Black vs. Hispanic vs. White vs. Asian) ANOVA was conducted to determine whether Race had any interaction with Type of Technology. The analysis did not reveal main effects or interaction, indicating that there were no significant differences in the performance among students of different races.

Supplemental analysis was also performed on Gender. To examine the interrelationships between independent variables, a 2 (Type of Technology: experimental vs. control group) × 2 (Gender: male vs. female) analysis of variance (ANOVA) was conducted on midyear assessment scores. Descriptive statistics are presented in Table 7

Place Table 7 here

The analysis revealed a main effect for Gender ($F(1, 1903) = 4.10$ $p < .05$) with the female students performing significantly better than the males, with the females in the experimental group receiving the highest mean. Finally, a two-way interaction (Type of Technology × Gender) was not significant.

Fall Semester Grades

To examine the interrelationships between the independent variables, a 2 (type of technology: experimental vs. control group) × 3 (student level: F vs. R vs. A) analysis of variance (ANOVA) was conducted on fall semester grades. Descriptive statistics are presented in Table 8.

Place Table 8 here

The analysis revealed a main effect for Type of Technology ($F(1, 919.49) = 5.65$ $p < .001$), with the experimental group performing significantly higher than the control group. Also, the analysis yielded a main effect for Student Level ($F(2, 3943.48) = 24.22$ $p < .001$). Post hoc Scheffe tests indicated that students on Level F scored significantly higher than both those on Level A ($p < .001$) and Level R ($p < .05$). Students on Level R scored significantly higher than Level A ($p < .001$).

A two-way interaction (Type of Technology × Student Level) was also significant ($F(2, 1114.18) = 6.84$, $p < .001$). A one way ANOVA and then post hoc Scheffe tests revealed that there was a significant difference only between Level F experimental and control groups ($p < .05$). Levels R and A didn't differ significantly between experimental and control group.

Supplemental analysis was performed on Race. 101 students in experimental group and 166 students in control group reported their races. A 2 (type of technology: experimental vs. control) × 4 (Black vs. Hispanic vs. White vs. Asian) ANOVA was conducted to determine whether Race had any interaction with Type of Technology. The descriptive statistics is shown in Table 9.

Place Table 9 here

The analysis only revealed a main effect for Type of Technology ($F(1, 773) = 3.91$ $p < .05$), with the experimental group ($M = 68.33$, $SD = 13.09$) performing significantly higher than the control group ($M = 62.69$, $SD = 14.59$), but there was no main effect for Race or interaction of the two variables. This analysis indicates that there were no significant differences in performance among students of different races.

Supplemental analysis was also performed on Gender. To examine the interrelationships between independent variables, a 2 (type of technology: experimental vs. control group) × 2 (Gender: male vs. female) analysis of variance (ANOVA) was conducted on fall semester grades. Descriptive statistics are presented in Table 10.

Place Table 10 here

The analysis revealed a main effect for Type of Technology ($\underline{F}(1, 1699) = 8.87$ $p < .005$), with the experimental group performing significantly higher than the control group. Also, the analysis yielded a main effect for Gender ($F(1, 1368) = 7.11$ $p < .008$) with the female students performing significantly better than the males. Finally, a two-way interaction (Type of Technology × Gender) was also significant ($F(1, 888) = 4.63$, $p < .03$). The analysis of interactions shows that the females in the experimental group

(M = 72.11) performed significantly better than all other groups i.e., the females in the control group (63.07), the males in the experimental group (63.62) and the males in the control group (62.17).

Spring Semester Grades

A 2×3 ANOVA was conducted on the students' spring semester grades. Descriptive statistics are presented in Table 11

Place Table 11 here

The analysis revealed a main effect for Type of Technology (F(1, 933.12) = 6.07 p < .05) with the experimental group performing significantly higher than the control group. It also showed a main effect for Student Level (F(2, 2714.92) = 17.65 p < .001). Post Hoc Scheffe tests indicated that students on Level F scored significantly higher than Level A (p < .001) and Level R (p < . 05) and students on Level R scored significantly higher than Level A (p = .001). Finally, a two-way interaction (Type of Technology × Student Level) was not significant.

Supplemental analysis revealed no main effects on group × race and no main effect on group × gender. But overall girls received higher spring semester grades than boys.

NYS Regents Mathematics A Exam Scores

A 2×3 ANOVA was conducted on the students' Regents scores. The analysis revealed no main effect for Type of Technology with the experimental group (M = 59.68, SD = 12.02, N = 92) outperforming the control group only by about three points (M = 56.99, SD = 12.75, N = 92). No main effects were revealed for Student Level either (see Table 12 for Descriptive Statistics).

Supplemental analyses did not reveal significant differences among students of different races or gender. Although all previous analyses indicated that girls outperformed boys, Regents scores actually revealed almost identical scores. More specifically, girls had a mean of 58.32 (SD = 11.97, N = 133) and boys had a mean of 58.27 (SD = 10.88, N = 89).

Conclusions and Discussion

Research has established that students who are taught through technology receive higher scores in mathematics than those who are taught without technology (e.g. Ellington, 2005; Heller, 2005; Khoju, Jaciw & Miller, 2003). Now we enter the era of having to determine the efficacy of new generations of technology as compared to old, and this was the purpose of the present study. More particularly, the student participants in the experimental group, received algebra instruction with the new TI-Nspire handhelds, while the control group received instruction through TI-84 graphing calculators – an existing technology. New generations of technology provide a variety of features that need to be tested empirically for their effectiveness.

In addition to the manipulation of technology, another component of our study was that all teacher participants were supported through year-long professional development. The reason for the inclusion of the professional development component was that in recent years many schools that demonstrated improvement in student learning have used: a) technology as integral to their transformation and b) have shown that effective professional development is needed to help teachers use technology to change how they teach core subject areas (Wolf, 2007). Further, in a recent article on improving impact studies of teachers' professional development Desimone (2009 p. 192) states: "Finally, we need more work that links professional development and changes in teaching practice to student achievement." In our case the changes in teaching practice came about through the incorporation of new technology, which was supported by professional development, which focused on effective use of new or existing technology within algebra. Now we can turn to the discussion of student achievement as measured by our dependent variables.

Student Achievement

The following research questions guided the study: was TI-Nspire supported instruction more effective than the instruction with graphing calculators in improving high school students' algebra achievement in different level courses as measured by (1) mid-year assessment; (2) fall semester grades; (3) spring semester grades; and (4) scores on NYS Regents Math A exam?

The results were as follows: (1) The midyear assessment scores indicated that the experimental group did significantly better than the control group; further, students in Level F – the high achieving students - did significantly better than students in Level A, who did significantly better than students in

Level R. (2) Fall semester grades indicated that the students in Level F who were in the experimental group did significantly better than the students in every other group. (3) Spring semester grades indicated that the experimental group did significantly better than the control group; and students in Level F did significantly better than students in Level A, who did significantly better than students in Level R. (4) There were no significant differences between control and experimental groups or between different levels in the Regents exam scores.

Our results confirmed our expectations that differences between our two groups would be more apparent when students were assessed by their teachers/schools and less apparent when they were assessed through the state–wide test. This result is consistent with research on large scale assessments (e.g., Stancavage, Shepard, McLaughlin, Holtzman, Blankenship, Zhang, 2009).

Discussion on Gender

The next research question examined whether there were any gender differences in students' achievement. The results on gender were as follows: (1) Midyear assessment indicated that girls did significantly better than boys (with girls in the experimental group receiving the highest mean score); (2) fall semester grades indicated that girls in the experimental group performed significantly better than all others (i.e. boys/experimental, boys/control, girls/control); (3) spring semester grades indicated that there were no significant differences between boys and girls (although girls had higher means); (4) Regents scores indicated no significant differences between girls and boys performance either.

So, in two earlier assessments girls did significantly better than boys and in the two later assessments boys and girls did equally well. These results should not surprise us as we review the history of gender differences in mathematics. In the 70's and 80's, boys typically performed better than girls from adolescence on (Fennema, 1974, 1980; Leder, 1985; Peterson & Fennema, 1985). Later studies have challenged this trend by showing that this trend has declined (Barker, 1997; Hyde, Fennema, & Lamon, 1990; Knodel, 1997). And some other studies have shown no gender differences in mathematics achievement (e.g., Bronholt, Goodnow, & Conney, 1994). Finally, Geist and King (2008) in the most recent study stated that data from the National Assessment of Educational Progress show a gap of only two points between girls and boys and that is a development that occurred during the last decade. Therefore, our results are consistent with the current literature on gender differences.

Discussion on Race

The final research question examined whether there were any race differences in students' achievement. Results indicated that there were no significant differences among races (Hispanic, Black, White, Asian) in any of the assessments. These findings do not confirm any of the existing literature which indicates that the academic achievement gap, particularly the mathematics achievement gap, between Black students and their White counterparts has been well documented (Strutchens, Lubienski, McGraw, & Westbrook, 2004; U.S. Department of Education, 2003). Although mathematics educators know that the achievement gap exists, we do not fully understand the complexities of why it exists and this analysis is beyond the scope of the present study. Many educators attempt to come up with methods to assist teachers in eliminating the race gap. One suggested method is by having class work that is challenging but not so difficult that it lies outside an individual's ability (Gonzales, Blanton, & Williams, 2002). We can only speculate as to why, but the teachers of our students were able to eliminate race differences when teaching integrated algebra with technology.

Limitations

The quasi-experimental design of this study implies that strong causal inferences are not warranted because students were not randomly assigned to classes and classes were not randomly assigned to teachers. Research findings suggest that teachers who volunteer to participate in the experimental condition are probably motivated to change and try something new (Supovitz & Zeif, 2000). Further, there is evidence that the most qualified teachers are the ones who seek out professional development with effective features such as content focus (Desimone, Smith, & Ueno, 2006). In our study we controlled for teacher effects such as teacher age and teacher experience by running regression analyses and finding that these were not significant factors but we did not control for teacher motivation and attitudes.

Another limitation of our study was that we worked in a Title I school which has high mobility and dropout rate among the students. This led to about 30% data lost, and the possibility of under-representation of low achievers who often are more mobile and have higher dropout rates than their peers.

Implications and Future research

Teachers need to stay informed about the features and effectiveness of new technology. TI-Nspire technology was chosen for the experimental group because it represents an advance in the capabilities of a low-cost personal computing device that is reliable and easy-to-use, supporting a broad range of instructional models and advanced modes of assessment for teaching mathematics, and it is allowed on the state tests. The two new capabilities that TI-Nspire™ incorporates are: a) The ability to display multiple representations which are connected and in a single plane. That is, the multiple representation capability dynamically links graphical curves, axes, equations and tables in simultaneous displays, such that a change in one representation is transmitted to the others. This feature allows teachers to design new tasks for their students to address the NCTM standards focusing on connections between algebraic and geometric representations, on inquiry-based approaches to teaching and learning mathematics, etc. b) The ability to document content. This document-based content system is an organized presentation of multiple screens of mathematics, which can be saved, shared, annotated, and revisited, giving teachers new ways of assessing their students' understanding of mathematics and technology.

The present study compared the effects of two types of technology on students' achievement while providing the same type of professional development to the teachers in both, experimental and control groups. Future studies should look closer at the effects of professional development as part of the overall model that leads to improved student achievement. Desimone (2009) proposed a conceptual framework for studying the effects of professional development on teachers and students. It starts with the core features of professional development (i.e. content, active learning, coherence, duration, collective participation – and we believe that technology facilitates all those components), continues with increased teacher knowledge and skills as well as change in attitudes and beliefs, leads to change in instruction and ultimately improved student learning.

Another important question that this study raised is what the learning curve is for students at different levels of performance. Future studies should analyze if there is a predictable pattern to the students' learning curve, and it may even be possible to design teacher scaffolding to accelerate this learning curve. It may be that the degree and type of scaffolding students need changes according to their skill and knowledge profile. It may be that there is a skill hierarchy for mastery of TI-Nspire in classroom instruction. Research on these issues will allow to refine the professional development system supporting TI-Nspire in scalable, sustainable ways.

TI-Nspire has the ability to capture a progression of documents which may represent stages in a student's thinking about a problem-solving task. It also has the ability to support free-form answer entry using alphanumeric and algebraic notation. Further studies could explore how to exploit these capabilities for formative and summative assessment, with particular attention to deep understanding (of connected, contextualized concepts and principles) and “doing real work” (ill-structured problem solving) in mathematics.

References

Bannasch, S. (1999). The electronic curator: Using a handheld computer at the Exploratorium. Concord Consortium Newsletter. Retrieved August 10, 2003 from http://www.concord.org/library/1999fall/electronic-curator.html.

Barker, B. (1997). Girls' world or anxious times; What's really happening at school in the gender war? *Educational Review*, *49*, 221-227.

Bronholt, L. J., Goodnow, J., & Conner, G. H. (1994). Influences of gender stereotypes on adolescents' perceptions of their own achievement. *American Educational Research Journal, 31*, 675-692.

Cole, H., & Stanton, D. (2003). Designing mobile technologies to support co-present collaboration. *Personal and Ubiquitous Computing, 7*, 365-371.

Curtis, M. (2005). The rise of the handheld computer in schools. *Media and Methods, 41*(6), 14.

Danesh, A., Inkpen, K., Lau, E, Shu, K., & Booth, K. (2001). Geney: Designing a collaborative activity for the Palm handheld computer. Proceedings of CHI, Conference on Human Factors in Computing Systems. Seattle: WA.

Desimone, L. M. (2009). Improving impact studies of teachers' professional development: Toward better conceptualizations and measures. *Educational Researcher*, *38*(3), 181-199.

Desimonone, L. M., Smith, T., & Ueno, K (2006). Are teachers who need sustained, content-focused professional development getting it? An administrator's dilemma. *Educational Administration Quarterly*, *42*(2), 179-215.

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.

Fennema, E. (1974). Mathematics learning and the sexes: A review. *Journal for Research in Mathematics Education, 5*, 126-139.

Fennema, E. (1980). Sex-related differences in mathematics achievement: Where and why. In L. H. Fox, L. Brody, & D. Tobin (Eds.), Women and the mathematics mystique (pp. 76-93). Baltimore: The Johns Hopkins University Press.

Fleener, M. J. (1995). A survey of mathematics teachers' attitudes about calculators: The impact of philosophical orientation. *Journal of Computers in Mathematics and Science Teaching,* 14, 481-498.

Geist, E. A., & King, M. (2008). Different, not better: Gender differences in mathematics learning and achievement. *Journal of Instructional Psychology, 35*(1), 43-52.

Gonzales, P. M., Blanton, H., & Williams, K. J. (2002). The effects of stereotype threat and double-minority status on the test performance of Latino women. *Personality & Social Psychology Bulletin, 28*, 659–670.

Heller, J. L., Curtis, D. A., Jaffe, R., & Verboncoeur, C. J. (2005). *Impact of handheld graphing calculator use on student achievement in Algebra 1*. Oakland, CA: Heller Research Associates.

Inkpen, K. (2001). Designing handheld technologies for kids. Proceedings of CHI, Conference on Human Factors in Computing Systems. Seattle: WA.

Kastberg, S., & Leatham, K. (2005). Research on graphing calculators at the secondary level: Implications for mathematics teacher education. *Contemporary Issues in Technology and Teacher Education, 5*(1), 25-37.

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

Knodel, J. (1997). The closing of the gender gap in schooling: The case of Thailand. *Comparative Education, 33*(1), 61-86.

Laughbaum, E. D. (2003). Hand-held graphing technology in the developmental algebra curriculum. *Mathematics and Computer Education, 37*(3), 301-314.

Leder, G. C. (1985). Sex-related differences in mathematics: An overview. *Educational Studies in Mathematics, 16*, 304-309.

Lyublinskaya, I. (2003a). Connecting Mathematics with Science: Experiments for Precalculus. Emeryville, CA: Key Curriculum Press.

Lyublinskaya, I. (2003b). Connecting Mathematics and Science: Experiments for Calculus. Emeryville, CA: Key Curriculum Press.

Mandryk, R. L., Inkpen, K. M., Bilezkjian, M., Klemmer, S. R., & Landay, J. A. (2001). Supporting children's collaboration across handheld computers. Proceedings of CHI, Conference on Human Factors in Computing Systems. Seattle: WA.

National Center for Education Statistics. (2001). *The nation's report card: Mathematics 2000.* (No. NCES 2001-571). Washington, DC: U.S. Department of Education.

Niess, M. (2001). A Model for Integrating Technology in Preservice Science and Mathematics Content-Specific Teacher Preparation, *School Science and Mathematics, 101*(2), 102-109.

Peterson, P., & Fennema, E. (1985). Effective teaching, students engagement in classroom activities, and sex-related differences in learning mathematics. *American Educational Research Journal*, *22*(63), 309-335.

Roschelle, J., & Pea, R. (2002). A walk on the WILD side: How wireless handhelds may change computer-supported collaborative learning. *International Journal of Cognition and Technology, 1*(1), 145-272.

Roschelle, J. & Pea, R. (2002). A walk on the WILD side: How wireless handhelds may change computer-supported collaborative learning. *International Journal of Cognition and Technology, 1*(1), 145-272.

Roth, J. (2002). Patterns of mobile interaction. *Personal and Ubiquitous Computing 6*, 282-289.

Sharples, M. (2000). The design of personal mobile technologies for lifelong learning. Computers and Education, 34, 177-193.

Soloway, E., Norris, E., Blumenfeld, P., Fishman, B., Krajcik, J., & Marx, R. (2001). Log on education: Handheld devices are ready-at-hand. *Communications of the ACM, 44*(6), 15-20.

Stancavage, F., Shepard, L., McLaughlin, D., Holtzman, D., Blankenship, C., & Zhang (2009). Sensitivity of NAEP to the effects of reform-based teaching and learning in middle school mathematics. In *NAEP Validity Studies*.

Strutchens, M. E., Lubienski, S. T., McGraw, R., & Westbrook, S. K. (2004). NAEP findings regarding race and ethnicity: Students' performance, school experiences, and attitudes and beliefs, and family influences. In P. Kloosterman & F. K. Lester (Eds.), *Results and interpretations of the 1990 through 2000 mathematics Assessment of the national assessment of education progress* (pp. 269-304). Reston, VA: National Council of Teachers of Mathematics.

Supovitz, J. A., & Zeif, S. G. (2000). Why they stay away. *Journal of Staff Development*, *21*(4), 24-28.

Thomas, J. A., & Cooper, S. B. (2000). Teaching technology: A new opportunity for pioneers in teacher education. *Journal of Computing in Teacher Education*, *17*(1), 13-19.

Tinker, R. (1996). The whole world in their hands. Washington, DC: Department of Education. http://www.ed.gov/Technology/Futures/tinker.html

Tinker, R., & Krajcik, J. (Eds.). (2001). Portable Technologies: Science Learning in Context. New York: Kluwer Academic/Plenum Publishers.

U.S. Department of Education. (2003). *Status and Trends in the Education of Blacks.* (NCES 2003-034), by K. Hoffman & C. Llagas. Project Officer: T. D. Snyder. Washington, DC: National Center for Education Statistics. Retrieved July, 17, 2006, from http://nces.ed.gov/pubs2003/2003034.pdf

Vahey, P., & Crawford, V. (2002). Palm Education Pioneers Program: Final evaluation report. Menlo Park, CA: SRI International.

van't Hooft, M., Diaz, S., & Swan, K. (2004). Examining the potential of handheld computers: Findings from the Ohio PEP project. *Journal of Educational Computing Research*, *30*(4), 295-312.

Wolf, M. A. (2007). A Guiding Hand. *The Journal of Higher Education, 34*(7), 12-13.

Table 1: Racial and gender composition of control and experimental groups.

		Control Group	Experimental Group	Total
Level	Level F	32%	26%	29%
	Level A	48%	17%	37%
	Level R	20%	57%	34%
Race	African American	36%	49%	40.5%
	Hispanic	37%	35%	36%
	White	12%	12%	18.5%
	Asian or Pacific Islander	5%	9%	5%
Gender	Male	43%	45%	43.5%
	Female	57%	55%	56.5%

Table 2: Demographics for the Teachers in the Experimental Group (at the end of the 2007-2008 academic year)

Gender	Age	Teaching Experience (years)	Total sections taught	Number of experimental sections
F	23	1	5	1
F	26	3	5	2
F	30	10	2	1
M	37	13	1	1

Table 3: Demographics for the Teachers in the Control Group (at the end of the 2007-2008 academic year)

Gender	Age	Teaching Experience (years)	Total sections taught	Number of control sections
F	23	1	5	2
M	30	2	5	1
F	25	1	5	4
F	36	8	5	1

Table 4. Number of Collected Data Cases for Dependent Variables.

	Initial enrollment	Baseline assessment	Midyear assessment	Fall semester grades	Spring semester grades	NYS Math Regents
Control	197	134	133	178	171	147
Experimental	131	95	98	105	107	92

Table 5. Descriptive Statistics for Baseline Assessment Scores by Race.

Race	N	Mean	SD
White, not Hispanic	34	51.53	26.189
Hispanic	72	51.22	20.370
African America	81	48.37	20.667
Asian or Pacific Islander	15	49.20	24.583

Table 6: Descriptive Statistics for Midyear Assessment. Group × Course Level.

	STUDENT LEVEL			
	Level F	Level R	Level A	Total
Experimental Group	M = 74.76 SD = 13.14 N = 25	M = 49.26 SD = 17.30 N= 58	M = 52.60 SD = 12.17 N = 15	M = 56.28
Control Group	M = 63.87 SD = 19.21 N = 54	M = 39.00 SD = 26.37 N = 34	M = 54.96 SD = 19.74 N = 45	M = 54.50
Total	M =69.31	M = 44.13	M = 53.78	

Table 7: Descriptive Statistics for Midyear Assessment Scores. Group × Gender

	GENDER		
	Male	Female	Total
Experimental Group	M = 50/45 SD = 16.580 N = 44	M = 61.13 SD = 19.966 N= 52	M = 56.24
Control Group	M = 54.16 SD = 24.068 N = 57	M = 55.33 SD = 23.160 N = 70	M = 54.80
Total	M = 52.54	M = 57.80	M = 55.42

Table 8: Descriptive Statistics for Fall Semester Grades. Group × Student Level

	STUDENT LEVEL			
	Level F	Level R	Level A	Total
Experimental Group	M = 79.63 SD = 11.23 N = 27	M = 65.70 SD = 11.92 N= 60	M = 58.06 SD = 5.19 N = 18	M = 67.97
Control Group	M = 70.93 SD = 14.00 N = 56	M = 66.19 SD = 11.89 N = 36	M = 57.41 SD =13.46 N = 86	M = 64.35
Total	M = 75.28	M = 65.95	M = 64.17	

Table 9: Descriptive Statistics for Fall Semester Grades. Group × Race

	RACE				
	African American	Hispanic	White, not Hispanic	Asian or Pacific	Total
Experimental Group	M = 69.39 SD = 13.597 N = 49	M = 65.63 SD = 10.749 N = 35	M = 70.50 SD = 15.139 N = 12	M = 71.60 SD = 18.649 N = 5	M = 68.33
Control Group	M = 61.66 SD = 13.246 N = 59	M = 61.54 SD = 15.429 N = 61	M = 64.51 SD = 15.933 N = 37	M = 69.67 SD = 10.062 N = 9	M = 62.69
Total	M = 65.17	M = 63.03	M = 65.98	M = 70.36	

Table 10: Descriptive Statistics for Fall Semester Grades. Group × Gender

	GENDER		
	Male	Female	Total
Experimental Group	M = 63.62 SD = 10.86 N = 45	M = 72.11 SD = 13.59 N= 56	M = 67.87
Control Group	M = 62.17 SD = 14.5 N = 71	M = 63.07 SD = 14.7 N = 95	M = 62.62
Total	M = 62.73	M = 66.42	M = 64.57

Table 11: Descriptive Statistics for Spring Semester Grades

	STUDENT LEVEL			
	Level F	Level R	Level A	Total
Experimental Group	M = 73.96 SD = 14.61 N = 23	M = 62.88 SD = 11.90 N= 66	M = 59.72 SD = 8.13 N = 18	M = 64.73
Control Group	M = 67.08 SD = 13.00 N = 50	M = 61.89 SD = 11.32 N = 44	M = 55.00 SD =13.09 N = 77	M = 60.30
Total	M = 70.52	M = 62.39	M = 57.39	

Table 12: Descriptive Statistics for NYS Regents Exam. Group × Level

	STUDENT LEVEL			
	Level F	Level R	Level A	Total
Experimental Group	M = 62.6 SD = 9.17 N = 25	M = 59.14 SD = 13.30 N= 56	M = 55.82 SD = 10.05 N = 11	M = 59.19
Control Group	M = 59.21 SD = 13.01 N = 53	M = 55.97 SD = 11.71 N = 38	M = 55.57 SD = 13.10 N = 56	M = 56.92
Total	M = 60.29	M = 57.86	M = 55.61	