J·T·L·A

The Journal of Technology, Learning, and Assessment

Volume 6, Number 5 · January 2008

Examining the Relationship between Students' Mathematics Test Scores and Computer Use at Home and at School

Laura M. O'Dwyer, Michael Russell,
Damian Bebell, & Kevon Seeley

www.jtla.org

A publication of the Technology and Assessment Study Collaborative
Caroline A. & Peter S. Lynch School of Education, Boston College

Volume 6, Number 5

Examining the Relationship between Students' Mathematics Test Scores and Computer Use at Home and at School

Laura M. O'Dwyer, Michael Russell, Damian Bebell, & Kevon Seeley

Editor: Michael Russell
russelmh@bc.edu
Technology and Assessment Study Collaborative
Lynch School of Education, Boston College
Chestnut Hill, MA 02467

Copy Editor: Jennifer Higgins
Design: Thomas Hoffmann
Layout: Aimee Levy

JTLA is a free on-line journal, published by the Technology and Assessment Study Collaborative, Caroline A. & Peter S. Lynch School of Education, Boston College.

Preferred citation:

O'Dwyer, L.M., Russell, M., Bebell, D., & Seeley, K. (2008). Examining the Relationship between Students' Mathematics Test Scores and Computer Use at Home and at School. *Journal of Technology, Learning, and Assessment, 6*(5). Retrieved [date] from http://www.jtla.org.

J·T·L·A

Abstract:

Over the past decade, standardized test results have become the primary tool used to judge the effectiveness of schools and educational programs, and today, standardized testing serves as the keystone for educational policy at the state and federal levels. This paper examines the relationship between fourth grade mathematics achievement and technology use at home and at school. Using item level achievement data, individual student's state test scores on the Massachusetts Comprehensive Assessment System (MCAS), and student and teacher responses to detailed technology-use surveys, this study examines the relationship between technology-use and mathematics performance among 986 general education students, from 55 intact fourth grade classrooms in 25 schools across 9 school districts in Massachusetts. The findings from this study suggest that various uses of technology are differentially related to student outcomes and that in general, student and teacher technology uses are weakly related to mathematics achievement on the MCAS. Implications for improving methods for examining the relationship between technology use and standardized test scores are presented.

Examining the Relationship between Students' Mathematics Test Scores and Computer Use at Home and at School

Laura M. O'Dwyer, Michael Russell, Damian Bebell, Kevon Seeley
Technology and Assessment Study Collaborative
Center for the Study of Testing, Evaluation, and Educational Policy
Lynch School of Education
Boston College

Introduction

Over the past decade, standardized test results have become the primary tool used to judge the effectiveness of schools and educational programs, and today, standardized testing serves as the keystone for educational policy at the state and federal levels. During the same time period, substantial investments have been made in educational technology and the infrastructure to support its use. Since 1996 alone, state and district level agencies have invested over ten billion dollars to acquire and integrate computer-based technologies into American schools. Over this time period, the federal government has spent another three billion dollars on educational technology. As a consequence of these major investments, the average student-to-computer ratio has decreased dramatically over a twenty year period from 125:1 in 1983 to approximately 3.8:1 in 2006 (Market Data Retrieval, 1999, 2001; US Census Bureau, 2006, Bausell, & Klemick, 2007). Given this investment in educational technology coupled with the current focus on standardized tests, it is only natural that educators, policymakers, and the public consider the relationship between expenditures on educational technology and performance on standardized tests.

Since the early 1980s, the positive effects of educational technology have been documented in numerous formal and informal evaluation and research studies (Sivin-Kachala & Bialo, 1994; Coley, Cradler, & Engel, 1997; Mann, Shakeshaft, Becker, & Kottkamp, 1999). Several studies report that students enjoy classes and learn more in shorter periods of time when computer-based instruction is used (Kulik as cited in Chapman, 2000). Other research has found that when technology is used effectively, students develop stronger critical-thinking and problem-solving skills, and achieve higher-order levels of understanding than in non-technology enriched learning environments (Penuel, Yarnell, & Simkins, 2000). In addition, meta-analyses conducted to examine educational technology use and achievement issues suggest that specific student uses of technology have positive impacts on student achievement (Kulik, 1994; Goldberg, Russell, & Cook, 2003; Fletcher-Flinn & Gravatt, 1995; Waxman, Lin, & Michko, 2003).

While many studies have examined technology-related issues, observers note that research investigating the effects of educational technology on teaching and learning rarely link these uses to improved standardized test scores (for example, Cuban, 2001; Oppenheimer, 1998). The impetus for linking technology use to increased standardized test scores is fueled both by the large investments in educational technology that have been made at the federal, state, and local levels over the past decade or more, and the current mandates of No Child Left Behind for improving student achievement. While there are psychometric challenges in linking technology use with scores on standardized tests, improved standardized test scores remain the customary means for evaluating the benefits of educational innovations. These challenges stem from the reality that standardized tests (either state tests or off-the-shelf norm-referenced tests) are generally designed to broadly measure content areas and so are unlikely to be sensitive enough to detect the types of learning that have been found in previous studies to be affected by technology use.

Two notable studies that have attempted to examine the link between "business as usual"[1] educational technology use and standardized test scores were conducted by Wenglinsky (1998) and Angrist and Lavy (2002). More recently, the US Department of Education used an experimental approach to examine the impact of sixteen specific reading and mathematics software products on students' standardized test scores (Dynarski, Agodini, Heaviside et al., 2007).

In his study, Wenglinsky (1998) used fourth and eighth grade 1996 National Assessment of Educational Progress (NAEP) data to examine the relationship between technology-use and achievement. Using two measures of technology-use (use of simulation and higher order thinking software, and a measure of more general technology-use) and a nationally representative sample of students, Wenglinsky employed empirical measures of teacher characteristics (including professional development experiences), and aggregated measures of class size and school climate to estimate the impacts of technology-use. Wenglinsky concluded that both fourth and eighth grade students who used simulation and higher order thinking software had statistically significant higher scores in mathematics achievement. However, when considering general student technology-use, Wenglinsky found that computer use was negatively related to mathematics achievement for grades four and eight.

Using Israeli school data from a 1996 administration of a standardized middle school mathematics and Hebrew test, Angrist and Lavy (2002) examined the effects of educational technology on student achievement. In their study, the authors compared levels of academic achievement among students classified as receiving instruction in either high- or low-technology environments. Focusing on *access to technology* rather than *actual technology-use* to examine the impacts on achievement, the authors classified schools as "high-access schools" when schools were equipped with computers at a 10:1 ratio, meaning 10 students share 1 computer. Based upon their classification system, the authors found weak and, in some cases, negative relationships between technology and student test scores.

While both these studies attempted to estimate the impact of educational technology on standardized test scores, they are limited by: (1) the way in which students' and teachers' technology-use is measured, and (2) measures of achievement that are not specifically designed to capture the types of improved learning that occurs as a result of technology-use. Specifically, Angrist and Lavy's (2002) study equated access to technology with use of technology. Given that the present student-to-computer ratio in the United States is currently about 3.8:1, and given the current trend toward 1:1 computing environments (for example, New Hampshire, Vermont, Massachusetts, New York, and Maine each have 1:1 environments available), this study is not very useful for informing technology-related policy and practice decisions in the United States. Moreover, it is important to note that more recent research suggests that there are a variety of ways in which students and teachers use technology and these are not equivalent to simply having access to technology (Bebell, Russell, & O'Dwyer, 2004; O'Dwyer, Russell & Bebell, 2004).

Although Wenglinsky's analysis of NAEP data employed two measures of student computer use, measures of teachers' use of computers for instructional purposes were absent from the analyses. In addition, given that NAEP is designed to measure performance trends over time and that the questionnaires are intended to remain relatively stable over time, it can not be expected that the NAEP background questionnaire could be used to capture the rapidly evolving uses of technology for teaching and learning. Perhaps more importantly, the test scores used in these analyses provide incomplete information about each student's mathematics achievement. By design, NAEP yields test scores that are representative of a state and of the nation, but not of individuals or of classrooms. For this reason, the use of NAEP test scores to examine the relationship between technology-use and classroom-level performance may be problematic. Finally, both of these studies relied on either aggregate school level data or individual level data within classrooms and so did not take into account differences within and between classrooms and schools when modeling student outcomes. Given that characteristics of students within a classroom are likely to influence the attitudes and instructional practices of their teachers, and that these practices in turn affect all of the students in the classroom, it is important to examine the classroom as a hierarchical organization within which technology-use occurs.

More recently, the US Department of Education examined the impact of sixteen educational technology software products on first grade and fourth grade reading, sixth grade math, and algebra (Dynarski, Agodini et al., 2007). Using an experimental pre-post test design, the researchers randomly assigned 439 volunteer teachers to either the treatment or control conditions. In total, the study included 9,424 students. In addition to analyzing observational data in classrooms, the researchers created multilevel regression models in which they predicted students' scores on a variety of norm-referenced tests including the Stanford Achievement Test reading battery for first graders, the Stanford Achievement Test (SAT-10) reading battery for fourth graders, and the SAT-10 math battery for sixth graders (Dynarski, Agodini et al., 2007). Simply looking at the difference between the treatment and control group on the outcome measures, the researchers concluded that students' post-intervention scores on the norm-referenced tests did not differ significantly across the treatment and control conditions. However, the study found that the effects on test scores were correlated with the amount of time that the technology products were used in the fourth grade classrooms.

While this study is useful for evaluating the effectiveness of these sixteen educational technology software products, it tells us little about the relationship between the "business as usual" technology use that is taking place in our schools and classrooms everyday and students' standardized test scores. Additionally, it is questionable whether the outcome measures employed in this study were appropriate for examining the issue; in addition to the likelihood that these tests assess a domain too broadly to isolate the types of critical learning skills that have been found to be impacted by technology use, they are constructed in such a way that they are unlikely to be able to measure growth effectively. In broad terms, norm-referenced tests are constructed with the specific aim of comparing a sample to some norm group (usually a group that represents the larger population of all test-takers) and are comprised of items whose psychometric difficulty level hovers around the average (that is, the majority of the items are designed to have difficulty levels around 0.50 so they are most effective for discriminating among individuals). The effect of this design and construction is to limit the utility of these assessments for measuring changes in students' content knowledge.

In 2005, O'Dwyer, Russell, Bebell and Tucker-Seeley attempted to address some of the methodological and psychometric challenges faced in examining the relationship between technology use and students' standardized test scores (2005). Specifically, the authors examined the relationship between a variety of school and home technology uses and students' scores on the English/Language Arts portion of the fourth grade Massachusetts Comprehensive Assessment System (MCAS). In their study, the authors acknowledged the psychometric weakness of using standardized test scores as a measure of technology effectiveness, cautioned readers about the limitations of their findings, and called for the development of aligned learning outcome measures that could be used in future studies of the impact of educational technology on student learning. Their study employed multiple measures of how students use technology both at home and at school, used multilevel regression modeling to decompose classroom and student level effects, focused on student-level total test scores as well as sub-test scores on a criterion-referenced state test, and employed measures of prior achievement to control for pre-existing differences to examine the relationship between technology-use and achievement (2005).

After controlling for prior achievement and a proxy for student-level socio-economic status, their study found that students' use of technology to edit papers, to create presentations, and recreational use of computers at home (e.g., playing games, emailing friends, downloading music, etc.) were significant predictors of students' total test scores. While students' use of technology to edit papers was associated with higher test scores, use of computers to create presentations or for recreational purposes at home was associated with lower total test scores. When they examined sub-test scores on the English/Language Arts assessment, their study found that the variety of technology uses they examined were not equally useful predictors of achievement across writing and reading skills. Specifically, use of computers for editing was a significant positive predictor for writing and reading. Use of computers to create presentations was a negative predictor for writing, but not for reading. Finally, use of computers for recreation at home was not significantly associated with writing scores, but was a negative predictor of reading scores (O'Dwyer, et al., 2005).

Notwithstanding the psychometric limitations of using standardized test scores for examining the impact of educational technology in our schools and classrooms, student scores on standardized tests have in the past, and continue to remain the customary means for evaluating the benefits of educational innovations (McNabb, Hawkes, & Rouk, 1999). For this reason and building upon O'Dwyer et al.'s (2005) analyses of English Language Arts test scores, the present investigation examines the relationship between fourth grade mathematics achievement and technology use at home and at school. Using item level achievement data, individual student's test scores, and student and teacher responses to detailed technology-use surveys, this study examines the relationship between technology-use and mathematics performance among 986 general education students, from 55 intact fourth grade classrooms in 25 schools across 9 school districts in Massachusetts.

In today's era of educational accountability there have been an increasing number of calls for empirical research-based evidence that examines how these investments in technology are impacting teaching and learning. Therefore, despite the challenges in using a standardized test that assesses student learning across a broad domain, albeit one that is aligned with state frameworks, the study presented here provides insight into whether various types of technology-use are associated with students' mathematics performance. In addition, the findings raise important issues about how technology-use and student learning are measured when attempting to examine the relationship between technology-use and student learning in the area of mathematics. Note that throughout the research presented here, the term technology refers specifically to computer-based technologies and includes personal computers, LCD projectors, and Palm Pilots.

Sample

The research presented in this paper used data collected as part of the Use, Support, and Effect of Instructional Technology (USEIT) Study. The USEIT study was a three-year project conducted to examine how educational technologies are being used by teachers and students, what factors influence these uses, and how these uses affect student learning. During the 2001-2002 school year, information about district technology programs, teacher and student use of technology in and out of the classroom, and factors that are believed to influence these uses were collected through site visits and surveys. In total, survey responses were obtained from 120 district level administrators, 122 principals, 4,400 teachers, and 14,200 students from over 200 schools in 22 school districts in Massachusetts. The specific details on the sample, methodologies and analyses of the USEIT data are described in Russell, O'Dwyer, Bebell, and Miranda (2003). A subset of the USEIT sample was selected to participate in a second wave of data collection so that the relationship between mathematics achievement and technology-use could be examined.

Specifically, we purposively selected approximately 50 grade four classrooms from the original USEIT sample; all fourth grade teachers who completed the USEIT teacher survey during spring 2002 were stratified into three groups representing high, medium, and low levels of instructional technology-use. Within each group, a sub-set of teachers were recruited to participate in this study of the relationship between technology-use and achievement. The schools and districts were contacted in the fall 2002 and teachers and students were re-surveyed in spring 2003. Survey responses and achievement data from an additional district in which we were conducting related research were also incorporated into the sample. Thus, the current sample includes a total of 1,213 students and 55 teachers of intact classrooms from 25 elementary schools across 9 Massachusetts school districts. The sample of 1,213 students includes students who have been classified as English Language Learners (ELL), students with disabilities (SD), and students who are neither SD nor ELL. In order to reduce the possible confounding effects of specialized learning and language needs, this study examines only those students who are classified as non-SD and non-ELL students. Thus, the sample used for the analyses presented here includes 986 general education fourth grade students.

Table 1 (next page) displays demographic and aggregate achievement data for each of the 9 school districts and the student participants who were selected from the original USEIT study to participate in the current study.

Table 1: Mean Demographic and Achievement Characteristics for the Participating Districts

	District									Total Sample versus State	
	A	B	C	D	E	F	G	H	I	Total Sample	MA ('02–'03)
Total # of Elementary Schools	6	3	6	16	6	7	3	3	5	25	1,270
% White	89	86	96	81	85	64	87	81	91	84.3	75.1
% Free Lunch	3	5	6	5	14	24	19	3	2	8.9	26.2
Student : Computer Ratio	4.3:1	5.3:1	4.4:1	7.5:1	6.6:1	10.1:1	4.5:1	N/R	8.4:1	6.4:1	5.1:1
% Classes on Internet	100	100	100	66	100	58	100	N/R	72	86.9	82.8
% Grade 4 Mathematics Advanced	28	30	20	33	15	7	11	35	40	24	12
% Grade 4 Mathematics Proficient	38	35	36	36	32	23	33	32	38	34	28
% Grade 4 Mathematics Needs Improvement	31	29	39	25	41	50	40	25	19	33	43
% Grade 4 Mathematics Warning/Failing	2	5	5	6	12	19	16	7	4	8	16

Source: The district and school summary data has been adapted from the Massachusetts Department of Education web site (www.doe.mass.edu).

As reported first in O'Dwyer, Russell, Bebell and Tucker-Seeley (2005), the districts that participated in this study performed slightly higher than the state average on the grade 4 MCAS mathematics test. For example, across the districts 24% of the students scored in the advanced range compared to 12% statewide. Similarly, the average percentage of students classified as white was also slightly higher for the participating districts than the state average, and students in this sample had lower free or reduced priced lunch rates than the state average; 8.9% in the sample compared to 26.2% in the state. In terms of technology access, the district average student-to-computer ratio was slightly higher for the participating districts at 6.4:1 compared to the state average of 5.1: 1. From these summary statistics it is reasonable to infer that students in the participating districts were generally more affluent, higher performing in mathematics, and had less access to technology in school than the average district in Massachusetts.

Instruments

The relationship between use of technology and achievement was examined using data collected through student and teacher surveys and the mathematics portion of the state mandated MCAS test. Each source of data is described separately below.

The MCAS Fourth Grade Mathematics test

The Massachusetts Comprehensive Assessment System (MCAS) is a state-mandated test linked to the Massachusetts curriculum frameworks. Beginning in 1998, the paper and pencil tests were administered in grades 4, 8, and 10 and focused on English/language arts, science/technology, and mathematics. Currently, the MCAS has been expanded across subject areas and grade levels and now includes a third grade reading exam. Like some other state testing programs, the results are used to calculate school and district Adequate Yearly Progress (AYP) at all grade levels. Tenth grade MCAS results are also used to determine whether an individual student may graduate.

In the study presented here, students' fourth grade mathematics MCAS scores and mathematics subscale scores from the 2002-2003 MCAS administration were modeled as a function of student and teacher background information and technology-use measures. In all, the total mathematics raw score and raw scores on five MCAS mathematics reporting categories identified by the Massachusetts Department of Education were examined in this research. The outcome variables examined were as follows:

- Total math raw score;
- Number sense and operations component of the total mathematics score;
- Patterns, relationships, and algebra component of total mathematics score;
- Geometry component of the total mathematics score;
- Measurement component of the total mathematics score;
- Data analysis, statistics, and probability component of the total mathematics score.

The subscale scores were created by computing the sum of students' scores on each of the sub-domain items. In order to facilitate interpretation of the multilevel regression analysis models, the total raw score and five sub-domain scores were standardized to have a mean of zero and standard deviation of 1. Table 2 contains the reliability for the total MCAS mathematics raw score and for each of the five sub-domain scores (calculated prior to standardization). The Cronbach's alpha for the total mathematics raw score was high at 0.86 but the reliabilities of the sub-domain scores are generally lower, with the reliability for the data analysis, statistics, and probability component of the total mathematics score being the lowest at 0.32. The low reliabilities of the sub-domain measures were likely the result of two conditions: (1) the small number of items measuring each sub-domain, and (2) possible lack of unidimensionality. The magnitudes of the reliabilities have important implications for this research due to the fact that unreliability in the dependent variable is likely to increase the error in the prediction, making it more difficult to isolate statistically significant relationships.

Table 2: Reliability Measures for Achievement Scales

Outcome Measures	Number of Items	Reliability*
Total Mathematics Raw Score	39	0.86
Number sense and operations component of the total mathematics score	15	0.71
Patterns, relationships & Algebra Component of total mathematics score	8	0.49
Geometry component of the total mathematics score	5	0.44
Measurement component of the total mathematics score	4	0.41
Data analysis, statistics, & probability component of the total mathematics score	7	0.32

* Estimated using Cronbach's alpha

Despite the poor psychometric quality of some of the information provided by MCAS, we feel it is important to use these measures to examine the relationship between technology-use and MCAS performance given the concerns that policymakers and the public have about the effects of instructional strategies and tools on standardized tests. But we also believe it is important to emphasize that the value of such analyses are limited by the technical or psychometric quality of the test scores.

To account for students' prior achievement, students' scores on the third grade MCAS reading test (collected in Spring 2002) was included as a covariate when modeling the relationship between fourth grade mathematics achievement and technology-use. At the third grade, MCAS only assesses reading and so a measure of prior mathematics performance was not available. Students' third grade reading scores, fourth grade mathematics scores, and survey responses were combined with their teachers' survey responses allowing the relationship between achievement and technology-use to be examined as a function of both student and teacher characteristics while controlling for prior achievement, albeit reading achievement.

Technology-Use Surveys

Both teachers and students were administered a technology-use survey. The student survey included measures of socioeconomic status, students' access to technology in school, the types of technology-use that occur in school across subject areas, personal comfort levels with technology, access to technology at home, and use of technology at home for both academic and recreational purposes. The teacher survey included demographic measures, measures of several types of technology use in and out of the classroom, teachers' comfort level with technology, and teachers' attitude towards technology. The two survey instruments used in this study were refined and adapted from the original USEIT teacher and student surveys (Russell, O'Dwyer, Bebell, & Miranda, 2003)[2]. Teacher and student survey responses were linked using teacher and student names.

Table 3 (next page) summarizes the student level technology use scales and their constituent items used to predict each outcome variable. Principal components analysis was used to confirm the existence of unidimensional student recreational home use and academic home use scales, and to create standardized composite measures of these uses that have a mean of zero and a standard deviation of 1. The reliabilities of the student academic home use of technology composite and the teachers' use of technology for student accommodation composite are quite low at 0.54 and 0.45, respectively and so their regression coefficients will be interpreted in light of this.

To facilitate comparisons among the magnitudes of the multilevel regression coefficients, the student school use measures, the socioeconomic measures, and the measure of prior achievement were standardized to have a mean of zero and a standard deviation of 1.

Table 3: Student Measures Identified during Exploratory Data Analysis Phase

Measurement Categories	Constituent Items
Student school use of technology *(Entered into models individually)*	How often do you use a computer in school to work with spreadsheets/databases?
	How often do you use a computer in school for math?
	How often does your teacher use a computer for math?
Student recreational home use of technology *Cronbach's alpha = 0.74*	How often do you use your home computer to play games?
	How often do you use your home computer to chat/instant message?
	How often do you use your home computer to email?
	How often do you use your home computer to search the Internet for fun?
	How often do you use your home computer to create Mp3/music?
Student academic home use of technology *Cronbach's alpha = 0.54*	How often do you use your home computer to search the Internet for school?
	How often do you use your home computer to write papers for school?
Socioeconomic status measures *(Entered into models individually)*	About how many books of your own do you have at home, not counting school books or comic books?
	How many computers, if any, do you have at home?

Taking advantage of the power of multilevel models for including group characteristics to predict individual outcomes, measures of teacher characteristics were included to predict student achievement. The teacher variables and composites included in the models are shown in Table 4 (next page). As was the case with the student level measures, principal components analysis was used to confirm the existence of unidimensional measurement scales and to create composite measures with a mean of 0 and a standard deviation of 1. The reliabilities of the teacher scales were lower than optimal for the teachers' use of technology for class preparation (0.64) and for student accommodation (0.45).

Table 4: Teacher Use of Technology Scales and Beliefs about Technology Measure

Measurement Categories	Constituent Items
Teachers' use of technology for delivering instruction	How often do you use a computer to deliver instruction to your class?
Teacher-directed student use of technology during classtime *Cronbach's alpha = 0.89*	During classtime how often did students work individually using computers this year?
	During classtime how often did students work in groups using computers this year?
	During classtime how often did students do research using the internet or CD-ROM this year?
	During classtime how often did students use computers to solve problems this year?
	During classtime how often did students present information to the class/ using a computer this year?
	During classtime, how often did students use a computer or portable writing device for writing this year?
Teacher-directed student use of technology to create products *Cronbach's alpha = 0.77*	How often did you ask students to produce multimedia projects using technology?
	How often do you ask students to produce reports and term papers using technology?
	How often did you ask students to produce web pages, websites or other web-based publications using technology?
	How often did you ask students to produce pictures or artwork using technology?
	How often did you ask students to produce graphs or charts using technology?
	How often did you ask students to produce videos or movies using technology?
Teachers' use of technology for class preparation *Cronbach's alpha = 0.64*	How often did you make handouts for students using a computer?
	How often did you create a test, quiz or assignment using a computer?
	How often did you perform research and lesson planning using the internet?
Teachers' use of technology for student accommodation *Cronbach's alpha = 0.45*	How often do you use a computer to prepare or maintain IEPs using a computer?
	How often do you use a computer to adapt activities to students' needs?

Methodology

Multilevel regression modeling was used to analyze the relationships among prior achievement, technology use, and students' fourth grade mathematics MCAS scores. Since the characteristics of students within a classroom may influence the attitudes and instructional practices of their teachers, and these practices in turn affect all of the students in the classroom, it is important to examine the classroom as a hierarchical organization within which technology-use occurs. A hierarchical approach to analyzing the relationship between technology-use and achievement requires the analysis of individuals within classrooms (Raudenbush & Bryk, 2002; Goldstein, 1995; Kreft & de Leeuw, 1998).

The analyses presented in this research were conducted using two-level hierarchical linear regression models. In these models, individual student's MCAS scores were modeled at level-1 as a function of students' school and home technology-uses, socioeconomic status indicators, and grade 3 MCAS performance. These models allowed the total variability in fourth grade MCAS mathematics achievement to be partitioned into its within-classroom and between-classroom variance components, and allowed predictors to be added at each level that explained a proportion of both the within-classroom and between-classroom variance available. Although it may be considered more appropriate to model achievement as varying within-classrooms, between-classrooms within-schools, and between-schools, it was not possible to reliably do so with these data. In order to be able to examine differences between classrooms within schools independent of the differences between schools, more classrooms and schools than are available in this sample would be required. For this reason, the between-classroom variability was confounded with the between-school variability in the models presented in this research.

The hierarchical regression analyses were carried out in stages. When conducting a hierarchical analysis, the first step is to examine of the amount of total variability in the outcome measures that exist within and between classrooms. To this end, unconditional models were formulated which included only random school effects. To develop a better understanding of the technology-uses that may be associated with mathematics performance, the second stage of the analysis involved theory-driven, exploratory data analysis to identify student and teacher variables observed to be associated with each of the outcome measures. Measures of several different types of technology-use first described in Bebell, Russell & O'Dwyer (2004) were examined during the exploratory data analysis phase.

Using predictor variables identified during the exploratory phase, increasingly complex multilevel models were constructed to predict each student outcome measure. In total, six models for each outcome measure were formulated. The multilevel models were constructed such that the impact of different categories of predictor variables could be independently assessed. The categories of interest were: prior achievement, socioeconomic status, home technology-use, school technology-use, classroom level measures of student achievement and socioeconomic status, and finally, teacher technology-use.

Results

Figures 1 and 2 display the distribution and mean response for each of the student and teacher survey items examined in this study, respectively. In these figures, the use measures are represented on the original four point scale used on the technology-use survey. Figure 1 shows that students tend to use technology at school less frequently than they use it at home; both recreational and academic technology-uses at home are higher than school uses. At home, students report using their computer at home to play games more frequently than other uses. The low levels of use of technology in school for math supports previous findings from NAEP secondary data analyses (Hedges, Konstantopoulis, & Thoreson, 2003).

Figure 1: Distribution and Mean Items Responses for Student Uses of Technology

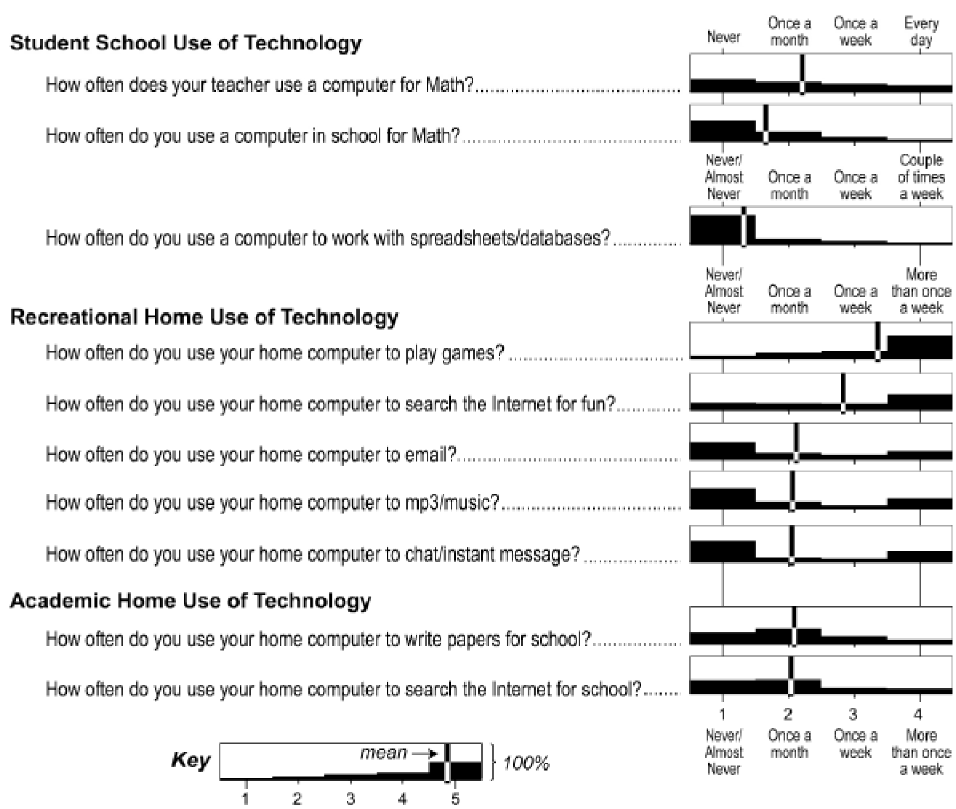

Figure 2 displays similar information for the items used to measure teachers' use of technology. The distributions show that teachers tend to use technology most frequently for preparation purposes outside of the classroom. Teachers also tend to have their students perform tasks using a computer more often than they have them create products using technology. Teachers report that they rarely use technology to deliver instruction in the classroom; on average, teachers only report using technology to deliver instruction several times a year.

Figure 2: Distribution and Mean Items Responses for Teacher Uses of Technology

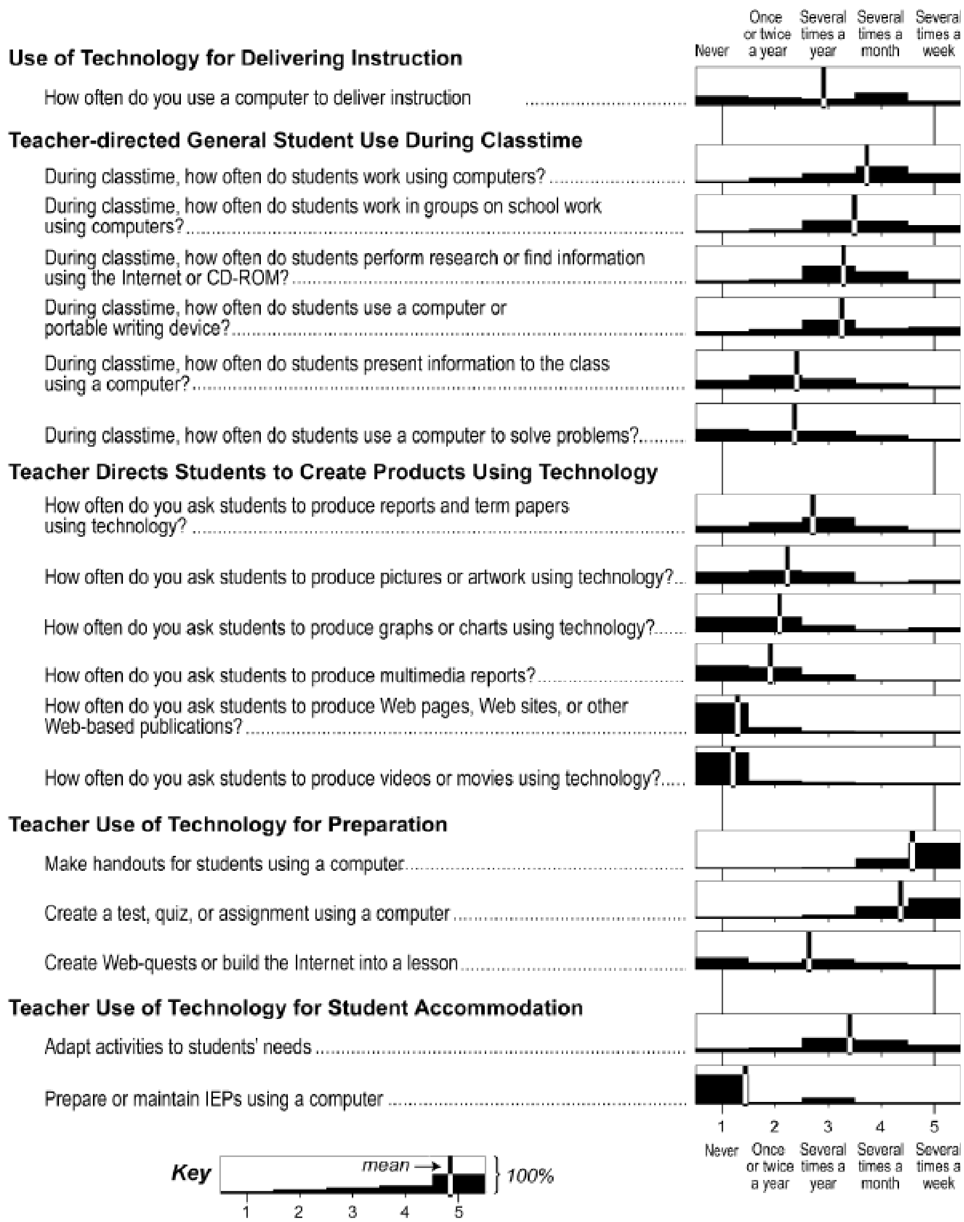

Table 5 presents the variance components for each of the standardized mathematics achievement score measures when the total variability in each was partitioned using the unconditional multilevel models. Although the majority of variability in each measure exists among students within classrooms, a significant proportion of the variability in achievement lies between classrooms for each outcome measure. The largest classroom-to-classroom differences occur for the total mathematics raw score and the number sense and operations component of the total score. For each of these measures, approximately 16% of the total variability in students' fourth grade mathematics scores exists between classrooms. The between classroom variance component is smallest for the patterns, relationships, and algebra component and the data analysis, statistics and probability component of the total mathematics score. Although the percentage of variability between classrooms for these two measures is very small, the unconditional models show that the variability between classrooms remains significant. The very small differences between classrooms for these measures will restrict the power of the level-2 models for using classroom level indicators to predict classroom-to-classroom differences.

Table 5: Unconditional Variance Components for the Math MCAS Outcome Measures

	Total Mathematics Raw Score	Number Sense & Operations Component of Math Score	Patterns, Relationships & Algebra Component of Math Score	Geometry Component of Math Score	Measurement Component of Math Score	Data Analysis, Statistics, & Probability component of Math Score
Percent of variance within classrooms	84%	83.4%	95%	86.8%	89.2%	94.0%
Percent of variance between classrooms	16%*	16.6%*	5%*	13.2%*	10.8%*	6.0%*

* The percentage of variability between schools is significant for $p < 0.001$.

During the second phase of the hierarchical analysis, characteristics measured at both the student and teacher levels were included in the unconditional models to explain some of the available variance. In all, six multilevel models were formulated for each of the six dependent variables. The models were constructed in a cumulative manner such that each model included additional categories of predictor measures. Model 1 included only third grade achievement to predict the fourth grade outcome measure. Model 2 included both prior achievement and indicators of socioeconomic status. Model 3 added students' use of technology at home for both recreational and academic purposes. In addition to the previous variables, Model 4 included measures of students' technology-use at school. Model 5, designed as an interim model, included measures of student achievement and socioeconomic status aggregated to the classroom level in the level-2 model. The sixth model incorporated measures of teachers' use of technology for predicting each of the outcome measures (Model 6). Each model will be discussed in turn.

Total Mathematics Raw Score

Table 6 (next page) contains six multilevel regression models constructed to examine the relationship between student and teacher characteristics and the total fourth grade mathematics raw score. Models 1 through 6 each shows that prior reading achievement is significantly positively related to fourth grade mathematics performance. The number of computers in a student's home, an indicator of the students' socioeconomic status, is also positively related to mathematics achievement and remains significant in each of the six models. The coefficients for students' use of computers at home, for either academic or recreational purposes, though negative are not significantly related to students' total mathematics test scores (Model 3). In addition, neither the frequency with which students use technology in school to work with spreadsheets/databases nor the frequency with which they or their teacher use computers for mathematics are significantly related to students' total mathematics scores (Model 4). In Model 5, the estimates of classroom mean socioeconomic status were small and non-significant. Classroom mean grade three achievement was positively and significantly related to fourth grade achievement.

Table 6: Total Mathematics Raw Score Model

Outcome = Math Raw Score	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6	
	Coeff.	Sig.	Coeff.	Sig.	Coeff.	Sig.	Coeff.	Sig.	Coeff.	Sig.	Coeff.	Sig.
Student Level Predictors												
How often do you use computers in school to work with spreadsheets/databases?							0.03	0.437	0.03	0.437	0.03	0.437
How often do you use a computer in school for Math?							–0.08	0.096	–0.08	0.096	–0.08	0.096
How often does your teacher use a computer for Math?							–0.01	0.728	–0.01	0.728	–0.01	0.728
Recreational home use					–0.07	0.076	–0.07	0.079	–0.07	0.079	–0.07	0.079
Academic home use					–0.04	0.43	–0.03	0.497	–0.03	0.497	–0.03	0.497
About how many books of your own do you have at home, not counting school books or comic books?			0.05	0.128	0.06	0.125	0.05	0.141	0.05	0.141	0.05	0.141
How many computers, if any, do you have at home?			**0.11**	**0.005**	**0.13**	**0.001**	**0.13**	**0.001**	**0.13**	**0.001**	**0.13**	**0.001**
Grade 3 Reading score	**0.30**	**0.000**	**0.29**	**0.000**	**0.28**	**0.000**	**0.28**	**0.000**	**0.28**	**0.000**	**0.28**	**0.000**
Teacher Level Predictors												
Teacher-mean student Grade 3 reading score									**0.62**	**0.000**	**0.67**	**0.000**
Teacher-mean number of books in student homes									–0.15	0.426	–0.20	0.230
Teacher-mean number of computers in student home									0.04	0.779	0.08	0.627
Teacher-directed student use of technology during classtime											0.08	0.362
Teachers direct students to create products using technology											**–0.16**	**0.024**
Teachers use technology for preparation											0.03	0.467
Teachers use technology to maintain IEPs											0.00	0.958
Teacher use of technology for delivering instruction											0.05	0.382

Table 6: Total Mathematics Raw Score Model (continued)

VARIANCE COMPONENTS		
	Available Variance	Percent Available
Between classrooms	0.15	16%
Within classrooms	0.77	84%
Total Variance Available	0.92	100%

Residual Variance and Variance Explained												
	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6	
	Res. Var.*	Var. Expl.**	Res. Var.*	Var. Expl.**	Res. Var.*	Var. Expl.**	Res. Var.*	Var. Expl.**	Res. Var.*	Var. Expl.**	Res. Var.*	Var. Expl.**
Between classrooms	0.15	0.0%	0.15	0.0%	0.15	0.0%	0.15	0.0%	0.09	39.6%	0.09	37.4%
Within classrooms	0.70	9.2%	0.69	10.6%	0.68	11.3%	0.68	11.4%	0.68	11.3%	0.68	11.3%
Total Variance Explained		8%		9%		9%		10%		16%		16%

* Residual Variance; ** Variance Explained

When classroom and teacher level characteristics were added to the model, classroom-mean prior achievement and the frequency with which teachers direct their students to create products using technology were significantly related to the total mathematics score (Models 5 and 6). The regression coefficients in Model 6 suggest that many of the teacher uses of technology included in the model were unable to predict the differences between classroom-mean total mathematics raw scores. Similarly, differences between classroom average raw scores did not appear to be attributable to socioeconomic status differences between classrooms; the classroom aggregate socioeconomic indicators were non-significant at level-2.

The percent of variability in the outcome explained by the models increased as more predictors were added to the models. Prior achievement and socioeconomic status indicators explained 9% of the total variability in the total mathematics score. Including home and school uses of technology in the model (Model 4) increased this amount by only one percentage point to 10%. Given that only predictors measured at the student level were included, Models 1 through 4 were unable to explain any of the variability in achievement among classrooms. It is interesting to note that although only 16% of the total variability in mathematics scores exists between classrooms as evidenced in the unconditional model, Models 5 and 6 explain more than 35% of the available between classroom variance. When considering the power of the model for explaining the total variability in the outcome, including classroom aggregate and teacher measures in the model increased the percent of total variability explained in the outcome from 10% to 16%.

Number Sense and Operations Subtest

Table 7 (next page) presents similar models for the number sense and operations component of the mathematics score. As was the case for the total score, prior achievement and the number of computers students report in their homes were positively and significantly related to students' number sense and operations scores. Students' use of technology at home for recreational purposes was significantly negatively related to this measure of achievement, suggesting that students who spend more time using computers at home for recreational purposes were likely to score lower on the number sense and operations component of the fourth grade MCAS mathematics test. Models 4 through 6 show that neither the frequency with which students use technology in school to work with spreadsheets/databases or the frequency with which they or their teacher use computers for mathematics were significantly related to this measure of students' mathematics ability.

In terms of the explanatory power of the models, prior achievement accounted for a substantial proportion of the total variability explained by the models. In Model 1, prior achievement explained 6% of the total available variance in the number sense and operations component of the total mathematics score. When other predictors were added to the models, the total variance explained increased to a maximum of 12% in Model 6.

Table 7: Number Sense and Operations Component of the Mathematics Subtest Model

Outcome = Number Sense and Operations Component of Math Score	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6	
	Coeff.	Sig.	Coeff.	Sig.	Coeff.	Sig.	Coeff.	Sig.	Coeff.	Sig.	Coeff.	Sig.
Student Level Predictors												
How often do you use computers in school to work with spreadsheets/databases?							0.00	0.955	0.00	0.955	0.00	0.955
How often do you use a computer in school for Math?							–0.05	0.340	–0.05	0.340	–0.05	0.340
How often does your teacher use a computer for Math?							0.01	0.733	0.01	0.733	0.01	0.733
Recreational home use					**–0.10**	**0.019**	**–0.10**	**0.020**	**–0.10**	**0.020**	**–0.10**	**0.020**
Academic home use					–0.04	0.425	–0.03	0.466	–0.03	0.466	–0.03	0.466
About how many books of your own do you have at home, not counting school books or comic books?			0.06	0.069	0.06	0.074	0.06	0.085	0.06	0.085	0.06	0.085
How many computers, if any, do you have at home?			**0.09**	**0.012**	**0.13**	**0.002**	**0.13**	**0.002**	**0.13**	**0.002**	**0.13**	**0.002**
Grade 3 Reading score	**0.27**	**0.000**	**0.26**	**0.000**	**0.25**	**0.000**	**0.25**	**0.000**	**0.25**	**0.000**	**0.25**	**0.000**
Teacher Level Predictors												
Teacher-mean student Grade 3 reading score									**0.58**	**0.001**	**0.60**	**0.001**
Teacher-mean number of books in student homes									–0.22	0.255	–0.26	0.147
Teacher-mean number of computers in student home									0.13	0.430	0.15	0.424
Teacher-directed student use of technology during classtime											0.04	0.635
Teachers direct students to create products using technology											–0.11	0.187
Teachers use technology for preparation											0.03	0.437
Teachers use technology to maintain IEPs											–0.02	0.839
Teacher use of technology for delivering instruction											0.04	0.537

Table 7: Number Sense and Operations Component of the Mathematics Subtest Model (continued)

VARIANCE COMPONENTS		
	Available Variance	Percent Available
Between classrooms	0.16	16.6%
Within classrooms	0.78	83.4%
Total Variance Available	0.94	100%

Residual Variance and Variance Explained												
	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6	
	Res. Var.*	Var. Expl.**	Res. Var.*	Var. Expl.**	Res. Var.*	Var. Expl.**	Res. Var.*	Var. Expl.**	Res. Var.*	Var. Expl.**	Res. Var.*	Var. Expl.**
Between classrooms	0.16	0.0%	0.16	0.0%	0.16	0.0%	0.16	0.0%	0.10	34.0%	0.11	26.4%
Within classrooms	0.72	7.4%	0.71	8.6%	0.71	9.8%	0.71	9.5%	0.71	9.5%	0.71	9.5%
Total Variance Explained		6%		7%		8%		8%		14%		12%

* Residual Variance; ** Variance Explained

Patterns, Relationships & Algebra Subtest

Table 8 (next page) presents analyses for the patterns, relationships, and algebra component of the mathematics test. In terms of variance structure, the percent of variance between classrooms is substantially smaller for this outcome variable than for the previous two outcomes; 5% for the patterns, relationships, and algebra component compared to approximately 16% for both the total raw score and number sense and operations component. The regression coefficients show that third grade achievement is a positive and significant predictor of this mathematics sub-domain score. The number of computers a student reports having in the home is also positively related to this achievement measure and is significant in Models 2 through 6. Unlike the model for the number sense and operations component of the total mathematics score, use of computers at home for recreational purposes is not a significant predictor for this outcome measure. Similar to the previous models, neither the frequency with which students use technology in school to work with spreadsheets/databases or the frequency with which they or their teacher use computers for mathematics are significantly related to students' total mathematics scores. Overall, including only student level predictors in the models explains only about 6% of the total variability among the patterns, relationships, and algebra scores.

When classroom-level predictors were added to the model, classroom mean prior achievement remains the only significant predictor at the classroom level and the model explains only 9% of the total available variance. Including teacher technology-use measures appears to be unable to improve the explanatory power of the model. The lack of power for explaining differences among classrooms is likely due to the small percentage of variability that exists between classrooms observed in the unconditional model. The frequency with which teachers direct their students to create products using technology is a significant and negative predictor of the differences among classrooms in Model 6 when included with several other teacher and classrooms predictors.

Table 8: Patterns, Relationships & Algebra Component of the Mathematics Subtest Model

Outcome = Patterns, Relationships & Algebra Component of Math Score	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6	
	Coeff.	Sig.	Coeff.	Sig.	Coeff.	Sig.	Coeff.	Sig.	Coeff.	Sig.	Coeff.	Sig.
Student Level Predictors												
How often do you use computers in school to work with spreadsheets/databases?							0.00	0.956	0.00	0.956	0.00	0.956
How often do you use a computer in school for Math?							–0.03	0.540	–0.03	0.540	–0.03	0.540
How often does your teacher use a computer for Math?							0.04	0.344	0.04	0.344	0.04	0.344
Recreational home use					–0.06	0.145	–0.06	0.140	–0.06	0.140	–0.06	0.140
Academic home use					–0.05	0.316	–0.05	0.323	–0.05	0.323	–0.05	0.323
About how many books of your own do you have at home, not counting school books or comic books?			0.03	0.497	0.03	0.450	0.03	0.497	0.03	0.497	0.03	0.497
How many computers, if any, do you have at home?			0.07	0.071	**0.10**	**0.027**	**0.10**	**0.021**	**0.10**	**0.021**	**0.10**	**0.021**
Grade 3 Reading score	**0.26**	**0.000**	**0.25**	**0.000**	**0.25**	**0.000**	**0.25**	**0.000**	**0.25**	**0.000**	**0.25**	**0.000**
Teacher Level Predictors												
Teacher-mean student Grade 3 reading score									**0.40**	**0.001**	**0.46**	**0.000**
Teacher-mean number of books in student homes									0.06	0.672	–0.03	0.861
Teacher-mean number of computers in student home									–0.05	0.579	–0.02	0.871
Teacher-directed student use of technology during classtime											0.09	0.161
Teachers direct students to create products using technology											**–0.13**	**0.017**
Teachers use technology for preparation											–0.01	0.896
Teachers use technology to maintain IEPs											–0.01	0.769
Teacher use of technology for delivering instruction											0.08	0.149

Table 8: Patterns, Relationships & Algebra Component of the Mathematics Subtest Model (continued)

VARIANCE COMPONENTS		
	Available Variance	Percent Available
Between classrooms	0.05	5.0%
Within classrooms	0.91	95.0%
Total Variance Available	0.96	100%

Residual Variance and Variance Explained												
	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6	
	Res. Var.*	Var. Expl.**	Res. Var.*	Var. Expl.**	Res. Var.*	Var. Expl.**	Res. Var.*	Var. Expl.**	Res. Var.*	Var. Expl.**	Res. Var.*	Var. Expl.**
Between classrooms	0.05	0.0%	0.05	0.0%	0.05	0.0%	0.05	0.0%	0.02	58.5%	0.02	63.1%
Within classrooms	0.86	5.8%	0.86	6.1%	0.85	6.6%	0.85	6.3%	0.85	6.2%	0.85	6.3%
Total Variance Explained		5%		6%		6%		6%		9%		9%

* Residual Variance; ** Variance Explained

Geometry Subtest

The models for the geometry component of the total mathematics score are detailed in Table 9 (next page). As was the case for the previous three outcome measures, prior achievement was significantly positively related to students' geometry subtest scores. The number of computers students' reported having in the home was also positively and significantly related to achievement in each of the six models. Student home use of technology for either recreational or academic purposes, though negatively related to achievement, was not a significant predictor of students' geometry scores. The frequency with which students reported using computers in school for math was significantly and negatively related to students' geometry scores. Given that prior achievement alone accounted for 5% of the total variability in geometry scores, adding socioeconomic indicators, and computer use at home and at school only improved the predictive power of the model by one percentage point; Models 2 through 4 only explain 6% of the total variability.

Adding class level predictors to the model increased the percent of variance explained to 10%. Classroom-mean prior achievement remained the only significant predictor at the classroom level in Models 5 and 6. Similar to the patterns, relationships, and algebra models, the frequency with which teachers directed their students to create products using technology was a significant and negative predictor of the differences among classrooms in Model 6 when included with several other teacher and classrooms predictors.

Table 9: Geometry Component of the Mathematics Subtest Model

Outcome = Geometry Component of Math Score	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6	
	Coeff.	Sig.	Coeff.	Sig.	Coeff.	Sig.	Coeff.	Sig.	Coeff.	Sig.	Coeff.	Sig.
Student Level Predictors												
How often do you use computers in school to work with spreadsheets/databases?							0.05	0.331	0.05	0.331	0.05	0.331
How often do you use a computer in school for Math?							**–0.09**	**0.037**	**–0.09**	**0.037**	**–0.09**	**0.037**
How often does your teacher use a computer for Math?							0.00	0.921	0.04	0.921	0.04	0.921
Recreational home use					–0.06	0.131	–0.06	0.134	–0.06	0.134	–0.06	0.134
Academic home use					–0.03	0.421	–0.03	0.457	–0.03	0.457	–0.03	0.457
About how many books of your own do you have at home, not counting school books or comic books?			0.04	0.252	0.04	0.231	0.04	0.267	0.04	0.267	0.04	0.267
How many computers, if any, do you have at home?			**0.09**	**0.024**	**0.11**	**0.011**	**0.11**	**0.010**	**0.11**	**0.010**	**0.11**	**0.010**
Grade 3 Reading score	**0.24**	**0.000**	**0.23**	**0.000**	**0.22**	**0.000**	**0.23**	**0.000**	**0.23**	**0.000**	**0.23**	**0.000**
Teacher Level Predictors												
Teacher-mean student Grade 3 reading score									**0.50**	**0.000**	**0.54**	**0.000**
Teacher-mean number of books in student homes									–0.16	0.289	–0.19	0.163
Teacher-mean number of computers in student home									0.11	0.430	0.13	0.352
Teacher-directed student use of technology during classtime											0.15	0.164
Teachers direct students to create products using technology											**–0.18**	**0.032**
Teachers use technology for preparation											0.01	0.770
Teachers use technology to maintain IEPs											–0.01	0.855
Teacher use of technology for delivering instruction											–0.02	0.823

Table 9: Geometry Component of the Mathematics Subtest Model (continued)

VARIANCE COMPONENTS		
	Available Variance	Percent Available
Between classrooms	0.12	13.2%
Within classrooms	0.82	86.8%
Total Variance Available	0.94	100%

Residual Variance and Variance Explained												
	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6	
	Res. Var.*	Var. Expl.**	Res. Var.*	Var. Expl.**	Res. Var.*	Var. Expl.**	Res. Var.*	Var. Expl.**	Res. Var.*	Var. Expl.**	Res. Var.*	Var. Expl.**
Between classrooms	0.12	0.0%	0.12	0.0%	0.12	0.0%	0.12	0.0%	0.09	30.8%	0.09	27.1%
Within classrooms	0.77	5.5%	0.76	6.3%	0.76	6.7%	0.76	6.9%	0.76	6.8%	0.76	6.9%
Total Variance Explained		5%		5%		6%		6%		10%		10%

* Residual Variance; ** Variance Explained

Measurement Subtest

The two-level regression results for the measurement component of the total mathematics score are contained in Table 10 (next page). As was the case for each of the previous models, prior achievement and the number of computers a student reports having in the home were positively and significantly related to students' measurement scores. Neither use of computers at home for academic or recreational purposes, or use of computers at school appears to be significantly related to this measure of achievement. In all, Models 1 through 4 which include only student level measures, explain less than 3% of the variability in the measurement scores that exists within classrooms and therefore, only 2% of the total variance.

At the classroom level, classroom-mean prior achievement and teachers' use of technology for preparation were significantly and positively related to this measure of student achievement. Similar to the models for the other outcomes, the frequency with which teachers directed their students to create products using technology was significantly and negatively related to student achievement. In total, the models did not explain a large percentage of the total variability in students' measurement scores; the most that was explained by any of the models is 6% (Model 6).

Table 10: Measurement Component of the Mathematics Subtest Model

Outcome = Measurement Component of Math Score	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6	
	Coeff.	Sig.	Coeff.	Sig.	Coeff.	Sig.	Coeff.	Sig.	Coeff.	Sig.	Coeff.	Sig.
Student Level Predictors												
How often do you use computers in school to work with spreadsheets/databases?							0.07	0.166	0.07	0.166	0.07	0.166
How often do you use a computer in school for Math?							–0.05	0.406	–0.05	0.406	–0.05	0.406
How often does your teacher use a computer for Math?							–0.06	0.076	–0.06	0.076	–0.06	0.076
Recreational home use					0.00	0.934	0.01	0.892	0.01	0.892	0.01	0.892
Academic home use					–0.02	0.604	–0.02	0.631	–0.02	0.631	–0.02	0.631
About how many books of your own do you have at home, not counting school books or comic books?			0.02	0.657	0.02	0.636	0.02	0.641	0.02	0.641	0.02	0.641
How many computers, if any, do you have at home?			**0.10**	**0.007**	**0.10**	**0.007**	**0.09**	**0.012**	**0.09**	**0.012**	**0.09**	**0.012**
Grade 3 Reading score	**0.14**	**0.002**	**0.14**	**0.003**	**0.14**	**0.003**	**0.14**	**0.003**	**0.14**	**0.003**	**0.14**	**0.003**
Teacher Level Predictors												
Teacher-mean student Grade 3 reading score									**0.51**	**0.000**	**0.52**	**0.000**
Teacher-mean number of books in student homes									–0.18	0.294	–0.21	0.208
Teacher-mean number of computers in student home									–0.03	0.851	0.01	0.974
Teacher-directed student use of technology during classtime											0.03	0.756
Teachers direct students to create products using technology											**–0.15**	**0.016**
Teachers use technology for preparation											**0.09**	**0.020**
Teachers use technology to maintain IEPs											–0.02	0.851
Teacher use of technology for delivering instruction											0.05	0.463

Table 10: Measurement Component of the Mathematics Subtest Model (continued)

VARIANCE COMPONENTS		
	Available Variance	Percent Available
Between classrooms	0.10	10.8%
Within classrooms	0.84	89.2%
Total Variance Available	0.94	100%

Residual Variance and Variance Explained												
	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6	
	Res. Var.*	Var. Expl.**	Res. Var.*	Var. Expl.**	Res. Var.*	Var. Expl.**	Res. Var.*	Var. Expl.**	Res. Var.*	Var. Expl.**	Res. Var.*	Var. Expl.**
Between classrooms	0.10	0.0%	0.10	0.0%	0.10	0.0%	0.10	0.0%	0.07	30.1%	0.06	36.6%
Within classrooms	0.83	1.9%	0.82	2.6%	0.82	2.4%	0.82	2.5%	0.82	2.4%	0.82	2.5%
Total Variance Explained		2%		2%		2%		2%		5%		6%

* Residual Variance; ** Variance Explained

Data Analysis, Statistics, and Probability Test

Table 11 (next page) presents the results for the data analysis, statistics, and probability component of students' total mathematics score. Similar to the previous models, prior achievement was significantly positively related to students' measurement scores. The number of books that students reported having in their homes was also significantly positively related to achievement while the frequency with which students report using computers at home for recreational purposes was negatively related to this measure of achievement. When classroom level predictors were added to the model, neither classroom-mean prior achievement, classroom socioeconomic indicators, or teachers' use of technology were significant predictors of the difference between classrooms in terms of the average data analysis, statistics, and probability scores. Table 11 also shows that these models were not very powerful for predicting the variability in students' data analysis, statistics, and probability scores; none of the models explained more than 10% of the variability in the outcome measure.

Table 11: Data Analysis, Statistics, & Probability Component of the Mathematics Subtest Model

Outcome = Data Analysis, Statistics, & Probability Component of Math Score	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6	
	Coeff.	Sig.	Coeff.	Sig.	Coeff.	Sig.	Coeff.	Sig.	Coeff.	Sig.	Coeff.	Sig.
Student Level Predictors												
How often do you use computers in school to work with spreadsheets/databases?							–0.05	0.332	–0.05	0.332	–0.05	0.332
How often do you use a computer in school for Math?							0.07	0.152	0.07	0.152	0.07	0.152
How often does your teacher use a computer for Math?							0.06	0.140	0.06	0.140	0.06	0.140
Recreational home use					–0.06	0.052	**–0.07**	**0.039**	**–0.07**	**0.039**	**–0.07**	**0.039**
Academic home use					–0.03	0.414	–0.03	0.392	–0.03	0.392	–0.03	0.392
About how many books of your own do you have at home, not counting school books or comic books?			**0.07**	**0.403**	**0.07**	**0.036**	**0.07**	**0.035**	**0.07**	**0.035**	**0.07**	**0.035**
How many computers, if any, do you have at home?			0.02	0.610	0.04	0.280	0.04	0.212	0.04	0.212	0.04	0.212
Grade 3 Reading score	**0.23**	**0.000**	**0.22**	**0.000**	**0.22**	**0.000**	**0.22**	**0.000**	**0.22**	**0.000**	**0.22**	**0.000**
Teacher Level Predictors												
Teacher-mean student Grade 3 reading score									0.25	0.057	0.27	0.060
Teacher-mean number of books in student homes									0.03	0.832	–0.01	0.949
Teacher-mean number of computers in student home									0.09	0.486	0.09	0.497
Teacher-directed student use of technology during classtime											0.03	0.752
Teachers direct students to create products using technology											–0.07	0.347
Teachers use technology for preparation											0.02	0.644
Teachers use technology to maintain IEPs											–0.06	0.249
Teacher use of technology for delivering instruction											0.04	0.577

Table 11: Data Analysis, Statistics, & Probability Component of the Mathematics Subtest Model (continued)

VARIANCE COMPONENTS		
	Available Variance	Percent Available
Between classrooms	0.06	6.0%
Within classrooms	0.91	94.0%
Total Variance Available	0.97	100%

Residual Variance and Variance Explained												
	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6	
	Res. Var.*	Var. Expl.**	Res. Var.*	Var. Expl.**	Res. Var.*	Var. Expl.**	Res. Var.*	Var. Expl.**	Res. Var.*	Var. Expl.**	Res. Var.*	Var. Expl.**
Between classrooms	0.06	0.0%	0.06	0.0%	0.06	0.0%	0.06	0.0%	0.04	29.6%	0.04	22.8%
Within classrooms	0.86	4.7%	0.86	5.0%	0.86	5.3%	0.86	5.5%	0.86	5.4%	0.86	5.4%
Total Variance Explained		4%		5%		5%		5%		7%		6%

* Residual Variance; ** Variance Explained

Discussion

To examine the relationship between technology uses at home and at school and students' mathematics achievement, the study presented here developed several statistical models in which multiple measures of technology-use were used to predict the performance of fourth grade students on the MCAS mathematics test. Recognizing that the MCAS mathematics test assesses several different mathematics sub-domains, the analyses focused on the effects of multiple uses of technology on students' total test scores as well as their performance within five specific sub-domains. To account for differences in prior achievement and to control for the relationship between socioeconomic status and student performance, the analyses also employed third grade MCAS Reading test scores and two measures of students' socioeconomic status as covariates. Finally, to separate the effects of individual factors and classroom level factors, the analyses employed hierarchical linear regression modeling techniques.

As expected, the analyses of total test scores and each sub-domain score indicate that prior achievement and SES are significant predictors of fourth grade MCAS mathematics scores. This relationship was consistent across all analyses. In addition, with the exception of geometry, all measures of students' use of technology in school included in the analyses did not show a significant (positive or negative) relationship with students' test scores. For geometry, the measure that focused on general use of computers in school for math showed a small negative relationship with geometry sub-domain scores. Similarly, teachers' use of computers was significantly related to students' performance for only one sub-domain, namely the measurement subtest. More specifically, the relationship between teachers' use of computers to prepare for instruction and students' geometry scores was small but positive, while the relationship between teachers directing students to use computers to create products and geometry scores was small and negative. Students' use of computers at home for academic purposes was not significantly related to test performance for any of the sub-tests. However, the relationship between students' use of computers at home for recreational purposes and the number sense and operations, and with the data analysis, statistics, and probability sub-domain scores were each small and negative.

As described in the tables, each of the statistical models accounted for a relatively small percent of the total variance in test scores. Specifically, the largest percentage of total variance explained by the models occurred for the total test score (16%) and the number sense and operations sub-domain scores (12%). Models for the remaining sub-domain scores accounted for 10% or less of the total variance, with the data analysis, statistics, and probability model accounting for the least amount of variance

(5%). In part, the low amount of variance accounted for by these models may result from the relatively poor reliability of the sub-test scores on the MCAS. As shown in Table 2 (page 13), only the total test score (0.86) and the number sense and operations sub-domain (0.71) scores had reliability indices greater than 0.5. In addition, the use of a covariate that focused on prior reading performance, rather than prior mathematics performance, may have decreased the variance accounted for by prior achievement.

Despite these two shortcomings, perhaps the most noticeable aspect of these analyses is how little variance is accounted for by any individual measure of student or teacher computer use or by the collective set of uses. Although approximately one third of the classrooms that participated in this study were selected because use of technology for instructional purposes was reported to be relatively high compared to all the classrooms that participated in the USEIT Study (recall that three groups of teachers representing high, medium, and low levels of instructional technology-use on their USEIT study responses were re-surveyed for the present study), this lack of explanatory power may result from the infrequent use of computers for mathematics that occurs in these "high-use" classrooms. As was seen in Figure 1 (page 18), very few students reported that they or their teachers used computers more than once per month for mathematics and even fewer students reported using spreadsheets or databases during the school year. Thus, the majority of 4th grade respondents reported using technology in their mathematics class very rarely, with most students stating they had used computers fewer than nine times (once per month) during the school year. Although many teachers reported they use computers to deliver instruction, this use did not typically occur during mathematics instruction or directly involve students. Thus, the potential for students' use of computers to impact student achievement in mathematics was relatively limited given the lack of student computer use reported in mathematics class.

Although this study employed multiple measures of student and teacher technology-use, attempted to control for prior achievement by employing prior year test scores as a covariate, and employed multi-level modeling techniques to account for individual level and classroom level factors that influence test performance, this study and future efforts to examine the relationship between technology use and achievement could be improved in several ways. First, rather than employing prior year reading scores as a covariate, it would be desirable to include test scores that are more closely aligned with the constructs measured by the current year test(s), namely mathematics.

Second, although multiple measures of teacher and student computer use were employed, many of these measures were still relatively vague. As an example, the most specific student use item asks students how often they use a computer to work with spreadsheets and/or databases. While this item is more specific than asking students how often they use a computer for math (which is another item in the survey), it clusters all potential uses of spreadsheets to explore mathematical concepts into one item. These potential uses may include recording data, creating and working with graphs, creating and exploring algebraic functions, performing basic arithmetic, exploring number patterns, or exploring statistical concepts. While it is unlikely that many of these specific uses are occurring in the classrooms included in this study given the relatively small amount of use for mathematics, in truly high-use settings grouping these multiple and distinct uses into a single measure is likely to obfuscate the effects of technology-use on student learning.

Similarly, unlike Language Arts tests, which are composed primarily of reading and writing skills and are absent content, mathematics includes a large body of content. As seen in the fourth grade MCAS mathematics test, at least five content areas are expected to be covered during the fourth grade. For each of these content areas, computers may be used in a variety of ways to help students develop their understanding. However, a given use of a computer to develop skill and knowledge in one content area may not be effective for another content area. As an example, building and exploring graphs may be highly effective for geometry and statistics, but not as useful for number sense or basic arithmetic. Thus, when measuring students' use of computers, it may be important to not only develop more precise measures of what students are doing with computers, but also what content students are learning as they use the computers. Although developing such a detailed technology-use instrument would require considerable time and attention and would likely require considerable time for students to complete, doing so would enable richer and more precise estimates of the effects of computer use on students' learning as measured by standardized test scores.

Finally, to better understand the effects of computer use on students' mathematical development, it may be necessary to survey a much larger body of classrooms to identify settings in which high levels of computer use occur. Alternatively, rather than sampling widely and then narrowing on high-use settings, a future study might begin by identifying settings where use is likely to be high and then sample from within those settings to identify classrooms with the highest levels of use. As an example, several districts and states have recently developed settings in which each student has their own laptop computer. While access does not equal use, these 1:1 settings may provide fertile ground from which truly high-use classrooms may be identified.

Ultimately, studies designed to examine the impact of educational technologies on student learning will continue to be inconclusive until methodologically sound large-scale longitudinal studies that examine instructional uses of computers to develop specific content knowledge and skills within classrooms are conducted. In particular, large-scale longitudinal studies are needed that (a) examine how all types of educational technologies are used by students, (b) accurately capture students' myriad uses of technology, and (c) look beyond state test scores or off-the-shelf norm-referenced tests as a way to assess the impact of technology. Until such studies are conducted, the conclusions drawn about the nature of the relationship between technology use and academic achievement will continue to be questioned by both critics of and advocates for educational technology.

Endnotes

1. Here, we use the term "business as usual" to describe the standard, every-day practices around technology use in schools.
2. Original USEIT surveys are available at http://www.bc.edu/research/intasc/researchprojects/USEIT/useit.shtml.

References

Angrist, J., & Lavy, V. (2002). New evidence on classroom computers and pupil learning. *The Economic Journal, 112*, 735–765.

Bausell, C.V., & Klemick, E. (2007, October). Tracking U.S. Trends. *Education Week, 26*(30), p. 42–44.

Bebell, D., Russell, H., & O'Dwyer, L.M. (2004). Measuring teachers' technology uses: Why multiple-measures are more revealing. *Journal of Research on Technology in Education, 37*(1), 45–63.

Coley, R., Cradler, J., & Engel, P.K. (1997). *Computers and classrooms: The status of technology in U.S. schools*. Policy information report. Princeton, NJ: Educational Testing Service Policy Information Center. (ERIC Document Reproduction Service No. ED412893).

Cuban, L. (2001). *Oversold and underused*. Cambridge, MA: Harvard University Press.

Dynarski, M., Agodini, R., Heaviside, S., Novak, T., Carey, N., Campuzano, L., Means, B., Murphy, R., Penuel, W., Javitz, H., Emery, D., & Sussex, W. (2007). *Effectiveness of Reading and Mathematics Software Products: Findings from the First Student Cohort*. Washington, D.C.: U.S. Department of Education, Institute of Education Sciences.

Fletcher-Flinn, C.M., & Gravatt, B. (1995). The efficacy of computer assisted instruction (CAI): A meta-analysis. *Journal of Educational Computing Research, 12*(3), 219–242.

Goldberg, A., Russell, M., & Cook, A. (2003). The effect of computers on student writing: A meta-analysis of studies from 1992 to 2002. *Journal of Technology, Learning, and Assessment, 2*(1). Available from http://www/jtla.org

Goldstein, H. (1995). *Multilevel statistical models*. London: Edward Arnold.

Haney, W. (1980). Units and levels of analysis in large-scale evaluation. *New Directions for Methodology of Social and Behavioral Sciences, 6*, 1–15.

Hedges, L.V., Konstantopoulis, S., & Thoreson, A. (2003). *Computer Use and Its Relation to Academic Achievement in Mathematics, Reading and Writing. NAEP Validity Studies*. Washington, DC: National Center for Education Statistics (ED).

Kreft, I., & de Leeuw, J. (1998). *Introducing multilevel modeling*. Thousand Oaks, CA: Sage.

Kulik, J.A. (1994). Meta-analytic studies of findings on computer-based instruction. In E. L. Baker & H. F. O'Neil, Jr. (Eds.), *Technology assessment in education and training* (pp. 9–33). Hillsdale, NJ: Lawrence Erlbaum Associates.

Mann, D., Shakeshaft, C., Becker, J., & Kottkamp, R. (1999). *West Virginia story: Achievement gains from a statewide comprehensive instructional technology program*. Retrieved December 1, 2003, from http://www.mff.org/publications/publications.taf?page=155

Market Data Retrieval. (2001). *Technology in education 2001*. Shelton, CT: Author.

O'Dwyer, L.M., Russell. M., Bebell, D. (2004, September 14). Identifying teacher, school and district characteristics associated with elementary teachers' use of technology: A multilevel perspective. *Education Policy Analysis Archives, 12*(48). http://epaa.asu.edu/epaa/v12n48/.

O'Dwyer, L.M., Russell. M., Bebell, D., & Tucker-Seeley, K. (2005). Examining the relationship between home and school computer use and students' English/language arts test scores. *Journal of Technology, Learning, and Assessment, 3*(3). Available from http://www.jtla.org

Penuel, B., Yarnall, L., & Simkins, M. (2000, September). Do technology investments pay off? *The evidence is in!* [Electronic version]. Leadership, v30 p.18.

Raudenbush, S.W., & Bryk, A. (2002). *Hierarchical linear models: Applications and data analysis methods*. Sage Publications. Russell, M., O'Dwyer, L.M., Bebell, D., & Miranda, H. (2003). *The USEIT study technical report*. Boston, MA: Technology and Assessment Study Collaborative.

Sivin-Kachala, J., & Bialo, E.R. (1994). *Report on the effectiveness of technology in schools*, 1990–1994. Washington, DC: Software Publishers Association. (ERIC Document Reproduction Service No. ED 371 726)

Wenglinsky, H. (1998). *Does it compute? The relationship between educational technology and student achievement in mathematics*. Princeton, NJ: Educational Testing Service Policy Information Center.

Author Biographies

Laura O'Dwyer is an Assistant Professor at the Educational Research, Measurement and Evaluation department in the Lynch School of Education at Boston College. Her research interests include examing the effects of contextual characteristics on individual outcomes, international comparative studies, and teacher professional development.Her work also examines the impact of school organizational characteristics on the adoption of technology as a teaching and learning tool. Laura O'Dwyer can be reached at odwyerl@mail.bc.edu.

Michael Russell is an Associate Professor in the Lynch School of Education at Boston College and directs the Technology and Assessment Study Collaborative (inTASC) which is housed within the Center for the Study of Testing, Evaluation, and Educational Policy at Boston College. His research interests lie at the intersection of technology, assessment, and educational policy.

Damian Bebell is an Assistant Research Professor at Boston College's Lynch School of Education and a Research Associate at the Technology and Assessment Study Collaborative. Damian has been actively involved in the study of measurement issues concerning teachers and student use of technology. Damian's work also examines the emergence of 1:1 computing environments.

Kevon Tucker-Seeley is a doctoral candidate in the Lynch Graduate School of Education's Educational Research Measurement and Evaluation program at Boston College. He is currently working at Education Development Center (EDC) as a research assistant. Kevon's research interests, in general, revolve around the cognitive and socioemotional development of pre-adolescent children. Specifically, his interests lie at the intersection of race/ethnicity, socioeconomic status (SES), and bio-ecological contexts and focus on the collective impact that these variables have on achievement outcomes in a formal school environment.

J·T·L·A

The Journal of Technology, Learning, and Assessment

www.jtla.org

Technology and Assessment Study Collaborative
Caroline A. & Peter S. Lynch School of Education, Boston College