JTLA

Examining the Relationship Between Home and School Computer Use and Students' English/ Language Arts Test Scores

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

www.jtla.org

Examining the Relationship Between Home and School Computer Use and Students' English/Language Arts Test Scores

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

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

Copy Editor: Rachel Kay
Design and Layout: Thomas Hoffmann

Preferred citation:

Note:

Abstract:

With increased emphasis on test-based accountability measures has come increased interest in examining the impact of technology use on students' academic performance. However, few empirical investigations exist that address this issue. This paper (1) examines previous research on the relationship between student achievement and technology use, (2) discusses the methodological and psychometric issues that arise when investigating such issues, and (3) presents a multilevel regression analysis of the relationship between a variety of student and teacher technology uses and fourth grade test scores on the Massachusetts Comprehensive Assessment System (MCAS) English/Language Arts test. In total, 986 fourth grade students from 55 intact classrooms in nine school districts in Massachusetts were included in this study. This study found that, while controlling for both prior achievement and socioeconomic status, students who reported greater frequency of technology use at school to edit papers were likely to have higher total English/language arts test scores and higher writing scores. Use of technology at school to prepare presentations was associated with lower English/language arts outcome measures. Teachers' use of technology for a variety of purposes were not significant predictors of student achievement, and students' recreational use of technology at home was negatively associated with the learning outcomes.

Examining the Relationship Between Home and School Computer Use and Students' English/Language Arts Test Scores

Laura M. O'Dwyer
University of Massachusetts, Lowell

Michael Russell, Damian Bebell, & Kevon R. Tucker-Seeley
Center for the Study of Testing, Evaluation, and Educational Policy
Lynch School of Education
Boston College

Introduction

Over the past 20 years, substantial investments have been made in educational technology. Since 1996, state and district level agencies have invested over ten billion dollars to acquire and integrate computer-based technologies into American schools. During this period, the federal government has spent another three billion dollars on educational technology. Recently, Quality Education Data, Inc. (QED) estimated that during the 2003–2004 school year, approximately $400 million was spent on software and an additional $220 million would be invested in the very latest technological trend: wireless networks that support portable wireless laptops and wireless-enabled handhelds (QED, 2003).

Proponents of educational technology argue that when technology is used effectively in the classroom, investments in technology lead to more highly developed critical thinking skills, stronger problem-solving skills, and higher-order levels of understanding (Penuel, Yarnell, & Simkins, 2000; Salpeter, 2000). Some observers also 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). These positive effects are documented in several hundred formal and informal evaluation and research studies conducted since the early 1980s (Sivin-Kachala & Bialo, 1994; Coley, Cradler, & Engel, 1997; Mann, Shakeshaft, Becker, & Kottkamp, 1999).

Critics, however, argue that few of these studies meet rigorous empirical methodological standards or directly link use of technology in the

classroom with improved standardized test scores (Angrist & Lavy, 2002; Cuban, 2001). In this vein, several observers have questioned the extent to which technology is impacting teaching and learning. For example, Stoll (1999) and Oppenheimer (2003) have criticized investments in educational technologies arguing that there is little evidence that they impact teaching and learning in a positive way. In essence, little has changed since a 1999 U.S. Department of Education sponsored conference where it was reported that "parents and teachers, school boards and administrators, governors and state legislatures, and Congress all want to know if the nation's investment in technology is providing a return in student achievement. Indeed, if resources are to be expended on technology, it is becoming a political, economic, and public policy necessity to demonstrate its vital effectiveness" (McNabb, Hawkes & Rouk, 1999, p. 1). 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, and standardized test scores remain the customary means for evaluating the benefits of educational innovations (McNabb, Hawkes, & Rouk, 1999).

Yet, some observers recognize that standardized test scores may not provide valid measures of the learning that occurs when students use technology. McNabb et al. (1999) contend that "...the tools [used to] measure basic skills don't evaluate how technology supports students in developing capacities to think creatively and critically and vice versa" (p.10). Similarly, Russell (2002) argued that most standardized tests attempt to measure a domain broadly, yet technology is often used in the classroom to develop a specific skill or information that may be tested by just one or two items on a standardized test. As a result, standardized test scores are often not sensitive enough to measure the changes in learning that may occur when technology is used to develop specific skills or knowledge. In addition, Russell argues that many of today's paper-based standardized tests may be underestimating the performance of students who are accustomed to working with technology simply because they do not allow students to use these technologies when being tested. Through a series of randomized experiments, Russell and his colleagues provide empirical evidence which suggests that students who are accustomed to writing with computers in the classroom perform between 0.4 and 1.1 standard deviations higher when they are allowed to use a computer when performing tests that require them to compose written responses (Russell, 1999; Russell & Haney, 1997; Russell & Plati, 2001). Becker (as cited in Kirkpatrick & Cuban, 1998) also maintains that standardized achievement tests fail to accurately capture technology's role even when gains are achieved and notes that when "standardized test scores rise, it's difficult to discern whether the rise was due to the students' work with computers or another

influence" (p. 6). Together, these shortcomings of standardized tests complicate efforts to examine the direct effects of technology use on student learning.

In the following pages we summarize a small number of previous investigations into the relationship between use of educational technology and student achievement. First, we present a few examples of smaller studies where specific programs or applications of technology have been examined in a research setting and summarize findings from meta-analyses of these smaller studies. Next, we examine three large-scale studies that employ standardized test scores to examine the relationship between educational technology use and student achievement. We then detail some of the psychometric and methodological challenges that are inherent in a study of the relationship between technology use and student achievement. Finally, we describe the methodology and present findings from a study intended to overcome some of these methodological challenges in order to examine the relationship between technology use in fourth grade classrooms and student performance on a state-mandated English/language arts test.

Small-Scale Specific Analyses

The majority of research that examines the impact of educational technology on learning focuses on discrete technology programs, software, or applications. In most cases, these studies employ small, non-representative samples of students who use technology in classrooms that receive atypically high levels of support. To measure student achievement, many of these studies use a measure that is developed by the research team and has not undergone extensive validation. As an example, one such study was conducted by McFarlane, Friedler, Warwick and Chaplain (1995). In this study, seven and eight year old students used computer-based probes and accompanying graphing software to collect data as a means to help them develop a better understanding of line graphs. This research was conducted in a small number of elementary classrooms, in which the researchers actively worked with the teachers to use the probes and software. The study reported the experience of introducing the new technology into the classroom and also included pre- and post-test scores for those students in these classrooms to examine changes in student learning.

Similarly, Ramirez and Althouse (1995) developed and examined the Palm Beach County Geographic Information System (GIS) Project which employed ArcView software to help students examine various environmental issues. The project involved ten teachers from six high schools who received training on the software and worked with the researchers to develop a two-semester curriculum that employed the software. Each teacher then implemented his or her curriculum. The study examined the

ways in which the software was used within each curriculum to help students develop an understanding of environmental sciences.

These studies are helpful in providing preliminary evidence about the potential impacts of technology uses on student learning. However, by themselves the narrow focus on a discrete use of technology, the active participation of the research team in helping teachers and students use the technology in the classroom, the small, non-representative samples, and the use of non-standardized measures of learning leave unanswered questions about the generalizeability of the findings from any one of these studies.

In an effort to estimate the effects of technology on learning reported across this body of research, several meta-analyses have been performed. Meta-analyses are particularly valuable for summarizing the methodological properties and variations across a wide variety of research studies on a given topic (Glass, McGaw, & Smith, 1981). Generally, meta-analyses examining educational technology use and achievement 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). However, it is worth noting that meta-analytic studies are potentially prone to methodological weaknesses which result from different criteria for selecting samples, different experimental and analytic methodologies, different experimental contexts, and possible study selection biases.

One of the most recent meta-analyses concerning technology and achievement is the work of Waxman, Lin, & Michko (2003) who retrieved nearly 200 educational technology research studies published between 1997 and 2003. Looking across the educational technology studies, Waxman et al. first found that the majority of quantitative studies published within those six years "did not meet the standard" for inclusion in their meta-analysis and that "the lack of quality, refereed quantitative studies points to a serious problem of research in the field" (Waxman et al., 2003, p.13). Of the 42 studies that were deemed appropriate for inclusion in the meta-analysis, only 25% were categorized as randomized experimental designs and 67% were categorized as quasi-experimental designs. Similarly, Waxman reported that the studies varied widely in their definitions and measurement of "technology" and "student achievement". Specifically, the researchers found that 30% of their final sample of research studies investigated the use of personal computers, 26% investigated networked labs, 5% investigated multimedia applications, and the remaining 39% used some variety of other technology resources. In terms of student academic/cognitive outcomes addressed across the 42 studies, 38% used a researcher-constructed test, 14% used authentic assessments, and 10%

used standardized tests. It is noteworthy that the majority of the studies included in the meta-analysis also focused on non-achievement based student outcomes, including affective outcomes (57%) and behavioral outcomes (83%). The overall results of Waxman et al.'s meta-analysis suggest that the various educational technologies examined relate positively to several student measures, including academic achievement and cognitive outcomes.

As demonstrated by the work of Waxman et al. (2003), educational researchers have explored the effects of educational technology on student achievement from a variety of perspectives. This is not surprising given that there are a wide variety of ways to use computers to support the learning process. In some cases, these uses focus on productivity tools, such as word processors, hypermedia authoring software, or spreadsheet creation. In other cases, the research has focused on software designed to help students develop an understanding of specific topics. In still other cases, the focus has been on information and data collection tools, such as CD ROM-based encyclopedias, the Internet, science probes, or Global Positioning System (GPS) use. Given the wide variety of computer-based tools that have been examined, this line of research is fractured by the focus on how particular tools can be used to improve specific types of student learning within specific populations. However difficult to generalize, the hundreds of small-scale studies highlight the challenges faced when measuring the effect of technology use in educational settings. Moreover, given that the findings from this large body of research are de-emphasized by policy-makers because of the relatively small sample sizes and use of non-standardized measures of learning, these studies also highlight the perceived need for systematic, large-scale research that employs measures of learning that are accepted by critics and policy-makers.

Large-Scale General Analyses

In recent years, a number of large-scale investigations have been conducted to examine the impact of educational technology on student achievement. As an example, Mann, Shakeshaft, Becker and Kottkamp (1999) employed a sample of 290 teachers and 950 fifth graders in West Virginia to examine the effects of the Basic Skills/Computer Education (BS/CE) program. With measures of teachers' and students' level of technology participation, teacher and student attitudes, teacher training, students' prior achievement, and socioeconomic status, the researchers used multiple measures of student achievement to examine the impacts of the technology program across the curriculum. The results suggest that the BS/CE program had positive effects in the schools in which it was most heavily integrated, resulting in the attainment of practical and statistically significant gains in Stanford-9 reading, writing, and math achievement scores (Mann et al.,1999). In addition, the researchers found that in

comparison to other intervention alternatives for raising test scores, the "...BS/CE [program] was...more cost-effective than other popular interventions including class-size reduction....[and was] especially successful with low income and rural students as well as with girls" (1999, p. 3).

Through a series of regression analyses, Mann et al. (1999) also found that approximately 70% of the variation in students' standardized test scores was associated with factors outside of the school including family characteristics, home life, and socioeconomic status. In other words, Mann et al. estimated that any school-based educational reform (technology or otherwise) has the potential to account for only 30% of the total variance in students' achievement scores. Given the 11% of variance in achievement gains explained by Mann et al., these results suggest that the BS/CE program may have had a particularly strong impact on achievement across the curriculum.

A second large-scale study that examined the general impacts of educational technology on students' academic achievement is the work of Angrist and Lavy (2002). In their highly publicized study, Angrist and Lavy (2002) used Israeli school data from a 1996 administration of a standardized middle school Mathematics and Hebrew test to examine the relationship between educational technology use and student achievement. In this study, the authors compared levels of academic achievement among students classified as receiving instruction in either high or low technology environments. Unlike Mann et al. (1999), who measured the frequency with which students used computers in school, Angrist and Lavy (2002) did not measure use, but instead focused on technology access using the ratio of students to computers in a school. In other words, the authors use students' access to technology rather than the extent to which students actually use technology as the independent variable in their analyses. Schools designated by the authors as high access schools were equipped with computers at a 10:1 ratio, meaning 10 students share 1 computer. The results of their analyses show none of the positive effects of technology that were reported by Mann et al. (1999). In fact, the results of Angrist and Lavys' analyses suggest that mathematics scores were lower in schools that were equipped with computers at a 10:1 ratio.

A third approach to measuring the impact of educational technology on student achievement was adopted by Wenglinsky (1998) who performed secondary data analyses of fourth and eighth grade 1996 National Assessment of Educational Progress (NAEP) data. Unlike the two large-scale studies discussed above (Mann et al., 1999, and Angrist & Lavy, 2002), Wenglinsky focused his research on the impact of general technology use and a specific type of student reported technology use (simulation and higher order thinking technologies and software) on mathematics achievement. Using a nationally representative NAEP sample, Wenglinsky employed empirical measures of teacher characteristics (including

professional development experiences), students' socioeconomic status, and aggregated measures of class size and school climate to measure the impacts of technology use. By matching classrooms on the above measures, Wenglinsky concluded that both fourth and eighth grade students who used simulation and higher order thinking software had statistically significantly higher mathematics achievement. However, when considering general student technology use, Wenglinsky found that computer use was actually negatively related to mathematics achievement for grades 4 and 8. It should be noted that Hedges, Konstantopoulis, and Thoreson (2003) have cautioned against placing too much emphasis on Wenglinsky's findings and have argued that "the design of the NAEP data collection precludes using such data to make even tentative conclusions about the relationship of achievement and computer use" (p. 1).

The findings from these large-scale studies, each of which employed standardized measures of student learning across large samples of students, present a mixed message regarding the relationship between technology use and student achievement. Whereas Mann et al. (1999) report positive effects, Wenglinsky (1998) found mixed effects depending upon the type of use examined, and Angrist and Lavy (2002) found that access to technology was not related to Hebrew achievement and was negatively related to mathematics achievement. Beyond the actual findings, these studies provide valuable lessons regarding the psychometric and methodological issues related to examining impacts of technology on student learning.

Psychometric and Methodological Issues

As described briefly above, much of the research on the effect of the use of educational technology on student learning is limited by: (1) the way in which students' and teachers' technology use is measured; (2) measures of achievement that are not specifically designed to capture the types of improved learning that are intended to occur as a result of technology use; and (3) reliance on either aggregate school level data or individual level data within classrooms which does not take into account differences within and between schools when modeling student outcomes.

Measuring Technology Use

While there is a strong desire to examine the impact of technology use on student achievement, research suggests that the impacts on learning must be placed in the context of teacher and student technology use (Bebell, Russell and O'Dwyer, 2004). In other words, before the outcomes of technology integration can be rigorously studied, the following conditions should be met: (1) there must be a clear understanding of how teachers and students are using technology; and (2) valid and reliable measures of these uses must be developed. Instead of developing measures of use, many research studies that examine the relationship between achievement and technology use assume that teachers' and students' *access* to technology is an indicator of technology use.

As an example, Angrist and Lavy (2001) use a school's student-to-computer ratio as a proxy for technology use, and designated schools in which one computer was available for every ten students as high-use settings. Given this approach, it is not surprising that such limited access to technology resources resulted in the observed negligible impact on student achievement. Similarly, Wenglinksi's work (1998) demonstrated that the type of technology use employed has consequences for the conclusions. Therefore, depending upon whether and how one measures use, the relationship between technology use and achievement seems to differ. Moreover, it is important to note that recent research suggests that there are a variety of ways in which students and teachers use technology and these are not equivalent to access (Bebell et al., 2004; O'Dwyer, Russell, & Bebell, 2004).

Aligning Learning Outcome Measures with Use

When examining the impact of 1:1 technology use on student learning, it is critical that the outcome measures actually assess the types of learning that may occur as a result of technology use and that those measures are sensitive enough to detect potential changes in learning that may occur. As mentioned earlier, since most standardized tests attempt to measure a domain broadly, standardized test scores often do not provide measures that are aligned with the learning that may occur when technology is used to develop specific skills or knowledge (Russell, 2002). As an example, some teachers use computers to develop students' spatial reasoning skills. While state tests may test these skills, they often include only one or two items that are specific to spatial reasoning. Given the limited information regarding spatial reasoning, it is unlikely that such tests could be used to detect changes in students' spatial reasoning skills that result from computer use.

Similarly, the constructs measured by standardized tests are not always aligned with the constructs developed through computer use. As an example, many advocates of technology believe that computer use may have a positive impact on students' problem-solving skills. Most mathematics tests include items that test students' mathematical problem-solving skills. Typically, these items take the form of word problems for which students must define a function that represents the relationship described, enter the appropriate numbers, and perform accurate computations. While it is important for students to develop these mathematical problem-solving skills, problem solving with computers involves more than simply decoding text to define functions. As Dwyer (1996) describes, when developing problem-solving skills with computers, "students are encouraged to critically assess data, to discover relationships and patterns, to compare and contrast, to transform information into something new" (p. 18).

As an example, some teachers have students use HyperCard and other multimedia tools in order to help students think strategically about how to assimilate, organize, and present information. As Tierney's (1996) study of Hypercard use in a small set of classrooms describes:

> Technology appears to have increased the likelihood of students' being able to pursue multiple lines of thought and entertain different perspectives. Ideas were no longer treated as unidimensional and sequential; the technology allowed students to embed ideas within other ideas, as well as pursue other forms of multi-layering and interconnecting ideas. Students began spending a great deal of time considering layout, that is, how the issues that they were wrestling with might be explored across an array of still pictures, video segments, text segments, and sound clips. (p. 176)

These findings are echoed by teachers in other schools. After studying technology use across classrooms in one school district, Russell (2000) wrote:

> In addition to exposing students to a larger body of information related to the topic of study, creating HyperStudio stacks also requires students to more carefully plan how they integrate and present this information. As one teacher explains, "First they do the research and identify what it is they want to include in their stack. They then create a flow chart that depicts how the pieces fit together. They sketch their stack on paper and then begin putting it into the computer." Through this process, students develop their planning skills and learn to anticipate how information will be received by their audience. (p. 11)

Despite the skill development enabled by HyperCard and other multimedia authoring tools, students who develop complex, high quality products using HyperCard do not necessarily perform well on current achievement tests. While studying the impact of computers on student learning in the Apple Classrooms of Tomorrow project, Baker, Herman, and Gearhart (1996) found that "…a sizeable portion of students who used HyperCard well to express their understanding of principles, themes, facts, and relationships were so-so or worse performers judged by more traditional forms of testing" (p. 198). Over the past decade these and similar findings have led proponents of computer use in schools to conclude that technology enables students to develop new competencies, "some of which were not being captured by traditional assessment measures" (Fisher, Dwyer, & Yocam, 1996, p. 5). Thus, when using standardized tests, or any measure of student achievement, it is important that the constructs measured by the instrument are aligned with the constructs developed through students' uses of technology. In some cases, this may require that researchers focus on sub-scale test scores that measure the constructs of interest rather than focusing only on the test's total score.

Adopting a Multi-level Analytic Approach

In addition to the previous methodological limitations, many studies of educational technology fail to properly account for the organizational processes that mediate the relationship between technology use and achievement. Over the past two decades, researchers have become increasingly aware of the problems associated with examining educational data using traditional analyses such as ordinary least squares analysis or analysis of variance. Since educational systems are typically organized in a hierarchical fashion, with students nested in classrooms, classrooms nested in schools, and schools nested within districts, a hierarchical or multilevel approach is often required (Robinson, 1950; Cronbach, 1976; Haney, 1980; Burstein,

1980; Bryk & Raudenbush, 1992; Kreft & de Leeuw, 1998). At each level in an educational system's hierarchy, events take place and decisions are made that potentially impede or assist the events that occur at the next level. For example, decisions made at the district level may have profound effects on the technology resources and support available for teaching and learning in the classroom. Similarly, within classrooms, teachers' uses of technology are likely to impact how students use technology during class-time.

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. Similarly, since decisions to make technology available in classrooms are typically made at the school or district level, it is important to examine the school system 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, and has at least three advantages over traditional analyses: (1) the approach allows for the examination of the relationship between technology use and achievement as a function of classroom, teacher, school, and district characteristics; (2) the approach allows the relationship between technology use and achievement to vary across schools; and (3) differences among students in a classroom and differences among teachers can be explored at the same time therefore producing a more accurate representation of the ways in which technology use is related to student learning (Bryk & Raudenbush, 1992; Goldstein, 1995; Kreft & de Leeuw, 1998).

Recognizing these challenges to estimating the influence of technology use on student achievement, this study attempts to overcome some of these obstacles. First, this study employs multiple measures of students' technology use inside and outside of school, and employs a multi-level analytic approach that includes student- and classroom-level modeling. In addition, this study uses both total test scores and sub-domain scores as measures of student achievement. Although the authors recognize that standardized test scores may not be suitable for isolating the types of improved learning that may occur as a result of technology use, we believe that it is necessary to use standardized test scores in this study for the following reasons. First, given the current era of heightened accountability for students, teachers and schools and the important role that standardized test scores play in the accountability movement, we believe that these scores remain a necessary starting point for understanding whether technology use is related to student learning outcomes. While mindful of the criticisms we levied at standardized test scores with regard to their

possible lack of sensitivity for detecting improved learning as a consequence of technology use, we remain aware that "[s]tandardized test scores have become the accepted measure with which policymakers and the public gauge the benefits of educational investments" (McNabb, Hawkes, & Rouk, 1999, p.6). Second, this study is exploratory in nature and so the results have the potential to guide future research that may overcome the remaining methodological obstacles that were faced here. For these reasons and in order to maximize the benefits of using standardized test scores, this study uses not only the total standardized test scores, but also sub-domain scores that have the potential to be more sensitive to technology use than the test as whole.

Specifically, this paper presents a hierarchical regression analysis of the relationship between a variety of student uses of technology and their performance on the Grade 4, state mandated paper-based English/Language Arts (ELA) Massachusetts Comprehensive Assessment System (MCAS) test. Using item level achievement data, individual student's test scores, and student and teacher responses to a technology-use survey, this study examines the relationship between technology use and performance among 986 regular students, from 55 intact fourth grade classrooms in 25 schools across nine school districts. Given increased interest in both student accountability and educational technology, the current research provides insight into how different types of technology use impact student achievement in the domain of English/Language Arts. It is important to 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 study presented in this paper employs data collected as part of the Use, Support, and Effect of Instructional Technology (USEIT) Study. The USEIT study was a three-year project that began during the Spring of 2001 and was conducted to better understand how educational technologies are being used by teachers and students, what factors influence these uses, and how these uses affect student learning. Phase I of the USEIT study was dedicated to understanding how technology was being used by teachers and students as a teaching and learning tool. During this phase (the 2001-2002 school year), information about district technology programs, teacher and student use of technology in and out of the classroom, and the factors that were associated with these uses were collected through site visits and surveys. In total, survey responses were obtained from 120 district level administrators, 122 principals, 4400 teachers, and 14,200

students from over 200 schools in 22 districts. The specific details on the sample, methodologies and analyses of the USEIT data is detailed in Russell, O'Dwyer, Bebell, and Miranda (2003).

Phase II of the USEIT study was concerned with how technology use influences student learning. Teacher survey responses from Phase I of the study were used to identify fourth grade teachers that reported various levels of instructional technology use. Specifically, all fourth grade teachers who completed the USEIT teacher survey during the Spring of 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 purposively recruited to participate in Phase II of the USEIT study designed to examine the relationship between technology use and achievement. Teachers in eight school districts were contacted in the Fall of 2002, and these teachers and their 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 Phase II sample. Thus, the current sample includes a total of 1,206 students nested within 55 classroom teachers from 25 elementary schools across nine Massachusetts school districts. Since teachers in the sample taught only one classroom each, the students may be considered nested within 55 teachers or nested within 55 intact classrooms. The sample of 1,206 students included students who had been classified as English Language Learners (ELL), students with disabilities (SD), and students who are neither SD or ELL. In order to reduce the possible confounding effects of specialized learning and language needs, this study examined only those students who were classified as non-SD and non-ELL students. Thus, the sample used for the analyses presented here included 986 fourth grade students in 55 teachers/classrooms in nine school districts.

Table 1 displays demographic and mean achievement data for each of the nine school districts that participated in this study of the relationship between technology use and achievement. The district and school summary data was collected from the Massachusetts Department of Education Web site (www.doe.mass.edu).

(Table 1 is shown on the following page.)

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

	District									Sample versus Massachusetts	
	A	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**Sample**	**MA** ('02–'03)
% 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 English/LA Advanced	20	25	14	23	13	9	10	31	39	20	10
% Grade 4 English/LA Proficient	59	53	50	54	51	47	49	50	49	51	45
% Grade 4 English/LA Needs Improvement	20	20	33	20	32	35	33	14	10	24	34
% Grade 4 English/LA Warning/Failing	1	2	3	4	3	9	8	5	2	4	10
% 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
Total # of Elementary Schools in the District	6	3	6	16	6	7	3	3	5		

Sample Characteristics									
	A	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**
Number of Participating Schools in Each District	1	1	2	6	3	3	3	1	5
Average Class Size for Each District	25	20	19	17	18	14	17	20	20
Number of Students Participants in Each District	99	59	115	168	128	84	154	184	215

As seen in Table 1, the districts that participated in this study performed somewhat higher than the state average on the grade four English/Language Arts and Math MCAS. The mean percentage of "white" students was also slightly higher for the participating districts than the state average. Additionally, the percentage of students receiving free or reduced price lunch in the sample (8.9%) was lower than the state average (26.2%). Teachers from between one and six elementary schools from each district participated in the study and the average class size across all nine districts was 19. 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 Math and English/Language Arts, and had less access to technology than the average district in Massachusetts.

Although the district averages show the similarities among the participating school districts and Massachusetts state averages, such an analysis may not be fully representative of the actual students in the current sample since the students and teachers included in this study represent a purposive sample of classrooms selected from Phase I of the USEIT study that were re-surveyed. As a consequence, the district averages in Table 1 may not be completely reflective of the current sample. In order to illustrate the wide variety of mean achievement and demographic characteristics across the schools included in this study, the minimum value, maximum value and range are presented across the 25 participating elementary schools in Table 2.

Table 2:   Min, Max and, Range for Participating Elementary Schools

	MIN	MAX	RANGE
School Size	151	563	412
% White	47%	99%	52%
% Free/Reduced Lunch	0%	48%	48%
Student : Computer Ratio	3:1	33.8:1	30.8:1
% Classes connected to the Internet	17%	100%	83%
Grade 4 ELA Advanced	2%	53%	51%
Grade 4 ELA Proficient	28%	71%	43%
Grade 4 ELA Needs Improvement.	0%	53%	53%
Grade 4 ELA Warning/Failing	0%	20%	20%
Grade 4 Math Advanced	0%	55%	55%
Grade 4 Math Proficient	9%	52%	43%
Grade 4 Math Needs Improvement	5%	72%	67%
Grade 4 Math Warning/Failing	0%	31%	31%

Table 2 illustrates the range among the participating elementary schools in terms of both demographic variables and student achievement as measured by the 2002–2003 4th grade MCAS test. Particularly interesting is the wide range for the schools' percentage of white students with a minimum percent of 47% and a maximum percent of 99%. A similar range is found for the percentage of students per school who receive free or reduced price lunches with a minimum percent of 0% and 48% for the maximum. With respect to technology, a similarly large range is found across the sample. At the tails of the distribution, one school had a reported student to computer ratio of 3 computers to 1 student while one school had 33.8 students for each computer. Similarly, the percentage of classrooms connected to the Internet ranged from a minimum value of 17% to a maximum value of 100%. Table 2 also illustrates the wide range of student performance levels across the twenty-five participating elementary schools. Specifically, the grade 4 Math and English/Language Arts MCAS performance level range is quite dramatic with a range of 67 percentage points.

Instruments

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

Technology Use Surveys

Both the teachers and the students included in this phase of the study were administered a technology use survey. The student survey gathered demographic information, and included measures of students' access to technology in school, the types of technology use they engage in at school as well as uses across subject areas. In addition, measures of personal comfort levels with technology, access to technology at home, and various uses of technology at home were included on the survey. The teacher survey also gathered demographic information, and included measures of technology use in and out of the classroom, teachers' comfort level with technology, and teachers' attitude towards technology. Both survey instruments were refined and adapted from the original Phase I USEIT teacher and student surveys (Russell, O'Dwyer, Bebell, & Miranda, 2003).[1]

Specifically, this study examined the relationship between the following student and teacher uses of technology and English/language arts test scores:

- Teacher Technology Uses
    - Teachers use of technology for delivering instruction
    - Teacher-directed student use of technology during classtime
    - Teacher-directed student use of technology to create products
    - Teachers' use of technology for class preparation
    - Teachers' use of technology for student accommodation
- Student Technology Uses
    - Student use of technology at school
    - Student recreational use of technology at home
    - Student academic use of technology at home

Where appropriate, composite measures of these technology uses were developed using groups of items from the teacher and student technology use surveys. The reliabilities of these measurement scales were examined and a discussion is presented in the Methodology section of this paper.

The MCAS English/Language Arts 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. The tests consist primarily of multiple-choice items but include some short response items and one extended writing item. Currently, the MCAS has been expanded across subject areas and grade levels and included a third grade reading exam. Like some other state testing programs, MCAS results are used to determine whether an individual student may graduate and for ranking schools and districts. It is worth noting that the MCAS is a controversial test and has been the subject of numerous scholarly debates regarding the validity of the testing program. On closer inspection, much of the criticism levied at the test focuses on how the test results have been used by policy makers, the media and school officials, and how the test has impacted teaching and learning in Massachusetts rather than on the psychometric characteristics of the test itself (Wheelock, Bebell, & Haney, 2000).

In the study presented here, students' fourth grade ELA MCAS raw scores and sub domain scores in writing and reading and literature were

used as the dependent variable. The language sub-domain score was found to be unsuitable for inclusion in the analyses as it lacked sufficient variability. The reading and literature sub-domain score was calculated by summing students raw scores across the short answer items relating to this content domain. The writing sub-domain score was derived from the extended writing exercise on the test that was scored on two criteria: topic/idea development and standard English conventions. The topic/idea development dimension was scored using a 6 point scale and the standard English conventions dimension was scored on a 4 point scale. All essays were scored by two raters and the scores awarded by the two raters were summed into a single score. The combined topic/idea development and standard English conventions score, which had a possible range from 4 to 20, was also used as the writing dependent variable. It should be noted that although a sixteen point range was possible on the writing sub-domain score, the actual range for the sample included in this study was only 12 points (8 to 20). The restriction of range on this variable may have resulted in the underestimation of the strength of the observed relationships.

In any investigation into the relationship between student achievement and instructional practices, technology-related or otherwise, it is important to take into account students' prior achievement. To this end, the authors included individual students' grade 3 reading raw scores collected from the 2001–2002 MCAS administration, as a control variable when modeling the relationship between fourth grade ELA achievement and technology use. Similar to the scores for the writing sub-domain, the third grade reading scores were quite high for the sample indicating a ceiling effect for assessing prior achievement. Due to this truncation of scores, the relationship between prior achievement and fourth grade test and sub-domain scores may have been underestimated in the models formulated in this research. Students' third grade reading scores, fourth grade ELA scores, and survey responses were combined with their teachers' survey responses into a single model allowing the relationship between measures of achievement and technology to be examined as function of both student and teacher technology characteristics while controlling for prior achievement.

Methodology

The analyses presented in this report were conducted using a two-level hierarchical linear regression model. In this model, 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. The general hierarchical model assumes a random sample of i students within J classrooms, such that Y_{ij} is the outcome variable (English/language arts achievement in this case) for student i nested within teacher or classroom j (Bryk & Raudenbush, 1992). Given that intact fourth grade classrooms were sampled, teachers' survey responses were considered analogous to measures of classroom practices and so individual students were considered to be nested within classrooms. The level-1 or student model may be expressed as follows:

$$Y_{ij} = \beta_{0j} + \beta_{1j} X_{1ij} + \beta_{2j} X_{2ij} + ... + \beta_{kj} X_{kij} + r_{ij}$$

In this model, achievement, Y_{ij} was modeled as a function of a linear combination of student predictors, X_{kij}. This model states that the predicted outcome was composed of a unique intercept β_{0j}, and slope for each classroom β_{kj}, as well as a random student effect, r_{ij}. The intercept represents the base achievement in each classroom and the random student effect was assumed to be normally distributed with a mean of zero and variance, σ^2. In the models used in this research, only mean achievement was allowed to vary between classrooms.

The variation in mean achievement across classrooms was modeled at the second level. The level-2 model may be expressed as follows:

$$\beta_{kj} = \gamma_{0k} + \gamma_{1k} W_{1j} + \gamma_{2k} W_{2j} + ... + \gamma_{P-1k} W_{P-1j} + u_{kj}$$

Each β_{kj} was modeled as a function of aggregate student characteristics (prior achievement and socioeconomic status indicators) as well as measures of teachers' use of technology and beliefs about technology, W_{pj}. Each γ_{pk} represented the effect of the predictors on the outcome. Each classroom had a unique random effect, u_{kj}, which was assumed to be normally distributed with a mean of zero and variance τ_{kk} for any k.

These models allowed the total variability in fourth grade MCAS ELA achievement to be partitioned into its within-classroom and between-classroom variance components, and allowed predictors to be added at each level that explain a portion of both the within-classroom and between-classroom variance available. Although it could be argued that achievement could be modeled as varying within-classrooms, between-classrooms within-schools, and between-schools, we make the case that the sample selection procedures adopted for this study make the two-level

model a more appropriate choice for the analyses conducted here. The sample of teachers included in this phase of the study was selected based on their responses to the survey administered during Phase I of the USEIT study and was not concerned with school characteristics. Though we recognize the importance of school level variables for supporting technology use, we do not feel that it would be appropriate to add a macro level above the level at which the sampling units were selected. For this reason, the between-classroom variability was confounded with the between-school variability in the models presented in this research.

Three outcome measures (Y_{ij}) were modeled in this research; the total ELA raw score, the writing component score, and the reading and literature subscale score. In order to facilitate comparisons among models, each of these outcome measures was standardized to have a mean of zero and a standard deviation of 1. The hierarchical regression analyses were carried out in a number of stages. When conducting any hierarchical analysis, the first step requires the examination of the amount of variability in the three outcome measures that exists within and between classrooms. In order to accomplish this, unconditional models, in which no predictors other than school membership were known, were formulated. To develop a better understanding of the technology uses that may be associated with ELA performance, the second stage of the analysis involved extensive theory-driven, exploratory data analysis to identify student and teacher variables observed to be associated with each of the outcome measures. Nuanced measures of technology use first described in Bebell, Russell & O'Dwyer (2004) were examined during the exploratory analysis phase.

Guided by past research and theory, exploratory multilevel models were formulated. In addition to identifying teacher/classroom level predictors, student level predictors were identified that were significantly associated with technology use at $p \leq 0.1$. Since the emphasis during this section of the analysis was on data exploration, a significance level of 0.10 was adopted as the criterion for significance in preference to the more stringent 0.05. Using predictor variables identified during the exploratory phase, increasingly complex multilevel models were constructed to predict each of the outcome measures. The total ELA raw score model was used as the archetypal model and guided the construction of the models for predicting the remaining two outcomes. In total, beginning with a model that included only prior achievement and ending with the most parsimonious model, 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 indicators, home technology use measures, school technology use measures, and teacher technology use. Table 3 contains the student level variables and composite

measures used in addition to prior achievement for predicting each of the three outcomes. Principal components analysis was used to confirm the existence of the student recreational and academic home use scales and to create standardized factor scores for these scales that have a mean of zero and a standard deviation of 1. The individual items used to characterize student use of technology at school and students' socioeconomic status were also 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 at school *(Entered into models individually)*	How often do you use a computer in school to send/receive email?
	How often do you use a computer in school edit papers using a computer?
	How often do you use a computer in school to find information on the Internet?
	How often do you use a computer in school to create a Hyperstudio or PowerPoint presentation
Student recreational home use of technology *Cronbach's alpha = 0.78*	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?

The teacher variables and composites included in the models are shown in Table 4. As was the case with the student level measures, principal components analysis was used to confirm the existence of measurement scales and to create standardized factor scores that have a mean of zero and standard deviation of 1. In addition to the variables in Table 4, the third grade measure of prior achievement as well as the two student socioeconomic status measures were aggregated to the classroom level. 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.

(Table 4 is shown on the following page.)

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

Measurement Scale	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

Results

To provide a sense of the degree to which teachers and students use technology for each of the measurement categories and scales, Figures 1 and 2 display the distribution and mean response for each of the individual student and teacher survey items that comprise the technology use measures examined in this study. 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. By far, students report using their computer at home to play games more frequently than any other use.

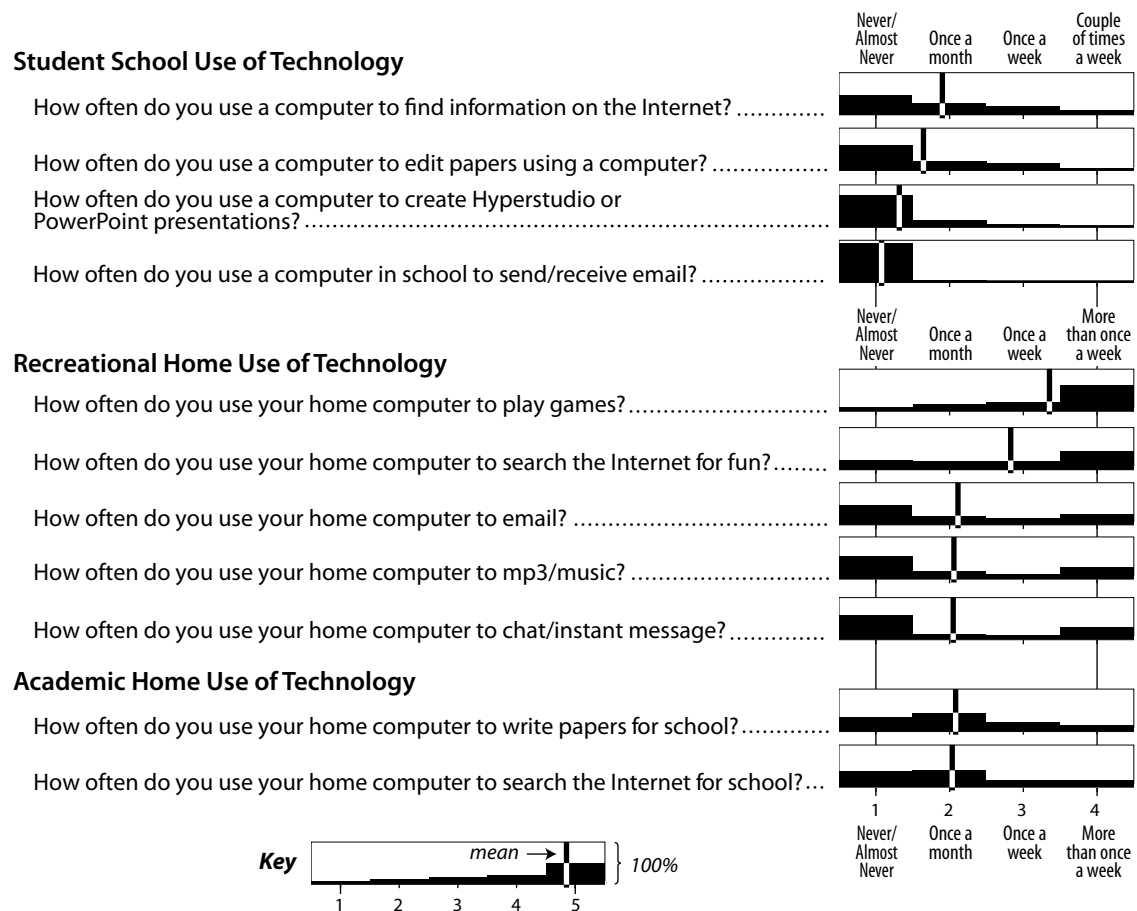**Figure 1:   Distribution and Mean Items Responses for Student Uses of Technology**

Figure 2 displays similar information for teacher uses of technology. The distributions show that teachers tend to use technology most frequently for preparation purposes. Teachers also tend to have their students perform tasks during classtime using a computer more often than

they have them create products using technology. Teachers report that they rarely (on average, several times a year) use technology to deliver instruction in the classroom.

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

Use of Technology for Delivering Instruction

How often do you use a computer to deliver instruction? ....

Teacher-directed General Student Use During Classtime

During classtime, how often do students work using computers? ....

During classtime, how often do students work in groups on school work using computers? ....

During classtime, how often do students perform research or find information using the Internet or CD-ROM?...

During classtime, how often do students use a computer or portable writing device? ....

During classtime, how often do students present information to the class using a computer? ....

During classtime, how often do students use a computer to solve problems? ....

Teacher Directs Students to Create Products Using Technology

How often do you ask students to produce reports and term papers using technology? ....

How often do you ask students to produce pictures or artwork using technology?

How often do you ask students to produce graphs or charts using technology? .....

How often do you ask students to produce multimedia reports? ....

How often do you ask students to produce Web pages, Web sites, or other Web-based publications? ....

How often do you ask students to produce videos or movies using technology?.....

Teacher Use of Technology for Preparation

Make handouts for students using a computer ....

Create a test, quiz, or assignment using a computer ....

Create Web-quests or build the Internet into a lesson ....

Teacher Use of Technology for Student Accommodation

Adapt activities to students' needs ....

Prepare or maintain IEPs using a computer ....
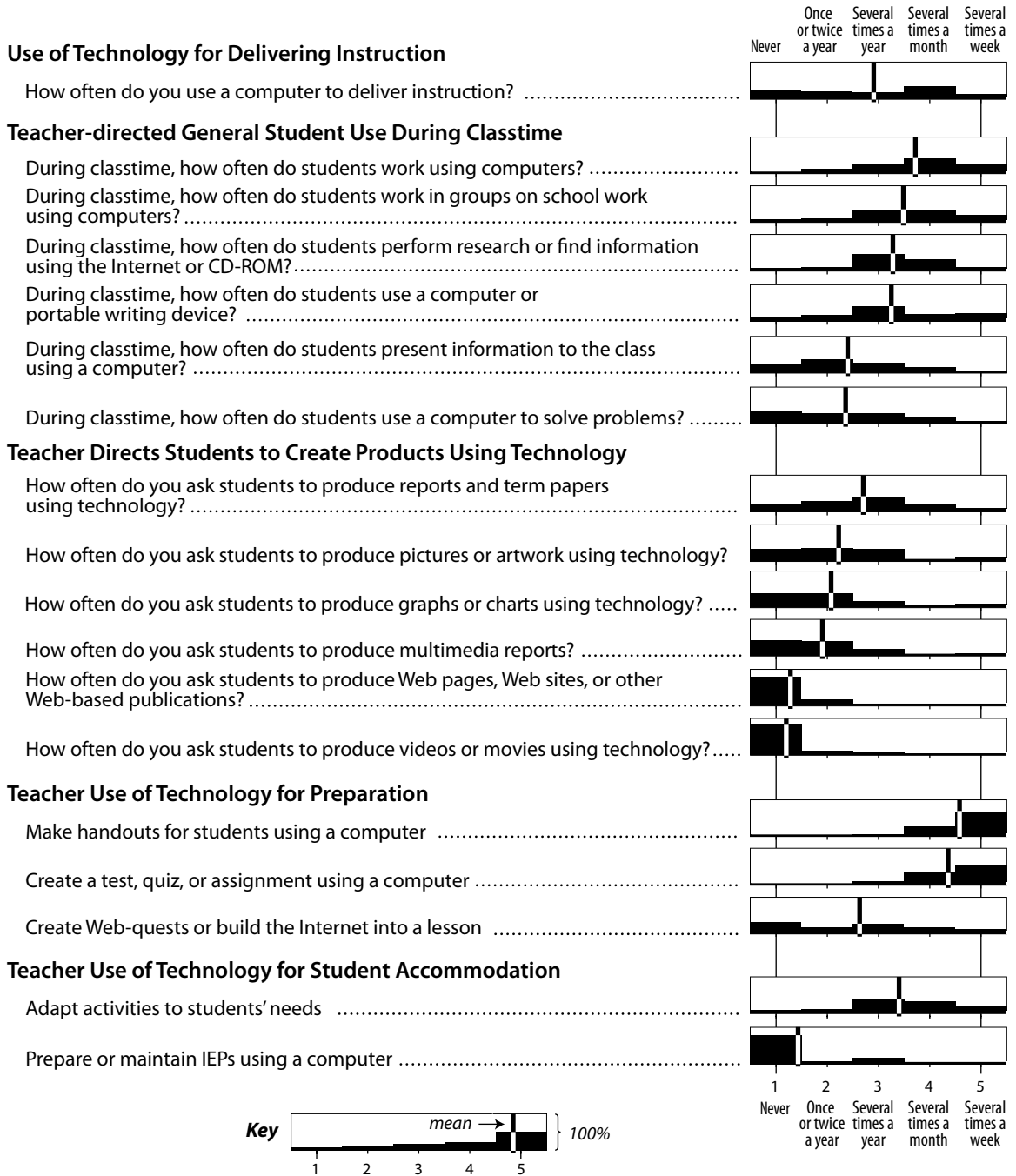
Key — mean → 100%
1 2 3 4 5

Table 5 presents the variance components for each of the three achievement score measures when the total variability in each is partitioned using the unconditional multilevel model. Although the majority of variability in each measure exists among students within classrooms, a significant proportion of the variability in achievement lies between classrooms. The largest classroom-to-classroom differences occur for the writing component measure; almost 20 percent of the total variability in the writing measure exists between classrooms. It appears that the smallest between classroom differences occur for the reading and literature component score (12.9%).

Table 5:   **Unconditional Variance Components for Three Standardized MCAS Outcome Measures**

	Standardized ELA Total Raw Score	Writing Component Score	Reading and Literature Component Score
Percent of variance within classrooms	83.6%	80.7%	87.1%
Percent of variance between classrooms	16.4%‡	19.3%‡	12.9%‡

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

Once the total variability in the outcome was partitioned, the hierarchical approach allowed characteristics measured at both the student and teacher levels to be added to the models in order to explain some of the available variance. Six multilevel models were constructed to estimate the impact of different categories of predictor variables independently. The first and simplest model included only third grade achievement to predict fourth grade scores. The second model included both prior achievement and indicators of socioeconomic status. The third model added students' use of technology at home for both recreational and academic purposes. In addition to the previous variables, the fourth model included measures of students' technology use at school. The fifth model built upon Model 4 and included aggregate student measures, teacher measures of technology use and beliefs about the impacts of technology at the second level in the multilevel equation. The sixth and final model was a scaled back version of Model 5 in which only predictors that were statistically significant ($p \leq 0.05$) were retained. Given that the ELA raw score is comprised of the Writing and Reading/Literature scores, the best fitting model (Model 6) for the ELA total score was applied to the two sub-scores.

Tables 6 through 8 present the multilevel regression coefficients and the percent of variance explained by each of the six models for all three

outcomes. Multilevel regression modeling is a generalization of ordinary least squares analysis in which each level in the hierarchy is represented by a separate regression equation. For this reason, the multilevel regression coefficients refer to specific levels in the hierarchical structure of the data and are interpreted in the same way as traditional regression coefficients.

ELA Total Score

As seen in Table 6, third grade reading achievement was a positive and significant predictor of fourth grade ELA scores in all six models. The two socioeconomic indicators were also significant predictors of total ELA scores in Models 2 through 6. Model 3 shows that students' academic use of computers at home was only weakly associated with the ELA raw score; the coefficient was close to zero and non-significant in Models 3 through 5. The frequency with which students reported engaging in recreational computer use at home was negatively and statistically significantly related to the total ELA raw score. When the four measures of student use of technology at school were included in the models, only the frequency with which students use a computer in school to edit papers and the frequency with which they use a computer to create a Hyperstudio or PowerPoint presentation were significant. Interestingly, students who reported higher frequencies of computer use for creating presentations (using Hyperstudio and PowerPoint) during school time tended to have lower ELA total test scores. When student variables aggregated to the classroom level and teacher use of technology measures were included in the model (Model 5), only the classroom mean third grade reading score was a significant predictor of the differences between the classrooms in terms of mean ELA raw score; teachers' use of technology did not appear to be able to predict differences between classroom mean ELA 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.

(Table 6 is shown on the following page.)

Table 6:   Standardized Total ELA Raw Score Model

Outcome = Standardized ELA raw score	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6	
Student Level Predictors	Coeff	Sig.	Coeff	Sig.	Coeff	Sig.	Coeff	Sig.	Coeff	Sig.	Coeff	Sig.
How often do you use a computer in school to send/receive email?							-0.02	0.819	-0.02	0.819		
How often do you use a computer in school to edit papers?							**0.10**	**0.019**	**0.10**	**0.019**	**0.10**	**0.018**
How often do you use a computer in school to find information on the Internet?							-0.01	0.801	-0.01	0.801		
How often do you use a computer in school to create a Hyperstudio or Powerpoint presentation?							**-0.09**	**0.019**	**-0.09**	**0.019**	**-0.09**	**0.011**
Recreational Home Use					**-0.09**	**0.009**	**-0.10**	**0.006**	**-0.10**	**0.006**	**-0.10**	**0.005**
Academic Home Use					0.00	0.966	0.01	0.876	0.01	0.876		
About how many books of your own do you have at home, not counting school books or comic books?			**0.11**	**0.000**	**0.11**	**0.000**	**0.11**	**0.000**	**0.11**	**0.000**	**0.11**	**0.000**
How many computers, if any, do you have at home?			**0.08**	**0.021**	**0.11**	**0.003**	**0.10**	**0.002**	**0.10**	**0.002**	**0.10**	**0.002**
Grade 3 Reading score	**0.38**	**0.000**	**0.37**	**0.000**	**0.37**	**0.000**	**0.36**	**0.000**	**0.36**	**0.000**	**0.36**	**0.000**
Teacher Level Predictors									Coeff	Sig.	Coeff	Sig.
Teacher-mean student Grade 3 reading score									**0.70**	**0.000**	**0.66**	**0.000**
Teacher-mean number of books in student homes									-0.12	0.435		
Teacher-mean number of computers in student home									0.06	0.727		
Teacher-directed student use of technology during classtime									0.00	0.977		
Teachers direct students to create products using teachnology									-0.05	0.513		
Teachers use technology for preparation									-0.02	0.662		
Teachers use technology to maintain IEPs									-0.07	0.253		
Teacher use of technology for delivering instruction									0.03	0.586		

	Variance Components			
	Available Variance	**Percent Available**		
Between Classrooms	0.16	16%		
Within Classrooms	0.79	84%		
Total Variance Available	0.95	100%		

	Residual Variance and Variance Explained											
	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6	
	Residual Variance	Variance Explained	Residual Variance	Variance Explained	Residual Variance	Variance Explained	Residual Variance	Variance Explained	Residual Variance	Variance Explained	Residual Variance	Variance Explained
Between Classrooms	0.16	0.0%	0.16	0.0%	16.00	0.0%	16.00	0.0%	0.08	48.8%	0.07	53.1%
Within Classrooms	0.68	14.7%	0.66	16.7%	0.65	17.4%	0.65	18.0%	0.65	18.0%	0.65	18.2%
Total Variance Explained		12%		14%		15%		15%		23%		24%

Bolded values are significant at p < 0.05.

The percent of variance explained by the models increased as more predictors were added to the models. Without predictors at the classroom level, Models 1 through 4 were unable to explain any of the variance between classrooms. When student uses of technology at school were included (Model 4), the model explains 15% of the total variability in the ELA raw score. However, most of the explanatory power of the model was due to the inclusion of the grade 3 reading score; prior achievement alone (Model 1) explained 12% of the total variability. The final model (Model 6) which included an aggregated measure of grade 3 performance, individual measures of prior achievement, two socioeconomic indicators, recreational use of technology, and two measures of school use of technology explained almost one quarter of the total available variance in ELA raw scores (24%).

Writing Scores

Table 7 presents similar models for the MCAS writing score. As was the case for the total score, prior achievement and the two socioeconomic indicators were positively and statistically significantly related to students' writing scores. Neither of the home use measures were significant predictors of writing scores; the coefficients for both measures were close to zero. Again, students who reported higher frequencies of computer use for editing papers during school time tended to have higher writing scores, while students who reported higher frequencies of computer use during school for creating presentations tended to score lower on the writing component of the ELA MCAS test. The teacher uses of technology included in these models do not appear to be significant predictors of student achievement. In terms of the explanatory power of the model, the combination of variables in Model 6 explained only 12% of the total variance in writing scores. Compared to the total ELA raw score model, the models for writing were not as powerful for predicting student achievement.

(Table 7 is shown on the following page.)

Table 7:   Standardized Writing Component Raw Score Model

Outcome = Standardized Writing component	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6	
Student Level Predictors	Coeff	Sig.	Coeff	Sig.	Coeff	Sig.	Coeff	Sig.	Coeff	Sig.	Coeff	Sig.
How often do you use a computer in school to send/receive email?							0.01	0.868	0.01	0.868		
How often do you use a computer in school to edit papers?							**0.09**	**0.042**	**0.09**	**0.042**	**0.08**	**0.048**
How often do you use a computer in school to find information on the Internet?							0.00	0.925	0.00	0.925		
How often do you use a computer in school to create a Hyperstudio or Powerpoint presentation?							**-0.11**	**0.027**	**-0.11**	**0.027**	**-0.11**	**0.02**
Recreational Home Use							**-0.03**	**0.421**	-0.03	0.395	-0.03	0.34
Academic Home Use							-0.02	0.592	-0.01	0.733	-0.01	0.733		
About how many books of your own do you have at home, not counting school books or comic books?			**0.13**	**0.001**	**0.14**	**0.001**	**0.14**	**0.001**	**0.14**	**0.001**	**0.14**	**0.001**
How many computers, if any, do you have at home?			**0.12**	**0.003**	**0.13**	**0.001**	**0.13**	**0.001**	**0.13**	**0.001**	**0.13**	**0.001**
Grade 3 Reading score	**0.23**	**0.000**	**0.22**	**0.000**	**0.22**	**0.000**	**0.21**	**0.000**	**0.21**	**0.000**	**0.21**	**0.000**
Teacher Level Predictors									Coeff	Sig.	Coeff	Sig.
Teacher-mean student Grade 3 reading score									**0.48**	**0.008**	**0.48**	**0.000**
Teacher-mean number of books in student homes									-0.24	0.215		
Teacher-mean number of computers in student home									0.24	0.178		
Teacher-directed student use of technology during classtime									-0.13	0.281		
Teachers direct students to create products using teachnology									0.08	0.353		
Teachers use technology for preparation									-0.08	0.155		
Teachers use technology to maintain IEPs									-0.07	0.397		
Teacher use of technology for delivering instruction									0.09	0.338		

Variance Components		
	Available Variance	**Percent Available**
Between Classrooms	0.19	19.3%
Within Classrooms	0.82	80.7%
Total Variance Available	1.01	100%

Residual Variance and Variance Explained												
	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6	
	Residual Variance	Variance Explained	Residual Variance	Variance Explained	Residual Variance	Variance Explained	Residual Variance	Variance Explained	Residual Variance	Variance Explained	Residual Variance	Variance Explained
Between Classrooms	0.19	0.0%	0.19	0.0%	0.19	0.0%	0.17	0.0%	0.15	22.3%	0.15	20.9%
Within Classrooms	0.77	5.1%	0.75	8.6%	0.75	8.5%	0.74	9.0%	0.74	9.0%	0.74	9.3%
Total Variance Explained		4%		7%		7%		7%		12%		12%

Bolded values are significant at p < 0.05.

Reading and Literature Scores

Table 8 present the results of analyses for the reading and literature component of the ELA test. As seen in Table 8, third grade achievement and both socioeconomic status measures were significant predictors of reading and literature scores. Similar to the previous models, academic home use of technology was not significantly related to this achievement measure, and recreational home use was negatively related to achievement. The frequency with which students reported using a computer in school to edit papers remained a positive predictor of reading and literature scores. Overall, including only student level predictors in the models explained about 18% of the total variability among reading and literature scores. When classroom-level predictors were added to the model, mean prior achievement remained the only significant predictor at the classroom level and the model explained 25% of the total variance in reading and literature scores. Including teacher technology use measures had no meaningful effect on the explanatory power of the models.

(Table 8 is shown on the following page.)

Table 8:   Standardized Reading and Literature Component Model

Outcome = Standardized Reading and Literature componet	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6	
Student Level Predictors	Coeff	Sig.	Coeff	Sig.	Coeff	Sig.	Coeff	Sig.	Coeff	Sig.	Coeff	Sig.
How often do you use a computer in school to send/receive email?							0.04	0.291	0.04	0.291		
How often do you use a computer in school to edit papers?							**0.09**	**0.044**	**0.09**	**0.044**	**0.08**	**0.045**
How often do you use a computer in school to find information on the Internet?							-0.04	0.392	-0.04	0.392		
How often do you use a computer in school to create a Hyperstudio or Powerpoint presentation?							-0.04	0.327	-0.04	0.327	-0.05	0.212
Recreational Home Use					**-0.08**	**0.018**	**-0.09**	**0.014**	**-0.09**	**0.014**	**-0.09**	**0.007**
Academic Home Use					-0.02	0.595	-0.02	0.670	-0.02	0.670		
About how many books of your own do you have at home, not counting school books or comic books?			**0.10**	**0.000**	**0.11**	**0.001**	**0.11**	**0.000**	**0.11**	**0.000**	**0.10**	**0.000**
How many computers, if any, do you have at home?			**0.07**	**0.045**	**0.10**	**0.009**	**0.09**	**0.009**	**0.09**	**0.009**	**0.09**	**0.008**
Grade 3 Reading score	**0.42**	**0.000**	**0.42**	**0.000**	**0.41**	**0.000**	**0.41**	**0.000**	**0.41**	**0.000**	**0.41**	**0.000**
Teacher Level Predictors									Coeff	Sig.	Coeff	Sig.
Teacher-mean student Grade 3 reading score									**0.61**	**0.000**	**0.62**	**0.000**
Teacher-mean number of books in student homes									-0.02	0.873		
Teacher-mean number of computers in student home									-0.02	0.911		
Teacher-directed student use of technology during classtime									-0.07	0.430		
Teachers direct students to create products using teachnology									0.00	0.980		
Teachers use technology for preparation									0.04	0.245		
Teachers use technology to maintain IEPs									-0.06	0.192		
Teacher use of technology for delivering instruction									0.02	0.670		

	Variance Components		
	Available Variance	**Percent Available**	
Between Classrooms	0.12	12.9%	
Within Classrooms	0.81	87.1%	
Total Variance Available	0.93	100%	

	Residual Variance and Variance Explained											
	Model 1		**Model 2**		**Model 3**		**Model 4**		**Model 5**		**Model 6**	
	Residual Variance	Variance Explained	Residual Variance	Variance Explained	Residual Variance	Variance Explained	Residual Variance	Variance Explained	Residual Variance	Variance Explained	Residual Variance	Variance Explained
Between Classrooms	0.12	0.0%	0.12	0.0%	0.12	0.0%	0.14	0.0%	0.05	54.5%	0.05	58.1%
Within Classrooms	0.66	18.1%	0.65	19.7%	0.64	20.4%	0.64	20.6%	0.64	20.6%	0.64	20.7%
Total Variance Explained		16%		17%		18%		18%		25%		25%

Bolded values are significant at p < 0.05.

Overall, the multilevel models identified a number of variables that were useful for predicting ELA MCAS scores. Not surprisingly, higher prior achievement was associated with higher fourth grade test scores. The socioeconomic indicators included in the models were each significant and positive predictors of the three outcomes. Students' recreational use of technology at home was negatively associated with the outcomes indicating that students who reported higher levels of recreational use at home tended to have lower ELA scores. Students who reported greater frequency of technology use at school to edit papers were likely to have higher total ELA raw scores and tended to score higher on the writing component. Use of technology at school to prepare presentations was associated with lower outcome measures for each of the three outcomes modeled. None of the teacher uses of technology included in these multilevel models were significant predictors of student achievement. In terms of explained variance, the models predicting the total ELA raw score and the reading and literature components were the most powerful.

Discussion

Since the mid 1980s, several hundred studies have examined the impacts of technology on teaching and learning. As described previously, the majority of these studies focus on specific uses of technology, employ relatively small samples, and measure achievement using a variety of instruments, most of which are non-standardized. Over the past several years, however, increasing emphasis has been placed on the use of standardized tests to measure student achievement when examining the efficacy of educational programs. With respect to technology, the three largest studies (at least in terms of sample size) that employ standardized test scores as the outcome measure have found mixed results. However, each of these studies suffers from one or more important methodological shortcoming. Among these shortcomings are: a) weak or non-existent measures of student use of technology; b) measures of technology use that treat use as a unidimensional construct rather than a multi-faceted set of constructs; c) failure to use a measure of prior achievement to control for pre-existing differences in achievement; d) use of total test scores as the outcome measure rather than focusing on the sub-scores that are most closely associated with the constructs developed through a given use of technology; e) use of analytic methods that do not consider the multi-level structure of educational settings; f) use of school-level rather than student-level measures of achievement; and g) failure to randomly assign participants, either at the individual or classroom/school level, to control and experimental conditions. The study reported here was designed to overcome some, but not all, of these shortcomings. Specifically, the study

employed multiple measures of technology use in school and at home, used hierarchical analytic methods to decompose classroom- and student-level effects, focused on student-level total test scores as well as sub-test scores, and employed third grade test scores to control for differences in prior achievement. The study was not able to randomly assign students or classrooms to control or experimental groups.

The results of the models showed that prior achievement and the proxies for social economic status included in the models were statistically significant predictors of students' fourth grade total English/language arts scores, as well as their writing and reading and literature test scores. In addition, 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 the total test scores. However, whereas 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. Perhaps more importantly, when examining sub-test scores, it appears that these uses of technology are 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.

For those interested in understanding the relationship between achievement and technology use in school and at home, these analyses provide evidence that instructional use of computers during the writing process has a positive relationship with students' performance on the essay portion of Massachusetts state-mandated standardized test. Since the MCAS writing test requires students to compose and revise an essay using paper and pencil, this finding is particularly noteworthy given previous research that provides evidence that the achievement of students who are accustomed to writing with computers is underestimated by paper-based written tests (Russell, 1999; Russell & Haney, 1997; Russell & Plati, 2001). In other words, despite being disadvantaged by producing text on paper rather than on a computer during testing, students who reported using computers for editing in the classroom still achieved higher scores than did students who use computers less frequently for editing in the classroom. This positive relationship between computer use for writing and achievement is consistent with findings from a recent meta-analysis of research on computers and writing (Goldberg, Russell, & Cook, 2003).

In contrast, the analyses presented above also provide evidence that students who report using computers to create presentations perform worse than expected on the written portion of the MCAS. Although speculative, this negative relationship may result from students spending less time writing during classtime and more time creating and revising multimedia products that contain relatively small amounts of written work. In essence, time spent creating presentations may detract from time available during class to develop students' writing skills.

Finally, these analyses provide evidence that students who spend more time recreating with a computer at home perform worse than expected (based on prior achievement) on the reading and literature portion of the MCAS. Again, although speculative, this negative relationship may result from students investing more time playing with a computer and less time reading at home.

From a methodological perspective, these analyses provide evidence that multiple measures of technology use, coupled with measures of specific constructs (in this case, writing and reading) yield a more nuanced understanding of how various technology uses affect student learning. In fact, it is interesting to note that the more general measures of use, which asked students and teachers how often technology is used during Language Arts, were not significant predictors of any achievement measures while some of the more specific measures of use were useful predictors. In addition, given that a significant portion of the explained variance resides at the classroom level, these analyses provide evidence that a multi-level approach to analyzing the effects of technology on learning is warranted. Finally, the fact that prior achievement (i.e., grade 3 MCAS reading scores) accounts for the largest percentage of variance in fourth grade test scores highlights the value of including measures of prior achievement when examining the effects of technology use on student achievement.

Yet, it should also be noted that despite the inclusion of measures of prior achievement and socioeconomic status, none of the analyses accounted for more than 25% of the variance in fourth grade test scores. This suggests that, although some of the factors included in the models were statistically significant predictors of achievement, there are likely to be other factors that were not measured that also contribute to student achievement. Despite including several measures of specific technology uses, some of these additional factors may relate to other uses of technology, while other factors are likely unrelated to technology use. As an example, neither the student or teacher survey included items that focused on specific uses of technology to develop reading skills. Similarly, due to the high-stakes nature of the MCAS tests, several companies have developed computer-based test preparation materials. Given the variety of software

that is available to schools that focus on developing decoding and comprehension skills coupled with the increasing availability of test preparation materials, it would be valuable to include measures of these technology uses in future research.

As discussed in the Instrumentation section, when considering the explanatory power of the multi-level models created as part of our analyses, it is also important to recognize that many students performed very well on the third grade reading test, thus creating a ceiling effect for the measure of prior achievement. Similarly, although the essay portion of the MCAS is scored on a 20 point scale, the scoring procedures for this essay results in scores that generally have a 12 point score range (from 8 to 20). Together, the third grade Reading test ceiling effect and the 4th grade essay item's restriction of range may have contributed to an underestimation of the relationship between prior achievement and current achievement, which in turn decreased the amount of variance explained by prior achievement. While it is not possible for researchers to alter the difficulty or scoring procedures for a standardized test, future research might be strengthened by including additional measures of prior achievement and/or writing skills in order to compensate for ceiling and restriction of range effects.

Similarly, the relatively low level of teacher use of technology for instructional purposes created a floor effect which may have reduced the explanatory power of the models. Although the selection procedures for this study began by grouping teachers into low, moderate, and high level of technology users based on surveys they had completed during Phase I of the USEIT study, this grouping was based on a composite of many types of uses. These types of uses included delivering instruction, preparation, communicating with colleagues, and asking students to use technology in the classroom and to create products. Moreover, this grouping was normative. As can be inferred by Figure 2, teachers who were grouped into the lowest group rarely used technology for any of these purposes while teachers in the middle group generally used technology for these purposes only a few times per year. Within the high group, few teachers reported using technology for these purposes on a daily basis. Thus, although normative comparisons allowed us to form three distinct groups, many of the teachers placed in the high-level group do not represent teachers who regularly use technology for a variety of purposes. As a result, the estimated effect of teachers' use of technology on student achievement reported here may not reflect the actual effects for teachers who make regular and robust use of computers in and out of the classroom.

In light of these shortcomings, future research on this topic could be improved by establishing criteria that define teachers who are low, medium

and high technology users and then broadly surveying teachers to identify those who fall into each level. Similarly, by expanding the demographic diversity of the classrooms included in the study, it is likely that the ceiling effect for the student achievement covariate would be reduced. Increased demographic diversity might also increase the relationship between home and classroom level factors and student achievement. Finally, and perhaps most importantly, given that state tests provide a general measure of student achievement and are comprised of multiple sub-domain scores each of which typically have lower reliability estimates, this line of research would be strengthened by including measures that are aligned with specific categories of student and teacher technology use in the classroom. As an example, if technology is used in multiple classrooms to develop students' geometry skills, then an extended test that focuses solely on geometry would be valuable to include in future research. Undoubtedly, each of these improvements requires considerable investment and would require that the research begin with a substantially larger pool of classrooms and would include survey instruments that provide more detailed information about the content and skills that technology is being used to develop. Similarly, achievement instruments that are aligned with these uses would need to be developed and validated prior to use. Until such investments are made, however, researchers will have little options other than the use of existing standardized tests and samples that may not represent the full range or intensity of technology use by teachers or students.

Despite these limitations, this study provides empirical evidence that students' technology use is significantly related to students' fourth grade performance on the Massachusetts Language Arts test, and that this relationship is significant after controlling for differences in prior achievement and social economic status. More importantly, these analyses provide evidence that different types of technology use affect achievement in different ways. Specifically, use for writing has a positive effect on writing achievement, while use of technology to create multimedia presentations and use of computers to recreate have negative relationships with writing and reading performance, respectively. Clearly, the findings from this study need to be replicated across multiple states, years, and at different grade levels before policymakers will have adequate evidence to inform policy decisions regarding the efficacy of educational technology. Nonetheless, the methodology employed for this study holds promise to provide policymakers and the general public with evidence, based on standardized measures of student learning, of the positive and negative effects of students' technology use.

Endnote

[1] Original USEIT surveys are available at www.intasc.org.

References

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

Baker, E., Herman, J., & Gearhart, M. (1996). Does technology work in school? Why evaluation cannot tell the full story. In C. Fisher, D. Dwyer, & K. Yocam (Eds.), *Education and technology: Reflections on computing in classrooms* (pp. 185–202). San Francisco: Apple Press.

Bebell, D., Russell, M., & *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.

Bryk, A. S., & Raudenbush, S.W. (1992). *Hierarchical linear models: Applications and data analysis methods.* Newbury Park, CA: Sage

Burstein, L. (1980). The analysis of multi-level data in educational research and evaluation. In D.C. Berliner (Ed.), *Review of research in education* (Vol.8, pp. 158–233). Washington, DC: American Educational Research Association.

Chapman, E. (2003). Alternative approaches to assessing student engagement raters. *Practical Assessment, Research & Evaluation*, 8(13). Retrieved June 7, 2004 from http://PAREonline.net/getvn.asp?v=8&n=13.

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)

Cromwell, S. (1999). Laptops change curriculum – and students. *Education World*. Retrieved November 6, 2003 from www.education-world.com/a_curr/curr178.shtml

Cronbach, L.J. (1976). *Research on classrooms and schools: Formulation of questions, design, and analysis* (Occasional paper). Stanford, CA: Stanford Evaluation Consortium, Stanford University.

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

Dwyer, D. (1996). Learning in the age of technology. In C. Fisher, D. Dwyer, and K. Yocam (Eds.), *Education and technology: Reflections on computing in classrooms* (pp. 15–34). San Francisco: Apple Press.

Fisher, C., Dwyer, D., & Yocam, K. (1996). *Education and technology: Reflections on computing in classrooms.* San Francisco: Apple Press.

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.

Glass, G., McGaw, B., & Smith, M. L. (1981). *Meta-analysis in social research.* Beverly Hills, CA; Sage Publications, Inc.

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* (I). 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).

Kirkpatrick, H. & Cuban, L. (1998). Computers make kids smarter – right? [Electronic version]. *Technos Quarterly*, v7, No. 2. Retrieved November 12, 2003, from http://www.technos.net/tq_07/2cuban.htm

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

McFarlane, A. E., Friedler, Y., Warwick, P., & Chaplain, R. (1995). Developing an understanding of the meaning of line graphs in primary science investigations, using portable computers and data logging software. *Journal of Computers in Mathematics and Science Teaching, 14*(4), 461–480.

McNabb, M., Hawkes, M., & Rouk, U. (1999). *Critical issues in evaluating the effectiveness of technology.* Secretary's Conference on Educational Technology: Evaluating the Effectiveness of Technology, Washington, DC. Retrieved November 6, 2003, from http://www.ed.gov/Technology/TechConf/1999.html

O'Dwyer, L. M., Russell, M. & Bebell, D. J. (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). Retrieved September 14, 2004 from http://epaa.asu.edu/epaa/v12n48/

Oppenheimer, T. (2003). *The Flickering Mind: The False Promise of Technology in the Classroom and How Learning Can Be Saved.* New York, NY: Random House.

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

Quality Education Data, Inc. (QED). (2003, September 16). *QED's Technology Purchasing Forecast 2003–2004*, 9th Edition. Retrieved November 11, 2003, from http://www.qeddata.com/technologypurchasingforecast.htm

Ramirez, M., & Althouse, P. (1995). Fresh thinking: GIS in environmental education. *THE Journal*, *23*(2), 87-90.

Robinson, W.S. (1950). Ecological correlations and the behavior of individuals. *American Sociological Review*, 15, 351-357.

Russell, M. & Haney, W. (1997). Testing Writing on Computers: An Experiment Comparing Student Performance on Tests Conducted via Computer and via Paper-and–Pencil. *Education Policy Analysis Archives.* Received on April 2003 at: http://olam.ed.asu.edu/epaa/v5n3.html

Russell, M. & Plati, T. (2001). Mode of administration effects on MCAS composition performance for grades eight and ten. *Teachers College Record*. Retrieved February 1, 2003 from http://www.tcrecord.org/Content.asp?ContentID=10709

Russell, M. (1999). Testing on Computers: A Follow-up Study Comparing Performance on Computer and On Paper. *Education Policy Analysis Archives*. Received on October 2, 2002 at: http://epaa.asu.edu/epaa/v7n20

Russell, M. (2000). *Wellesley Public Schools curriculum technology program evaluation report*. Wellesley, MA: Wellesley Public Schools.

Russell, M. (2002). It's time to upgrade: Tests and administration procedures for the new millenium. *Essays in Education,* 1. Retrieved March 2, 2004, from http://www.usca.edu/essays/vol12002/time_to_upgrade_revised.pdf

Salpeter, J. (2000, June 15). Taking stock: What does the research say about technology's impact on education? Interview with Jay Sivin-Kachala. *Technology & Learning – The Resource for Education Technology.* Retrieved November 7, 2003, from http://www.techlearning.com/db_area/archives/TL/062000/archives/kachala.html

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)

Stoll, C. (1999). *High-tech heretic.* New York: Random House.

Tierney, R. (1996). Redefining computer appropriation: A five-year study of ACOT students. In C. Fisher, D. Dwyer, & K. Yocam (Eds.), *Education and technology: Reflections on computing in classrooms* (pp. 169–184). San Francisco: Apple Press.

Waxman, H. C., Lin, M., & Michko, G. M., (2003). *A meta-analysis of the effectiveness of teaching and learning with technology on student outcomes.* Naperville, IL: Learning Point Associates. Retrieved April 22, 2004, from http://www.ncrel.org/tech/effects2/

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

Wheelock, A., Bebell, D. J., & Haney, W. (2000) What can students' drawings tell us about high-stakes testing in Massachusetts? *Teachers College Record* ID Number: 10634.

Author Biographies

Laura O'Dwyer is an Assistant Professor at the Graduate School of Education at the University of Massachusetts Lowell. Her research interests include the effects of organizational characteristics on individual outcomes, international comparative studies, and imputation procedures. Recently, her work has focused on international tracking practices, and the impacts of school organizational characteristics on the adoption of technology as a teaching and learning tool.

Michael Russell is an assistant professor in the Lynch School of Education and directs the Technology and Assessment Study Collaborative (inTASC) which is housed within the Center for the Study of Testing, Evaluation, and Educational Policy. 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 is also examining the emergence of 1:1 computing through collaborative research with Andover Public Schools and most recently New Hampshire's Technology Promoting Student Excellence program. In addition, Damian is currently examining the relationship between students various uses of technology and their performance on achievement measures.

Kevon Tucker-Seeley is a second-year doctoral student in the Lynch Graduate School of Education's Educational Research Measurement and Evaluation program at Boston College. He is currently working as a research assistant in the Center for the Study of Testing, Evaluation, and Educational Policy (CSTEEP). 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.

The Journal of Technology, Learning, and Assessment

www.jtla.org