March 2009

Berkshire Wireless Learning Initiative
Final Evaluation Report

Damian Bebell, Ph.D.
Rachel E. Kay
Technology and Assessment Study Collaborative
Lynch School of Education
Boston College

Study Methodology and Data Sources

Evaluation Opportunities and Program Objectives

Given the varied educational settings across the BWLI schools and the participation of two matched comparison settings, the BWLI provides a unique and valuable opportunity to document the effects of 1:1 computing on teaching and learning using both qualitative and quantitative approaches. Specifically, the research design aimed to capitalize on the opportunities inherent in the three-year technology deployment of the BWLI schedule by examining different grade levels (6–8) through a series of naturally occurring pre/post comparisons in addition to comparisons at the two comparison middle schools.

The evaluation sought to provide meaningful data concerning the immediate impacts of the technology on classroom practices for practitioners and administrators throughout the deployment period. The evaluation design also aimed to address a number of the more far-reaching goals of the program by examining the impacts of the technology on student achievement and on more nuanced educational impacts using both qualitative and quantitative techniques. Specifically, a series of teacher surveys, selected teacher interviews, student surveys, student drawings, analysis of existing school records and test scores, as well as qualitative classroom observations were used to document and track the impacts of 1:1 computing on teaching and learning practices across the five experimental settings. Student achievement measures were examined through the secondary analysis of student level MCAS test data in the three participating public BWLI schools and two comparison sites through a non-equivalent comparison group study. An additional student writing assessment was undertaken in Spring 2008 whereby 7th grade students were randomly assigned to complete an extended writing exercise using their school laptop or using traditional paper and pencil. This writing assessment provides an opportunity to assess the impacts of the BWLI program on students writing that may be missed using only a paper-based writing assessment.

Specifically, the evaluation of the program sought to measure how successfully the BWLI program achieved the following targeted outcomes:

1. Enhanced student achievement as shown through *test scores, teacher survey data and assessments aggregated at the classroom and school levels*;
2. Improved student engagement as shown through *teacher and student survey data, student drawings, teacher interviews, principal interviews, and classroom observations*;
3. Fundamental paradigm changes in teaching strategies, curriculum delivery, and classroom management as shown through *teacher and student survey data, teacher interviews, principal interviews, student drawings, and classroom observation*; and
4. Enhanced capabilities among students to conduct independent research, and collaborate with peers as shown through *teacher and student survey data, teacher interviews, principal interviews student drawings, and classroom observations*.

Table 3, below, describes each of the BWLI data collection procedures linked to the targeted project outcomes that they address.

Table 3: BWLI data collection procedures and targeted project outcomes

Procedure	Description	Outcome(s)
Student Survey	Web-based student surveys were given to all participating BWLI and comparison group students both before and during the laptop implementation	1, 2, 3, 4
Teacher Survey	Web-based teacher surveys were given to all participating BWLI and comparison group teachers before their students experience 1:1 computing (Pre) and again near the end of each 1:1 school year (Post).	1, 2, 3, 4
Student Drawing	A student drawing exercise asked BWLI students to reflect on "writing in school" through an open ended drawing exercise before they experience 1:1 computing (Pre) and again near the end of each 1:1 school year (Post).	2, 3, 4
Classroom Observation	Trained researchers and college students conducted pre-arranged visits to observe and record technology practices in 1:1 classrooms during the second year of program implementation.	2, 3, 4
MCAS Analysis	Research team analyzed 2006-2008 item-level MCAS results for each participating BWLI and comparison group student to determine the impacts of various technology practices (as measured via surveys) on standardized test performance.	1
Computer Writing Assessment	1:1 BWLI students were randomly selected in Spring 2008 (Year 3) to participate in an extended computer-based writing assessment to determine the impacts of technology practices on writing length and quality.	1
Teacher Interviews	At various intervals across the deployment of 1:1 student laptops, samples of BWLI teachers participated in short informal interviews regarding their progress, attitudes and results related to the program.	2, 3, 4
Principal Interviews	At various intervals across the deployment of 1:1 student laptops, BWLI principals participated in short informal and formal interviews regarding their progress, attitudes and outcomes related the program.	1, 2, 3, 4

Thus, over the course of the three year study, the research team employed a series of teacher surveys, selected teacher interviews, principal interviews, student surveys, student drawings, and qualitative classroom observations to document and track the impacts of 1:1 computing on teaching and classroom practices.

As the first part of the program implementation, all teachers were provided with laptops in late summer 2005 and have been regularly offered a variety of training and professional development opportunities. In early January 2006, each of the 7th grade students across the five participating schools (n=633) received Apple iBook G4 laptops for use during the remaining first year of the BWLI implementation. During this first year implementation, 7th grade students and teachers both reported substantial increases in the frequency and variety of technology use across the curriculum. At the first months of the second and third year of the laptop implementation all 6th, 7th, and 8th grade students across each participating school (n=1,700) were provided iBook G4 laptops for the majority of the 2006–2007 and 2007–2008 school year.

It should be noted that each of the participating BWLI schools varied the deployment and management of the laptop program to best suit the needs of their own distinct educational community. For example, each school subscribed to their own nuanced resource allocation, training and professional development opportunities, policies, and technical and curricular support. Even though the BWLI program raised the overall level of technology so that every student and teacher had a laptop computer in school, the differences in the deployment and management across the five settings remained substantial throughout the implementation period. For example, even though every school had policies and practices established so that students could bring their school laptops home for use, the actual policies and practices varied across each setting during the implementation period. Additional differences in the deployment and educational environment across the settings will be further explored in the results and discussion sections of this paper.

Given that the BWLI is modeled as a pilot program funded partially by legislatures interested in the utility and application of 1:1 computing, an integral component of BWLI is the current three-year research study funded by the state. The following sections of this report provide an additional overview of the specific research tools used in the current study, a summary of the project findings to date, and a discussion of these findings. After a summary of the studies methodological approach and instrumentation, the report focuses on BWLI and comparison group *teacher* results including teachers' use of technology across the curriculum, and how they saw the 1:1 program impacts their own teaching and their students. The report next focuses on the examination of the BWLI and comparison group *student* results including a detailed summary of students' use of technology across schools and over time. Finally, the current document presents analyses concerning the relationship between students' use of technology and student achievement as measured by various components of the Massachusetts Comprehensive Assessment System (MCAS) as well as in a randomized computer writing exercise.

Student and Teacher Survey Response Rates

Student Survey

As more fully described in the BWLI evaluation plan (Bebell & Russell, 2005), all participating BWLI students completed a web based survey designed to focus on the frequency of many student technology uses both in and out of the classroom and across the curriculum. Students were additionally asked to report on the frequency of their teachers' use of technology across major curricular areas (Math, Reading/ELA, Social Studies, and Science) in addition to a handful of demographic items and a brief attitudes and beliefs inventory.

Given that the first year of the program focused on the 7th grade, grade seven students across the BWLI schools completed pre-1:1 laptop surveys in December 2005/January 2006. The overall response rate from the pre-laptop student survey was 574 surveys; or 90.4% of the 635 7th grade students. After approximately five months of 1:1 computing, these 7th grade students completed a post-laptop survey in early June 2006. Across the BWLI schools, the overall response rate from the student survey was 524 surveys, or 82.5%. Upon completion of the second year of the program (June 2007) when students across all grade levels had access to laptops, the Year 2 online survey was collected from 1,839 of the potential 1,898 students resulting in a 96.8% response rate. Table 4, below, shows the BWLI response rates across the BWLI schools for the first three administrations of the student survey (December 2005/January 2006, June 2006, and June 2007).

Table 4: Student survey response rates for Year 1 and Year 2 data collection

	# of Participating BWLI Students	# of Survey Responses	BWLI Response Rate
Dec. 05/ Jan. 06 (7th grade only)	635	574	90.4%
June 2006 (7th grade only)	635	524	82.5%
June 2007 (Grade 6–8)	1898	1839	96.8%

Thanks to the diligence of school staff and survey technology resources, nearly *every* eligible student across the five BWLI schools completed third year and final student survey in June 2008. Like past survey administrations, students across the two comparison schools were also solicited to participate in the survey. The June 2008 student survey response rates are presented below for each of the BWLI and comparison schools in Table 5.

Table 5: Year 3 (2007–2008) student survey response rate

School Name	Student Population	Survey Responses	Response Rate
South Middle School	697	556	79.7%
North Middle School	790	553	70.0%
Total Comparison Schools	**1487**	**1109**	**74.6%**
Conte Middle School	318	318	100%
Herberg Middle School	699	691	98.9%
Reid Middle School	644	643	99.8%
St. Mark	84	84	100%
St. Joseph	41	26	63.4%
Total BWLI Schools	**1786**	**1762**	**98.7%**

A total of 1,109 students completed the Year 3 survey from the two comparison schools yielding a combined response rate of 74.6%. The response rate was higher across the BWLI schools with 1762 out of 1786 eligible students completing the survey resulting in a 98.7% response rate. Across the five BWLI schools, response rates ranged between 100% at St. Mark and Conte to 63.4% for the small (n=41) 8th grade class participating in the program at St. Joseph. Survey response rates in this range are unusually high and represent the efforts of teachers, principals and administrators at each school. In addition, daily email updates were sent to school leaders containing lists of un-surveyed students and a web site was established where response rates could be monitored across schools in real time. Such exemplary student response rates ensure that later statistical analyses will be highly representative of the population of all BWLI students (with the exception of the small population of St. Joseph participants).

Teacher Survey

As more fully described in the BWLI Evaluation Plan (Bebell & Russell, 2005), every teacher participating in the BWLI program was to be surveyed prior to and during the three years of the program. To this end, 7th grade teachers across the four BWLI schools (as well as the two comparison group schools) were asked to complete pre-1:1 laptop surveys in January 2006 and again across all grades in June 2006, June 2007, and June 2008.

In the current report, results from the June 2008 teacher survey are compared to past survey administrations. Results from the first (January 2006) teacher survey again reflect a time when teachers had recently received their own laptops but no students had yet been issued computers. The May 2006 teacher survey administration asked teachers to focus on their first year experiences in the BWLI program with 7th grade students having had access to computers for the last five months of the year but still found all the 6th and 8th grade students without computers. It should be noted

that the Year 1 teacher survey response rates were not optimal, but efforts focused on the 7th grade teachers who were most involved with the first year of the project implementation. An overview of the number of teacher survey responses across BWLI schools and survey administration is presented below in Table 6 for Year 1 and Year 2 data collection.

Table 6: Number of teacher survey respondents by school and survey administration for Year 1 and Year 2

School Name	Jan. 06	May 06	June 07	Total
Comparison	**12**	**9**	**43**	**64**
Conte Middle School	20	9	44	73
Herberg Middle School	10	20	56	86
Reid Middle School		21	35	56
St. Mark/ St. Joseph	3	3	8	14
Total BWLI Schools	**33**	**53**	**143**	**229**

As Table 6 shows, fewer teachers responded to the teacher survey in its earlier administrations with only 33 7th grade BWLI responses from the January 2006 (Year 1—7th grade pre student laptop) survey. A total of 53 BWLI school teachers were surveyed again in June 2006 after their 7th grade students had 1:1 access to laptop computers (Year 1—7th grade post student laptop). In June 2007, the first year of student access across all grade levels, a school-wide survey of 143 BWLI teachers was completed. The response rates from the June 2007 teacher survey are presented below in Table 7.

Table 7: Year 2 (June 2007) teacher survey response rate

School Name	Teacher Population	Survey Responses	Response Rate
South Middle School	68	42	61.8%
North Middle School	73	1	1.4%
Total Comparison Schools	**141**	**43**	**30.5%**
Conte Middle School	44	44	100%
Herberg Middle School	61	56	91.8%
Reid Middle School	55	52	94.5%
St. Mark	5	5	100%
St. Joseph	3	3	100%
Total BWLI Schools	**168**	**160**	**95.2%**

For the Year 2 school-wide teacher survey, a total of 43 teachers completed the survey from the two comparison schools yielding a combined response rate of 30.5%. Due to a scheduling conflict, North Middle School chose not to participate in the teacher survey, dramatically lowering the overall response rate for the comparison schools' teachers. Fortunately, the response rate was substantially higher across the BWLI schools with 160 out of 168 teachers completing the survey resulting in a 95.2% response rate.

Like the 2008 student survey, nearly *every* BWLI teacher completed the Year 3 teacher survey in June 2008. The online teacher survey focused on capturing the variety and extent of teachers' technology use, teachers' attitude toward technology, teaching, and learning, as well as teachers' beliefs on the effects of the BWLI program and is one of the main foci of the current report.

The June 2008 teacher survey response rates are presented in Table 8.

Table 8: Year 3 (2007–2008) teacher survey response rate

School Name	Teacher Population	Survey Responses	Response Rate
South Middle School	80	49	61.3%
North Middle School	73	39	53.4%
Total Comparison Schools	**153**	**88**	**57.6%**
Conte Middle School	42	42	100%
Herberg Middle School	59	59	100%
Reid Middle School	58	58	100%
St. Mark	5	4	80%
St. Joseph	3	0	0%
Total BWLI Schools	**167**	**163**	**97.6%**

A total of 88 teachers completed the survey from the two comparison schools yielding a combined response rate of 57.6%. Again, the response rate was substantially higher across the BWLI schools with 163 out of 167 teachers completing the survey resulting in a 97.6% response rate. It is unfortunate that the two Pittsfield parochial schools (St. Joseph and St. Mark) resulted in the less than 100% perfect response rate, with respective response rates of 0% and 80%.

Despite the lack of St. Joseph responses, response rates in this range are substantially above average and represent the combined efforts of teachers, principals, and administrators across each school and district. In addition, daily email updates were sent to school leaders which contained lists of un-surveyed teachers and a web site was established where response rates could be monitored in real time. Again, such exemplary teacher response rates ensure that later statistical analyses will be highly representative of the total population of BWLI teachers. Because the surveyed sample of teachers so closely approximates the total population, a wide variety of teachers completed

the final Year 3 (June 2008) survey. Amongst the 163 teacher survey respondents, a substantial portion taught or worked with students outside of the classroom or in subject areas beyond the original scope of the evaluation (foreign languages, physical education, etc.). Table 9, below, shows the number of Year 3 teacher survey respondents comprising the current sample.

Table 9: Characteristics of Year 3 (June 2008) teacher survey sample by school

	All Subject Areas		Primary Subject Areas	
School Name	**# of Respondents**	**% of Sample**	**# of Respondents**	**% of Sample**
Conte Middle School	42	26	29	27
Herberg Middle School	59	36	41	38
Reid Middle School	58	36	35	32
St. Mark	4	3	4	4
Total	**163**	**100%**	**109**	**100%**

Table 9 shows that of the 163 teachers across the four BWLI schools who completed the BWLI teacher survey, 109 were classified as teaching one of the focused primary subject areas (English/ Language Arts, Math, Science, Social Studies). In addition to the 109 primary subject classroom teachers the survey was also completed by an addition 54 teachers and educators who taught subjects including: physical education, special education, reading intervention, foreign languages, and health.

To simplify the interpretation of the results, the current report presents survey results from the 109 primary subject teachers. In other words, the research team focuses the current report on data reflecting the impacts of the BWLI program on the core academic curriculum. Interestingly, the differences between the survey responses of the 109 primary subject classroom teachers and the other subject area specialists concerning the impacts of the BWLI program were largely negligible.

Student Drawings

As demonstrated in past research studies of technology programs, student drawings provide a unique and valuable reflective tool for both the participants and observers to perceive the impact of technology from the students own viewpoint (Russell, Bebell, Cowan, & Corbelli, 2003; Russell, Bebell, & Higgins, 2004). Although student drawings may be an unusual tool for collecting information about students and their classrooms, student drawings continue to provide a rich descriptive examination of students' perspective that may be neglected by more traditional data sources. Participating students across all BWLI schools contributed student drawings both before and after implementation of the laptop program.

In each case, teachers distributed the blank drawing forms to their students which prompted them to:

> Think about the work you do in your classroom.
> In the space below, draw a picture of yourself writing in school.

For each student drawing, a trained member of the research team coded a number of dichotomous features that have been pre-selected using an emergent analytic coding process established through prior student drawing and 1:1 research studies (Bebell, 2001; Bassette, 2008). The specific features coded in the drawings fall into four broad categories:

1. Student Characteristics (what the students were doing),
2. Technology Present (type of technology depicted),
3. Student Demeanor (whether the student was depicted positively, negatively, or neutral), and
4. Other Features (presence of teacher or other students, classroom decorations, multi-frame drawing).

With minimal training these drawing categories and codes can be applied to student drawings with both a high degree of inter-rater and intra-rater reliability so that the codes assigned to the drawings are stable across different raters and over time. Before the team of researchers commenced coding, 30 drawings from a prior 1:1 research study were randomly selected and coded by the research members to document inter-rater reliability. Across the raters, the inter-rater reliability estimate (i.e. percent agreement) was consistently greater than 95% which we consider more than sufficient for the current investigation and moved forward with coding the BWLI drawings.

Across the three years of the study, over 3,500 drawings were collected and analyzed from BWLI students. For the current analyses, baseline drawings were examined from the first cohort of 7th grade students in December 2005 before laptops were deployed. These students again were asked to submit drawings at the end of their first half-year of 1:1 computing in June 2006, and the exercise was replicated with the subsequent cohort of exiting 7th grade students in June 2008, who at that time had approximately two years of experience participating in the 1:1 laptop program. The response rates from this inquiry were generally quite good, with 91% of all eligible students completing drawings in December 2005, 84% in June 2006 and 85% in June 2008.

As more described in the results section of this paper, an internet database has been created to share common or interesting examples of drawings. In addition, codes and their definitions used in the current evaluation can be accessed on the BWLI web site at: www.bc.edu/bwli.

Classroom Observations and Teacher Interviews

Throughout the three-year implementation of the 1:1 program, members of the research and evaluation team made regular visits to each participating school to conduct classroom observations, informally interview teachers, and both formally and informally interview school principals as well as other building level and district level leadership. Although a formal classroom observation study

was eliminated from the original evaluation due to post-proposal budget cuts, the research team leveraged their available resources to regularly observe 1:1 classrooms across the study settings. The information provided from the classroom observations served multiple purposes in the overall evaluation of the 1:1 initiative. First, the classroom observations served to document the variety and types of specific technology use occurring across a variety of subject areas and settings. These recorded observations were reflected upon with project leadership and the schools themselves during the course of the study to guide, monitor and facilitate implementation. Second, through these observations and school visits the evaluation team developed a better sense of how quickly technology practices occurred within the schools as well as levels of student engagement, pedagogical practices, and school climate. For example, the classroom observations documented meaningful examples of the real world day-to-day challenges and complications inherent with the transformation of established teacher and learning practices. Finally, at the end of the data collection period, the observation notes and experiences prompted exploratory data analyses of the teacher and student surveys for meaningful trends and relationships.

Thus, classroom observations were used as one of many tools to help facilitate a formative understanding of the varied ways in which the laptops and technology were being used throughout the school day. To facilitate this aspect of the classroom observations, during the first year of the project, two education majors from the Massachusetts College of Liberal Arts (MCLA) were recruited and trained to carry out classroom observations focusing on the 7th grade classrooms where students had been recently provided laptops. In Fall 2005, the two MCLA students completed formal training and National Institute of Health certification on conducting classroom research. In Spring 2006 (during the first months of student 1:1 computing), a total of 11 complete lessons or class periods were formally observed and documented by the MCLA student team. All of these pre-scheduled observations with voluntary teachers occurred at the Conte Middle School due in large part to its close geographic proximity to the MCLA campus. Since the classroom observations were chiefly concerned with documenting technology practices, this approach was feasible from both the perspective of the somewhat apprehensive classroom teachers who had just begun their 1:1 experience and the research studies formative and summative objectives. Given this, the classroom observations collected during Spring 2006 were not presented as a representative sample of 1:1 classrooms at that time, but rather volunteer teachers who offered to share a technology-rich lesson with an observer.

During each of the second and third years of the program, members of the Boston College research team made approximately 45 classroom observations. Although the resources allocated for these Year 2 and 3 school visits were minimal, the school visits continued to remain a rich source of information in addressing the potential impacts of the BWLI as they had during the first year. These research team visits were conducted across all BWLI schools and generally included classroom observations as well as interviews with teachers, principals, and other staff. It is challenging to summarize the conditions of the actual observations that were conducted given that the conditions of the observations fluctuated widely. In some cases, classes were observed when the teacher knew in advance of the visit. In other situations the researcher would "drop in" on classes about to begin or already underway without the teachers' direct advance notice. Some observations lasted

for full instructional periods and others were only for segments of a class. In many cases, informal interviews with the observed teachers were arranged to address or clarify any specific queries from the observation. In many cases, the informal teacher interview also allowed the teaching staff an opportunity to express their thoughts and beliefs about the program, their students, or the technology programs in more candid way than the formal surveys. Although these sentiments and expressions were not formally or systematically analyzed, they provided the research team with valuable insights and perspectives to further explore using the other research tools/inquiry methods, particularly the data analyses of the survey. Specific examples and formative results detailing how the classroom observations were used during the evaluation period can be found in the April 2007 report, *Berkshire Wireless Learning Initiative Evaluation Implementation Notes and Concerns* (Bebell, 2007).

Principal/School Leadership Interviews

In addition to the classroom observations and teacher interviews, interviews were also conducted with each BWLI principal as well as with various support staff since the time of the original deployment. Although much of the data was used for formative reflection during the deployment period, the school principal interviews also served as an important formative data point reflecting the experience of school-level leadership concerning many of the program outcomes. In other words, the principals were all asked to reflect on their experiences not only implementing the program, but were also asked to think critically about the impacts of the program.

Principal interviews were conducted in person with each building leader approximately three to four times during each of the three years of the program and again on two occasions following the final year of the original deployment timeline.

Most interviews lasted between 20 and 30 minutes and were often conducted following classroom observations. After each round of principal interviews, field notes were summarized across the schools and follow-up questions, clarifications, and queries were also handled via email and phone.

The content and nature of the principal interviews continuously changed and adapted to suit the needs of the evaluation. For example, in some cases, specific interview scripts were used to elicit specific responses and allow for cross-school comparison while in other cases the interviews were more loosely structured to allow principals more latitude in the interviews content and to specifically express specific practices or outcomes within their setting.

In late 2008, after the completion of three years of 1:1 computing, a final principal/school leader interview was conducted in each of the participating schools. This final formal interview followed a scripted interview protocol to ensure that every respondent would address the same questions and issues. A portion of the final interview content consisted of a series of discrete short response questions based on the teacher survey which concerned the impacts of the program at each school. The interview also included a number of opportunities for the school leader to reflect more widely on their experiences in a series of open response questions concerning the deployment, management, obstacles, and impacts of the BWLI program. Summarized results from the principal/school leadership survey are shared in the results and discussion sections of this paper.

Exploring the Impact of 1:1 Computing on Student Achievement

School and Student Level MCAS Analyses

Given that the first targeted outcome of the BWLI program was to "enhance student achievement" a central component of the evaluation is the investigation of 1:1 participation and technology use on student test scores. Despite the massive investments that federal, state, and local agencies devote to educational technology expenditures and initiatives, there remain only a few empirical studies that explore the relationship between students' and teachers' use of technology and their performance on measures of student achievement. This lack of rigorous research is particularly conspicuous at a time when nearly all educational ideas and programs are being evaluated by policy makers and educational leaders for their impact on standardized local and state level assessments.

One of the biggest reasons for this lack of research is the inherent complexity and difficulty involved in measuring emerging technology practices in 1:1 settings and associating them with valid measures of student achievement. Past research focusing only on participation in a 1:1 program may fail to account for important differences and distinctions between how individual teachers and students use technology. For example, one classroom teacher and their respective classes may effectively use in a way that may be positively impacting student achievement. However, this educational practice is not shared with the other schools teachers whose students show no difference on measures of their achievement during this same period. When both classrooms are merged and considered equally across a 1:1 setting, the evidence of improved student achievement is marginalized. In other words, by only looking if average scores or grades improve in laptop *settings*, a lot of potentially valuable data may be overlooked. Given the exploratory and varied nature of 1:1 initiatives, the current evaluation of student outcomes used three years of data collected from student surveys that was matched and merged with students' individual state assessment results to explore how *students* varied technologies practices have impacted student achievement. Given that there was no commonly shared assessment across all five of the BWLI schools, the Massachusetts Comprehensive Assessment System (MCAS) was deemed the best available measure outcome for student achievement.

One of the challenges in determining the impact of educational technology (or anything else for that matter) on student achievement is that student achievement means different things to different people. To many classroom teachers, student achievement can mean how well their students interact in class and on content specific teacher-created tests. To policy makers, student achievement is increasingly defined as a students' performance on a state-sanctioned standardized assessment. In addition to the use of standardized tests in summarizing academic performance for policy discussions, standardized tests, such as the MCAS also offer a convenient and potentially meaningful measure of student achievement that is shared across all public schools within a state. Although the subject offerings and grade requirements have changed under different federal and state administrations and policies, the MCAS has been the de facto measure of student achievement in Massachusetts since its debut in 1998. As the state's most important and publicized metric of school quality, student MCAS scores are taken seriously by school and district leadership as well as teachers and students within each school. As we will later discuss in the later sections of this paper, the MCAS is

by no means the perfect outcome measure for our current research, but given a lack of alternatives, we feel the MCAS is highly suited for this purpose since it is the most widely accepted measure of student achievement for both education and policy communities available.

The current investigation will employ MCAS results to show at least one aspect of the impacts of the BWLI program and 1:1 computing on students and schools. In our exploration of student achievement, we will consider results from a variety of perspectives and approaches including both school averages and individual student results for those 7th and 8th grade students who attended BWLI schools and experienced 2 years of 1:1 student computing. Specifically, the current investigation will address:

- trends in schools overall MCAS performance over time compared to the comparison schools and state trends during this same period, and
- which, if any, of students technology uses in school or at home have impacted student-level performance on various MCAS measures (while statistically controlling for students pre-BWLI academic performance using prior MCAS performance)

School level MCAS results and performance indicators from 1998 to 2008 were accessed from the Massachusetts Department of Education while student level data was provided to the research team directly from the BWLI and comparison schools for all participating students for 2008, 2007, and 2006. To facilitate our analyses of how different types of use impacted student test performance, grade 7 and 8 student survey results across all BWLI and comparison students were merged with the item-level MCAS data. Grade 6 results were not analyzed as 6th grade students had the least exposure to the BWLI program and 1:1 computing. Again, because only public school students participated in the MCAS as a mandatory state assessment, only results from the three BWLI public schools (Conte, Herberg, and Reid) and the two comparison schools (North and South) were available for these analyses. Thus, for each grade level, a new data set was created by the research team that included both the student's MCAS results and demographic information as well as their BWLI student survey on technology use and practices. So, for each 7th and 8th grade student who completed the MCAS and the BWLI survey, the relationship between various technology uses and various outcome/student achievement measures could be examined. Since nearly all students in the BWLI public schools completed both the MCAS and the BWLI student survey in Spring 2008, we are capable of performing our investigation with a highly representative sample of BWLI student participants. Table 10, on the next page, summarizes the MCAS subject tests schedule used by the state across various years of the BWLI program.

Table 10: Summary of student level MCAS administration results across subjects and BWLI student laptop deployment schedule

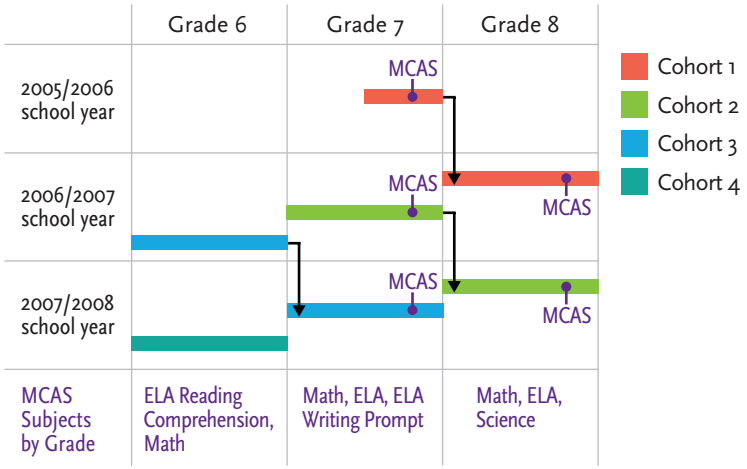

Table 10 provides a simple list of what outcomes measures were available for each student cohort (grade 7 and grade 8) in our student level analyses of achievement. In the final year of the BWLI program, grade 8 students completed the Spring 2008 MCAS testing after nearly two full school years participating in the 1:1 student laptop program (throughout grade 7 and 8). The grade 8 MCAS tested students in English language arts (ELA), math, and science. In ELA, there were 36 multiple-choice items and 4 open-response items split between two strands, Reading and Literature, and Language. In math, there were 30 multiple-choice items, 4 short-answer items, and 5 open-response items. These were split between five content strands: Number Sense and Operations; Patterns, Relations, and Algebra; Geometry; Measurement; and Data Analysis, Statistics and Probability. In science, students had 34 multiple-choice items and 5 open-response items. They were divided into four strands: Earth and Space Science, Life Science, Physical Sciences, and Technology/Engineering.

Table 10 also shows that in the final year of the BWLI program, grade 7 students also completed the Spring 2008 MCAS after nearly two full years of 1:1 computing (throughout their 6th and 7th grade). Grade 7 students were tested in ELA and math. In ELA, there were 36 multiple-choice items and 4 open-response items. Additionally, the ELA grade 7 test includes an extended writing exercise which is scored by the state and yields two scores: a topic and idea development score and a standard English conventions score. In math at Grade 7, there were 29 multiple-choice items, 5 short-answer items, and 5 open-response items. For both the 7th grade ELA and math exams, the content strands were the same as those reported above for Grade 8.

In addition to looking at school level trends in student performance both before and during the laptop implementation period, the current investigation will use these student-level MCAS data at each grade level with student survey responses to determine the relationship between technology practices and student achievement. These student level analyses will employ linear regression techniques and use students most recent MCAS scores (Spring 2008) as the dependent or out-

come variable. Students individual MCAS scores from the pre-BWLI test administration (for example, when the 8th grade students were in Grade 6) will be employed as a independent variable to account (or control) for their prior level achievement. For both of the outcome measures and control variables, MCAS raw scores will be used for each subject area of the assessment. Since the students in this analysis are clustered within schools, it would make sense to do multilevel regression analysis as it can be assumed that students within the same school may have correlated errors (Luke, 2004). However, since there are a limited number of BWLI and comparison public schools (n=5), there would not be sufficient power to analyze the data reliably.

Computer Writing Assessment

Like the rest of the MCAS, the grade 7 ELA open-ended writing assessment is completed entirely on paper and pencil, which research now suggests may be unfairly underestimating the true writing ability of students who have grown accustomed to composing and editing text on a computer (Russell & Plati, 2001; Russell, 1999; Russell & Haney, 1997). In other words, the mode of the state test administration (paper and pencil) may be having a negative impact on student performance for those very students who have adopted and are using the new learning techniques ushered by the 1:1 laptop program. In response to this growing body of literature suggesting that paper-based assessments, like the MCAS, may be poorly suited for accurately measuring the writing quality of students who may be more accustomed and experienced using Word processors and technology, a study of students computer writing ability was incorporated into Year 3 of the BWLI evaluation. Specifically, in Spring 2008, 7th grade students across all participating BWLI schools were randomly assigned at the classroom or student level (depending on school) to complete a mock-MCAS open response assessment using either their BWLI laptops or the traditional paper/pencil required by the state.

Given that each of the three public schools had a great deal of experience administering and interpreting the aforementioned 7th grade ELA MCAS open-response writing prompt, a publicly-released writing prompt from a previous (2006) MCAS ELA assessment was selected. Since the public schools were quite familiar with this assessment, adapting it to allowed computer access for some test takers was not a major challenge. However, given each schools' past experience with official and practice MCAS sessions, interviews with ELA teachers and school leadership were conducted to ensure that each schools 7th grade student population had no prior familiarity or experience with the specific writing prompt to be used in our computer writing assessment. It was determined through this interview process that no 7th grade students in any of the BWLI settings in 2008 had any prior experience with the 7th grade ELA writing prompt originally included on the 2006 MCAS:

> After-school activities give students an opportunity to do something they enjoy outside the classroom. Some students play sports, while others might enjoy acting or other activities.
>
> Your school wants to add some new after-school activities and is taking suggestions from students. In a well-developed composition, suggest a new after school activity and explain why students might enjoy it.

The open ended writing assessment provided students with a 45-minute session to draft a written composition; then, in a separate 45 minutes administration session, students completed a final composition based on their draft. Again, the chief difference between the actual MCAS administration of the ELA writing assessment and the BWLI computer writing study conducted in Spring 2008 was that in our study students were randomly assigned to complete the assessment using either their BWLI laptop or using paper and pencil. Before the writing study, rosters containing the names of all of their 7th grade students which were collected and collated by student homeroom and ELA teacher in each BWLI school.

Although the original study design was to randomly assign testing conditions (computer vs, paper) at the student level, the task of simultaneously re-assigning hundreds of students to different testing classrooms resulted in the two Pittsfield public schools randomly sampling their intact homerooms. In other words, in these settings in-tact homerooms of 7th grade students were randomly assigned to one of the two potential testing conditions. Although of weaker methodological design, the random assignment of intact classroom was not seen as a major drawback to the overall study design since (1) student assignment to the homerooms in these schools was largely random or alphabetic and (2) the research team could additionally control statistically for differences in students' non-laptop writing performance using the state 2008 MCAS results. Indeed, since all participants in our computer writing study would be taking the actual paper-based MCAS writing assessment later in the school year, a higher proportion of students (approximately two-thirds) were randomly assigned to complete the BWLI computer writing study in the laptop setting. The breakdown of student assignment to the computer and paper testing conditions for the BWLI computer writing study is detailed across the four participating schools in Table 11, below.

Table 11: Student assignment to testing conditions for the Spring 2008 BWLI computer writing study across BWLI schools

School Name	Condition	# Students	% of 7th grade population
Conte Middle School	Computer	50	69%
	Paper	26	
Herberg Middle School	Computer	144	90%
	Paper	45	
Reid Middle School	Computer	85	67%
	Paper	56	
St. Mark	Computer	30	83%
	Paper	14	
Total	**Computer**	**310**	
	Paper	**141**	

As Table 11 shows, a total of 310 BWLI 7th graders were assigned to complete the computer writing study in laptop conditions (or 68.7% of all 451 participating students) while 141 students completed the assessment using traditional paper/pencil conditions (31.3%). It should be noted that despite the fact that St. Mark students and teachers had no prior experience with the state assessment, they were enthusiastic participants in the computer writing study. Thus, a total of 451 7th grade students completed the computer writing study in Spring 2008, after students had laptop access for most of their 7th and 6th grade school years. Given no make-up period was allowed, the percent of students participating in the BWLI assessment was diminished by student absences and testing exemptions. However, the large majority of the 7th grade population participated in the writing study.

Testing materials and directions were adapted from the state's for both the paper and computer testing classrooms and distributed across each school prior to the assessment. Specifically, proctoring teachers and students within each setting were provided paper directions, scripted instructions and student rosters. Directions for the paper-learning conditions were nearly identical to those of the state, while the directions for the laptop conditions specified that all spell-checking, grammar-checking and other automated features of Microsoft Word (students most commonly used writing program) be temporarily turned off and/or removed for the duration of the assessment to ensure the only substantive difference between the computer and paper environments was the mode of test administration. All other testing circumstances (time, resources, directions, scoring criteria, etc.) were held exactly the same to ensure a realistic and meaningful testing experience as well as to ensure valid results. Copies of all directions and materials used in the BWLI computer writing study are available at www.bc.edu/bwli.

The completed copies of student paper-and-pencil essays were collected and sent in pre-paid mailers to the evaluation team while computer created student essays were uploaded to teacher computers (via teacher drop boxes, file transfers, or as email attachments) and then either burned to CDR and mailed to the evaluation team or sent via email. In addition, USB-Flash drives were provided to classroom proctors in the laptop settings for saving electronic versions of student essays directly from students' computers.

To eliminate any later bias inherent in the paper or computer formatted essays when being scored, a team of six trained education undergraduate students at Boston College were recruited and trained to electronically input each of 141 paper essays (including all student mistakes, typos, etc.) into Microsoft Word. Once all the paper essays were successfully input into an electronic format, a second recruited team of 8 Boston College undergraduate students completed formal training and reliability testing for use of the states coding rubric to score the 7th grade essays.

Training for the eight essay scorers took place over two days and closely followed the same detailed instructions and procedures used by the state's testing contractor outlined in "Ensuring Technical Quality: Policies and Procedures Guiding the Development of the MCAS Tests" (Massachusetts Department of Education, 2008) as well as the additional scoring information found at the online resource NCS Mentor for MCAS ELA Composition: http://www.ncsmentor.com/. In addition to the general descriptions, the training also provided anchor papers and benchmark

papers for each category. These anchor and benchmark papers provide concrete examples of each performance level. The anchor and benchmark papers were first introduced to raters during the common scoring training session and were available to essay scorers throughout the scoring process. In addition, to ensure accurate and consistent scoring each trained essay scorer completed a qualifying set of ten example essays during the training session. Before an essay scorer could begin rating the BWLI essays, they first needed to meet or exceed the state's minimum accuracy rate, as described below:

> To be eligible to score operational MCAS responses, scorers were required to demonstrate scoring accuracy rates of minimum 70 percent exact agreement and at least 90 percent exact-or-adjacent agreement. In other words, exact scores were required on at least seven of the Qualifying Set responses and either exact or adjacent scores were required on a total of nine of the 10 responses; scorers were allowed one discrepant score, as long as they had at least seven exact scores.
>
> (Massachusetts Department of Education, 2008)

After completion of the training, each of the eight scores met or exceed the state requirements for scoring open-response essays, Thus, every essay scorer in the BWLI computer writing study received the same training offered by the state and the testing contactor and were qualified to use the same rating criteria.

After this training, each member of scoring team was provided a random sample of student essays to score. Each essay was scored on two dimensions: Topic Development and Standard English Conventions. The scale for Topic Development ranged from 1 to 6 and the scale for English Conventions ranged from 1 to 4, with one representing the lowest level of performance for both scales. The rubrics and range of scores for the Topic Development score and the Standard English Conventions scores and are summarized on the next page in Table 12.

Table 12: MCAS 7th grade open response scoring rubric for Topic/Idea Development and Standard English Conventions

Score	Topic Development	English Standards
1	• Little topic/idea development, organization and/or details • Little or no awareness of audience and/or task	• Errors seriously interfere with communication; **and** • Little control of sentence structure, grammar and usage, and mechanics
2	• Limited or weak topic/idea development, organization, and/or details • Limited awareness of audience and/or task	• Errors interfere somewhat with communication; and/or • Too many errors relative to the length of the essay or complexity of sentence structure, grammar and usage, and mechanics
3	• Rudimentary topic/idea development and/or organization • Basic supporting details • Simplistic language	• Errors do not interfere with communication; and/or • Few errors relative to the length of the essay or complexity of sentence structure, grammar and usage, and mechanics
4	• Moderate topic/idea development and organization • Adequate, relevant details • Some variety in language	• Control of sentence structure, grammar and usage, and mechanics (length and complexity of the essay provide opportunity for students to show control of standard English conventions)
5	• Full topic/idea development • Logical organization • Strong details • Appropriate use of language	
6	• Rich topic/idea development • Careful and/or subtle organization • Effective use of language	

Each essay was independently scored by two raters and assigned an official score in each area; the two scorers' scores for each area are combined to report a total score range from 2 to 12 for Topic Development and a total score range from 2 to 8 for Standard English Conventions. Returned essays and their scores were provided back to each of the schools within six weeks of the exercise. Student scores were also input into our master data set where open-response scores from the computer writing exercise could be examined with official MCAS results and survey results.

In theory, the random assignment of students to either testing condition should serve to eliminate potential pre-existing differences among the two groups of students and that any differences observed between the two groups in the computer writing study are a sole result of the testing format. However, to even better control for pre-existing differences between the computer and paper students, 7th grade open response scores from the "official" paper-and-pencil ELA MCAS are used in the our statistical models as a covariate to represent students writing ability using traditional paper and pencil. In other words, with official writing data for each student from the 2008 paper-based state assessment contrasted with the results from the 2008 BWLI computer writing study, we

can more fully explore the impacts of students use of computers in the MCAS writing assessment. Using a general linear model, we statistically "controlled" for paper-tested writing scores while evaluating the impact of a very similar writing assessment where student employed their laptop and word processor rather than paper/pencil.