

FEBRUARY 17, 2009



## Laptops for Learning Baseline Survey: **Scanning Procedures and Estimate of Data Accuracy**

**Rachel E. Kay and Damian Bebell**

Technology and Assessment Study Collaborative  
Lynch School of Education  
Boston College  
Chestnut Hill, MA 02467



FEBRUARY 17, 2009

# Laptops for Learning Baseline Survey: Scanning Procedures and Estimate of Data Accuracy

**Rachel E. Kay and Damian Bebell**

Technology and Assessment Study Collaborative  
Lynch School of Education  
Boston College

## Introduction

The current paper documents the process and procedures used by Boston Public Schools and the Boston College evaluation team to collect and digitize the paper surveys collected during the baseline data collection of the Laptops for Learning (L4L) evaluation. In addition, the current report also presents an examination of the estimated accuracy of the scanning process by comparing a sample of teacher surveys digitized using the scanning process to hand-entered responses from the original paper surveys.

## Background of the L4L Baseline Teacher Survey

As part of their overall technology plan, Boston Public Schools launched the Laptops for Learning program in the spring of 2008. Through this program, every teacher within the school district received a dual-platform Macbook laptop. Included in this initiative was an evaluation of technology use by teachers and their students across the four-year program using teacher surveys. Although future surveys for this evaluation will be universally delivered via the web, the Use, Support, and Effect of Instructional Technology Study (USEIT) survey was delivered on paper to teachers as they first received their laptops in May of 2008. The survey covered a wide variety of topics relevant to student and teacher technology use, such as:

- Technology resources,
- Technology support,
- Professional development,
- District vision for technology,
- Technology beliefs,
- Pedagogical beliefs, and
- Demographic information

The survey consisted of 334 items across 4 double-sided pages. All survey questions were closed-ended and were completed by filling in an oval on the response sheet. In other words, there were no open-ended response questions. The purpose of delivering this baseline survey was to gain information about how teachers were using computers and their attitudes toward technology *before* the distribution of the L4L laptops. Data collected through the baseline survey will be compared with results from future surveys, after teachers have had individual laptops for one, two, and three years to determine how uses and beliefs change over time and in response to the program.

## Distribution and Collection of L4L Baseline Surveys

The baseline L4L survey was delivered to approximately 5000 teachers across 143 schools comprising the entire Boston Public School system. Specifically, teachers were expected to complete the survey before they received their project laptop during a series of implementation and deployment sessions held in spring 2008. To accommodate teacher schedules, facilitate distribution, and maximize response rates, each school was provided discretion on how to distribute and collect the baseline teacher survey. To best accommodate the wide range of settings and resources available across the 143 schools, the district provided both paper copies of the survey as well as an electronic PDF of the survey that teachers could alternatively print themselves and complete at their own discretions. In addition, some schools chose to distribute the surveys using color and/or black and white copies printed directly from the PDF whereas other schools created their own photocopies using the original paper survey. Some of the surveys were printed one-sided, while others were printed double-sided again resulting in both 4 and 8 page surveys. Lastly, although some teachers completed the survey using pencil, many completed the survey in pen or other writing instrument. Across all these variations of the teacher survey, just over 2200 baseline surveys were collected in spring 2008.

## Digitizing/Scanning the Paper Surveys

The completed teacher surveys collected during the laptop deployment/training session were returned to Boston Public Schools central offices where teacher survey responses could be digitized for empirical analyses. The digitization and scanning of the individual teacher surveys was completed using a flatbed Fujitsu fi-5750c scanner and Cardiff Teleform software for Windows. Company specifications detail that the Fujitsu scanner can scan up to 90 pages per minute including both single and two sided pages and a wide variety of documents, such as photos or pages with taped on receipts<sup>1</sup>.

Before any scanning of actual surveys occurred, the Teleform scanning software was configured and programmed to accommodate the custom-made surveys being scanned. This programming was conducted by a member of the Boston Public School staff and included the identification of item numbers, their geographic placement on each page, and then the assignment of a value for

each potential response. After this information was programmed, several test surveys were scanned to ensure the programming and digitization process was accurate.

Because the surveys had been reproduced in different ways across different school settings, each four-page double-sided survey (or eight-page single-sided survey) had to be placed in the scanner individually. This meant that any staples or paper clips were manually removed from each survey. Then each survey was placed on the scanner where the user selected the one-sided or two-sided templates from the Teleform software. After each survey was scanned, it was then necessary to confirm that the image had been successfully scanned and saved using the Teleform software. Given there was only one scanner and the varied formats of the survey, approximately 50 surveys could be scanned in a successful hour. In some cases, the scanning process took substantially longer such as when printed text was not perfectly aligned on the survey page or when pages were sized irregularly. In these cases, surveys were scanned one page at a time.

The second part of the digitizing process involved verifying the results using the scanned images of the surveys and the Teleform Verifier Software. This software application identified any response within the survey that was questionable or unclear such as items with more than one response marked or items misaligned from the survey master form. Specifically, flagged results were highlighted by the scanner and required the scanning operation to confirm the correct scanned response. In cases where an individual survey response remained in question, a correction was made manually. Additionally, the staff using the verification software were instructed to be alert for any areas of missing responses due to the printed survey being misaligned, in which case those responses were also input manually.

There were four staff people involved in the scanning and verifying processes including three Boston Public Schools district-level staff and one member of the Boston College evaluation team. Scanning and verifying was completed in approximately one month with resulting total of 2201 complete teacher surveys successfully scanned. Approximately 25 teacher surveys were not successfully scanned due to missing pages and other related problems. Raw data for all scanned surveys was delivered to the evaluation team in comma separated value (CSV) format and imported into SPSS for subsequent analyses.

## Determining the Accuracy of the Scanning Process

To examine the accuracy and precision of the digitized survey data, a sample of 28 original paper surveys were selected at random from the 2201 baseline surveys. This random sample of 28 surveys were manually input by a trained Boston College student using a pre-made Microsoft Excel template and then compared to the scanned survey responses digitized using the aforementioned electronic scanning process. An additional two surveys were originally selected for this analyses, but were subsequently not included. Specifically, a positive match from the scanned survey files could not be found for one of the surveys, and was eliminated from this analyses. A second survey was eliminated since the last pages of the survey were misappropriated within the many boxes of

teacher surveys and could not be reasonably located to be hand input, although the file was present in the electronic database. Thus, 28 surveys were available to input by hand for comparison with the original scanned results. A trained Boston College student then commenced to input by hand the responses across the 28 teacher surveys. The matching results from the SPSS file were cut and paste into two worksheets of an Excel spreadsheet where responses from the hand-input results and electronically input results could be compared. A formula was developed and applied within the Excel spreadsheets to match and compare corresponding individual item responses across the 28 surveys. Any survey response that did not match between the original scanned results and the hand-entered results were flagged and analyzed.

A preliminary analyses of the matched survey responses showed that a small concentration of the survey items exhibited a proportionally high number of errors compared to the remaining survey content. A second manual scorer was trained and employed to hand-verify the results for those survey items that contained the greatest proportion of mismatched results. The second hand-scorer found that less than 3% of the entries of the first hand-scorer were incorrect. In such cases where the mismatch responses were attributed to the hand-scored results, corrections to the original human rater were made.

## Accuracy of the L4L Baseline Data File

In total, there were 28 surveys with 334 individual survey items compared, resulting in a total of 9352 possible data points to match. In 392 of these 9,352 total comparisons, the responses from the electronically scanned surveys and the hand-inputs survey did not match. In other words, across all of the 28 teacher surveys we examined, 95.8% of survey items were deemed “correct”. Conversely, when looking across the entire 8-page baseline teacher survey we can estimate that approximately 4% of the survey data may be incorrect due to the variation in the survey forms and scanning process.

To further investigate the nature and degree of errors in the data set attributed to the scanning process, the research team examined the matched data files for potential patterns or trends in the mismatched survey responses. In other words, were the scanning errors concentrated in particular sections of the survey or were they found equally distributed throughout the 8 page document? Indeed, our analyses revealed that there were three survey questions which accounted for more than a third of the total errors. In fact, errors were found to be concentrated in sections of the survey where multiple survey items were located across a single row on the page. For example in Figure 1 (next page), the items on the right half of the page, under “How often used?” had a greater percentage of errors than any other items on the survey.

**Figure 1: Example of survey item with greater likelihood of scanning related errors**

16. Indicate how many of the following devices you have access to and how often you use these devices:

	How many?						How often used?				
	0	1	2	3-5	6-7	8+	Never	Daily	3 X's a week	Once per week	Less than weekly
Desktop computers in your classroom	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Desktop computers in the lab/media center	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Desktop computers in the library	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Printer in your room	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Laptop computers permanently in your classroom	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Shared laptop computers	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Alphasarts in your classroom	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Alphasarts that are shared by other teachers	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
PDAs (Palm Pilots)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Although, the overall accuracy rate of the scanning process we estimate to be 95.8%, we know that for the majority of the survey items, the actual accuracy rate is somewhat higher. For example, if those 31 survey items which contained the greatest proportion of discrepancies were eliminated from our analysis, a total of 303 survey items would remain, or in other words, 91% of the overall survey content had a 97.5% accuracy rate.

From this description and examination of the baseline L4L teacher survey, it is clear that the scanned paper surveys were by no means foolproof, however, given the constraints and challenges in providing a paper based survey to teachers across 143 schools, our estimated accuracy rate of over 95% across all survey items (and even greater when the small number of most problematic survey items were removed) suggests that the electronic data representing the paper-based baseline surveys are an acceptably accurate snapshot of teachers responses.

Additional information about the Laptops for Learning Initiative and corresponding research can be accessed at: [www.bc.edu/L4L](http://www.bc.edu/L4L).

## Endnotes

- 1 Information obtained from <http://www.fujitsu.com/us/services/computing/peripherals/scanners/production/fi-5750c.html>