Beyond Self-Report: Using Technology to Evaluate the Use of Technology in Schools

A paper presented at the annual conference of the American Educational Research Association (AERA)
April 13, 2007
Chicago, IL

Jonathan D. Becker, J.D., Ph.D.
Hofstra University

Charol Shakeshaft, Ph.D.
Hofstra University

Dale Mann, Ph.D.
Teachers College, Columbia University (Emeritus)

NOTE: DO NOT CITE WITHOUT PERMISSION FROM THE AUTHOR(S)

Beyond Self-Report: Using Technology to Evaluate the Use of Technology in Schools

ABSTRACT

I. INTRODUCTION

While “technology integration” is a widely accepted term of art, it is not well-defined in the worlds of educational research or practice. Not surprisingly, then, it is even less well “measured” or documented. Some of the more widely cited, large-scale studies in the field of educational technology use have utilized a survey research design wherein teachers report their own use of computers and/or the degree to which they direct students to use technology (Wenglinsky, 1998; Becker, 2000; O’Dwyer, Russell, & Bebell, 2004). Those studies advance our understanding of technology use in schools, but only incrementally.

The main argument for better methods is, of course, that previous methods are not very good. We have better ways to measure gas mileage than asking drivers what they feel yet for the most part that is how we have been measuring computer use in schools, so-called "smile check" data. One practitioner assesses the limits of these data as follows:

> Workshop evaluation forms are a necessity, but self-evaluation is not always reliable. Teachers indicated they were 'highly accomplished…but observations indicated otherwise. Teachers reported that they had

> acquired knowledge in the workshop but, upon returning to the classroom, forgot a lot of details and could not remember how to start.[1]

Similarly, a program evaluator summarizes the disadvantages of self-report data as follows: "Low response rate; no control over misunderstanding or misinterpretation of the questions; missing data, or inaccurate responses; not suited for people who have difficulty reading and writing; not appropriate for complex or exploratory issues."[2] Yet, despite their limitations, teacher self-report is still the most common measure of technology use. Realistically, it is likely to continue to be used.

This paper discusses a comprehensive, federally-funded effort to advance school-based program evaluation by refining measures of technology integration by teachers and use by students. Specifically, the Office of Technology and Information Systems of the West Virginia Department of Education was awarded a grant under the federal “Evaluating State Education Technology Programs” (ESETP) to study the Technology Model Schools (TMS) initiative; a program whereby technology integration specialists (TIS) were placed in schools on a full-time or part-time basis to work directly with teachers on technology integration. The ESETP request for proposals asked applicants "…to determine whether the program implemented produces meaningful effects on student achievement or teacher performance…" (*Competitive Preference Priority*, p 35127). The grant was awarded to West Virginia based on a proposal to search for student achievement outcomes associated with the TMS program, but to initially focus on

---

[1] Zelia Frick, Guilford County, NC., "Lessons Learned from Action Research: Evaluation from the Trenches", SEIR-TEC News Wire, ibid, p 10.
[2] Anna Li, "Thinking Beyond Surveys" SEIR-TEC News Wire, ibid., p 10.

the “teacher performance” aspect. That is, the initial goal was to concentrate on documenting the intermediate, teacher and instructionally related gains that are intrinsically worthwhile---especially technology integration.

Furthermore, the unique aspect of this particular ESETP grant was the implementation of novel research methods. Specifically, the originality of the proposal was in the effort to use technology to measure technology integration and use. This paper discusses those efforts. The next section contains an explanation of the research methods. Subsequent sections examine the successes of those efforts and the underlying data analysis issues, and the discussion section explores implications for educational research and policy.

II. METHODS

In addition to the typical pre-post, self-report questionnaires, and in order to triangulate the data, the two technology-based methods of documenting technology integration and use were the installation of metering software on classroom computers and random-interval web-based surveys triggered by the activation of pagers, a modified version of the Experience Sampling Method (ESM) (Hektner, Schmidt & Csikszentmihalyi, 2006; Csikszentmihalyi & Hunter, 2003).

A. The Modified Experience Sampling Method (ESM): Pagers + Web-Based Questionnaires

The Experience Sampling Method (ESM) was originally developed to “detect variations in emotional states over time” (Csikszentmihalyi & Hunter, 2003, p. 186). The design has since been expanded and used in educational research to study, among other areas, student stress (Verma, Sharma & Larson, 2002), teacher motivation and job satisfaction (Bishay, 1996), and student affective experiences of studying (Asakawa & Csikszentmihalyi, 1998). ESM involves asking participants to answer questions at multiple random moments over time whenever an electronic timing device (e.g. a pager, a watch, a personal digital assistant, etc.) prompts a response. The "unique advantage of ESM is its ability to capture daily life as it is directly perceived from one moment to the next, affording an opportunity to examine fluctuations in the stream of consciousness and the links between the external context and the contents of the mind" (Hektner, Schmidt & Csikszentmihalyi, 2006, p. 6). The questions can ask respondents about anything ranging from activities to feelings to their physical context.

For this study, the ESM was modified a bit. Typically, the ESM involves the electronic triggering of responses several times per day during each day of one normal week. That strategy is particularly advisable where attitudinal data are being collected, but in this study, the objective of the ESM was to document behaviors or activities (i.e. teacher and student use of computers) over the course of the majority of the school year. Therefore, teachers were equipped with pagers that were activated once or twice a day at random times every other week. When the pagers were activated, teachers were to complete a short web-based questionnaire at their earliest convenience. Additionally, one randomly selected student was to complete a different, even shorter web-based survey.

Setting up the paging process required attention to a number of details, including selecting a paging company, getting the pagers into the teachers' hands, "training" the teachers on how to use the pagers and respond to the survey, and scheduling the pages. One of the early issues to overcome was how to activate 100+ pagers without having to actually dial over 100 different phone numbers at random times throughout a given day. Fortunately, the pager service provider maintained a web-based paging "scheduler." Furthermore, each of the pagers was programmed into eight different group numbers. Thus, only eight different numbers needed to be programmed through the web-based scheduler.

One minor modification from the original research plan was made with respect to the pagers. The web-based scheduling program only allowed pages scheduled in advanced to be scheduled on the hour once per day. Therefore, while the hour on any given day was randomly selected, the pagers were only activated on the hour. Given that the teachers were paged many times over the course of a school year, this minor modification should be noted only as a potential minor limitation of this study.

Getting the pagers to the teachers and getting them up to speed presented another set of complications. For the treatment schools, the liaison to the teachers was the Technology Integration Specialist (TIS). In those schools, pagers and pager manuals were sent to the TIS with a set of directions on how to instruct the teachers to use the pagers and respond to the survey. Those directions came as part of an "evaluation kit"[3]. Ultimately, the TISs were extremely helpful in getting teachers up and running in those schools.

In the control schools, a few different approaches were taken. The first step was to call the principals of the schools to ask for cooperation. If the principals were agreeable, and had a "technology contact" in the school, the evaluation kit was sent to the school directly. In some cases, the principal was agreeable but did not offer the technology contact. So, those schools were visited by a member of the research team to work with the teachers directly.

The web-based questionnaires were designed and tested with reasonably parallel items for teachers and students. Specifically, using SNAP Survey software and the hosting capabilities of a cooperating university, two questionnaires (teacher and student) were developed for use in the study. The questionnaires were comprised of 3-5 questions, including straightforward inquiries about whether teachers and/or students were using computers at the time the pagers were activated.

Determining the frequency of paging was partly empirical and partly political---what would teachers accept. It was critical not to alienate the teacher participants since we needed their cooperation over two years (and that cooperation was voluntary and not compensated). Utlimately, the mutually accepted plan was 5 to 7 randomly determined pages during alternate weeks.

[3] Also included in the evaluation kit were instructions on how to create a desktop icon/link to the web-survey, the TrueActive monitoring software with installation instructions, and a CD-ROM with a multimedia presentation describing the study and the research methods.

B. Documenting File Activity with Metering Software

In addition to the ESM, this study involved a less well-established means of using technology to measure technology use by teachers and students. Specifically, file activity on individual computer workstations was documented through metering software. Several such commercially-available programs existed at the time of the program and were typically used by organizations wishing to know what was occurring on employee workstations. After a review of a number of available options, TrueActive Monitor (TAM) version 5.0 was selected.

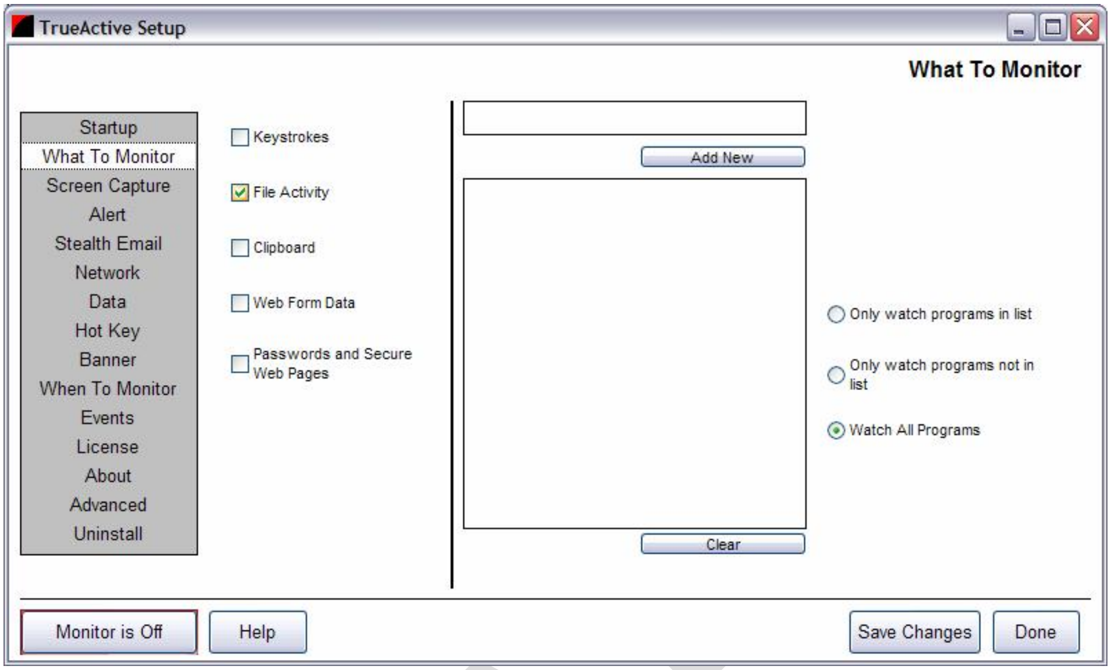
TrueActive Software, Inc. (formerly known as WinWhatWere Corp.) was founded in 1991 as a custom software development company. Originally, the company wanted a program they could use to assist in debugging their clients' software. So, they developed a computer monitoring tool with the ability to record and view every activity on a Windows-based computer. This in-house monitoring tool was so useful that it was released to the public in 1993 under the name "WinWhatWhere." Ultimately, with the help of their customers, WinWhatWhere evolved into a full-featured, award-winning computer monitoring solution called TrueActive.

TrueActive Monitor (TAM) version 5.0 had the capacity to capture an extraordinary amount of finely-grained data about the functioning of any computer workstation. For example, TAM could be configured to capture every keystroke made by any user of a device. Also, TAM could capture screenshots at regular intervals and store those image files for later viewing. While those sorts of data were beyond what was necessary for this particular study, TAM is customizable and could be configured to meet a number of different monitoring needs. Figure 1 is an image of the TAM setup screen. For this study, only file activity was recorded.

Data on file activity were automatically forwarded to the researchers through the Stealth E-mail function of TAM. Each installation was customized to have the workstation forward the data as an attachment to an e-mail every seven days. Those data files were sent to an e-mail address set up exclusively for this study.

Getting TAM installed on hundreds of workstations required a great deal of teamwork and cooperation across the state of West Virginia. As with the pagers and associated web-based surveys, the technology integration specialists were very helpful in installing TAM in the schools in which they worked. Each TIS was given a comprehensive installation guide with the exact specifications of how the software was to be configured. For the control schools, WVDE staff and members of the research team either contacted the district to request assistance or they personally installed the program during visits to the school.

Figure 1. TrueActive Setup Screen



III. THE PILOT STUDY

Both methods were pilot tested with two teachers in a single school in April 2004 before full-scale launch in academic years 2004-05. Members of the research team visited the school and met with the technology integration specialist, the principal and the teachers that agreed to participate in the pilot studies.

One significant deviation from the original research plan was initiated after meeting with the pilot school participants at the beginning of the pilot study. The utility of the data captured by TAM was directly related to the ability to link those data with a particular user (student and/or teacher). TAM 5.0 automatically captures the "name" of the workstation and the "user" of the workstation. The workstation ID would have been useful information, as would have been the "user." However, the "user" information was relatively static as the assumption built in to the system was that a workstation was used by a single user. In schools and classrooms, though, that is rarely the case. Furthermore, as was the case in the pilot school and like many other schools across the state, every workstation had the same user ID---"cl" which stands for "Compass Learning"---that must be used in order to access that software from the server which hosts the Compass Learning system.

Therefore, a problem arose for workstations in computer labs. Where even only a certain number of teachers and their associated students in a given school were participants in the study, without unique user IDs, it was nearly impossible to be able to isolate the data captured by TAM and attribute them to the students and/or teachers in the study. It was not feasible to get student ID numbers and especially not every time a student logged onto a computer and then logged off. Therefore, a decision was made to monitor only classroom computers and not workstations located in computer labs.

Another issue that arose around TAM was confidentiality. There were good reasons for educators to be squeamish and skeptical about monitoring software, but we were clear at all stages that TAM 5.0 was being used ONLY for research purposes. For any computer on which TAM was installed, there was the option of displaying a splash advisory message when the computer was turned on. Thus, in the pilot school, on all of the monitored computers, the following customized splash advisory was added:

> This computer has a meter that records the applications that are accessed. The data are collected as part of study funded by the US Department of Education and conducted by the West Virginia Department of Education. All data are confidential; no school or individual names will be reported in the research. Anyone using this computer expressly consents to such monitoring.

One concern here was that the advisory message only appeared when the workstation was first turned on. So, a teacher or student who arrived at a computer that had been turned on earlier in the day would not see the message and may not have been aware of the presence of the monitoring software. It is relevant, though, that all computers on the state network were already essentially being monitored by the West Virginia Department of Education. Also, there were "Acceptable Use" policies at the State, county and school levels.

At our initial meeting with the teachers in the pilot school, they agreed wear the pagers, to respond to the Web-based survey accordingly, and to use their roll books to select students to complete the surveys with the understanding that they should progress down the list with no repeats until the end. Also, we placed shortcut icons linked to the Web-based surveys on the desktops of each of the classroom workstations and two of the computers in the computer labs. This provided convenient, direct access to the survey without anyone having to remember a URL.

Another issue that surfaced during our initial visit related to teacher identification numbers. The Web-survey asked teachers for an ID number and students for their teachers' ID number. Since there was no common teacher ID number provided by the state of which teachers were readily cognizant, we agreed to use the last four digits of the teachers' Social Security numbers.

One of the concerns raised by the teachers at the initial meeting was that they might be in the middle of a whole-group lesson when paged and that during such an occasion, it would be difficult to stop everything and get to a computer. So, the technology integration specialist agreed to prepare paper versions of the web surveys (with larger print for the students). The teachers could then use the paper versions to respond as nearly as possible to the pager requests and would post the results to the web site as soon thereafter as practical.

Also, the teachers agreed that when they were absent, they would need to (a) check the pager for missed messages and times and (b) complete the web survey as "not available". The substitute would not be expected to respond to any queries.

The TAM monitoring and paging began a few weeks after the initial visit to the pilot school. Table one lists the dates and times of each of the 24 times the pagers were activated during the pilot study. Additionally, the table indicates the instances when teachers actually completed the Web-based survey (denoted as an "X" in the table). There was an initial problem with teacher two's pager which was solved after the first week.

Table 1. Pilot Study Paging Schedule

Date	Time	Teacher 1		Teacher 2	
		Teacher	**Student**	**Teacher**	**Student**
May 3	12:00	X	X		
May 4	9:00	X	X		
	10:00				
May 5	2:00				
May 6	11:00				
May 7	11:00				
	2:00				
May 8					
May 9					
May 10	8:00	X			
May 11	1:00	**SCHOOL**	**CLOSED**	**ELECTION**	**DAY**
May 12	1:00	X	X	X	X
May 13	1:00	X	X	X	X
May 14	10:00				
May 15					
May 16					
May 17	9:00	X			
May 18	NONE				
May 19	9:00				
	10:00	X			
May 20	2:00				
May 21	8:00				
	10:00				
May 22					
May 23					
May 24	10:00				
	12:00	X			
May 25	NONE				
May 26	8:00				
May 27	10:00				
	1:00				
May 28	10:00				
	1:00				

At the conclusion of the pilot testing period, the two teachers were interviewed by phone to debrief about their experiences with the pagers and the Web-based surveys. The teachers were quite clear that whereas they were very skeptical and hesitant in the beginning, they found the whole experience with the pagers rather unobtrusive and simple. In fact, one unintended dynamic was that the students in their classes began to anticipate and look forward to the pagers being activated; the students were eager to get their turn at the computer.

Certainly, one reason for the teachers' report of "unobtrusiveness" is that they only actually completed a survey less than half the times they were paged. The single explanation for that low "response rate" (as demonstrated in the table above) is simply that the teachers never received the pages; had they done so, they would have responded.

The only explanation reached as to why teachers did not receive our pages is related to signal detection issues within the school. During our initial visit to the school, it became clear that pagers (and cell phones for that matter) could not obtain a signal in certain parts of the school; in particular, the computer lab (which is located in a subterranean part of the school building) seemed off limits to satellite communications. The teachers' classrooms were reasonably good communication points, but there were certainly parts of the school building where pagers could not be reached. Furthermore, if an attempt was made to reach a pager outside of the range of satellite communications, it is not the case that the pager would then be activated as soon as it obtained a signal; the attempt at communication was simply lost.

The pilot study was deemed a reasonable success; data from TAM arrived as expected and responses from the pager-associated Web-surveys proved reliable. As a result of the pilot study, though, a couple of methodological decisions were made for the larger study moving forward. First, as fears over anonymity waned and concern over using any part of the Social Security number increased, the Web-based surveys associated with the ESM were modified to ask for the teachers' last name and the first initial of their first names. Second, the scenario for handling teacher absences become too cumbersome for teachers. So, moving forward, teachers were told to ignore the pages or even to turn off the pages if they were out of school on a normally scheduled school day.

III. RESULTS: HOW WELL DID THE METHODS WORK?

The data generated by the metering software and the web-based surveys were extensive and full of rich detail on technology use by teachers and students in West Virginia. However, the focus of this paper is on the methodology itself. Thus, what follows is a summary of the effectiveness of the methods.

A. The Modified Experience Sampling Method (ESM): Pagers + Web-Based Questionnaires

This part of the research design proved most successful, and surprisingly so. Originally expected technical hurdles and teacher reluctance both proved manageable. The state’s mountainous topography was expected to defeat even the strongest signals with the closest repeaters. However, signals appear to have reached all of the schools involved. It is possible that the reason teachers did not respond to every page is that individual attempts to activate the pager were stifled by a lost signal on a given day at a given time. But, as a whole, in each of the two years of the study, at least one teacher from each of the schools involved received a page and responded at least once.

Additionally, while there was resistance from individuals about the pagers, for the most part teachers were cooperative and took their role seriously. Furthermore, teachers reported that the students loved when the pager was activated. The students listened for the ring or vibration and enjoyed completing the web-based survey.

Year One

In the first year of the study, 109 teachers (61 teachers in treatment schools and 48 teachers in control schools) were equipped with pagers. From November 15, 2004 through April 15, 2005, the pagers were activated 41 times. In the end, 1,598 usable[4] pager-triggered web survey responses were received from teachers. Those 1,598 responses came from 109 different teachers (i.e. every teacher responded at least once). The range in the number of responses from individual teachers was one to 41 (i.e. a couple of teachers responded once, and one teacher responded to all 41 pages).

Students sent 1,311[5] usable pager-triggered web survey responses. Those 1,311 responses came from 102 different classrooms. The range in the number of responses from individual classrooms was one to 34 (i.e. a couple of teachers only once asked a student to respond, and one teacher had students respond 34 different times).

---

[4] By "usable", we mean a few things: first, the response came on a day we were expecting a response; some teachers responded on a day during a week when no pages were scheduled. Second, the response came at a time close to when it was expected. If a response came well before 9 a.m. or well after 2:00 p.m., we did not use the response as we only scheduled the pagers to go off between 9 and 2. Finally, there were some instances when the same teacher submitted two or more responses nearly simultaneously with the exact same responses. In those instances, we assumed that the teacher clicked "submit" at the end of the survey more than once. We discarded all but one response in those cases.

[5] There are a couple of reasons why this number is smaller than the number of teacher responses. First, because of a necessary change in the Web-survey for the students, the student web-survey data start with December 1, 2004 (compared to November 15, 2004 for teachers). Second, seven teachers never had any students respond to the web-survey.

Year Two

Since the major focus of the second year of the study was “continuing effects,” this form of data collection was limited to certain teachers. Also, teachers in three such schools refused to participate in this form of data collection for a second year. Thus, significantly fewer teachers were equipped with pagers in the second year of the study than in the first year. Fifty-one teachers were paged 40 times from December 2005 through March 2006. Over the course of that time, 33 different teachers sent 319 usable random-interval, pager-triggered web survey responses. The range in the number of responses from individual teachers was one to 32.

In addition to teacher surveys, 297 usable pager-triggered web survey responses arrived from students. Those 297 responses came from 33 different classrooms. The range in the number of responses from individual classrooms was one to 31 (i.e. a few teachers only once asked a student to respond, and one teacher had students respond 31 different times).

B. Documenting File Activity with Metering Software

Throughout the study, it was necessary to troubleshoot the use of data on computer file activity captured automatically and unobtrusively by TAM. As one example, the TAM developers had to revise the TAM 5.0 system to run on Windows 98 computers; the functionality of TAM 5.0 could never be supported by Windows 95. Classrooms where Windows 95 computers were being used were encouraged to upgrade the operating system if possible.

Additionally, the stealth e-mail function which sends the data automatically to the researchers had to be reprogrammed to allow the data to be sent every seven days. As originally developed, the longest possible time span between e-mails was 60 hours. While the software was reprogrammed to allow for the data to be sent every seven days, upon initial implementation, many of the computers were sending data every seven minutes. This was slowing down the workstations and bombarding the researchers with data.

Eventually, the data from the metering software arrived, and in bunches. Between October 1, 2004 and April 1, 2005, TAM was installed on approximately 287 computers. Over the course of the two years of the study, data was ultimately received at one point or another from 184 computers in 25 schools. The difference between 287 and 184 arose for a number of reasons, including the fact that some classroom computers were never turned on at all.

Also, the "seven minute" problem was never entirely resolved. That is, despite multiple maintenance attempts, some computers remained mis-configured and continued to send data every seven minutes over an extended period of time. Thus, in each of the two years of the study, almost 35,000 e-mails with a data file attached were forwarded to the researchers.

IV. DATA ANALYSIS ISSUES

Given the relative success of both methods of data collection, the researchers were awash in data. Thus, data analysis required extraordinary amounts of preparation in the form of data management. Data files were cleaned, merged, and aggregated where necessary. This section of this paper does not present a comprehensive set of findings; rather it is an exploration of the issues that arose around data analysis.

A. The Modified Experience Sampling Method (ESM): Pagers + Web-Based Questionnaires

Quantitative analyses of data produced through the ESM require attention to the sample and/or the unit of analysis. “At one level (what some have called the ‘beep’ or ‘response-level’), the sample is comprised of a collection of moments in time in the lives of several individuals” (Hektner, Schmidt & Csikszentmihalyi, 2006, p. 84). In fact, the responses come in at this level and, therefore, the initial databases are constructed such that each “case” or row in the dataset(s) represents one individual’s response to a single page. Additionally, those response-level data can be aggregated into a person- or respondent-level dataset. The fields or variables in the aggregated dataset can represent means or counts of appropriately-coded response-level data. Typically, in reporting the results of an ESM-based study, analyses of both datasets are conducted. “Whether the research question is one about situations or about persons, analyses appropriate to the question can be conducted at the response level or person level” (Hektner, Schmidt & Csikszentmihalyi, 2006, p. 89).

. Another key decision point for the analysis of ESM-generated data is how to handle low frequency responders. Whereas the majority of the teachers were as cooperative as we could have hoped, there were certainly some who were barely helpful. We could not know if the infrequent were different from the frequent responders, but because our interest was in documenting technology use across the days and weeks of the work life of teachers, we excluded the less cooperative group whose technology use was *ipso facto* poorly documented. More specifically, for any analyses, we eliminated responses from teachers who responded less than five times. In the first year, this eliminated 35 responses from 18 different teachers (bringing the totals to 1,563 responses from 91 teachers). In year two, this strategy eliminated 24 responses from 13 different teachers (lowering the total to 295 responses from 13 teachers).

The initial analyses of the ESM-generated data serve as good examples of the issues discussed above. To repeat, many of the research questions for the large-scale evaluation for which the ESM was being used focused on teacher and student use of computers. Therefore, when the pagers were activated and when teachers turned to the web survey on their desktops, the first question they saw was: "At the date and time your pager was most recently activated, were ANY OF YOUR STUDENTS using a computer?" A subsequent question was similar in nature but asked the teachers about their own use of computers.

The data from these questions were coded as "0=no" and "1=yes." Thus, for analyses of those questions, data were aggregated to the teacher level (the "means" function was used to aggregate the yes/no questions) and then weighted by the number of times any particular teacher responded. That is, using the first year as an example, the data were aggregated from a dataset containing 1,563 cases (i.e. responses) to one containing 91 cases (i.e.

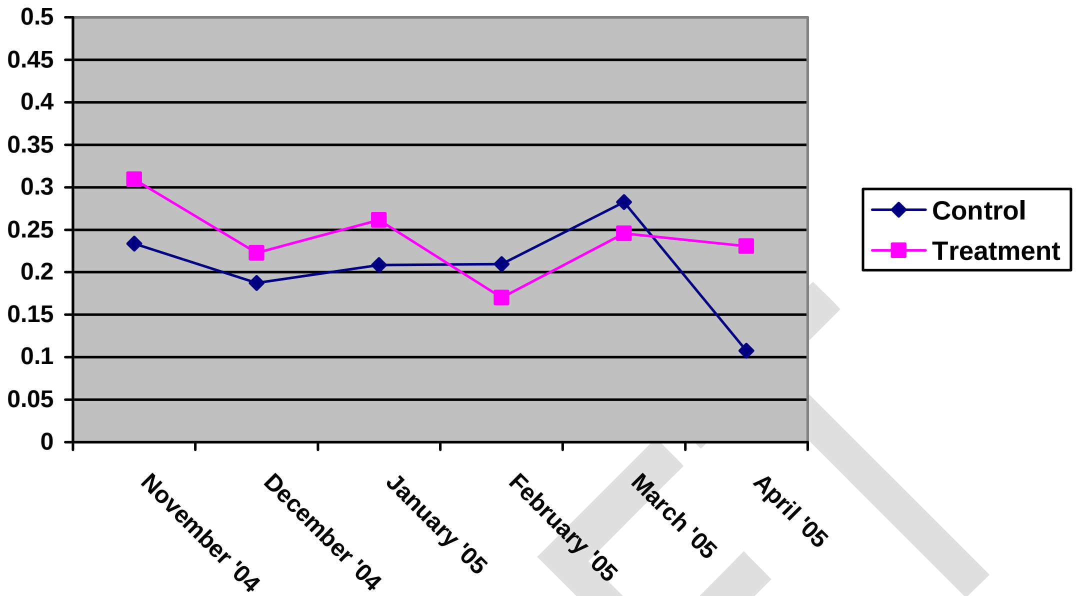
teachers). Also, giving the more cooperative and diligent teacher responses greater weight[6]---for both treatment and control groups--- assumed that data from the most diligent teachers are also the most reliable and increased our ability to discern patterns in teacher work. Those processes yielded a finding that teachers in the treatment group reported concurrent student computer use 23% of the times they responded to the electronic signal as compared to 21% of the time for the control group teachers[7].

ESM-generated data are also typically reported chronologically. That is, the data allow for an examination of changes in the “measured” feelings or activities over time. Toward that end, when the data were aggregated to the teacher-level, the variable for the month during which the survey was taken was aggregated using the “maximum” function. Then, the data file was split according to the “month” variable and overall means were computed. The resulting analyses, again using data from the first year, yielded the graph in figure two. The graph paints a picture of patterns in teacher-reported student computer use.

[6] The assumption also introduces some threats to the validity of the analysis. All the sampled teachers knew this was a study of technology and the non-responders may have obscured their non-use by failing to cooperate. Second, pagers and web surveys, while common, may have discouraged participation from techno-phobic teachers. Weighting by response, along with excluding the infrequent responders might result in our overestimating technology use.

[7] Those differences were statistically significant (p<.001) and in year two, the treatment group teachers responded affirmatively 21.3% of the time compared to 14.7% of the time for control group teachers; again a statistically significant difference (p<.001).

Figure 2. Representative graph of longitudinal analyses



The data were rich enough that more sophisticated multivariate analyses could be computed, but the research questions guiding this particular study were answered with the sorts of simpler inferential statistical analysis summarized above.

B. Documenting File Activity with Metering Software

Working with and making sense of the data generated by TAM were more complicated procedures. Those difficulties resulted mostly from the sheer number of data files that needed to be detached from e-mails, imported by the TAM software, and exported into a spreadsheet for cleaning and analysis. In fact, there were more cases or observations (i.e. rows) than could fit into a single Microsoft EXCEL worksheet. The data ultimately fit into a single EXCEL workbook with nine worksheets of approximately 65,000 rows and a 10th worksheet with over 20,000 rows. By the end of the first year, we had a dataset with 564,457 observations (file activations). Those data came from 207 different computers in 26 different schools. The range of file activity on any single computer was from one file activated[8] to 13,507 files. The 136 megabyte file served as the raw data.

Figure three is a screenshot of those raw data. The first column contains the name of the workstation, while column “B” contains the username[9]. Column “E” contains the data on the elapsed time for which a file was active, and Column “H” contains the name of the files that were activated.

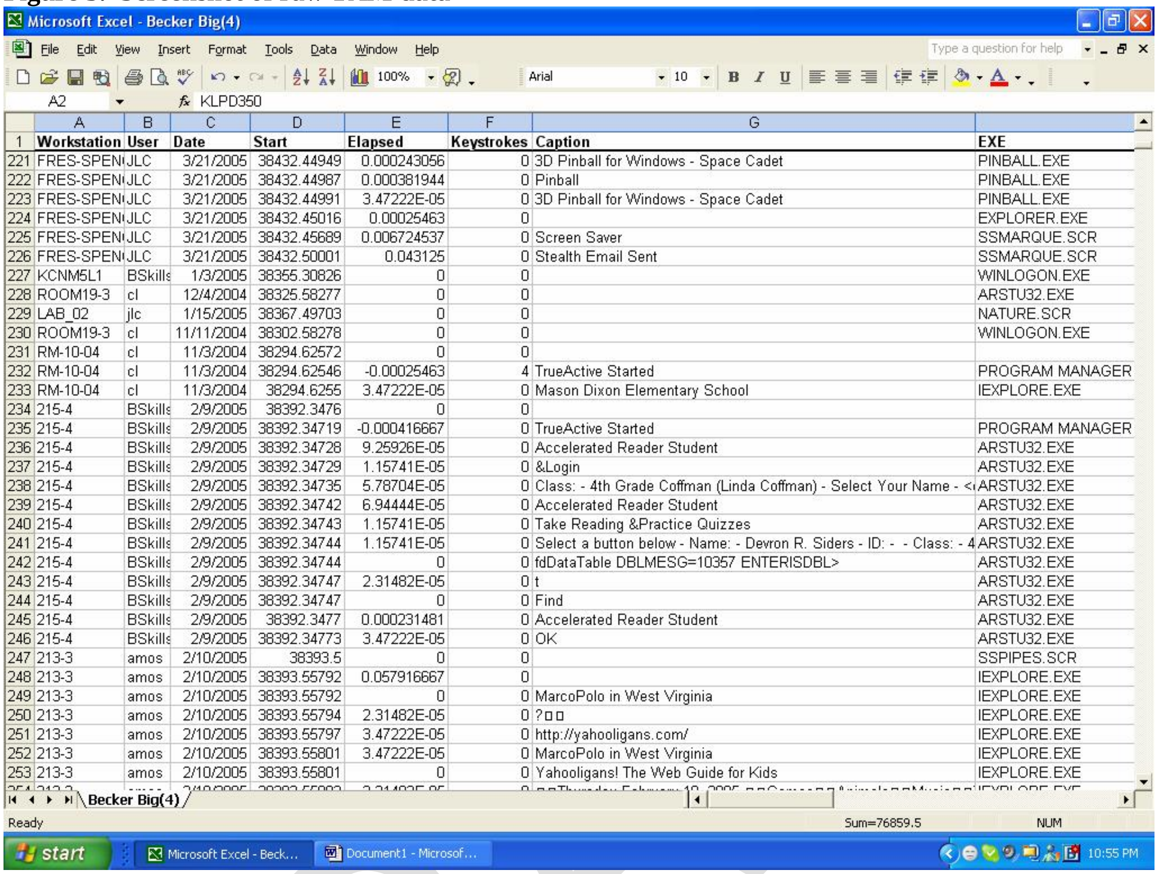
Preparation of the dataset for analysis included coding the computer names for schools, classroom teacher and study condition (treatment/control). This process was particularly time consuming as the research team had accumulated a list of computer names associated to teacher classrooms, but there were numerous data points that did not match

---

[8] There were definitely computers on which the TAM software was installed, but from which no data on file activity was ever forwarded to the researchers. It is very difficult to know exactly how many computers fit this description, but there were at least 93 computers about which we knew the name of the computer and its classroom location but from which data were never received. This does not necessarily mean that those 93 computers sat idle for the whole year; they may have, for instance, been mis-configured and thus never forwarded the data via the TAM stealth e-mail. Also, if the computer was running Windows 95 it was incompatible with TAM.

[9] Note the “cl” designation in, for example, rows 230 through 233. This generic username used on machines utilized for student use of Compass Learning software prevented the researchers from monitoring computers in labs.

Figure 3. Screenshot of raw TAM data



	Workstation	User	Date	Start	Elapsed	Keystrokes	Caption	EXE
221	FRES-SPEN	JLC	3/21/2005	38432.44949	0.000243056	0	3D Pinball for Windows - Space Cadet	PINBALL.EXE
222	FRES-SPEN	JLC	3/21/2005	38432.44987	0.000381944	0	Pinball	PINBALL.EXE
223	FRES-SPEN	JLC	3/21/2005	38432.44991	3.47222E-05	0	3D Pinball for Windows - Space Cadet	PINBALL.EXE
224	FRES-SPEN	JLC	3/21/2005	38432.45016	0.00025463	0		EXPLORER.EXE
225	FRES-SPEN	JLC	3/21/2005	38432.45689	0.006724537	0	Screen Saver	SSMARQUE.SCR
226	FRES-SPEN	JLC	3/21/2005	38432.50001	0.043125	0	Stealth Email Sent	SSMARQUE.SCR
227	KCNM5L1	BSkills	1/3/2005	38355.30826	0	0		WINLOGON.EXE
228	ROOM19-3	cl	12/4/2004	38325.58277	0	0		ARSTU32.EXE
229	LAB_02	jlc	1/15/2005	38367.49703	0	0		NATURE.SCR
230	ROOM19-3	cl	11/11/2004	38302.58278	0	0		WINLOGON.EXE
231	RM-10-04	cl	11/3/2004	38294.62572	0	0		
232	RM-10-04	cl	11/3/2004	38294.62546	-0.00025463	4	TrueActive Started	PROGRAM MANAGER
233	RM-10-04	cl	11/3/2004	38294.6255	3.47222E-05	0	Mason Dixon Elementary School	IEXPLORE.EXE
234	215-4	BSkills	2/9/2005	38392.3476	0	0		
235	215-4	BSkills	2/9/2005	38392.34719	-0.000416667	0	TrueActive Started	PROGRAM MANAGER
236	215-4	BSkills	2/9/2005	38392.34728	9.25926E-05	0	Accelerated Reader Student	ARSTU32.EXE
237	215-4	BSkills	2/9/2005	38392.34729	1.15741E-05	0	&Login	ARSTU32.EXE
238	215-4	BSkills	2/9/2005	38392.34735	5.78704E-05	0	Class: - 4th Grade Coffman (Linda Coffman) - Select Your Name - <	ARSTU32.EXE
239	215-4	BSkills	2/9/2005	38392.34742	6.94444E-05	0	Accelerated Reader Student	ARSTU32.EXE
240	215-4	BSkills	2/9/2005	38392.34743	1.15741E-05	0	Take Reading &Practice Quizzes	ARSTU32.EXE
241	215-4	BSkills	2/9/2005	38392.34744	1.15741E-05	0	Select a button below - Name: - Devron R. Siders - ID: - - Class: - 4	ARSTU32.EXE
242	215-4	BSkills	2/9/2005	38392.34744	0	0	fdDataTable DBLMESG=10357 ENTERISDBL>	ARSTU32.EXE
243	215-4	BSkills	2/9/2005	38392.34747	2.31482E-05	0	t	ARSTU32.EXE
244	215-4	BSkills	2/9/2005	38392.34747	0	0	Find	ARSTU32.EXE
245	215-4	BSkills	2/9/2005	38392.3477	0.000231481	0	Accelerated Reader Student	ARSTU32.EXE
246	215-4	BSkills	2/9/2005	38392.34773	3.47222E-05	0	OK	ARSTU32.EXE
247	213-3	amos	2/10/2005	38393.5	0	0		SSPIPES.SCR
248	213-3	amos	2/10/2005	38393.55792	0.057916667	0		IEXPLORE.EXE
249	213-3	amos	2/10/2005	38393.55792	0	0	MarcoPolo in West Virginia	IEXPLORE.EXE
250	213-3	amos	2/10/2005	38393.55794	2.31482E-05	0	?□□	IEXPLORE.EXE
251	213-3	amos	2/10/2005	38393.55797	3.47222E-05	0	http://yahooligans.com/	IEXPLORE.EXE
252	213-3	amos	2/10/2005	38393.55801	3.47222E-05	0	MarcoPolo in West Virginia	IEXPLORE.EXE
253	213-3	amos	2/10/2005	38393.55801	0	0	Yahooligans! The Web Guide for Kids	IEXPLORE.EXE

.

Additionally, the data needed to be aggregated in order to conduct meaningful analyses. First, we aggregated the data by application, i.e., we summed the total number of minutes the metered computers had accessed the applications. From the initial aggregation, it became clear that the vast majority of files activated by the computers were system files or other files from non-instructional applications (including screensavers). Those data were not particularly useful or relevant given the research questions guiding this study. The dataset used for analysis was created by aggregating by date, by computer, and by application. Since computers accessed program files multiple times on a given day, the number of minutes the files were active on a given computer on a given day was summed. Thus, in the final dataset, each observation could be thought of as "use of an application on a computer on a day." The dataset contained 19,746 such observations. Figure four depicts the structure of that dataset.

Figure 4. Structure of aggregated TAM dataset

Workstation	User	Date	Activity	Elapsed Time	Count
KA2RZV7	cl	4/28/2004	C:\Compass\LEXP.EXE	0:39:59	1
23MMXWD	cl	4/28/2004	C:\PROGRAM FILES\INTERNET EXPLORER\IEXPLORE.EXE	0:03:00	48
KA2RZX6	cl	4/28/2004	C:\PROGRAM FILES\RENAISSANCE LEARNING\MATHFACTS IN A FLASH\MATHFACT.EXE	0:00:13	6
23MMXWD	cl	4/28/2004	C:\PROGRAM FILES\TAM\TAMRPT.EXE	0:01:08	18
23MMXWD	cl	4/28/2004	C:\PROGRAM FILES\TAM\TAMSET.EXE	0:00:48	6
23MMXWD	cl	4/28/2004	C:\WINDOWS\EXPLORER.EXE	1:42:03	66
KA2RZV7	cl	4/28/2004	C:\WINDOWS\IBM.SCR	0:39:59	2
KA2RZV7	cl	4/28/2004	G:\CURRIC\WRKSTATN\WIN\START.EXE	0:00:01	1
23MMXWD	cl	4/28/2004	Hofstra University - Microsoft Internet Explorer	0:00:09	4
23MMXWD	cl	4/28/2004	http://people.hofstra.edu/faculty/Jonathan_D_Becker/	0:00:04	1
23MMXWD	cl	4/28/2004	http://people.hofstra.edu/faculty/Jonathan_D_Becker/WV_Surveys.html	0:00:08	7
23MMXWD	cl	4/28/2004	http://people.hofstra.edu/faculty/Jonathan_D_Becker/WV_Surveys.html - Microsoft Internet Explorer	0:00:00	1
23MMXWD	cl	4/28/2004	http://people.hofstra.edu/faculty/Jonathan_D_Becker/wvstudents1.htm	0:00:02	1
23MMXWD	cl	4/28/2004	http://people.hofstra.edu/faculty/Jonathan_D_Becker/wvstudents1.htm - Microsoft Internet Explorer	0:00:02	2
23MMXWD	cl	4/28/2004	http://people.hofstra.edu/faculty/pages_faculty/fac_alpha.html	0:00:08	1

Three other complications arose in analyzing these data. First, the TAM software was installed on computers throughout the fall and winter of 2004; some considerably later than others. And, since the variable of interest is elapsed time, data from the earliest months when not all computers were up and running with TAM were eliminated. In other words, only data from times when all possible file activation data was being sent could be used. There was reason to believe that the installation of TAM was complete across the state by the beginning of 2005. Thus, analyses were based on the months of January, February and March of 2005.

Outlier observations provided the second complication. For example, one observation indicated that on one day for one computer, “LEXP.exe” (the executable file for *Compass Learning*) was activated one time for a total of 6,874 minutes. Since there are only 1,440 minutes in a day, that observation was nonsensical. On the other end, there were hundreds of observations where, on one day for one computer, a file (e.g IEXPLORER.exe, the executable file for Internet Explorer) was activated multiple times for a total of 0.75 minutes. Since it was unlikely that anything productive happened on the Internet for less than one minute, those observations proved troubling as well. Thus, the only observations included were those where, on a given day on a given computer, a particular file was active for more than two minutes. Further, observations where the file was active for any number of minutes that exceeded the mean by more than two standard deviations (the standard for excluding “outliers”) were excluded.[10]

Finally, it is worth repeating and emphasizing that these data were ONLY from classroom-based computers. Each classroom under study had from two to six computers, but data from other sources indicated that much of the student

[10] For example, in analyzing data on use of Microsoft Word, the analyses included observations of days/times on a given computer that MS Word was accessed for at least 2 minutes and not more than 132 minutes (2 S.D. above the mean).

use of computers occurred in computer labs. In particular, we knew from our observations in the schools that when students used the state-provided, BS/CE supported learning solution (*Compass Learning* or *Riverdeep*), they were much more likely to do that as a whole group in a lab than in the classroom. Thus, the data were reported very specifically as representative of classroom-based computer use.

Early in the second year of the study, all legal rights to TAM were sold to another company. As a result, the support needed to render the TAM data useable disappeared; programmers and developers of TAM were critical links in the chain of data management. The company that acquired the rights to TAM never utilized or supported the technology. Thus, the data retrieved by TAM over the course of the second year of the study remained in their most raw form and were never analyzed.

The cumbersome nature of the data and legal issues notwithstanding, TAM data yielded a number of very interesting and important findings. For example, by far, the three most frequently activated (potentially) instructional applications were: Internet Explorer, Accelerated Reader and Microsoft Word. Considering just MS Word (the most frequently active program in classrooms across West Virginia), the data showed that (in just the treatment school classrooms, for example) 34 computers from 12 different schools accessed MS Word a total of 95 different instances/days for an average of 9.75 minutes per instance. Those sorts of finely grained measures of computer use in schools are surely unprecedented and represent a significant advance in educational technology research and evaluation.

V. DISCUSSION

The Larry Cubans (2001) and Hank Beckers (2000) of the world have debated the impact of technology on teaching and learning largely by either observing classrooms or asking teachers to self-report. Those works made important contributions to our understanding of technology use and integration, but too much time and money is earmarked for educational technology to rely on those limited modes of data collection. The research methods used in the study and reported in this paper advance our ability to comprehend and detail integration and use of technology in schools and classrooms across the country.

If the merits of these research methods are to be measured by the quantity of data collected, then they were an overwhelming (pun intended) success. In each year of the two-year study, hundreds of thousands of data points were obtained. More likely, though, the success of the methods lies in the ability of the data to yield findings directly relevant to the research questions. And, by that measure as well, the methods proved highly effective. The ESM-generated data yielded meaningful portraits of computer use over the course of the school year. The TAM data produced rich and finely grained findings about computer use in classrooms.

Each form of data collection has great promise and utility moving forward, particularly with appropriate modifications. To maximize reliability and cooperation, the ESM might be conducted in a more intense manner over a shorter period of time. For example, consistent with more typical uses of the ESM, teachers might be paged four or five times over the course of a day during each day of a typical week. That process might be repeated every 6 or 7 weeks. Also, it is worth thinking about an electronic signaling system that does not rely on satellite technology and that, therefore, might be more reliable. Popup messages on an Internet-connected computer or computing device might be worth considering.

Data collection with TAM proved tricky and at times burdensome, much more so than the modified ESM. However, in the long run, this mode of data collection holds the most promise. Particularly in instances where students and/or teachers log into workstations with unique identification numbers, metering programs generate truly distinctive and important data about computer use. The ability to know which programs are used and for how long on any given day greatly advances the “measurement” of amounts of computer use. Researchers would no longer have to rely on teachers retrospectively reporting broad categorical estimates of the frequency of student computer use.

Consider, for example, the case of ubiquitous or 1:1 computing programs. Since each laptop or computing device is essentially assigned to a teacher or student, placing metering software on those machines would allow researchers or evaluators to know exactly the programs for which the machine is being used most frequently. Furthermore, data from the software would allow researchers or evaluators to know nearly exact durations of use disaggregated by the various programs on the machine. Those data could very easily be assigned to the designated user of the machine.

Finally, while this paper is written from a research methods perspective, there is potentially a meaningful policy implication for the use of metering software. The arena of educational technology policy is particularly contested terrain. That is to say, in an arena in which resources (e.g. hardware, software, support, etc.) are particularly limited, how do policymakers reconcile significant issues of distributive justice (i.e. how do they decide who gets what and how do they decide that?) The modal approach to those matters is to spread the resources equally amongst everyone. Or, another common decision point is to try to get the resources to those who “need” it, where need is equated with socioeconomic disadvantage.

However, a likely more (cost-)effective approach would be to develop policies or programs that aim to get the limited resources into the hands of those who will best take advantage of them. In this case, "need" might be defined as "will do great things, but only if they get the resources." Furthermore, determinations of ability to take advantage of the resources, particularly the hardware and software, could be made by utilizing metering software. In other words, workstations could be metered/monitored to determine which machines are being used most frequently and in ways most consistent with the school and/or district curriculum.

In addition to giving new meaning to the concept of "acceptable use," this approach would allow limited resources to be directed to where they would be used in the most cost-effective ways. West Virginia's teachers and students use computers more than their counterparts in virtually every other state However, even there, the data showed plenty of computers where the files that were activated most frequently were those necessary to run the screen saver. On some computers, those were the only files that were ever activated. Clearly, those machines could have been used more efficiently/effectively somewhere else.

Thus, TAM, or other similar programs, could be implemented both to advance the cause of educational technology research and evaluation, but also to make important educational policy decisions.

REFERENCES

Asakawa, K. & Csikszentmihalyi, M. (1998). The quality of experience of Asian American adolescents in activities related to future goals. *Journal of Youth and Adolescence,* 27, 141-63.

Becker, H.J. (2000). Findings from the teaching, learning, and computing survey: Is Larry Cuban right? *Educational Policy Analysis Archives*, *8 (51).* Retrieved November 7, 2002, from http://epaa.asu/epaa/v8n51.

Bishay, A. (1996). Teacher motivation and job satisfaction: A study employing the Experience Sampling Method, *Journal of Undergraduate Science*, vol. 3, pp. 147-154.

Csikszentmihalyi, M. & Hunter, J. (2003). Happiness in everyday life: The uses of experience sampling. *Journal of Happiness Studies*, 4(2), 185-99.

Cuban, L. (2001). Oversold & Underused: Computers in the Classroom. Cambridge, MA: Harvard University Press.

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 March 16, 2007 from http://epaa.asu.edu/epaa/v12n48/.

Verma, S., Sharma, D., & Larson, R. (2002). School stress in India: Effects on time and daily emotions. *International Journal of Behavioural Development,* 26 (6), 500-508.

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