# Handhelds for Formative Assessment in the Elementary School

Kristin Redington Bennett, Ph.D.
Ann C. Cunningham, Ph.D.
Wake Forest University

## Background

Over the last 15 years one-to-one computing in the classroom has become a prominent topic for educators. The need to ensure all students are able to function efficiently and effectively in academic settings coupled with the underlying economic goals of global competitiveness drive the interest in providing one-to-one computer access in American schools. The advent of lightweight, portable, easy to use, and increasingly inexpensive laptops and handheld computing devices make one-to-one options more enticing. While the business world embraces devices that pack as much computing and telecommunication power as possible into a lightweight handheld machine, many educators opt for laptops to fulfill their one-to-one computing needs. There are excellent reasons to support both types of computing tools for one-to-one initiatives, and the device certainly must fit the needs of the user and the computing context. But the reality is that handheld computing devices are here to stay and are growing more flexible, more powerful, and less expensive with each product generation. The convergence of basic computing applications, internet, and cell phone communication may appear to open Pandora's Box for the educational community, but it is important to recognize and embrace the value of these new, powerful, and inexpensive tools for classroom use, now and in the future.

So far, the focus of one-to-one computing in the educational community has been predominantly on K12 administrator and student use. Although there is a clear need for these groups to use handheld computing tools, there is an obvious and pressing need for

classroom teachers to develop skills with handhelds as well. The most obvious reason is to develop their understanding of the potential for classroom instruction and to guide student use. However, the latest generation of handheld devices and software has the potential to support the tasks and needs of classroom teachers in ways that are not yet fully recognized. Many capabilities of newer handheld computing devices (often called Smart Phones) support teacher use. For example, increased communication via phone and email, scheduling capabilities with calendar features, web access, and simple document creation with word processing and spreadsheet applications all support teacher productivity and professional practice (NETS*T V). Survey development software tools specifically developed for a handheld operating system are also available, which is helpful as the demands of NCLB have made assessment an increasingly important task for teachers (NETS*T IV). Faced with the need to ensure that their students are prepared for standardized tests, it is now imperative that teachers possess the knowledge and skills necessary to implement appropriate assessment. And appropriate classroom assessment now tends to utilize formative measures with greater frequency, especially in the early grades and with learner groups at risk of not passing state-mandated standardized tests. What was once primarily viewed as an informal way to check what students know prior to beginning an instructional unit has now become a critical method for evaluating and documenting individual student progress toward achievement of academic standards. The most efficient way to capture, document, and analyze formative data on individual student performance over time requires technology, and handheld computing devices provide the most productive way to capture this data in the classroom without wasting full group instructional time. Handheld computing devices and appropriate software tools

have the potential to revolutionize classroom-based formative data collection and analysis.

**Literature Review**

The effectiveness of any instructional program often turns in part on the use of effective assessment materials and tools. While teachers strive to create meaningful learning situations involving the 'whole child' they often turn back to summative, standardized assessment strategies, which are a poor fit to instructional programs (Means, Penuel, & Quellmalz, 2001). Improving classroom assessment practices has proven to be challenging. Assessment, especially formative assessment, is given little attention in classrooms where teachers rarely have adequate time to plan or implement assessment activities in a methodical manner or learn new strategies for assessment from peers or experts (Black & William, 1998b; Darling-Hammond, Ancess, & Falk, 1995).

Student performance expectations established by state education agencies in response to No Child Left Behind legislation provide an impetus for a greater focus on formative assessment in the classroom. Assessment must be viewed as an integral part of instruction because it "not only provides information about learning, it drives learning" (Lewis, 2006, p. 29). Well-designed formative assessment instruments that capture performance data on individual students is a vital component of student success on formal benchmark assessments. Interim assessment is necessary to diagnose student needs for additional instruction (Lewis, 2006, p. 17). What is needed is a method that teachers can easily employ in the classroom to make formative assessment timely, "provide(s) quality feedback, strengthen(s) teaching and learning, and involve(s) students in accountability" (Lewis, 2006, p. 30). While experts in the field of psychometrics report

that this type of assessment practice is valuable, they also admit that it is not currently being implemented in K-12 classrooms. The Center for Research on Evaluation, Standards, and Student Testing (CRESST) co-director, Eva Herman states "These types of assessments are often nonexistent in curriculum materials, in teachers' repertoires, and not often found in benchmark tests. According to Lewis (2006), teachers do not have time to devise these types of assessments on their own (p. 29). A need exists to increase knowledge and understanding of formative assessment techniques in teacher preparation programs.

Handheld computers can facilitate teachers' implementation of continuous, authentic, and reflective assessment in the classroom. Because of their portability, handheld computers ease the data gathering process because of their unobtrusiveness, and ease the aggregation of data because of their technological features. Many researchers have heralded the myriad potential uses of handhelds in field investigations (Hsi, 2000; Soloway et al., 1999). Ongoing, in the moment, formative assessment made easier by the portability of handhelds means that assessment can occur in any location at any phase of the learning process. Further, their affordability means that they are feasible option for widespread use in our schools.

**Theoretical Framework**

This research project is framed within the contexts of situated learning and action research. Knowing that this project was multi-faceted in terms of technology training, pedagogical content, and logistics it was obvious that the framework for this research must implicitly and explicitly reflect the complexity. Herrington and Oliver (1999) compiled a list of "nine characteristics of a situated learning framework, namely: an

authentic context; complex authentic activities; multiple perspectives; expert performances; coaching and scaffolding; opportunities for collaboration, reflection, and articulation, and authentic assessment" (p. 402).

The framework provided by that of situated learned complements the methodological framework of action research. Cohen and Manion (1989) suggest that action research can be difficult to define; however, they also indicate that this method in the research literature has tangible features:

> *Action research is situational – it is concerned with diagnosing a problem in a specific context and attempting to solve it in that context. It is usually collaborative – teams of researchers and practitioners work together on a project; it is participatory – team members themselves take part directly or indirectly in implementing the research, and it is self-evaluative – modifications are continuously evaluated within the ongoing situation (p.217)*

Action research is often represented as a four-phase cyclical process where a plan is formulated, that plan is acted upon and the outcomes observed, and reflection is undertaken to understand the processes, strengths, and weaknesses of the effort. The conscious analysis or reflection stage of action research distinguishes it from the casual 'plan, act, sense and re-plan' cycle by which we operate in everyday life (Tripp, 1990).

Recent studies have asserted that undergoing action research as part of their education programs can benefit undergraduate preservice teachers (Price, 2001; Rock & Levin, 2002). With the benefits of action research being demonstrated in these studies and other studies (Burnaford & Hobson, 1995; Johnson & Button, 2000; Sax & Fisher, 2001) of graduate-level teacher education students, Fueyo & Koorland (1997) have called

for action research to be included as part of preservice education programs. Using action research strengthens professional development opportunities and has strong relevance for preservice and inservice teachers, especially as related to widely-accepted models of teaching and learning in technology education.

This paper will describe the preliminary findings from the first phase of an ongoing research effort that explores how handheld computers can support teachers in the elementary grades with the development and utilization of formative assessment instruments to inform and improve instruction.

**Research Methods**

A two-year pilot study to test the value of handheld computing devices and specialized survey software for the collection of formative assessment data in elementary classrooms was implemented with two groups of pre-service teachers at a small, private university in the southeast United States. Recognizing that changes in teachers' formative assessment habits need to start in pre-service programs, the researchers sought to explore the following questions:

1. Are the hardware and software tools appropriate for formative assessment data collection?
2. Do teacher candidates recognize the value of these tools for the classroom?

The participants in both the design process and the field trial were all teacher candidates in the final semester of the elementary education program at a small, private university in the southeastern United States. As the pilot study occurred over two years, there were two groups of participants with similar backgrounds and experiences. These participants were a convenience sample (Group 1 n=17, Group 2 n=15) who were all

enrolled in a required course “Teaching Children with Special Needs”, the final class in the elementary education course sequence. This course includes a field experience which required the participants to be in a local elementary school every week. All participants successfully completed their student teaching experience in the prior semester, and the participants all had extensive technology experience in part due to the ubiquitous resources of a one-to-one laptop university and through their teacher education coursework which includes a required technology course. Further, several of the participants had prior experience using handheld computing devices and the survey software as part of another unrelated pilot study.

Both groups of participants were issued their own iPAQ handheld devices with associated accessories, and one portable keyboard. They received training on the handheld through a collaborative training session designed by members of the Information Systems Department and faculty members from the Department of Education. The first training session included an introduction to the handhelds and their general features, as well as features specifically designed for students at the university. During year one of the pilot, the participants were given time to experience (play with) the handhelds over a period of two weeks before being introduced to the field data collection software, *DataInHand* ™ in another training session. A change made for the second year of the pilot included the early distribution and training on the handhelds at the end of the fall semester in order to increase the time for participants to familiarize themselves with the hardware. Additionally, group two had two ‘working sessions’ scheduled during the semester. These sessions were designed to work on hardware and software barriers, redesign of surveys, and group reflection.

The researchers provided training on the survey creation software in a full group session attended by all participants. *DataInHand* ™ is software developed by the Research and Development team in the Information Systems division at the university. The features of this software include a simple interface that permits survey development on the computer (Likert scale, yes/no, short response questions) and simple transfer and retrieval of completed surveys to and from the handheld device. Results can be reported in HTML, Microsoft ® Word, and Microsoft ® Excel formats. More information on this free program is available at http://datainhand.wfu.edu.

***Survey Design***. In order to engage the participants with the inquiry process, survey design was both an individual and collaborative endeavor. During the first year of the pilot study, participants selected one student identified with special needs from their student teaching experience to work with one-on-one for two hours each week. Participants were asked to work with school personnel to establish growth goals for the chosen student and were guided to address 'the whole child' in their goals rather than establishing purely academic instructional goals. Once the educational context was established, the participants were asked to design a survey instrument to assess the effectiveness of their interactions with the student they selected as well as any changes they noted during their sessions. As a group, the participants decided that they would create a survey that was consistent across the group and allowed individuals to add additional questions specific to his/her child's special needs. Once the survey questions were determined, all participants entered it into the *DataInHand* ™ program while in the university classroom. Participants were then asked to perform several trial runs of the survey with fictional information to practice the process of downloading, uploading, and

reporting. It was agreed that the participants would use this survey in the field for two weeks, and then come together to modify either the survey or the process. The researchers involved with this project as well as support personnel from the Information Systems division were made available to answer questions about hardware or software during this time.

After the two week trial period, the survey was modified to include fewer open-ended questions. Further, when the participants shared their individually created 'additional questions' several participants modified their survey to include questions particularly relevant to their individual tutee. The participants then utilized this modified survey for the remainder of the semester, repeating the administration of the survey and collection via handheld computer on a weekly basis.

Several aspects of the survey content, creation, and data collection were modified for the second year of the pilot study. One important change included the placement of all participants in a new elementary school partnership. None of the participants completed their student teaching in this school, and the classroom setting for this field placement was 100%. The change in setting, as well as analysis of the data from year one led to several changes in the survey content, creation, and data collection. Participants were given a survey adapted from the year one study to include fewer open ended questions and more Likert-type items. The survey also included more questions on the affect of the both the tutor and the tutee during the tutoring sessions. Participants were also given the option to include questions that specifically related to their tutee. The researchers suggested the possibility of including North Carolina Standard Course of Study objectives as assessment items, but left the final decision up to the participants.

***Data Sources*** - This pilot project collected a wealth of data from the participants and their students. Each participant completed a pre and post survey consisting of Likert-Type and open-ended response questions. The pre survey asked about computer use and habits, cell phone use, and handheld use/experience. The post survey asked participants to respond to questions about their use of the handheld device and the software tools during the pilot study as well as indicate their attitudes toward the value of the tools for classroom use. The survey is a living document that represents valuable data in its changes and adaptations by the groups and individuals. The data garnered from the weekly administration of the assessment provides an enormous amount of information about the weekly successes, setbacks, and other information about both the participants and their students. Finally, as part of the final project in the course, students were asked to use the data downloaded from their surveys to reflect on their experience with the students they worked with and how the use of the handheld assessment informed their weekly interactions. In year two, the final project of the course also asked participants to reflect on the design process and the actual implementation of that design. They were also asked to discuss the positive and negative aspects of using the handheld for documenting and assessing progress.

**Findings**

The data collected from the first year of the pilot study lead to small, but significant, changes in the second year of the pilot study. Preliminary analysis of the data as well as anecdotal evidence identified several significant findings that affected the design of the second pilot study and included:

- Hardware problems often made the technology burdensome rather than helpful
- Surveys need to be short and questions must permit input via the stylus rather than the keyboard
- Although open ended questions yield rich data, they make the process taxing.
- Knowing the right questions to ask is KEY to data driven instructional change
- It takes several weeks of data to be able to draw significant conclusions
- Having to do assessments each time forces teachers to focus on the work product and process, and helped teachers develop a greater understanding of what students knew and could do.

This evidence resulted in earlier distribution of hardware to the second group of participants with the idea that more familiarity with the tools would reduce the frequency of hardware problems. The feedback on challenges associated with the survey and the new placements in EC classrooms resulted in revising the surveys to reflect state performance standards and the state assessment scale which removed the need to enter data with a keyboard.

The first research question, “Are the hardware and software tools appropriate for formative assessment data collection?” was addressed with seven questions that required the participants to rate their response on a scale from 1-4 (1=strongly disagree, 2=disagree, 3=agree, 4=strongly agree). The Appendix provides the questions posed to the participants upon conclusion of each pilot study.

Responses from participants in both groups were analyzed in terms of means and modes in order to show more complete data for each small group. Participants in the 2006 pilot study agree more often than participants in the 2007 pilot study that the handheld was a useful tool for collecting data in the field, although the mean for both groups is skewed more toward disagreement. In response to question 2, participants in the 2006 group agree more often that *DatainHand* software made it easier to capture data in the field than respondents in the 2007 group, but once again, the mean for both groups is skewed more toward disagreement. In response to question 3, the means of both groups is skewed more toward disagreement, but the mode for the 2006 group indicated that more respondents agreed that the software and hardware facilitated completion of course requirements. Although the means for both groups' responses skew more toward disagreement to question 4, the mean for the 2007 group is higher and the mode indicates that this group agreed that the tools facilitated processing field observations in a timely manner. In response to question 5, the majority of respondents in both groups agreed that when used consistently, the tools facilitated data analysis. However, their responses to question 6 indicate that both groups did not agree that using the tools to collect data in the field was worthwhile. Both groups were consistent in their responses to question 7 revealing that their understanding of classroom technology use was enhanced by this experience, and these responses had the highest means for both groups. Figure 1 shows the comparison of the means for the responses from both the first pilot group (2006) and the second pilot group (2007). Figure 2 shows the comparison of the modes for the responses from both groups.

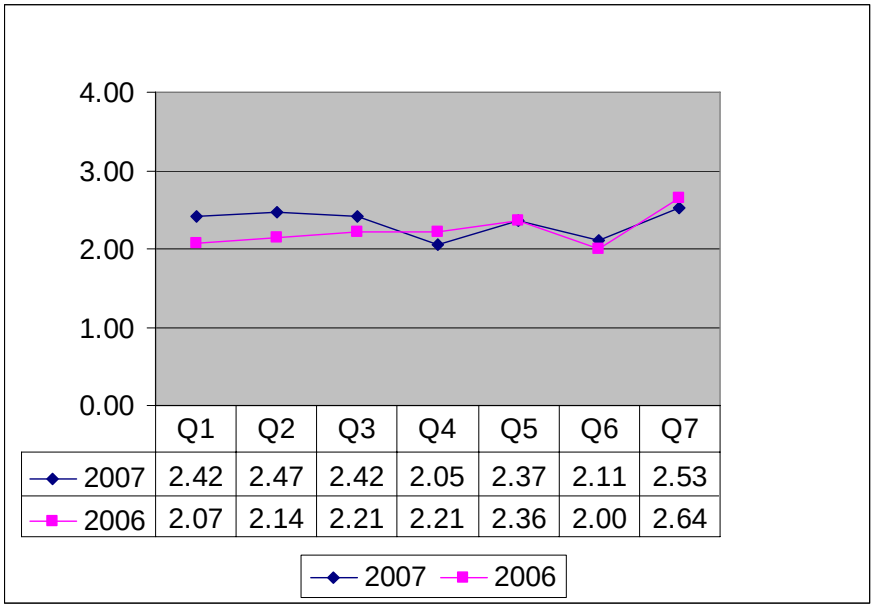



*Figure 1.*

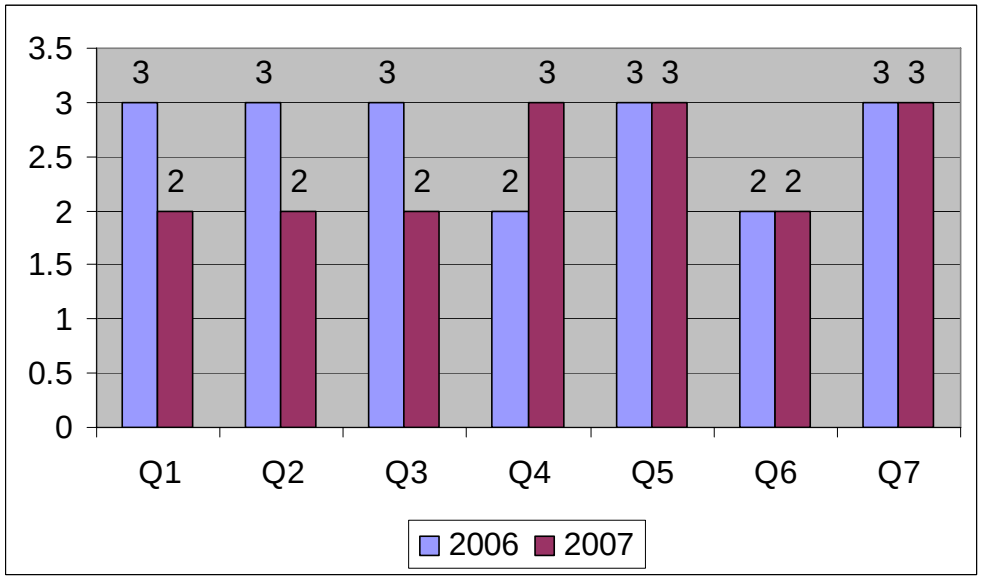
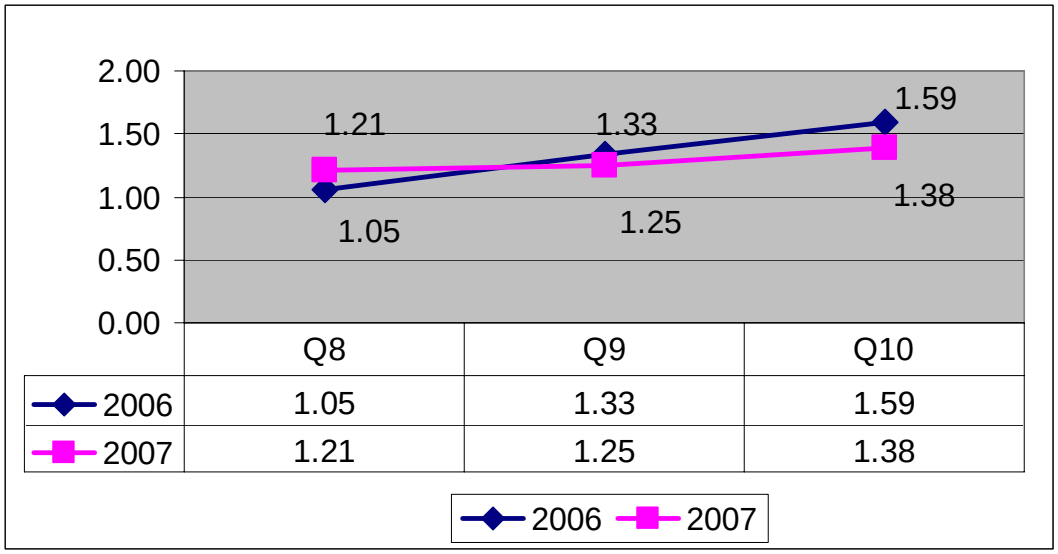



*Figure 2.*

The second research question "Do teacher candidates recognize the value of these tools for the classroom?" was addressed through responses to three yes/no questions (8, 9, and 10) on the post experience survey. Respondents were asked to circle the option that best reflects their future use or beliefs about handheld computers for the following questions: I

am likely to use a handheld device in the future; I think the handheld is a valuable classroom tool for teachers; I think elementary students (grades 3-5) would benefit from using handheld devices in the classroom. The means of both groups for question 8 were close and indicate that few respondents in either group thought they were likely to use a handheld device in the future. Although the means rose slightly for both groups in response to question 9, results indicate that few respondents agreed that a handheld is a valuable tool for a classroom teacher. Even though the respondents did not experience elementary students using handheld devices during the pilot study or during other field experiences planned by the teacher education program, the means for both groups in




response to question 10 were higher than means on the other two questions, and the mode

*Figure 3*

for the 2006 group indicates that most respondents from that group thought that elementary students would benefit from using handheld computing devices. Figure 3 provides a comparison of means on all three questions for both groups in the pilot studies and Figure 4 provides the comparison of modes for both groups in the pilot studies.

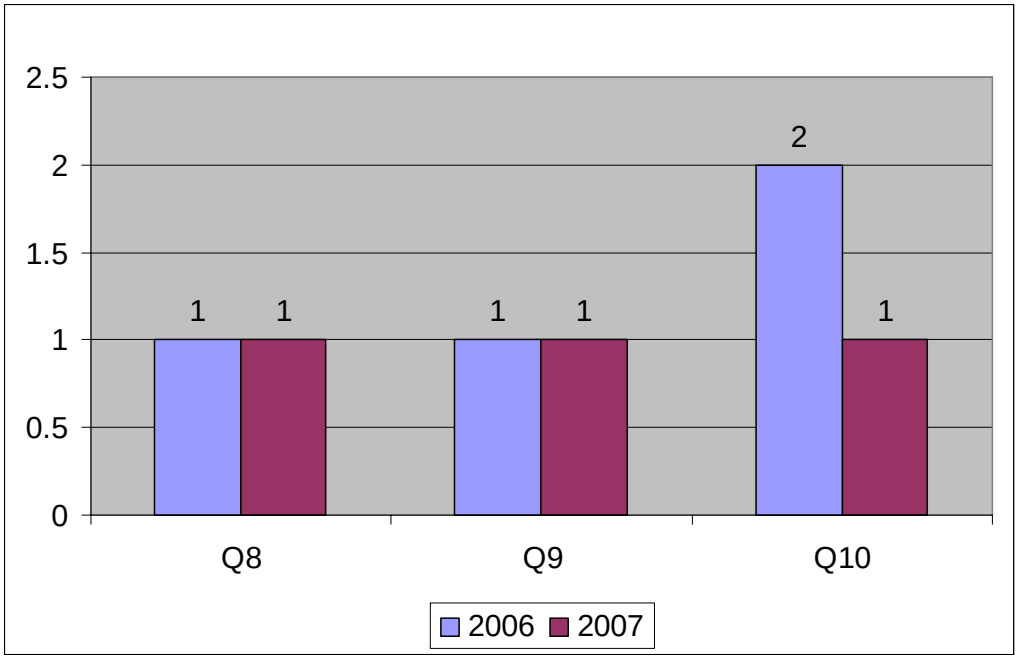



*Figure 4.*

Both research questions were addressed in the open-ended questions asked in the survey instrument, and within the reflection portion of the final assignment in year two. Analysis of the participants' responses to the open-ended questions reveals similar results to those found in analysis of the Likert-type responses to the pre and final surveys of participants.

In general the qualitative analysis of responses related to question one, "Are the hardware and software tools appropriate for formative assessment data collection, " revealed that the participants understood the purpose of using the handhelds, but for a

variety of reasons found the iPaq to be burdensome for data collection. Several students indicated that once given the choice, they preferred to use their laptops and the familiar Microsoft Excel® program. The following quotes are representative of the general attitudes towards the appropriateness of using the handhelds for formative assessment data collection:

- "The iPaq was intended as a useful technological tool to make data collection simpler and more efficient, and I feel in that it succeeded. However, I didn't find it so useful that I would prefer it over other methods of data collection."
- "I began by entering the information as soon as I was done by answering the survey questions. I lost everything during spring break because I didn't charge it. Instead of re-creating the survey, I just used Excel. I liked the layout better anyway."
- "The additional comments section was the most helpful. This is where I would record things that I wanted to remember to do next time, or things that I wanted to remember about the current session. I liked having these details, but with the problems I had with the iPaq it would have been just as easier (and maybe faster) to write these into a notebook."

More so than the results from the closed-ended questions of the survey, the open-ended questions related to the second research question, "Do teacher candidates recognize the value of these tools in the classroom," clearly indicated that the participants did recognize the value of formative assessment in the classroom. The following quotes are indicative of the teacher candidates' discovery and awareness of the purpose for formative assessment:

- "I saw correlations between his mood and focus level, and the use of the survey caused me to look at other factors that played into focus and progress. The assessment format definitely helped me to know the correct way to conduct and improve my tutoring sessions. Through using the survey, my awareness of student progress and the need to make adaptations became especially clear when I looked at the data on the excel spreadsheet. I am not sure that without the reflection my sessions would have had gradual improvement over time."
- "For instance, Elijah struggled with complex operations and fractions. Seeing this documented continuously reminded me to spend extra time in these areas. This was the most valuable part of having the written survey results. As well, when days didn't go well, I could look back and see that there were days that went well, and that was encouraging (and I could see what worked that day)."
- "During the first half of the semester, I used the original tutoring survey yet felt that I was not evaluating anything directly pertaining to my sessions with Elan. I later created a new survey to better track Elan's progress with measures of motivation, mood, and distractibility. I was better able to assess my effectiveness with this survey."

**Discussion**

Quantitative indices from participants about the appropriateness of the hardware and software for data collection in the field proved to be consistent across both pilot studies despite any changes implemented in the second year. The qualitative analysis

affirmed these findings, with comments about the lack of reliability, speed, and user-friendliness. Notably, the most frequent responses were disagree or agree which made the means for each response for both groups cluster more toward “disagree” on most of the questions. The means for question 7 for both groups was the highest of the seven questions and shows that the participants’ knowledge of technology in the classroom has grown. The responses for both groups are very close and a comparison of the modes indicates that the responses from the 2006 pilot group are higher than the responses from the 2007 group. Analysis of the means shows a slightly higher rating on all questions except 4 and 7 from the 2007 group.

Responses from participants about the value of handheld devices for classroom use were not surprising considering the challenges participants faced with hardware and software. It was surprising to see that respondents believed that the device was more beneficial to elementary students than teachers, but that could be due to the perception that loss of formative assessment data collected by a teacher would have a greater impact than loss of data collected during an instructional lesson would be to an elementary student.

Open-ended responses and reflection about the value of these tools for formative assessment were, in general, very positive. The participants actively sought to reflect and transform their practices as well as their data collection instruments as a result of seeing trends in their work with tutees.

**Implications**

Although most of the responses indicate a neutral attitude toward the use of handhelds for formative assessment, the lack of more vehement negative responses is

positive in light of the chronic hardware challenges the participants faced in both pilot studies. The handheld computing devices did not always hold a charge and failure to maintain the charge resulted in complete loss of data for many of the participants. Faulty keyboards and surveys that required substantial data entry thwarted participants in the first pilot study. Using the same unreliable hardware tools, the second pilot group had to overcome the challenges of working in an unfamiliar classroom environment with students they had not known prior to this field experience. A variety of negative conditions unrelated to formative assessment, survey creation or data collection are likely contributors to the less than enthusiastic responses from the participants. Positive results that weren't captured in the responses were the ease of use of the survey creation tool and the simplicity of capturing and exporting results of data collection to the computer. The respondents' acknowledgement that their understanding of classroom technology use was enhanced by this experience is also positive and relevant.

Future studies should include better hardware, post-experience measurements should be more specific, and the field experience should be more solidly framed around formative assessment and teacher inquiry to maintain the focus on the classroom value instead of the tools. Larger teacher preparation programs or programs with teacher candidates who have less experience with technology should consider expanding this type of study beyond one semester to allow participants to become comfortable with the tools prior to collecting data in the field.

If a formative assessment system generates information about student achievement, it needs to include within the system effective ways to communicate results.

The best technology and most accurate assessments in the world are wasted if the results aren't communicated in a way that can be used to improve learning as a whole.

**References**

Black, P., & William, D. (1998a). Assessment and classroom learning. *Assessment in Education, 5*(1), 7–74.

Cohen, L. & Manion, L. (1989). *Research Methods in Education, 3rd Edition.* New York: Routledge.

Darling-Hammond, L., Ancess, J., & Falk, B. (1995). *Authentic assessment in action: Studies of schools and students at work*. New York: Teachers College Press.

Fueyo, V. & Koorland, M. A. (1997). Teacher as researcher: A synonym for professionalism. *Journal of Teacher Education, 48,* 336-344.

Herrington, J., & Oliver, R. (1999). Using situated learning and multimedia to investigate higher order thinking. *Journal of Educational Multimedia and Hypermedia,* 8(4), 401-421.

Hsi, S. (2000, April). *Using handheld technologies to connect Web-based learning to outdoor investigations.* Paper presented at the National Association for Research in Science Teaching Annual Meeting.

Johnson, M. & Button, K. (2000). Connecting graduate education in language arts with teaching contexts: The power of action research. *English Education, 32* (2), 107-126.

Lewis, Anne (2006). Celebrating 20 years of research on educational assessment: Proceedings of the 2005 CRESST Conference (Technical Report No. 698). Los Angeles, CA: Center for the Study of Evaluation & National Center for Research on Evaluation, Standards, and Student Testing Conference. [Online available: http://www.cse.ucla.edu/products/reports_set.htm].

Means, B., Penuel, W. R., & Quellmalz, E. (2001). Developing assessments for tomorrow's classrooms. In W. Heinecke & L. Blasi (Eds.), *Research Methods for Educational Technology.* (Vol. Volume One: Methods of Evaluating Educational Technology.). Greenwich, CT: Information Age Press.

Milner, J. (2006). Learning Centers: Assessment and Accountability [Online document]. Retrieved September 29, 2006 from http://thejournal.com/the/learningcenters/center/?msid=14&1d=17865&clp=2.

Penuel, W.R. & Yarnall, L. (2005). Designing handheld software to support classroom assessment: An analysis of conditions for teacher adoption. *The Journal of Technology, Learning, and Assessment*, *3*(5). Available from http:www.jtla.org

Price, J. N. (2001). Action research, pedagogy and change: The transformative potential of action research in pre-service teacher education. *Journal of Curriculum Studies, 33* (1), 43-74.

Rock, T. C. & Levin, B. B. (2002). Collaborative action research projects: Enhancing preservice teacher development in professional development schools. *Teacher Education Quarterly, 29* (1), 7-21.

Sax, C. & Fisher, D. (2001). Using qualitative action research to effect change: Implications for professional education. *Teacher Education Quarterly, 28* (2), 71-80.

Soloway, E., Grant, W., Tinker, R., Roschelle, J., Mills, M., Resnick, M., et al. (1999). Science in the palm of their hands. *Communications of the ACM, 42*, 21–26.

Tripp, D. (1990). Socially critical action research. *Theory in Practice, 29* (3), 158-166.

# Appendix A

Post-Experience Survey
Questions regarding use and value of handheld tools for field-based data collection

| |
|---|
| 1. The handheld is a useful tool for collecting data in the field. |
| 2. DatainHand software made it easier to capture data in the field. |
| 3. The handheld and DatainHand facilitated completion of course assignments. |
| 4. The handheld and DatainHand allowed me to collect and process field observations in a timely manner. |
| 5. Using a handheld device and a consistent data collection tool facilitated data analysis. |
| 6. Using these tools to collect data in the field was worthwhile. |
| 7. My understanding of how to use technology in the classroom has been enhanced by this experience. |

Figures