Education Tech Research Dev (2011) 59:865–884
DOI 10.1007/s11423-011-9210-9

DEVELOPMENT ARTICLE

Integrating physical activity data technologies into elementary school classrooms

Victor R. Lee · Jonathan M. Thomas

Published online: 19 August 2011

Abstract This paper describes an iteration of a design-based research project that involved integrating commercial physical activity data (PAD) sensors, such as heart rate monitors and pedometers, as technologies that could be used in two fifth grade classrooms. By working in partnership with two participating teachers and seeking out immediate resources in the classrooms and elementary school site, we devised a set of technology-supported learning activities in which students pursued investigations related to the distances that they walk, the relationship between heights and footsteps taken, and variations in heart rates among twins and with adults. In addition, we assessed the students' knowledge before and after the PAD technology supported learning intervention using both a written assessment and interviews. Results from the written assessments indicated that the newly designed activities indeed covered the intended content related to measures of center and averages. Results from the interviews suggested that students who participated in the unit designed to incorporate PAD technologies more reliably accessed knowledge related to measures of center and averages in scenario-based problems than their counterparts who followed a traditional unit. Practical lessons related to the use of this technology with elementary school children that were learned from this design activity are also summarized.

Keywords Sensors · Physical activity data · Probeware · Elementary school · Measures of center · Design-based research

Introduction

Over the past few decades, there has been a sustained interest among educational technologists in sensor-based learning technologies (Pea et al. 1999; Resnick 1998; Tinker and

V. R. Lee (✉) · J. M. Thomas
Department of Instructional Technology and Learning Sciences,
Utah State University, Logan, UT 84322-2830, USA
e-mail: victor.lee@usu.edu

Krajcik 2001). Sensor-based learning technologies originally found a niche in what were called "micro-computer based labs" (Redish et al. 1997; Tinker 1996). These were desktop computer-based learning environments that involved attachable probes with which students could explore phenomena in the physical sciences. This technology was considered a powerful tool that could enable students to measure and ultimately visualize (by way of computer-generated data representations) quantities associated with motion, force, pH, and temperature (Linn and Hsi 2000; Redish et al. 1997; Tinker and Krajcik 2001). Over time and with further improvements in computer technology, these devices were no longer tethered to a desktop computer and became recognized and ultimately relabeled as 'probeware' (Linn and Hsi 2000). While this new version of the technology was still computer-based, the sensors became handheld and mobile, thus allowing new learning possibilities and new contexts for use (e.g., Nemirovsky et al. 1998; Resnick et al. 2000) that moved beyond the desktop.

This interest in sensors as learning technologies has continued into current day (e.g., Struck and Yerrick 2010; Zucker et al. 2008). For example, a recent report prepared by the National Science Foundation Task Force on Cyberlearning placed repeated emphasis on sensor technologies as a key area for future educational technology research and development (Borgman et al. 2008). The task force's interest in sensor technology stems from both the positive findings related to science and mathematics learning from earlier literatures (e.g., Linn et al. 1987) and what also appear to be the increased availability, affordability, and diversity of new sensor-based devices.

Yet with all of the interest in sensor technology, there has been relatively minimal consideration of the potential for sensors that are physiological in nature (Tinker 2000). This lack of attention should not be seen as being due to lack of applicability or availability of such devices. Sensors that read information from human bodies have been around for decades. For example, medical and health sciences regularly use sensors to detect heartbeats or electrical signals in the human body as part of their research practice (Janz 2002). Outside of physiological research, a growing number of people are using physiological sensors even more informally (McClusky 2009). Avid bicyclists are attaching sensors to their bikes in order to track their effort and distance, and fitness enthusiasts are wearing heart rate monitors at the gym. Even modern gaming consoles are getting involved in the use of body sensing technologies. For example, in 2010, video game company Electronic Arts released a second version in their *EA Sports Active* game series that collects and utilizes data about players' body motions and heart rates as part of game play.

Given the growing interest and the availability of body-sensing technologies, we see an opportunity for the educational technology community to explore new forms of teaching and learning that involve this class of tools. This paper is a report of one design-based research effort (Brown 1992; Collins 1992; The Design-Based Research Collective 2003) in which we sought to develop and support activities using such physical activity data (PAD) technologies in the context of two elementary school classrooms. We view our work as being situated within the aforementioned bodies of work that have examined the use of technologies that involve sensors, the modest body of research that has considered passive acquisition of inspectable information for educational uses, and the paradigm of design-based research. The primary goal of this paper is a description of the considerations we made and the activities that we ultimately developed involving PAD devices with respect to a particular elementary school site. That description will be accompanied with some assessment of the learning that took place in the instruction we developed using PAD technologies.

What are PAD devices?

In our work, we are deliberately leveraging a class of technologies that have already been developed for adults to use when monitoring their own physical activities. Because they involve use during and for physical activities, we refer to these as PAD technologies (Lee and DuMont 2010). Canonical examples include pedometers, heart rate monitors, accelerometers, and distance trackers. The benefits of these PAD technologies relate to their suitability for their intended tasks. For example, because the only thing that a pedometer needs to do is count steps, a well-designed pedometer has a high level of accuracy and portability so that it is easily used by an avid walker (Bassett and Strath 2002). And because these devices are also associated with active movements and full-day use, many pedometers have a long battery life and are physically quite durable. Thus, pedometers have the potential to be used as a technology tool to support mathematical and scientific inquiry (Rye et al. 2005; Sun et al. 2010).

It is important to note that commercial devices are not the only option for using sensor technology to capture PAD. In fact, science education supply companies are one possible source for such technologies. In our exploration of one such company's devices (i.e., Vernier, described in Trotter (2008)), we found that while their range of sensor-based offerings were excellent, there were only a few tools that could be used to detect physiological information, and there were serious limitations with respect to their usability. For example, the Vernier heart rate monitors use a pair of connected metal rods that must be held by both of a student's hands. For a recording to be made, a student must be within a few feet of a stationary recording device. Furthermore, in our tests of this setup, we found that it was very easy to lose the signal between the sensors and the recording device.

In contrast, the commercial PAD devices we ultimately have committed to using were designed with a more active context in mind. Because these tools came out of athletic settings, they are made for active use in which a full range of movement must be allowed. Therefore, they are commonly self-contained units that are worn on one's person. A typical commercial heart rate monitor, for example, will involve a heart rate detection strap worn around one's chest and a sport wristwatch that records the information from the detection strap. This close proximity of the device to the user's body allows the sensor and recording device to stay in constant contact, regardless of what the individual is doing. Furthermore, accessing the data from these devices was, by design, a very simple and straightforward process, as athletes do not want to spend an extensive amount of time transporting their data or running transformations on them. In our tests of various commercial PAD devices, we have found that data extraction can take place in just a few steps with little to no involvement with proprietary software.

Passive acquisition of information during activity

One of the potentials that we see with PAD-based technologies is that they enable users to passively acquire a substantial amount of physiological information while focusing attention on other activities. That is, a PAD device will operate in the background and record naturally occurring changes to its immediate environment while a user is focused on another task. This information, when harnessed, can be made into an object of inspection and reflection. One early and oft-cited inspiration for this kind of information gathering and use comes from the human–computer interaction work of Hill et al. (1992) who described the informational potential that *wear* introduces into document processing. To illustrate, consider the wear that appears on a favorite cookbook. Sections that are propped open and reread frequently begin to loosen the book spine and the page edges inevitably become dirtier from where the reader's fingers (and foodstuffs) repeatedly make contact with the paper. As this wear accumulates, the savvy chef can quickly return to this recipe in the book with the aid of the wear that has taken place. She can look along the paper page edges of the book, see where are the dirtier areas, and flip the book open to that section. The cracked binding, having accumulated a memory of being left open to specific recipes, will give way to the precise page that this chef seeks. Overall, this allows the chef to find the recipe she desires much more quickly than if she were to flip open the book to the index and search through an extensive list of names or ingredients to find the recipe she was looking for.

Using that type of interaction as a model for how humans inevitably take advantage of the opportunities generated by wear, Hill et al. (1992) describe an analogous technology intervention in which the amount of time spent viewing or modifying portions of a text-based document are automatically recorded and then re-represented as a vertically organized histogram embedded within the scroll bar of a document window. The time and editing efforts themselves become 'wear' for the computational medium, with increased wear leaving a larger mark on the document. Here, the information traces left by regular document-use activities, be they reading or editing, are converted into a source of information without the user needing to take any additional steps to record this information. The user can then use that information to return to areas that have required a great deal of attention previously or recognize 'wear' landmarks in their document editing. The low overhead associated with leveraging wear is captured nicely in the following quote:

> Wear is gradual and unavoidable change due to use. As a source of useful information, wear is particularly appealing since it is a by-product of normal activity and thus essentially free. No extra effort, nor scheduling of additional tasks are required to get its effects. (pp. 6–7)

In our view, this potential for taking information that is created 'for free' during activities and harnessing it in a potentially consequential way represents a powerful, and perhaps even opportunistic, move for designers and educational technologists.

With the proper equipment, information can be obtained from these activities and used strategically for other goals in the future, with the implication that some of those goals may support learning. This capability has been recognized explicitly by Roschelle and Pea (2002), who in a seminal paper about wireless internet learning devices, have discussed how simple use of the devices can leave information traces regarding technology use. For example, messages sent or queries made are automatically stored in log files and these can then be returned to students as an object of inspection for review after a specific learning experience that involves the use of such a device (e.g., a field trip to a museum in which students are provided with wireless mobile devices.) Roschelle and Pea describe this conversion of passively acquired information into an object of inspection as the conversion of 'act' to 'artifact' and specifically highlight it as a unique feature of this class of technology that could support the design of learning activities.

Also related is the interest that has been generated in other passively acquired activity data by educational technology researchers. Educational data mining activities have begun to treat usage logs that are automatically generated in the course of student search or computer-guided inquiry activities as a source of data to understand student behaviors and activities in technology enhanced learning environments (Buckley et al. 2006). While this information is not necessarily returned to the students, it still provides a valuable source of passively acquired information for educational researchers and technologists and thus helps to make the case that approaches for passively obtaining use information has potential benefit for educational technology research.

With PAD-devices, our approach tends more to the former examples in that it is explicitly focused on using technology to passively obtain information about students that can and will be returned to the students and used as an object of individual and communal inspection during a learning activity. We illustrated this in an earlier design iteration (Lee and DuMont 2010), where we worked with a pilot group of high school students to capture the PAD related to a variety of games and activities, such as playing the basketball game H.O.R.S.E., running on treadmills, and playing a game of Frisbee. Our results from that intervention were very encouraging, capturing video cases of students reflecting on the data collected from their activities and posing substantive questions about how best to characterize central tendency and distribution of the data they had collected. Moreover, the fact that the data came from them and was situated in activities that were already intimately familiar to them seemed to mediate and bootstrap their understandings of the data displays they created and examined. The students in that design study were able to self-correct their analysis strategies by considering what they knew about their efforts in the games and activities and about their perceived selves as athletic individuals.

The fact that the acquired data were meaningful and engaging for students and that the topic of measures of center was such a fruitful area of discussion and consideration for high school students ultimately prompted us to consider if PAD technologies could be used as a learning tool in the classroom. Specifically, we sought to explore if PAD devices could function as a tool to help students learn about measures of center when that content is first introduced at the elementary school level. To consider that possibility, we pursued a second iteration of design-based research in which we explored ways to integrate PAD technology into the classroom in the service of supporting student learning about measures of center.

Design-based research (DBR) and orienting questions

Design-based research (e.g., The Design-Based Research Collective 2003) is an emerging research paradigm that has generated increasing interest in the educational technology community (Bannan-Ritland 2003; Shelton and Scoresby 2011). It is distinguished from other approaches to development and research in that it is iterative in nature, situated in real environments, and adopts a holistic approach toward technology-enhanced learning environment design (Wang and Hannafin 2005). Stated another way, according to a DBR perspective, it is not sufficient to simply bring in a new technology and test its effects. Instead, there must be some acknowledgement that the learning context is dynamic and hosts a number of stakeholders and participants who have existing routines and needs. For a learning technology to succeed in such a complex environment, the researcher often must be directly involved in the design, development, and implementation of a new learning technology as part of a new learning experience. He or she must also take responsibility for noting the decisions and adaptations that needed to be made in order to get the learning technology to 'work' as intended. These decisions are often part of the craft knowledge of design, but one of the promises of design-based research is that it can help make the implicit decisions and heuristics that influenced instructional design explicit (Edelson 2002). Design-based research also provides a means for small-scale investigations of learning or technology use to be gradually scaled to larger audiences (Lamberg and

Middleton 2009) and eventually allows for large scale evaluations of interventions (Bannan-Ritland 2003).

Given that view of how our research and design activities were conducted and that we are only beginning to move our research from understanding how the technology is used by small groups of students to how it could be used in intact classrooms, the reporting of our work in this paper has a descriptive flavor in that we recount some of the highlights of our design experiences. This is because our goal is to describe some our decision-making and describe the outcomes of that decision making in the form of narrative accounts of classroom activities that involved technologies integrated in designed classroom activities. We draw from our own observations from having been present for all implementations of the technologies and activities and from a set of video recordings we made of all classroom activities.

Our guiding assumption in this work was that the technology we were considering had special affordances for convenient acquisition of physical activity information, and while it had not been used in this context to support learning of measures of center previously, we believed it would be possible to thoughtfully establish ways of integrating the technology into the classroom to support teaching and learning activities. However, that assumption was based on positive prior experiences in a very different learning context (Lee and DuMont 2010) and required additional design iterations for verification. The current effort involves our second design iteration—we began to scale the technology in such a manner that it could be used with intact classrooms of students after analyzing how much smaller groups of students engaged with similar technologies.

Given that background, the questions that we consider throughout this paper are:

1. *Is it indeed feasible to bring commercial PAD technologies in as a useful learning resource at the elementary school level?* These are, after all, tools that were designed and built for adults to use and were intended for a context in which fitness progress is monitored. Monitoring and tracking improvements in physical activity are a different set of concerns than would normally be the case for elementary students. Our initial efforts involved older students who were closer to adulthood than the present audience and suggested it should be possible to use these technologies with younger students, but the usability of this technology for elementary schools still remained an open question in this design iteration.
2. *If it is possible to use these technologies, what approaches might be fruitful for establishing a meaningful integration in a classroom setting?* This is essentially a design research question related to some possible forms of instruction. Given that design-based research assumes the researcher/designer must be considerate of the context and the participants who ultimately enact designed instruction, we wished to determine what kinds of activities could be executed in a given school context.
3. *How does learning in a PAD based unit compare to a more traditional unit given the topical focus on measures of center?* Our minimal hope would be that the use of the technology could, at the very least, do no harm to student learning. Ideally, we were hoping to see some indications of improvements in learning that could be attributed to the use of the technology and design integration.

Development and testing site

To answer the above questions, we collaborated with a local, public elementary school in the Mountain West region of the United States. The school had two classes of 25–30 students for each grade level, and the fifth grade classes each had laptop carts and a one-to-one laptop to student ratio.

The two fifth grade teachers with whom we collaborated at this school were both veterans, each having taught for at least 10 years prior to their involvement in this project. Both reported to the researchers that they were comfortable using technology in the classroom, and we confirmed this by informally observing them using the technology resources in their classrooms prior to introduction of any learning intervention. One common use we observed during math class was to give students time to work on supplemental web-based materials that accompanied their regular textbook series.

These two classes followed a consistent weekly schedule. Math and language arts were taught daily in the early mornings, and the later mornings and afternoons saw rotations between art, library, social studies, science, writing, physical education, and other activities. Mornings were also time for a daily snack that was provided, on a rotating basis, by parents. As we will discuss below, snacks figured prominently in some of the learning activities.

The existing approach for teaching measures of center

In order to understand how our approach differed, it is important to note what was the norm for instruction related to measures of center at our school site. In an analysis of international elementary mathematics textbooks, Cai et al. (2002) have noted that in many commercial mathematics curricula in the United States, the emphasis is on algorithmic memorization and less on a model of intuitive understanding of what various measures of center mean with respect to the data they represent. Those findings accurately characterized the instructional materials at our field site. The commercial textbook series the participating school used was the Pearson *Envision Math* (EM) series (Charles et al. 2009). EM is designed such that upper grades elementary students are taught about and learn to compute elementary statistics. It makes extensive use of paper and pencil-based materials, including numerous worksheets and practice exercises dedicated to computing and illustrating the procedures for computing measures of center. EM also includes web-based modules in which students are presented with multimedia demonstrations that show and narrate the procedures for computing measures of center. The data that students are given are from a range of imagined contexts. For example, in one of the EM web-based modules, the data that are used in the instruction are the measured lengths of five snakes. In the paper-based materials, the students are given either a decontextualized list of 3–5 numbers for their examples and for computation exercise problems (e.g., a math problem taken from one such worksheet asks students to find the mean of the following group of numbers: 8, 13, 90, 17) or they are given other imagined data contexts in which a small set of numbers are situated (e.g., given a list of numbers associated with visits to a park over 5 days, determine the median number of visits to the park). Using these sorts of data, there were four mathematical computations that students were expected to learn: the arithmetic mean, median, mode, and range.

PAD and data visualization technologies introduced

We now turn to a brief overview of the technologies that we introduced into our field site. These include pedometers, heart rate monitors, and the *TinkerPlots* (Konold and Miller 2005) data visualization software.

Pedometers

Pedometers are also known as step-counters and are used to keep track of the number of steps that an individual takes during the day. They often come as small plastic clamshell containers that are clipped to the waistband of a user and will passively sense movement and impact as feet collide against pavement throughout the day. They will report a running total and can be reset at any time to start a new count. The more sophisticated models will track caloric expenditure and consider stride length in calculations for distance walked. As a sensor technology, it is fairly simple in that it detects a slight environmental change based on an individual's movement and then records it (Bassett and Strath 2002). While it is a simple technology, the feedback it provides has been shown to greatly increase the amount of walking done by users (Bravata et al. 2007). There is a growing awareness of educational potential in pedometer technology that is just now beginning to be explored by educational researchers (Rye et al. 2005; Sun et al. 2010).

We have observed that in order for students to use pedometers properly, they must wear them on their waistbands, and their clothing must not be excessively loose or baggy. We had a few incidents with students who were getting very low readings, and they could be attributed to oversized clothing that those students were wearing that day. Also, some students distorted their pedometer readings when the device was just attached to their pockets or were slightly angled. In both cases, the accuracy declined substantially. In using this PAD device with elementary students, we found it was necessary to be very specific about where and how to wear pedometers and to establish some peer and teacher checks to make sure the devices were positioned so that they would generate useful information.

Heart rate monitors

Heart rate monitors are tools that involve strapped sensors worn around one's chest that detect electrical signals that naturally are emitted during heartbeats. These work best when the sensor straps are being placed directly on the skin and below the sternum, and the sensors require a small amount of water between the skin and the sensor in order for the electrical signal to be detected. There are a number of models with a variety of features, some of which are reported in Lee and DuMont (2010). For this work, we opted to use the Garmin *ForeRunner* 305 which has a sport watch that communicates wirelessly with the chest strap heart rate monitor in which live heart rate information is displayed and recorded on a wristwatch and is then transferred to a computer.

As are most PAD devices, the *ForeRunner* is designed for adults. In order to make this usable for children, we had to make two small modifications to the technology. First, the chest straps were too large, even at the smallest setting, to fit around a young adolescent's chest. Because continuous body contact is necessary for heart rate to be detected, this proved a small challenge. However, adding a small knot to the strap took care of shortening the elastic bands and ensuring a tight fit. Also, the ideal situation is for the straps to be worn against the skin. We found that the straps work just as well if they are worn over clothing as long as the clothing is a little bit wet. In order to minimize the occurrence of uncomfortable situations with students needing help putting on chest straps, we simply provided some spray bottles with which the students moistened spots on their shirt and then attached the chest strap sensors. This proved effective and also prevented any uncomfortable situations.

TinkerPlots dynamic data visualization software

The last technology tool we introduced was *TinkerPlots*, a software environment that visually represents individual data points as randomly placed dots on a screen that will reposition based on parameters set by a user (Konold and Miller 2005). For example, *TinkerPlots* could begin with the entire class set of footstep data and then sort and order that data by number of footsteps in an automatically rendered histogram. *TinkerPlots* is quite powerful as a data exploration tool for children (Bakker et al. 2006; Lehrer et al. 2007) and was designed with students' intuitions about data representations in mind (Konold 2007; Konold et al. 2002). There are a number of configurations of data that can be shown in *TinkerPlots* ranging from cells to histograms to scatterplots. In *TinkerPlots*, it is easy to add in new data sets by creating a new database and typing in the values for each new datum. Sample screenshots of data rendered in *TinkerPlots* are shown in Figs. 2 and 3.

One challenge that we faced was that *TinkerPlots* is a tool designed for individuals or small groups of users at a single workstation. It can be more challenging to use it as a tool for the entire class. What we, as a research design team, needed to do was to create a way for data to be easily amassed from the class. One solution that we developed was a class website using PHP in which students could enter their individual data values. Those numbers could then be stored as a class set of information that could be exported as comma separated values with metadata about the individual inputting the information being attached. Given this simple formatting provided by this resource, it was simply a matter of exporting the entire class's information from the website and importing it directly into *TinkerPlots*. For other activities that focused on specific individuals, we also produced a Python script allowing us to extract the heart rate data from an individual's *ForeRunner* and further enabling us to rapidly import data into *TinkerPlots*.

Integration of technologies, activities, and the learning environment

One of our goals in this paper is to demonstrate, by way of described examples, how there may be a number of situations in which seemingly familiar activities or routines could be profitably converted into PAD collection opportunities. As we will describe below, we were opportunistic with respect to existing resources and routines in the school in which we worked. For example, snack time was a very important routine in the school and so we used snacks on multiple occasions as part of data investigations with the students. One class of students also had twins and thus that motivated the question as to whether or not twins had the same heart rates when doing the same activities, which we then explored. Also, a regular walk at school for the students involved travelling each week from their classroom to the school library. That walk was turned into a footstep data collection opportunity with pedometers.

Following initial visitations at the field site and one-on-one consultation with the participating teachers, Mrs. Caldwell and Mrs. Dehring,[1] we organized three major clusters of instructional activities that involved the collection and use of PAD, taking advantage of the above-mentioned resources at the school. In planning and designing these activities, we also sought ways for students to see a need for, and consequently learn about measures of center.

[1] All names are pseudonyms.

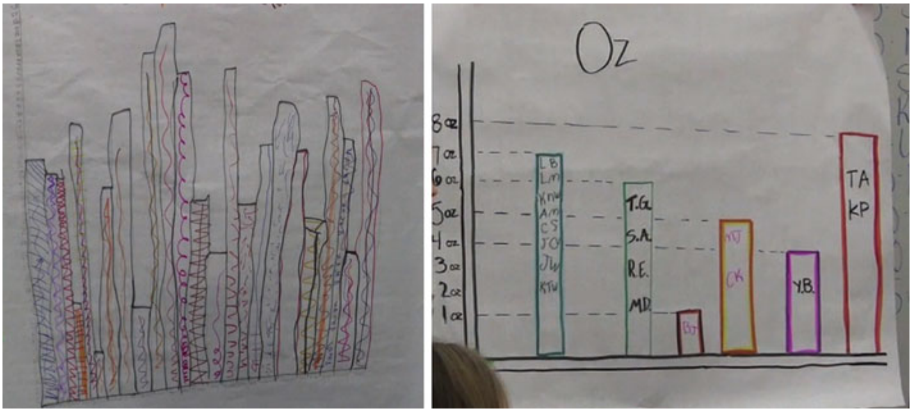

Fig. 1 Two data charts that students created in Mrs. Dehring's class to show footsteps to the library (*left*) and ounces of juice that students drank (*right*) after the walk to the library

Activity cluster 1: Pedometer data collection and analysis

The first major cluster of activities spanned across 3 days and involved students each receiving a pedometer to use. With the pedometers, the students were tasked with determining how far it was from their classroom to the school library. In this task, the students were already aware that the measurements that they were going to make would be dependent on every student starting from the same spot. One student suggested that, even though they were all in the same class, they begin and end at the same spots. The decision was made to have each student clear and start their pedometers at the doorway of their classroom and then check and record their values once they reached the doorway of the library. Also, as this library-walk activity was taking place near morning snack time, we arranged to have students pick up a glass of juice on their walk back as part of their morning snack, but refrain from drinking it so we could also use the amount of juice in their cup as an object of measurement.

Following collection of the data, the students all made records of their measurements and submitted those to the teacher. The following 2 days, students were given the numerical values of both footsteps and ounces of juice from the entire class as a long list and were tasked with inventing ways of showing the data in some graphical format so that someone else would be able to see what was typical for their class and also how different the numbers were. This was a task the students took on enthusiastically. They invented some unconventional displays, which can be seen in Fig. 1.

Because the emphasis on our unit was not on data representation, we did not engage in a more concerted effort to help students refine their design of data representations (Lehrer and Schauble 2000, 2004). In retrospect, that is an area in which we think our unit could benefit from some improvement. However, even from these charts, the classes were able to transition into discussions about what was being shown as typical. Furthermore the students began to become attuned to the idea that numbers (or initials in the case of the ounce chart in Fig. 1) that appeared the most (the mode) and the number that was in the middle of the vertical axis (the median) were both good candidates.[2] With a little prompting from the teacher, the students derived a way of using the arithmetic mean, or more colloquially, the

[2] The method for determining a median given information on a student-invented chart led to an extended and engaged discussion in Mrs. Dehring's class that we will discuss in a future report.

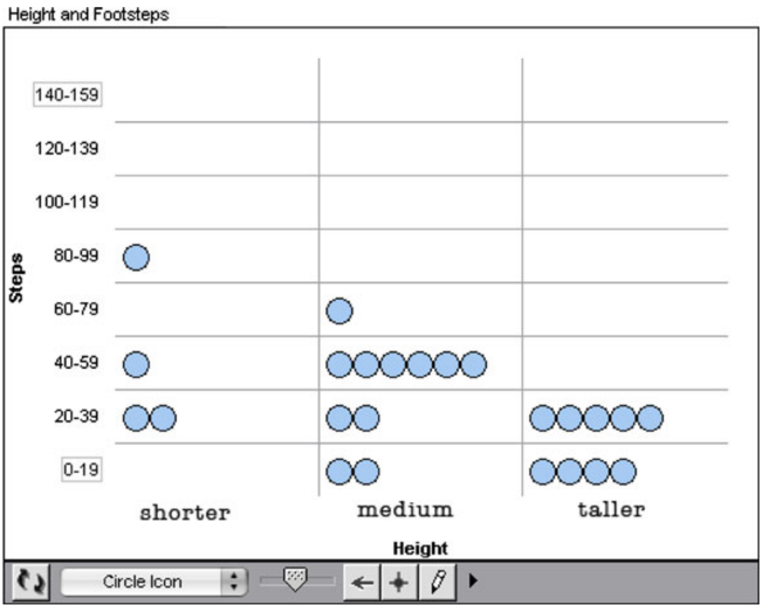

Fig. 2 Reproduced plot of student data comparing height against number of footsteps recorded for travelling the same distance

'average' as a measure of center. The mean, median, and mode were then officially designated as the three measures of center that often are used when discussing data. However, devising better ways of using information from PAD devices and encouraging and supporting students' sense of what constitutes a powerful data representation remains work to be done in a future iteration.

Activity cluster 2: Introducing data visualization software

Over the next 3 days, the students were then introduced to *TinkerPlots*. Students collected some additional data that could be input into *TinkerPlots*. In Mrs. Dehring's class, the data came from pedometer data obtained from another walk around the school. In Mrs. Caldwell's class, the students repeatedly hypothesized that the number of footsteps that students needed to take to get to the library might depend ultimately on the height of individual students. Given the interest in that idea, the class decided that the additional data they would collect and input into *TinkerPlots* was their height information, in the categories of "shorter", "medium", or "taller". To do this, they used individual students as reference points and compared how tall they were to the reference students. The results of this information were plotted with their library walk data and appeared to confirm their suspicion. Figure 2 shows what the class produced, and it indeed is suggestive of a negative correlation between height and number of measured footsteps.

Activity cluster 3: Exploring and investigating heart rates

Following introduction to *TinkerPlots*, the students were then introduced to a *ForeRunner* heart rate monitor and given instructions and a demonstration on how to use it. The students explored some of the *ForeRunner*'s capabilities and recorded some measurements that they could see on their watch displays. From these values, they tried to determine what was a typical heart rate for a few different activities, such as sitting still and jumping. Following an initial introduction to the technology, students began to brainstorm ideas that they might be interested in exploring with the heart rate monitors. The following day, they would collect data using the heart rate monitors, and one more day after that would be devoted to examining their numerical results and coming up with conclusions to share with the class. In total, this set of activities lasted 3 days.

The students planned investigations in each class, but there needed to be some strong guidance from the teacher and researchers to determine what would be a fruitful investigation. This guidance was necessary for a variety of reasons. First, earlier work with this technology suggested that investigations with an uncertain outcome might promote more substantive discussions of data (Lee and DuMont 2010). Our concern for this study was that investigations which had easily predicted outcomes would lead to disengagement or lead to students misreading their data to conform with their anticipated conclusions (e.g., Kuhn 1989). Second, there were safety concerns. For example, students in Mrs. Dehring's class wanted to blindfold and scare students by shoving them unexpectedly, and were very insistent that this would be the best activity for the class to pursue. However, that was not seen as a safe nor appropriate activity to do. Instead, after some discussion and negotiation, the class settled on comparing their heart rates during snack time and regular class-time in order to see if their heart rates were different, as they would be sitting at their desks quietly for both.

Mrs. Caldwell's class was enthusiastic about two investigation ideas and therefore split in half to perform each investigation. In that class, there were two sets of fraternal twins. The group of students who pursued this investigation wanted to see if the fraternal twins would have the same average heart rates over the same activities. The other group of students had noticed on the first day of working with their heart rate monitors that some of the older students seemed to have lower average heart rates than the younger students. They wanted to find out if, as people age, their heart rates decreased. They recruited staff members throughout the school of varying ages to test this hypothesis. Some of the data plots, rendered in *TinkerPlots*, from these investigations are shown in Fig. 3.

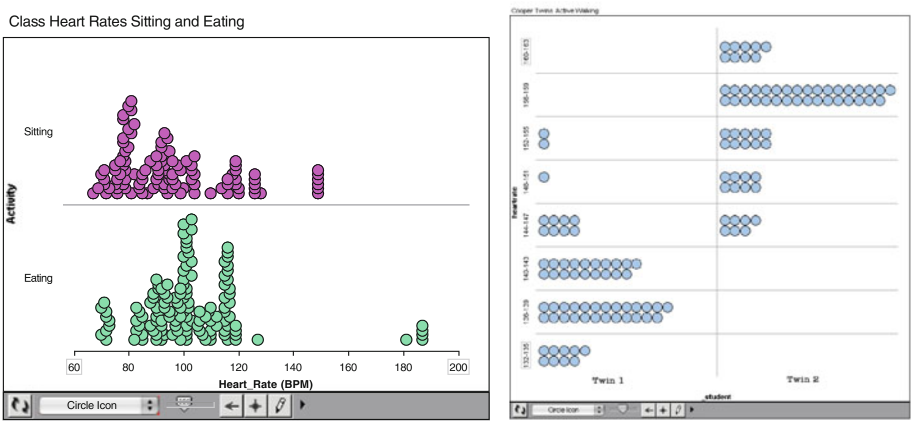

Fig. 3 Data plots from student investigations involving heart rates

Through their investigations, the first class discovered that eating, albeit only involving slightly more movement than sitting by itself, appeared to raise students' heart rates. The second class discovered that the twins seemed to have different heart rates while doing the same activities, and also that resting heart rates tend to increase with age (contrary to their initial hypothesis).

In retrospect, the selection of the culminating investigation in Mrs. Caldwell's class went much more smoothly than in Mrs. Dehring's class because we were more deliberate in seeding ideas and establishing clear parameters for acceptable investigations. Even though the age heart rate and twin investigations were seeded by the research team and Mrs. Caldwell, the students were still very enthusiastic about their investigations. Moving forward, one of the lessons learned is that, even with an orientation towards student autonomy and student-directed activities, seeding and establishing careful boundaries on investigations can still foster engagement and participation, a point that is also iterated by others who have engaged in design of instructional activities (Kanter 2010).

An initial investigation into the effects of instruction

While our description above makes it clear that both classes participated in the PAD-based unit, we arranged for Mrs. Caldwell and Mrs. Dehring's to stagger their classes' involvement with our project. Mrs. Caldwell agreed to use her regularly planned, EM textbook-based instruction on measures of center and data displays for 2 weeks and execute that to the best of her ability. At the same time, Mrs. Dehring's class completed our designed unit without any prior instruction about means, medians, modes, or data analysis. We then took 1 week off to complete post-unit assessments. Following that, we immediately brought the technology and activities into Mrs. Caldwell's class and helped her implement it as an 'add-on' to her unit so that the students could also be given a chance to use the devices and software we were providing. In Mrs. Caldwell's class, since the content was already covered in the traditional unit, the experimental unit was modified slightly to focus on completion of the data collection and analysis activities. Explicit targeting of the measures of center content that they had learned before was not included in Mrs. Caldwell's version.

We developed two assessment instruments for use with the students: a written test and a structured interview. The written test used items taken from state assessments and was focused on reading displays of data and also on computation or identification of measures of center. Two representative sample items are shown in Fig. 4. For post-tests, the same items as the pre-test were used although the values and answers were changed slightly and reordered. All students were given the written assessment. Students in Mrs. Dehring's class

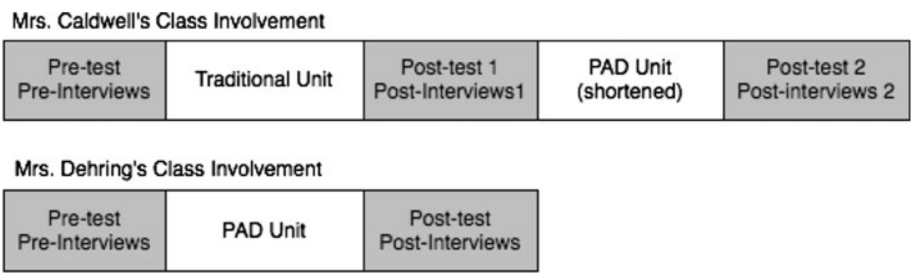

Fig. 4 Illustration of data collection sequence and classroom involvement

Table 1 Sample written test items

Measure of center questions	*Data question*
Use this set of data: 65, 73, 45, 87, 55, and 95	*Students are given a bar graph, tally marks, a table, and a line graph.* Which of the methods shown here are displaying the same set of data?
A. Find the mean of the set of data above	A. A & D
	B. B & C
B. Find the median of the set of data above	C. A & C
	D. A & B

took the pre-test before the experimental unit and the post-test after it. In Mrs. Caldwell's class, the students took the pre- and post-tests before and after the traditional unit, and also took one additional, similarly modified version of the post-test after they experienced the add-on experimental unit (Table 1).

The interviews were conducted with eight students from each class, selected independently by each teacher to be representative of the different achievement levels of their students. Each teacher provided two of their top students, three of their middle-performing students, and three of their lower performing students. Students in Mrs. Dehring's class were interviewed two times, once during the week before the new unit was introduced and once in the week after the unit was completed. Students in Mrs. Caldwell's class were interviewed three times, once before their traditional unit was enacted, once after, and once again the week after they experienced a simplified version of the experimental unit (Fig. 4).

The interview protocol involved showing students two lists of numerical data representing the heights of bean plants that students had hypothetically grown[3] in two other fifth grade classes. During the interview, the students were provided with pencil, paper, and a calculator. The students were asked to figure out what single number could be an effective way of describing how high bean plants tended to grow in each class. This question would consistently elicit the mode. Following this, the students were asked if there were any other ways in which they could provide a single number to describe how tall bean plants tended to grow. If the student could not come up with an answer, they were asked if it would be appropriate to try to figure out a number that would tell us about the 'middle'. If a student was familiar with and understood the median measure, this prompt would tend to lead to its computation. Following that questioning, the interviewer would ask if there were any other ways they could think of to come up with a number that showed how high plants in the hypothetical classes tended to grow. If the student did not offer a response, they were asked if they were familiar with averages and if thinking about averages would be appropriate in this context. This question was intended to get students to compute the arithmetic mean, as that is often described informally, and labeled in students' textbooks, as the 'average'.

The students were then posed with another task in the interview that involved reconstructing a set of data that would produce a specific average value. For example, they were told that the average price of a pack of gummy bears, as determined from seven different stores in the area, was $1.50. They were asked to list prices for the gummy bears from the seven stores that could produce an average price of $1.50. If students chose to assign the same price to each store (e.g., $1.50 for all seven stores), then they were asked to complete

[3] The data sets were [11, 13, 15, 13, 12, 14, 10, 12, 10, 12, 10, 10] for one hypothetical class and [14, 13, 13, 17, 9, 16, 16, 13, 14] for the other.

the task again but to assume that some of the stores had different prices than the others. After a set of varied numbers was produced, the student was asked if the values they selected would produce an average price of \$1.50. This particular line of questioning was adapted from a similar task in other reported research where students were asked to reconstruct a data set but instead used a bag of chips with a different average price (Mokros and Russell 1995).

The same interview protocol, with the same values and scenarios, was used for the first two interviews with students from each class. Students in Mrs. Caldwell's class, who did a third interview after the experimental intervention, a similar protocol was used, but the numerical values and scenarios were changed slightly. Instead of bean plants, the students were given numbers representing the number of pepperoni slices on medium pizzas from two different fictional pizza parlors. For the data reconstruction problem, the target price of \$2.00 was used instead of \$1.50, and rather than gummy bears, that average price was assigned to a hypothetical gallon of lemonade.

Data analysis procedure

Written assessment analysis

Written tests were analyzed by a single scorer who was blind to the identities of the student and the conditions to which the students were assigned for the study. This scorer followed a strict rubric in which a full point was given for each correct answer and a half point was given for each answer that was incorrect but showed signs of evidence that the student had taken appropriate solution steps. For example, a student might not report a median in a situation where there was an even number of data points (e.g., [4, 9, 6, 5, 2, 4]), but they may have ordered the data points sequentially and attempted to count toward a central data point. That would have received a half point on the written assessment. For an arithmetic mean problem, the student would have received a half point if they added all the data together but did not divide the sum by the total number of data points. The maximum possible score on the written assessment was 24 points.

Interview analysis

The interviews were viewed and scored by a single analyst. The scoring of the interviews was similar to the written test in that each line of questioning was scored with one point if the student produced the correct response. If the student did not produce the correct response but showed evidence of taking appropriate computational steps toward the correct response in the interview, even if they had not written anything down, then the student would be given a half point. For the data reconstruction task, if the student produced a set of data that indeed produced the requested average, then they were scored with a full point. If, despite prodding from the interviewer for a set of numbers that produced the desired *average*, the students went with an approach in which the median or the mode of the data they reconstructed was consistent with the requested value, they were scored with a half point.

Because some of the student responses were more ambiguous in their interview responses, a second analyst was asked to independently review and score 30% of the interview responses. The responses and accompanying video reviewed by the second analyst were randomly assigned. The scores from each analyst for each question were compared, and this resulted in $k = 0.93$, indicating a high level of inter-rater agreement.

Results

Written assessment results

The written assessments that both Mrs. Caldwell and Mrs. Dehring's classes had taken were identical to each other and administered on the same days. On the pretest, students in Mrs. Caldwell's class had an average score of 6.78 points (SD = 3.35). Students in Mrs. Dehring's class had an average score of 8.11 points (SD = 2.52). Levene's test was run and the variance in the two classes appeared to be homogeneous ($p = 0.11$). After 2 weeks of either the traditional or experimental units, both classes made learning gains. On the post-test, Mrs. Caldwell's class had an average score of 14.00 points (SD = 4.11) while Mrs. Dehring's class had an average score of 14.14 points (SD = 4.84). The gains in each class, as computed and compared in the R statistics package, were not significantly different from each other ($t = 0.13$, df = 43.36, $p > 0.80$). As noted above, Mrs. Caldwell's class completed the experimental unit activities after they had completed the traditional instructional unit. They were again given the written assessment following the PAD unit, and on that they scored an average of 15.33 points (SD = 4.03), which was not significantly different from the result immediately after instruction for either class ($p > 0.2$ in both cases).

These results suggest that the same amount of content was indeed being covered and learned in both units. Despite this being the first iterations of a classroom-based experimental unit, the students in the experimental class were doing just as well as those in the traditional unit. At a minimum, we "did no harm" to the experimental students. We suspect that with additional refinements to the instructional activities, particularly those related to data representation, we might be able to raise students written assessment scores even higher.

Interview results

A more striking difference between the classes is seen in the analysis of the interview data. Prior to instruction, the students from Mrs. Caldwell's class scored an average of 1.25 points on the interview assessment (SD = 0.75). The students in Mrs. Dehring's class scored an average of 0.875 (SD = 0.64). Levene's test was run on both groups and the variance also appeared to be homogeneous ($p = 0.21$). Following initial instruction in Mrs. Caldwell's class, the students scored an average of 1.81 points on the interview (SD = 0.60). In Mrs. Dehring's class, the students scored an average of 2.625 points (SD = 1.19). The mean gains for Mrs. Caldwell's class was 0.56 (SD = 0.56) and for Mrs. Dehring's class, 1.75 (SD = 0.96). Mrs. Dehring's students, who participated in the experimental instruction, had a significantly greater gain on the qualitative tasks in which they were asked to reason with actual, contextualized data ($t = -3.01$, df = 11.28, $p < 0.05$). Students in Mrs. Caldwell's class, after they completed activities from the experimental unit, scored an average of 2.75 (SD = 0.93), which was a significant improvement over the results from the traditional unit ($t = -2.43$, df = 11.89, $p < 0.05$).

We take these results to suggest that the experimental unit strengthened students' ability to reason about and use the various measures of center in problems that used more complex and contextualized data. This is suggested by the gains shown by Mrs. Dehring's class over Mrs. Caldwell's when the two classes completed their separate units, and also by the gains in Mrs. Caldwell's class after they completed activities from the experimental unit following traditional instruction. Because there was a significant change in the more complex interview problems in Mrs. Caldwell's class that was not accompanied by a significant change in the same class's written assessment scores, we believe that the activities of the experimental unit were especially well-suited for supporting students in working with more complex data than what is traditionally provided by their textbooks.

Discussion

In this paper, we had set out to determine if it would be feasible to use PAD-technologies with elementary school students within a formal school setting. This represented a new direction for research and development related to sensor technologies because they have not traditionally involved students using themselves and their own physical activities as sources of data. The technologies we had in mind, while initially successful with older students, were anticipated to pose challenges when used with a younger audience and as part of a more structured form of classroom instruction. One of our goals was to understand, by way of being partners in the design and implementation of instruction, what some of those challenges may be and how they could be handled in future iterations.

We saw this commitment to PAD technologies to be sensible given their ability to passively acquire data that can then be made into an object of reflection for the purposes of learning new content. Earlier we posed three questions that asked: (1) whether it was feasible for PAD devices, designed for adults, to be used with elementary school students, (2) how the technology could be meaningfully integrated into the classroom, and (3) what differences their might be in learning when a PAD-based unit is compared against a more traditional one.

With respect to the first two questions, we found that students could indeed use the technology as part of a designed unit, but there were certainly a number of practical lessons learned, including the following:

- Greater support should be included with respect to students' activities that involve converting the data they collected into representations of that data. We did not see dramatically different written test learning gains between the control and experimental classes, and that appeared to be related to a limited growth in understanding of data representations and displays.
- The existing activities, routines, and participants in a given classroom or school may prove to be useful resources in student-driven investigations with PAD. The students in this project were able to explore issues related to their snack times, related to their heights and walking, and related to the presence of twins in their classes, just to name a few that this single school site offered.
- Clear constraints and guidelines on student initiated activities need to be determined beforehand and potential student-driven learning activities that can foster engagement may need to be seeded. For example, one must consider whether it is prudent to allow students to deliberately frighten one another for the purposes of a student-designed investigation.
- Some minor physical adjustments in how PAD technology is to be used and very explicit guidance regarding how the devices should be worn need to be made in order to accommodate this younger age group. However, those accommodations do not negatively impact the potential for commercial PAD technology as a tool for data collection in schools.

With respect to our third question, regarding what students learned, we were encouraged to see from the interview data that there may be some added value to the instruction that we provided. That is, students seemed to do better in when asked to reason about situations with more complex data and actual problems. We believe this is attributable to the fact that students were intimately involved in working with larger amounts of data that were naturally messy because they were from real sources, rather than arbitrary numbers provided by a textbook author. This raises some interesting theoretical questions regarding the role that data familiarity and data complexity play in students' learning about and comfort with using measures of center in reasoning tasks. Also, the written tests showed no significant difference between Mrs. Caldwell's and Mrs. Dehring's class, despite Mrs. Caldwell's class having received almost twice as much instructional time. We take this to suggest that what we have developed can at least match what is being done already in schools. When both interview responses and written tests are considered, it seems that the PAD-based unit has the potential to support a stronger understanding of the measures of center content. This evaluation of learning is limited, however, in that it was restricted just to two classrooms with different teachers and only a small fraction of the students were interviewed. We also could have led a more systematic effort to study the experience for the teachers. Regardless, the results we have seen so far are encouraging, and we are eager to explore these issues further in the future.

Ultimately, what we have presented here represents one more step in a design-based research program. Our hope is that this paper makes the case that this under-explored breed of PAD sensor technology can be effectively integrated into the classroom and that it has some potential for supporting, and perhaps enhancing, teaching and learning activities. More work remains to be done to refine the designed instruction and the tools that were used, and some of the questions that are being raised still remain to be answered. For example, we raised a question earlier about how data familiarity and data ownership might play an important role in the learning outcomes that we had seen. We also could and should explore how well PAD-based classroom instruction can be completed with different versions of PAD devices. The affective component of the learning experience, both for the teachers and the students in this project, was not explored. It remains an open question as to whether the overhead associated with using this class of technology is too high for teachers and students in other elementary schools to use. And, as we have sought to explore the use of the technology with younger students, it remains also an open question as to "how low can we go?" The developmental differences between a third-grader and a fifth grader may be so dramatic that these technologies are simply not sensible for third-graders. Regardless, from the work that we have discussed here, we are optimistic that there are a number of compelling questions and promising opportunities out there for educational technologists and researchers to explore.

Acknowledgments Tom Caswell, Mischy DuMont, and Justin Wright provided valuable assistance during this study. Thanks go to the participating teachers and students. This work was supported in part by funding from the US National Science Foundation under Grant No. DRL-1054280. The opinions expressed herein are those of the authors and do not necessarily reflect those of the National Science Foundation.

References

Bakker, A., Derry, J., & Konold, C. (2006). *Using technology to support diagrammatic reasoning about center and variation.* Paper presented at the seventh international conference on teaching statistics, Salvador, Bahia, Brazil.

Bannan-Ritland, B. (2003). The role of design in research: The integrative learning design framework. *Educational Researcher, 32*(1), 21–24.

Bassett, D. R., & Strath, S. J. (2002). Use of pedometers to assess physical activity. In G. J. Welk (Ed.), *Physical activity assessments for health-related research* (pp. 163–177). Champaign, IL: Human Kinetics Publishers.

Borgman, C. L., Abelson, H., Dirks, L., Johnson, R., Koedinger, K. R., Linn, M. C., et al. (2008). Fostering learning in the networked world: The cyberlearning opportunity and challenge. Retrieved January 10, 2010, from http://www.nsf.gov/pubs/2008/nsf08204/nsf08204.pdf.

Bravata, D. M., Smith-Spangler, C., Sundaram, V., Gienger, A. L., Lin, N., Lewis, R., et al. (2007). Using pedometers to increase physical activity and improve health. *Journal of the American Medical Association, 228*(19), 2296–2304.

Brown, A. L. (1992). Design experiments: Theoretical and methodological challenges in creating complex interventions in classroom settings. *Journal of the Learning Sciences, 2*(2), 141–178.

Buckley, B. C., Gobert, J. D., & Horwitz, P. (2006). Using log files to track students' model-based inquiry. In S. A. Barab, K. E. Hay, & D. T. Hickey (Eds.), *Proceedings of the seventh international conference of the learning sciences* (Vol. 1, pp. 57–63). Mahwah, NJ: Lawrence Erlbaum Associates.

Cai, J., Lo, J., & Watanabe, T. (2002). Intended treatments of arithmetic average in U.S. and Asian school mathematics textbooks. *School Science and Mathematics, 102*(8), 391–403.

Charles, R. I., Caldwell, J. H., Canvanagh, M., Chancellor, D., Copley, J. V., Crown, W. D., et al. (2009). *enVisionMATH 5*. Glenview, IL: Pearson.

Collins, A. (1992). Toward a design science of education. In T. O'Shea & E. Scnalon (Eds.), *New directions in educational technology* (Vol. 96, pp. 15–22). Berlin: Springer.

Edelson, D. C. (2002). Design research: What we learn when we engage in design. *Journal of the Learning Sciences, 11*(1), 105–121.

Hill, W. C., Hollan, J. D., Wroblewski, D., & McCandless, T. (1992). Edit wear and read wear. In *Proceedings of the CM CHI '92 conference on human factors in computing systems: Text and hypertext* (pp. 3–9). New York, NY: ACM Press.

Janz, K. F. (2002). Use of heart rate monitors to assess physical activity. In G. J. Welk (Ed.), *Physical activity assessments for health-related research* (pp. 143–161). Champaign, IL: Human Kinetics Publishers.

Kanter, D. E. (2010). Doing the project and learning the content: Designing project-based science curricula for meaningful understanding. *Science Education, 94*(3), 525–551.

Konold, C. (2007). Designing a data analysis tool for learners. In M. C. Lovett & P. Shah (Eds.), *Thinking with data* (pp. 267–292). New York: Lawrence Erlbaum.

Konold, C., & Miller, C. (2005). *TinkerPlots. Dynamic data exploration. Statistics software for middle school curricula*. Emeryville, CA: Key Curriculum Press.

Konold, C., Robinson, A., Khalil, K., Pollatsek, A., Well, A. D., Wing, R., et al. (2002). *Students' use of modal clumps to summarize data.* Paper presented at the Sixth International Conference on Teaching Statistics, Cape Town, South Africa.

Kuhn, D. (1989). Children and adults as intuitive scientists. *Psychological Review, 96*(4), 674–689.

Lamberg, T. d., & Middleton, J. A. (2009). Design research perspectives on transitioning from individual microgenetic interviews to a whole-class teaching experiment. *Educational Researcher, 38*(4), 233–245.

Lee, V. R., & DuMont, M. (2010). An exploration into how physical activity data-recording devices could be used in computer-supported data investigations. *International Journal of Computers for Mathematical Learning, 15*(3), 167–189.

Lehrer, R., Kim, M.-J., & Schauble, L. (2007). Supporting the development of conceptions of statistics by engaging students in measuring and modeling variability. *International Journal of Computers for Mathematical Learning, 12*, 195–216.

Lehrer, R., & Schauble, L. (2000). Inventing data structures for representational purposes: Elementary grade students' classification models. *Mathematical Thinking and Learning, 2*(1&2), 51–74.

Lehrer, R., & Schauble, L. (2004). Modeling natural variation through distribution. *American Education Research Journal, 41*(3), 635–679.

Linn, M. C., & Hsi, S. (2000). *Computers, teachers, peers: Science learning partners*. Mahwah, NJ: Lawrence Erlbaum Associates.

Linn, M. C., Layman, J. W., & Nachmias, R. (1987). Cognitive consequences of microcomputer-based laboratories: Graphing skills development. *Contemporary Educational Psychology, 12*(3), 244–253.

McClusky, M. (2009). The Nike experiment: How the shoe giant unleashed the power of personal metrics. *Wired, 17*, 81–91.

Mokros, J., & Russell, S. J. (1995). Children's concepts of average and representativeness. *Journal for Research in Mathematics Education, 26*(1), 20–39.

Nemirovsky, R., Tierney, C., & Wright, T. (1998). Body motion and graphing. *Cognition and Instruction, 16*(2), 119–172.

Pea, R. D., Tinker, R., Linn, M. C., Means, B., Bransford, J., Roschelle, J., et al. (1999). Toward a learning technologies knowledge network. *Educational Technology Research and Development, 47*(2), 19–38.

Redish, E. F., Saul, J. M., & Steinberg, R. N. (1997). On the effectiveness of active-engagement microcomputer-based laboratories. *American Journal of Physics, 65*(1), 45–54.

Resnick, M. (1998). Technologies for lifelong kindergarten. *Educational Technology Research and Development, 46*(4), 43–55.

Resnick, M., Berg, R., & Eisenberg, M. (2000). Beyond black boxes: Bringing transparency and aesthetics back to scientific investigation. *Journal of the Learning Sciences, 9*(1), 7–30.

Roschelle, J., & Pea, R. D. (2002). A walk on the WILD side: How wireless handhelds may change computer-supported collaborative learning. *International Journal of Cognition and Technology, 1*(1), 145–168.

Rye, J. A., Zizzi, S. J., Vitullo, E. A., & Tompkins, N. O. H. (2005). The pedometer as a tool to enrich science learning in a public health context. *Journal of Science Education and Technology, 14*(5/6), 521–531.

Shelton, B. E., & Scoresby, J. (2011). Aligning game activity with educational goals: Following a constrained design approach to instructional computer games. *Educational Technology Research and Development, 59*(1), 113–138.

Struck, W., & Yerrick, R. (2010). The effect of data acquisition-probeware and digital video analysis on accurate graphical representation of kinetics in a high school physics class. *Journal of Science Education and Technology, 19*(2), 199–211.

Sun, Y., Rye, J. A., & Selmer, S. (2010). Using pedometers in elementary science and mathematics methods courses. *Journal of Mathematics Research, 2*(4), 70–78.

The Design-Based Research Collective. (2003). Design-based research: An emerging paradigm for educational inquiry. *Educational Researcher, 32*(1), 5–8.

Tinker, R. (1996). *Microcomputer-based labs: Educational research and standards*. Berlin: Springer.

Tinker, R. (2000). *A history of probeware*. Retrieved January 2, 2010, from http://www.concord.org/publications/detail.

Tinker, R., & Krajcik, J. S. (2001). *Portable technologies: Science learning in context*. New York: Kluwer.

Trotter, A. (2008). 'Probeware' on increase in schools' science labs. *Education Week, 27*(29), 1–14.

Wang, F., & Hannafin, M. J. (2005). Design-based research and technology-enhanced learning environments. *Educational Technology Research and Development, 53*(4), 5–23.

Zucker, A. A., Tinker, R., Staudt, C., Mansfield, A., & Metcalf, S. (2008). Learning science in grades 3–8 using probeware and computers: Findings from the TEEMSS II project. *Journal of Science Education and Technology, 17*(1), 42–48.

Victor R. Lee is an Assistant Professor in the Department of Instructional Technology and Learning Sciences at Utah State University. He does research in external representations, science curriculum materials, cognition, and the use of physical activity data devices as learning technologies.

Jonathan M. Thomas obtained his MS from Utah State University in Instructional Technology and is Manager of Instructional Technologies at the University of Utah.