An exploration into how physical activity data-recording devices could be used in computer-supported data investigations

Victor R. Lee & Maneksha DuMont
Department of Instructional Technology & Learning Sciences
Utah State University
Corresponding Author: victor.lee@usu.edu

1

Abstract

There is a great potential opportunity to use portable physical activity monitoring devices as data collection tools for educational purposes. Using one such device, we designed and implemented a weeklong workshop with high school students to test the utility of such technology. During that intervention, students performed data investigations of physical activity that culminated in the design and implementation of their own studies. In this paper, we explore some of the mathematical thinking that took place through a series of vignettes of a pair of students engaged in analyzing some of their own activity data. A personal connection to the data appeared to aid these students in recognizing their own errors, and ultimately helped them move from a point-based analytical approach for making sense of the data to an aggregate one. From our observations of this designed learning experience, we conclude that physical activity data recording devices can afford students the opportunity to reason with personally relevant data in meaningful ways.

Introduction

Popular culture is embracing a technology-driven data revolution (Wolf, 2010). This is easily seen on the internet, where web-surfing data are used to infer people's preferences, community statistics and map programs are posted on blogs as mash-ups, and opinion polls are regularly being administered and shared on social-networking sites. But beyond the computer screen and the web browser, there are many other venues where data are now being increasingly collected and analyzed. Parks and health clubs, for example, are being filled with novice runners and athletes who are collecting and analyzing data. Those data come in the form of numbers about the duration of exercise, the rate of heartbeats, the tempo of footsteps, the specific paths that are run, and the calories that are burned. This data collection often goes unnoticed because the devices that are used to gather and record this physical activity data are relatively inconspicuous. They take the form of chest straps and armbands, tank tops and "smart" shoes. Information from these items is being stored on media players and watches, and is being shown and processed on treadmills and smart phones. The idea behind the development and marketing of these physical activity oriented technology tools is that by providing personally relevant information to the masses, the way in which we exercise or recreate can be radically changed, and enhanced (McClusky, 2009).

Our view, and the primary thrust of this paper, is that while the collection and presentation of these data so far have their most profound influence on practices associated with personal health and wellness (Bravata, et al., 2007), there are also tremendous research and educational design opportunities if these same devices were to be considered as *learning technologies*. Specifically, students could use such devices to collect, and eventually analyze, personally relevant data about their own physical activities. The potential content learning could span across biology, physics, mathematics, and statistics. Although we are aware of (and quite eager) to explore these potentials in the future, the goals of this paper are far more modest. Here, we seek simply to describe and label the class of technologies under consideration and provide a brief case study that illustrates how a pair of students engaged in the description and analysis of their own physical activity data obtained with the aforementioned technology.

This paper is organized in two parts. First, we will more formally introduce the technologies we are considering in this work as Physical Activity Data (PAD) devices. In our discussion of this technology, we will compare and contrast PAD devices with similar, often portable, sensor-based technologies that have previously been documented in the educational technology literature. For the reader interested in exploring or using PAD technology in the future, we also provide a brief overview of a few commercial devices or device suites. Following that discussion, we describe a first-iteration design experiment (Brown, 1992; Collins, 1992) in which we created a weeklong learning experience for a group of secondary school students. From that learning experience, we report on a case of two students who sought to determine how their heart rates compared across two activities. In their exploration of the physical activity data they had obtained about themselves, these students showed evidence of strategy shifts for describing tendencies and distributions that appeared to be mediated by their own familiarity with themselves as active individuals and their recollection of the specific physical activities. Over the course of transcript excerpts, we describe a migration from a point-based view of activity data to an aggregate-one.

A new class of devices

As stated above, we designate the technologies we have described as PAD devices. They are generally wearable pieces of computerized equipment that originate from an athletic context and are intended to facilitate the provision of feedback to an athlete based on readings from sensors. Because PAD devices relay information from sensors, there are some parallels to be drawn with earlier research and development work in science and mathematics education that has also involved sensor-based technologies. In the sections below, we briefly describe some related technologies and then discuss how PAD devices are both different and comparable.

Microcomputer-Based Labs and Probeware

Some of the earliest uses of sensor technologies as a learning technology can be traced to Microcomputer-Based Lab (MBL) work that began in the late 1970s (Tinker, 2000). In its most basic form, the original MBLs involved a desktop computer, an attached sensor, and software to convert the readings from the sensor into an immediately rendered display of information. A canonical MBL activity would be the examination of cooling curves. In such an activity, students would take a beaker of heated water in an ice bath, place the temperature sensor in it, and then watch as the temperatures were plotted over time on a computer-rendered graph. Because the temperature would be plotted in real-time, the observing students could tie their immediate observations to the precise changes and fluctuations in the computer display. The simultaneous display of data representations and immediate feedback had been noted as an especially potent feature of MBL as a learning technology. It has also been identified as a source of significant gains in students' comprehension of graph data (Linn, Layman, & Nachmias, 1987; Mokros & Tinker, 1987).

This benefits of simultaneously displayed data has also been further demonstrated in detailed case studies presented by Nemirovsky and colleagues who detailed how experiences related to physical movement and representation of those experience as a data plot can be fused into a new, embodied understanding of mathematics and science (Nemirovsky, 1994, 2002; Nemirovsky, Tierney, & Wright, 1998). In those case studies, the emphasis has been on the use of motion sensors, in which the position of an object with one sensor is determined relative to another sensor (typically attached directly to a computer). Through Nemirovsky and colleagues' analyses, we have vivid images of knowledge change in action as a result of students drawing connections between the dynamically generated graphs and the movements that provided the data. Dynamic explorations of motion, along with measurements of speed and force, have previously and continue to form the basis for many successful college-level MBL-based interventions (Redish, Saul, & Steinberg, 1997).

Although automatic rendering of data displays has proven key in several studies of MBL-like technology, it is worth noting that this is not necessarily the sole beneficial feature or advantage that has come with the use of sensor technology. Sensors, particularly those that have been used as instruments of measurement, have also been successfully embedded into project-based science curricula (Krajcik, Czerniak, & Berger, 1999) with data collection being a primary goal. For example, in a multi-week middle school unit on water quality, the design-research team developed a set of activities in which students obtained readings of pH, temperature, and dissolved oxygen levels in nearby bodies of water which they then successfully analyzed and organized into various inscriptional forms (Wu & Kracjik, 2006). Longitudinal analyses of learning from the water quality and related curricular units in which sensors have been embedded have revealed that students tend to make substantial content and process learning gains in science

(Marx, et al., 2004). Granted, in those curricular units, the sensors are but one technological component, and thus is would not be appropriate to attribute those gains simply to the sensor technologies that were used. Content gains, specifically with respect to science content that was taught with the aid of sensors, or "probeware", as part of a carefully designed science curriculum have also been documented and extended elsewhere (Linn & Hsi, 2000; Metcalf & Tinker, 2004).

Constructionist uses of sensors

Another strand of sensor-based research involves work that came out of a more explicit constructionist tradition (Papert, 1980). In following with a constructionist tradition, the use of sensors in these projects has been in the context of designing and programming personally-meaningful projects (Resnick, Martin, Sargent, & Silverman, 1996). For instance, Resnick and colleagues have documented several cases of students who created novel scientific instruments in which sensors were embedded with 'crickets' – tiny, programmable, computational devices that can be connected to everyday objects (Resnick, Berg, & Eisenberg, 2000). In one case, a student designed and built a device that would capture photographs of birds visiting a birdfeeder throughout the day. Other cases involve children's use of temperature sensors to produce graphs that represented trips during a cold New England day in order to get hot chocolate.

Constructionist uses of sensors are qualitatively different from MBLs largely because the emphasis is on engineering experiences. Unlike MBLs, which are pre-built and treat sensors as existing instruments, constructionist devices and toolkits that embed sensors allow and encourage students to create and invent new configurations. This tradition of embedding and customizing the use of sensors continues in comparable constructionist-oriented projects, such as the GoGo Board (Sipitakiat, Blikstein, & Cavallo, 2004), and more broadly in development projects in which sensor programming is integral to the functioning of students' robotics activities (Bers, 2008).

PAD Devices

There are a number of similarities and differences between PAD devices and the other traditions of sensor use in learning settings. Clearly, there is a similarity in the nature of the technology infrastructure, by which a device responds to changes in the environment and then represents those changes in such a way that the technology provides feedback to a user. There is also similarity in that all of the considered technologies lend themselves quite well to explorations of science and mathematics. They are portable and thus allow for at least some degree of user mobility when the devices are being used. However, unlike the devices that follow an explicitly constructionist approach, PAD devices are already commercially produced, packaged, and distributed. They do not, by design, encourage tinkering or modification. And in contrast to sensors used in MBL-inspired activities, PAD devices collect and store data specific to the user of the device. The emphasis in MBL and modern probeware technologies has largely been on explorations in the physical sciences, despite some early efforts to expand focus to areas such as biology or physiology (Tinker, 2000)[1].

---

[1] It is worth noting that science education supply companies do still produce some devices that students can use to get readings about their selves. However, these represent a tiny fraction of the sensors that are promoted and sold by such companies and because of design limitations, are not easily used outside of the science classroom.

PAD devices are distinguishable from the technologies discussed above in at least two fundamental ways: they explicitly generate data from and about individual users, and they originate from athletic and wellness contexts. The latter feature is notable in that PAD devices are expressly designed for rougher use and mobility. Heart rate monitors (Janz, 2002; Leger & Thivierege, 1988) are a canonical example of a PAD device; they detect changes in electrical current during individual heartbeats, and then calculate the beats per minute based on counts over a set time interval (Karvonen, Chwalbinska-Moneta, & Saynajakangas, 1984). Pedometers, or step counters, are another example. The simplest of these include a coiled spring and adjoining lever arm that increases the step count each time the wearer's foot makes impact with the ground (Bassett & Strath, 2002). More sophisticated step counting devices have been developed that read the signature of motion from a typical foot step, marking initial foot contact with a surface and then measuring the duration of contact time with the floor or other flat surface. These devices can also do approximations of distance and speed based on the duration of floor contact time (Dyanstream Innovations Inc., 2003).

These are by no means the only such devices (e.g., accelerometers that are used to detect changes in speed when bicycling, GPS devices that track routes travelled during runs), but they are among the most pervasive, affordable (often sold for less than $100), and easily associated with physical activity contexts. In the sections that follow, we present a brief description and overview of three commercially distributed PAD devices. These descriptions are presented in order to describe the range of devices and features that are available.

Polar Electro's F6

The F6 is one of the most basic heart rate monitoring devices offered by Polar Electro, a Finnish based company that specializes in the development and marketing of cutting edge heart rate monitor technologies. They are the most widely recognized brand among fitness enthusiasts and exercise physiologists (Janz, 2002). As a company, Polar Electro targets endurance athletes, especially runners and bikers. The F6 is a low-cost, mass consumer offering that consists simply of a watch and an elastic chest strap worn at the base of one's sternum.

When running or biking, the F6 device records information about heart rate and reports duration of exercise, maximum, minimum, and average heart rates. A user can specify targeted heart rate zones, and the system will provide feedback to a user if they are below or above a target zone, so as to optimize a user's workout and pacing. Specific information about the range of heart rates and the amount of time the wearer was "in the zone" is stored on the wristwatch. This information is manually transferred to a PC computer using an infrared attachment or through "SonicLink" technology which communicates the information through a static sound that is detected by a computer microphone, translated, and then uploaded into the Polar Electro user community website. There, users with accounts can view profiles of individual workouts on a calendar, write comments that are stored on an online database, set specific goals, and receive general exercise tips from a staff of personal trainers. Log records of individual measurements during a workout are not made accessible.

Garmin ForeRunner 50

Garmin is a US-based company best known for their work in GPS navigation systems. Their first mass-market commercial PAD device release was called the *ForeRunner* 50. Like the Polar F6, it comes with a sport watch that records information sent from the heart rate monitor

chest strap. As an added feature, it has a small plastic device called a "FootPod" that attaches to one's shoelaces and tracks steps taken based on a foot's contact time with the floor. The FootPod wirelessly transmits this information to the wristwatch, where it is stored as steps taken, cadence, and a few other measurements related to distance. These data can then be downloaded using a USB antenna called an "ANT Stick" that plugs into a Mac or PC. Watches in range that have been linked to a specific computer will automatically send updates via the ANT stick to the computer so new workout records are uploaded without any direct involvement from the user.

The data are viewable in a free software tool distributed by Garmin called *Garmin Training Center*. This software maintains a history of the device's usage, and for each workout, stores log records of individual measurements collected at 5-second intervals during individual workouts. Records include timestamps, heart rate information, and step information. The data for each individual workout are tagged in a single XML file that can be directly accessed by a user.

Nike+iPod

Shoe and athletic apparel giant Nike and Apple offer a PAD device designed for runners to use with their iPod portable music players. The system, dubbed Nike+ (with an elliptical reference to iPod) integrates a new line of shoes with an accelerometer embedded in the sole. Like the Garmin FootPod, this device tracked the amount of contact time with the solid surface, and using that information, determined the distance and steps traveled. It also calculates the calories used during the exercise, determined through computations that consider a user's body weight. While this device set intended to be used with a special line of Nike shoes, it is possible to secure the accelerometer onto a standard pair of athletic shoes, though the accuracy of the accelerometer is thought to decrease when it is used in that alternate fashion as contact time is harder to determine when the sensor is not fully secured.

Data are transferred by manually connecting the iPod device to a Mac or PC using iTunes software. As part of the syncing of the iPod device, the Nike+ information is updated and automatically uploaded to a Nike running website where graphical displays of information are provided. These displays include high and low resolution whimsical graphs showing the runner's speed over the duration of the exercise. Samples of speed are recorded in 10-second intervals by default. XML log files with the individual measurements are stored on the iPod and on the linked computer.

Describing and Analyzing Data

Up to this point, we have discussed portable technological means for students to gather data from their own physical activities. While the features designed into the technology offer many possibilities, the larger question of what students can learn from the technology remains. Because PAD devices collect data, a natural extension of the technology would be to use the data gathered by it to support student learning about how to create and analyze data. Those data practices are often referred to as *data modeling* (Lehrer & Romberg, 1996).

Our present concern is on how students describe and analyze physical activity data. One of the features of PAD technology is that a large volume of data can be obtained in a relatively short time. While a large volume of data can be helpful for separating 'signal' from 'noise' (Konold & Pollatsek, 2002), the associated trade-off is that such large data sets quickly become unwieldy. This realization was made centuries ago, and ultimately led scientists, mathematicians, and everyday practitioners to invent and refine what we see now as elementary statistical

techniques for simplifying and describing data with one or a few numerical values (Bakker & Gravemeijer, 2006). Among the most recognized techniques refer to central tendency. In formal mathematics education, ideas of central tendency are often reduced to the add-and-divide algorithms for computing the arithmetic mean (Cai, Lo, & Watanabe, 2002) and the counting and segmenting algorithms associated the median or mode.

Based on quantitative and qualitative measures, the algorithmic emphasis does not appear to produce rich understandings of central tendency (Cai, 1998). Consider the mean as a canonical case: when given situations that involve thinking flexibly about how to determine a central value, students often fall back on the add-and-divide algorithm without considering the numerical situation being modeled. This often leads to inaccurate results or procedures that are inappropriate to the task at hand. This behavior is observed across all levels (Watson & Moritz, 2000), including elementary and middle school students (Mokros & Russell, 1995), undergraduates (Pollatsek, Lima, & Well, 1981) and even teachers (Jacobbe, 2008). Attempts have been made to explain why intuitive understandings and application of the mean concept poses such great difficulty (Mevarech, 1983; Strauss & Bichler, 1988). The current consensus appears to be focused on three major issues. First, measures of center have specific properties and attributes but students rarely get a chance to deeply explore them in traditional forms of instruction (Strauss & Bichler, 1988). Second, students are rarely given authentic situations that demand a meaningful description of central tendency (Mokros & Russell, 1995). Finally, students' pre-existing intuitions for reasoning about data sets are not being adequately considered or utilized as resources (Konold, et al., 2002).

It is possible for computational software to mitigate the latter issue. Software tools such as *Tabletop* (Hancock, Kaput, & Goldsmith, 1992), *TinkerPlots* (Konold & Miller, 2005), and *Fathom* (Finzer, 2005) all provide dynamic visualizations from which students can make intuitive descriptions of distributions. In essence, these tools are built with the capability of "showing" centrality and tendency through creative uses of position and space. They also often have support tools and features that are explicitly developed with students' pre-formed intuitions for numerical description in mind. For example, in *TinkerPlots*, a 'hat' display similar to what many students will draw in interview tasks, can be generated and overlaid on a distribution on the computer screen through the press of a button (Konold, 2007). Given that such software tools exist and can be effectively incorporated into learning activities that engage students in data modeling (e.g., Lehrer, Kim, & Schauble, 2007), our aim is to explore what could happen when students are given physical activity data to describe and analyze. In that situation, the second issue is addressed by providing students with an authentic context, namely something about their selves, that they must explain.

Student investigations of data with PAD devices

We designed a weeklong learning experience in which individual students collected and analyzed physical activity data gathered with individual Garmin *ForeRunner 50* devices (described above). While students of varying ages could use the PAD technology, we chose to work with high school students. This age group was selected as they were physically closer in age and size to the target age group for PAD devices (i.e., adults) and also because the emphasis on data modeling cases has been on younger students (e.g., Lehrer, Giles, & Schauble, 2002) leaving the older age group virtually unstudied. In light of the fact that conceptual difficulties with central tendency persist through high school (Watson & Moritz, 2000) and beyond

(Pollatsek, et al., 1981), we were also hoping to find with this group of students compelling cases of learning within this age band to add to the literature.

For this design experiment, four volunteers were recruited to participate in a weeklong workshop (i.e., 5 weekdays). Molly[2] was one of the student participants who was about to enter her 9th grade year. According to some early screening and survey information, we discovered that Molly typically got B grades in math and science classes. When asked about her physical activities, she reported that she hated her gym class and did not do any physical activities besides riding her bike around town or occasionally swimming. Molly's best friend and neighbor, Sarah, was the other female participant. Sarah was about to enter the 10th grade and described herself as very athletic. She had been a competitive figure skater for several years, was a fast runner (i.e.., she was one of the girls with the shortest mile-run times in her cohort), and generally earned A's in math and science. Molly and Sarah are the students we focus upon in the case described below.

The other two participants attended the same high school as Sarah, but only knew each other in passing. Stan was set to enter the 12th grade. He was a technology enthusiast, avid bicyclist, and would regularly compete at local track meets. He struggled in school mathematics (earning a C the previous year) but considered himself to be mathematically capable relative to his peers. His low grade, he believed, was a result of his struggle with assignment completion. Jeff was to enter his 9th grade year. He earned A's in math and science, and thought of himself as being well above average in those areas. PE classes, on the other hand, were an area where he considered himself to be below average. According to his PE teacher, Jeff would routinely be assigned to the less athletically competitive group of students who needed additional review on athletic skills.

These four students gathered during their summer break from school to participate in our design experiment, which was facilitated by the first author. The workshop was organized so that the first days allowed students familiarize themselves with the physical resources they had at their disposal (athletic equipment, facilities, etc.) and the technologies they would be using. The remaining time was devoted to investigations of PAD data that they collected using their selves as research participants. The students participated in two investigations planned by the researcher and then designed a pair of investigations of their own. A summary of the workshop schedule is provided in Table 1.

Day	Activities
1	• Introductions • Group discussion of health and physical activity and how one tracks information related to physical activity • Presentation and overview of PAD devices (*ForeRunner* 50s) • Walking tour of facilities
2	• Introduction to *TinkerPlots* software • Review of data gathered from PAD devices worn during walking tour from day 1 using *TinkerPlots* software. • Collection of PAD data from sitting at computers • Introduction of planned investigation: Comparing Frisbee disc throwing and

---

[2] All names are pseudonyms

	HORSE
	• Data collection of Frisbee disc throwing and HORSE data
3	• Analysis of data gathered from previous day's time sitting at computers
	• Analysis of previous day's data gathered from Frisbee disc throwing and HORSE
	• Collection of PAD data from sitting at computers
	• Brainstorm of design of new, student-developed investigations
4	• Data collection for two student-developed investigations
5	• Data analysis of data collected while sitting at computers on days 2 and 3.
	• Analysis of data collected during the two investigations from the previous day
	• Debrief group discussion and student commentary about the technologies used throughout the week.

Table 1: Table summarizing activities during the five-day summer workshop with the PAD devices.

Investigations in more detail

In total, the students participated in four data investigations, two of which were planned by the researcher and two that were student designed. In the first researcher-planned investigation, the students examined data of their heart rates while sitting at the computers over two days for roughly thirty-minute stretches of time. During those times, the students were being introduced to the software (day 2 data) that they would be using or completing an earlier investigation (day 3 data).

The second researcher-planned investigation was a comparison of heart rates when participating in two physical activities: repeated throwing of a Frisbee flying disc to a partner and continuous play of the basketball game HORSE (in which each player chooses a location on a basketball court and attempts to make a basket. If that player succeeds, the other player needs to make the same shot or be penalized with one of the five letters of 'horse'). All students participated in the activities. The students helped plan the data collection strategy by designating how students should be paired, the amount of time that should be allotted, and the specific rules for each activity.

Finally, student pairs were asked to develop questions to be answered in the design of their own investigations. The planning took place on the third day of the workshop and data were collected by all students on the fourth day. The researcher was on hand to make suggestions to help the students in the development of their investigations. Jeff and Stan planned one of the student-designed investigations. It compared distances and heart rates in different running environments, namely the track and the treadmill. The other, planned by Molly and Sarah, compared heat rates on two different cardiovascular machines that both involve cyclical motion: elliptical trainers and stationary bicycles. Again, all four students participated in the data collection process and data analysis for both investigations.

All data analysis activities took place in the *TinkerPlots* software environment (Konold & Miller, 2005). In *TinkerPlots*, students work with a database of "cases" that are each represented

by a single, randomly placed point in a plot window. Students can select specific case parameters around which cases can be organized, sorted, or ordered. This sorting is shown through dynamic animations in which all points reposition themselves on the screen to different cells. The cells can be resized such that the number of partitions of the data are increased or decreased by dragging a single case point vertically or horizontally. *TinkerPlots* also allows creation of modifiable text boxes in which annotations about the data can be added or questions can be written and posed. For all data investigations, the facilitator wrote 4-6 open-ended questions in a text box for the students to answer based on their review of the collected data. These questions were intentionally worded so that the answers to them were unclear and would require discussion by the pairs in order to determine what was being requested. For example, a common question posed to the students across all the different investigations was "Whose heart was working harder?" In responding to the question, the students could be concerned with trends in heart rates, central tendencies, specific student attributes (such as age or sex), or duration of activity. All those considerations could lead to unique and valid answers. Other questions required students to make predictions or explain specific patterns.

An illustrative case of students engaging with physical activity data

Videorecordings of each pair at work on a single workstation were collected for each day except for day 4, when students were using athletic facilities around campus to collect data for the student-designed investigations. All activity data files were stored, as were any *TinkerPlots* files that students had used or modified. Open-ended written responses regarding students' academic and athletic backgrounds were collected at the beginning of the week, and worksheets designed to elicit students' ideas related to heart rates were given each day. All of these data sources were considered in order to understand the students' experiences, but the videorecordings are the primary data source discussed below. In the case we consider here, Molly and Sarah are in the midst of answering a set of questions related to the Frisbee and HORSE investigation and struggle with how to determine in which activity their hearts were beating more. At the time of the workshop enactment, the researcher had noted this episode as an especially memorable period during the week because the pair of students were reasoning out loud and having an extended discussion (over 20 minutes) around a single problem. The second author independently reviewed all video recordings, selecting episodes in which there was extended discussion and explicit reasoning, and she also flagged this episode as the one involving the most revisions and discussions of the students' mathematical thinking.

With this episode jointly and independently selected, the two authors individually reviewed the full day's video and each constructed written narrative summaries of the selected video clip. The narrative summaries contained descriptive analysis of the specific computational strategies that the students were using along with some reasoned account as to why those strategies had changed over the observed period of time. In subsequent meetings, the authors compared their narratives while cross-referencing with the video footage, discussed alternative interpretations, and then formulated a single summary that formed the basis for what follows.

Background on Frisbee/HORSE Investigation

On the latter half of day 2, the researcher asked the group of students whether Frisbee or HORSE would make their hearts work harder. Almost immediately, the students stated that they

thought playing HORSE would make their hearts work more. When pressed for a justification, the students stated reasons such as minimal walking during Frisbee, more jumping in HORSE, and greater effort required to throw an object upward rather than horizontally. The researcher suggested they try to find a way to do a comparison, and solicited suggestions from the students on how to collect data. For several minutes, Stan and Sarah were the most vocal and offered ideas for how to collect the data. They decided that 15 minute blocks of time doing the activity would be appropriate to collect enough data points and students could each do one activity in pairs and then switch.

The group decided that pairings (rather than the entire group playing together) would be best so that the students would all remain active. Sarah's immediate recommendation was to pair with her friend, Molly. Stan commented that the pairings would need to account for the different genders and deal with the differences in athleticism. By this point in the workshop, the entire group had already come to the consensus that Stan and Sarah were the athletic students and Jeff and Molly were not athletic. After weighing a few different options, they decided that Molly and Jeff should be one pair because they were comparably non-athletic and Stan and Sarah should be the other one since they were both athletic. Stan and Sarah would first play HORSE for 15 minutes while Molly and Jeff spent 15 minutes on Frisbee, and then the groups would switch. The facilitator volunteered to tell both groups when to start and stop each activity so they could concentrate on playing rather than timekeeping. Both pairs would be doing their activities outdoors, with one playing on a basketball court and the other on an adjacent field. A five-minute water break would be set in between the two blocks of time so students had time to recover.

During the activity, the facilitator supervised the students' continuous play and kept time for the 15-minute intervals. For both activities, there was some added walking and jogging as students retrieved the basketball or Frisbee when it was poorly thrown. Students would periodically look at the *ForeRunner* watches and check their heart rates, and sometimes would announce the number displayed to their partner.

Data were easily collected by the students with one exception. While completing their turns at throwing the Frisbee, Stan realized that he had not set his *ForeRunner* to record his heart rate. It was displaying the numbers it was receiving from the sensor, but it did not write those to memory. This realization was made about 11 minutes into the second shift, and he quickly started recording after he made the realization. As a result, he had far fewer recordings for his Frisbee time than HORSE.

The next day, the researcher transferred all activity data from the wristwatches to the computers and prepared a set of questions specific to the data for students to discuss. These included:

1. Between playing Frisbee and HORSE, which makes your heart work harder? How much harder does it work? How did you figure that out?
2. If you did the same comparison of HORSE and Frisbee everyday for a week, would you expect one of them to make your heart work harder? Why?
3. Which one involved more walking? If you were to compare HORSE and Frisbee again, which do you think would involve more walking?
4. As far as you can tell from the information here, whose heart was beating the most? How did you figure that out?

When the students returned on day 3, they worked for about 20 minutes in pairs looking at data that was obtained from sitting at the computers the day before. Sarah and Molly worked together at one computer and Jeff and Stan worked together at another. In the "Sitting" data, Jeff's heart rate appeared to be unusually low relative to the others. Upon noticing that, all the students questioned Jeff, who explained that this was a result of a mild medical condition. He then added a bit of related trivia: endurance athletes, like Lance Armstrong, are known to have very low heart rates. The students commented about this, Sarah especially took note as she was an athlete herself. Immediately after examining the "Sitting" data from the previous day, the pairs opened a *TinkerPlots* file that contained the data from their Frisbee/HORSE activities. Each case in that data set was a single heart rate reading from their *ForeRunner*s and had additional fields defined that identified the specific student, the student's gender, the time of day, the cumulative distance traveled, and the name of the activity. The case below is based on the period of time that Molly and Sarah worked to figure out which activity caused each person to work harder and who worked hardest overall.

The case is presented as three sequenced vignettes illustrating strategies they used to determine whose heart was beating the most across the two physical activities.

Molly & Sarah make determinations based on min and max values

The first vignette illustrates an early preference of Molly and Sarah to rely on perceived maximum and minimum values in the data set for making determinations. For this vignette, the pair attempted to determine whose heart was beating the most by using a display they had already produced to answer another question. The plot they created was categorized by student and by activity, and the data points were ordered by distance (Figure 1). Some time earlier, the pair had selected an option that presented the number of cases in each cell in the top right corners.
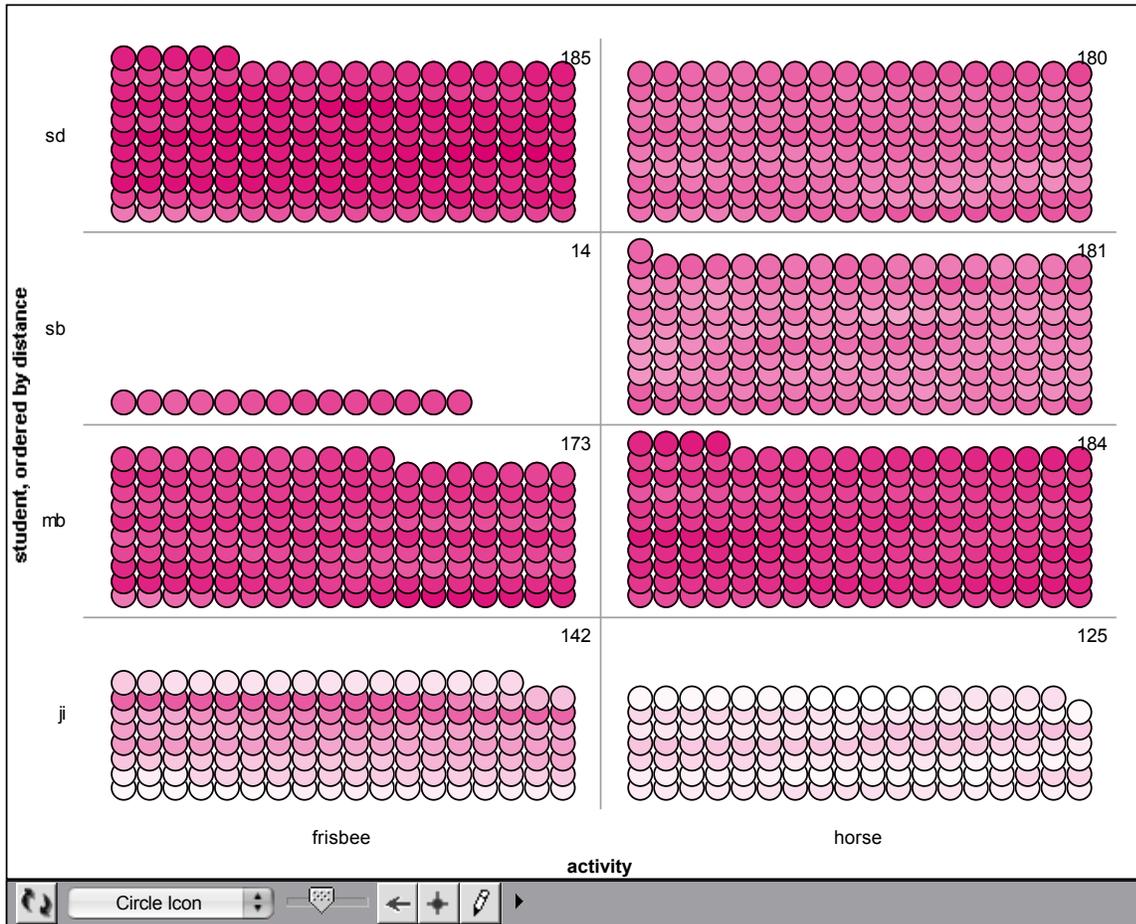
Figure 1: Initial plot display from *TinkerPlots* (Konold & Miller, 2005) created by Molly and Sarah during the Frisbee and HORSE investigation.

Sarah and Molly were trying to determine which activity caused each student to work harder. They began with Figure 1, which shows the relative value of each individual heart rate recording by giving the data points different colors. High heart rate readings had a darker color and low heart rates were closer to white. Sarah controlled the mouse and proceeded to select the bottom right point or top left point in each cell. Together, Molly and Sarah read the associated heart rate values as measures of the lowest and highest points of exertion, respectively. They did not seem aware that the data were still ordered by distance, not heart rate value, and thus they continued to select the corner points of each cell. Those points did not necessarily contain the minimum or maximum heart rate readings. Regardless of that oversight, we will refer to the points they selected as minimum and maximum as that was what they perceived them to be.

Prior to Sarah clicking on Molly's minimum and maximum values, Molly declared immediately that she worked harder at HORSE. In response to that declaration, Sarah quickly checked the maximum and minimum values associated with each of Molly's activity cells, confirmed the maximum in HORSE was higher, and agreed with Molly's claim. When they proceeded onto data for Stan (they opted to check the girls first, then boys after), they immediately noticed that he had fewer data points in this display. This led to an incorrect inference by Sarah based on the minimum values.

S: Stan worked harder at…wait…oh cause his wasn't recording at Frisbee.
M: Oh his wasn't recording at Frisbee?
S: No. Actually his lowest was 145 in Frisbee and his lowest at horse was 123. So we could assume he was working harder at Frisbee.
M: Yeah.
S: Even though we don't know his highest.

Sarah checked the bottom left data points in each of Stan's activity cells, saw that the heart rates were 145 for Frisbee and 123 for HORSE. From that, she inferred that because Stan's lowest value for Frisbee was higher than the lowest value for HORSE, his highest for Frisbee must be higher than for HORSE. Essentially she assumed the distribution of Stan's values would be the same in both activities and since the Frisbee data began at a lower heart rate value, it would therefore also terminate at a lower heart rate value. Molly tried to interject, seeing a flaw in this logic.

M: What if his lowest is just as high?
S: Because obviously if it was recording he would have had higher ones.
M: Ok.

Although the transcript of her utterance ("What if his lowest is just as high?") is open to some interpretation, we are confident from review of the video and from having been in the room with the students that Molly was indicating the possibility that 145 was not necessarily Stan's true Frisbee low. We believe Molly was making the point that the 145 was just as high as Stan's other high points, and the missing values could have instead been lower – perhaps even as low as 123. Molly's concern was dismissed as Sarah quickly explained her rationale (what was represented in the data set was representative of the entire time), and Molly accepted Sarah's explanation. They then proceeded to look at Jeff's data points. What is important to note is that Sarah was pushing toward a strategy of making determinations about greater physical exertion on the basis of single bounding data point values. This occurred in her readings of other people and in her inferences about Stan's activities. Eventually, Molly went along with this and the pair determined that Sarah and Jeff both worked harder at Frisbee. As they concluded this examination, Sarah declared:

S: We all worked harder at Frisbee except for you. You worked harder at HORSE. Because you didn't do anything at Frisbee.
M: He *[Jeff]* threw it right to me, I didn't have to do anything.
S: Oh that's awesome. And you made him go get it and stuff!
M: Yeah *[laughing]*.

At this point in their session, Molly and Sarah justified the idea that Molly had done less work when playing Frisbee on the basis that Molly recalled that Jeff had consistently threw the Frisbee directly at her, meaning that she didn't have to move around as much to catch or return the Frisbee. The approach of taking two boundary values and comparing them seemed to do the trick for the girls, and having accepted these values, they could find justifications for using those values based on experience or intuitions about the data sample.

Molly and Sarah count data points to determine overall exertion

Satisfied with their initial, albeit incorrect, assessment of who worked harder at which activity, the girls proceeded to think about the question concerning whose heart was beating the most overall. In response to this question, the girls adopted an entirely different approach but with the same plot display. After a brief discussion of what was being asked, Sarah again offered the next suggestion.

S: Oh, we have to add them up again *[referring to the numbers in the top right of each cell]*.

M: No, cause that's just saying how many *[motions over the points]* – Oh yeah, we do, huh?

S: Mm hmm.

Molly again began to object to the method Sarah proposed, but quickly agreed to it. The pair then proceeded to add the number of heart rate recordings made during each activity to get a total number of heart rate recordings for each student. To begin, Sarah suggested they "add up 185 and 180". Those values corresponded to the number of heart rate recordings captured for Sarah's readings during Frisbee and the number of heart rate recordings captured during HORSE. This strategy would not give the result they expected. Hearing this, the facilitator (VL) migrated over to the girls' workstation and asked them to clarify what they are doing.

S: *[to Molly]* So add up 185 and 180.

VL: *[from behind Sarah and Molly]* Where are you getting these numbers from?

S: We're just adding up the heart rate of the counts for each person.

VL: So what-so what's getting added up? You said you're adding up the heart rate of the counts? What does that mean?

S: The count of the heart rate.

M: Yeah.

VL: And what are they trying to-

M: How many different cases are in each.

VL: Oh, ok. I see.

From this discussion, it seemed that Sarah and Molly both at least partially understood that they were tabulating the total number of cases, not the heart rate measurements themselves. Coincidentally, each cell's values would have been a reasonable reading for a heart rate. That may have contributed to Sarah's apparent confusion when she was asked what was being computed. Molly interjected the correct response ("how many different cases are in each"), and that appeared appropriate at the moment so the facilitator left the pair alone. The two continued to do their planned computations and produced 365 for Sarah, 357 for Molly, 267 for Jeff, and 195 for Stan. After adding all the totals for the cases, Molly made the following observation:

M: So your heart rate was higher.

S: Huh! *[she turns to face Molly]*. And I thought we had said that the most- like more athletic people the less.

M: It's true.

S: Was my heart beating the most or was it reading the most? Was it like the watch? Was I getting more readings than you guys? Well obviously because these are the readings.

M: Yeah.

S: How do we know whose heart was beating the most?

M: We don't have the heart rates on here [*motioning across a row of data points*]. Unless-

S: [*pointing at the screen*] No, we don't.

While she had described the individual values appropriately as the count of cases moments earlier, during this transaction Molly interpreted Sarah's 365 total to mean that Sarah's heart would have been beating the most of everyone in the workshop. Sarah appeared surprised by this. She stopped what she was doing at the computer to directly engage Molly and reconsider what she knew about heart rates. Recall that earlier we had mentioned that outside of this workshop, Sarah was a competitive figure skater and was recognized by the group as one of the athletic students. Also, recall that earlier, Jeff suggested to the group that more athletic individuals have lower heart rates. We understand this background knowledge as having conflicted with the observation Molly had made. At this point in their analysis, the girls were concluding that Sarah must have had the highest heart rate, and that ran counter to what ideas had been previously accepted by the group about athletic people and heart rates. Sarah began to question what exactly they were computing ("How do we know whose heart was beating the most?"). Molly then reiterated that the values they were reading were not heart rates, and at that point, Sarah appeared to understand what they had been doing wrong.

In reflecting on this second vignette, we wish to make two observations. First, in order to make any sort of comparison between samples, Molly and Sarah seemed to desire a compact description of each person's heart rates. In that regard, they were doing a very sensible activity of trying to take a large set of numbers and reduce them to something simpler. They attempted to produce such a simplification by adding up the numbers along each student row, but realized that answer was in conflict with ideas that they already had.

That conflict is our second observation. In this context and with this familiar data, prior knowledge about one's self – in this instance, Sarah's knowledge of her own athleticism – mediated acceptance of the results. We contend that the data collected here were personally relevant as they were firsthand data the students had themselves collected (Hug & McNeill, 2008), and the numbers therein were also in some way a reflection of themselves as individuals. Each person, in a sense, 'owned' the data they had collected. Had they simply accepted that this was an error and continued with their activities, there would have been no obvious consequences. This workshop was an extracurricular activity which had no grades associated with it.

At the moment when Sarah began to sit upright and evaluate what was being said, we suspect the girls were interested in the answer because of what it would tell them about their previous day's physical engagement. If they were not, they could have, for instance, provided the nonsensical answer and moved on. Or, they could have taken the computation for Stan's data to be confirmation that athletic individuals have lower heart rates. But something more seemed to be at stake. When the girls realized that Sarah had not a lower, but the highest heart rate, the computational method was called into question and the meaning of the numbers they had used was fully reconsidered. Sarah was especially involved in this reevaluation, we believe, because

her position as an athletic individual was being called into question. When that aspect of her identity had become momentarily problematic, Sarah shifted how she looked at the numbers they were using and immediately realized that those numbers were associated with the wrong objects and realized that they needed a new strategy.

Molly and Sarah begin to consider the density of data points

The third vignette follows Sarah and Molly to their new strategy. As will be illustrated below, some aspects of the minimum and maximum value strategy from the first vignette bleed into this episode until the girls attended to different objects in the plots.

Shortly after confusing the number of cases with the number of heartbeats, Sarah and Molly were at an impasse. After Sarah realized that the numbers they were seeing were not telling them whose heart was beating more, the girls reset their plot display and then asked Stan and Jeff how they created the display that they were viewing on their separate computer. Stan called back the attributes and orderings, and the girls then produced the plot in Figure 2. In this plot, the rows correspond to each of the four students, with Sarah's readings on top, Stan's next, Molly's below that, and Jeff's at the bottom. The horizontal axis was labeled for the value of the heart rate measurements. The data points in each row were separated by activity. The purple (top set in each row) corresponded to HORSE. The green (bottom set in each row) corresponded to Frisbee.
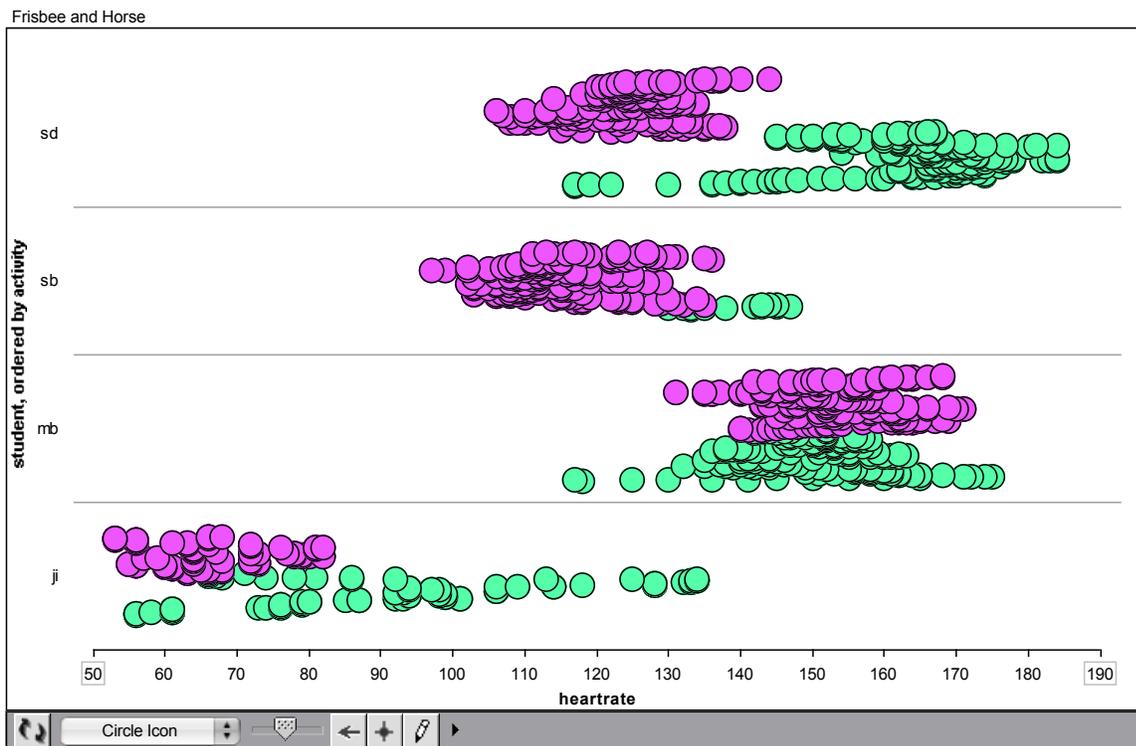
Figure 2: Heart rate plot, vertically separated by student and ordered by activity from *TinkerPlots* (Konold & Miller, 2005) prepared by Molly and Sarah. Sarah's heart rate data are in the top row. Stan's data are in the second row. Molly's data are in the third row. Jeff's data are at the bottom.

Once they had generated their plot, the girls quietly talked amongst themselves to reason about the meaning of the colors. Molly then traced her index finger from the horizontal axis up to the rightmost points in each row. During this, the pair verbally confirmed that Sarah did work harder in Frisbee than HORSE because her maximum was higher. Doing the same for Molly's data, however, introduced a different conclusion.

S: You worked kind of even on both of them, but you worked a little bit harder on Frisbee. Your highest was 175.
M: I thought I worked hardest on HORSE!
S: Oh yeah!
M: *[Molly groans and puts head in hands]* Now I'm confused.

At that moment, it appeared that both girls had returned to their single boundary point strategy. Instead of selecting the minimum value for each student and making inferences from that as they did in the first vignette, they were selecting the maximum heart rate value, which was the rightmost point in set for the rows. However, the conclusion that Sarah reached conflicted with what Molly had concluded about herself earlier, based on her read of the data in the previously used representation and her recollection of the previous day's activities – that she had a higher heart rate in HORSE. In response to the groaning, the facilitator moved over again to the girls' workstation to engage with them.

VL: I just heard groaning. What happened?
M: Well now it's saying I worked harder on Frisbee but I worked harder on HORSE.
VL: How does it say you worked harder on Frisbee?
M: The green is more and it's Frisbee.
VL: So how is the green saying?
S: Well just cause it's higher doesn't mean you worked harder – cause you have more in this section than you have down here. Like you have more like, I don't know
VL: Finish what you're saying Sarah, I think you're onto something
S: Um…
VL: *[prompting Sarah]* She had more?
S: She had more heart rate readings in this section than she did here. Is that what I'm trying to say?
VL: Can you draw in the sections you're talking about? It's just hard for me to tell which one you're talking about.
S: Okay, this section right here *[draws circle over center of top set]* has a lot compared to that one *[draws circle over righthand portion of bottom set]*. Or just cause this right here *[pointing at bottom right]*, she only has a little bit of readings in this section, but she has more cumulative points…I don't know what I'm trying to…
M: What about right there? *[points to dense region in green bottom set of data points directly below top drawn circle]*
S: I don't know.
VL: What do you think Molly? Which one, based on this would you think showed you worked more?

M: I think it shows I that worked harder in Frisbee but I didn't really.
VL: How does it show you worked harder in Frisbee?
M: Because it goes farther *[points to bottom right drawn circle]*.
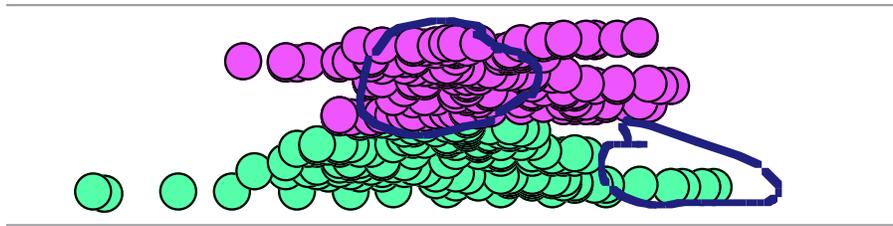
Figure 3: Close-up of data from Molly's row with Sarah's drawn circles overlaid on the plot, reproduced in the *TinkerPlots* software (Konold & Miller, 2005).

When Molly shared her thinking that the plot indeed shows she worked harder in Frisbee, she was attending just to the position of the rightmost data points ("The green is more and it's Frisbee"). Sarah immediately offered a possible explanation for this by saying "Well, just cause it's higher doesn't mean you worked harder". She hinted that the density of points might be important, but began to abandon that thought. The facilitator prompted her to finish and draw in the exact regions in the plot she was referencing. Following that request, Sarah drew in the regions she thought were supporting the statement she was trying to make and reiterated "just cause this right here, she only has a little bit of readings in this section, but she has more cumulative points".

We saw this transaction as evidence that Sarah was on the verge of making an important distinction related to data analysis: the central tendency is an entirely different measure than the maximum. While values on the tails may be important, they do not necessarily determine the trend for the entire set of data. The overall trends are generally better captured by reference to the central tendency. Granted, central tendency would not be the only measure that a statistician would want when describing a distribution (e.g., Lehrer, et al., 2007 describe the importance of characterizing error and variability in statistics education), but is a substantial leap for Sarah in light of some of her earlier strategies and interpretations of the data. This realization, however, was not shared by Molly, who at the end of the excerpt maintained that the data show that she worked harder in Frisbee "because it goes farther". Note also that, like Sarah had in the vignette, the problem arose out of a conflict between how Molly had viewed her athletic activity from the previous day and what the data were telling her.

The girls then spent the next minute quietly staring at the screen while Molly clicked various data points. She then verbalized her concern with Sarah's readings of the data.

M: The only reason I'm confused by what you're saying, is cause right here *[points at dense region below top drawn circle]*…
S: It has like the same amount
M: Yeah it has like the same amount as right here *[points at top drawn circle]*
S: So you worked about even in both of them
M: I did? Yeah, it's true, I did.

Molly reconsidered the data. She then reasoned about the where the greatest density of points lay with respect to heart rate value for each activity. In response to Sarah's statement "so you

worked about even in both of them", she agreed that the centers are about the same for each activity. Sarah's attention had moved from the endpoints of the Frisbee clump to something more like its center. Molly proceeded to write down her response to the original question, but then Sarah made another realization.

> S: Whoa, whoa, see, we can't say that you worked harder here *[pointing at bottom set]* just because it's higher. But then, again it's lower in that one *[pointing to leftmost point in the bottom set]* so we could also say you worked less hard on this one, but that wouldn't be accurate. So basically you worked about the same.
>
> M: Yeah
>
> S: Cause it doesn't matter how far they go.

As part of her interjection, Sarah was explaining that the approach that Molly was taking of looking at max and min values would not work because they would lead to opposing conclusions. That strategy was contingent upon which point is selected, and the conclusion one draws can be crafted to fit that point. Instead, Sarah implied that the two effectively cancelled each other out. She generalized this in her final statement "Cause it doesn't matter how far they go". At that point, she seemingly removed herself from the interpretations she had made in the first vignette as she attended more to the density of points.

For Sarah, this is an important discovery akin to what Konold, et al. (2002) have described as students' references to 'modal clumps'. Modal clumps are essentially plot areas in which the mode of a distribution resides, determined by the presence of several points. Modal clumps have previously been described with histrogram-like representations, but the idea is easily extended here. The clumps exist here where there appears to be greater density. Sarah has realized that there is a way to describe the distribution by attending to where there is a greater density of points and that choosing individual values when considering the aggregate limits what one can say.

Molly agreed with this interpretation, in part because she changed her belief about how much exertion she put out during the two activities. Beyond her initial detection of the problems with discounting dense regions, and her own agreement with Sarah at the end of the above excerpt, Molly later seemed to even more fully appropriate this analytic approach by referencing the dark regions in the plot, where many data points overlap and the closely positioned edges give the appearance of a darker region.

> M: Wait, that's true, that's true…cause look at the dark spots. There's a lot more than that, and there's a lot more than that *[pointing at a darker region, then at a lighter region in Sarah's row]*.

By articulating how the increased density manifests itself through changes in apparent darkness, Molly demonstrated that she too has changed how she viewed the data and has shifted her attention to densities instead of extreme values[3].

---

[3] Because of the nature of the data and how it was collected, we cannot make claims about broad changes to how Molly understood distributions or centers. We do wish to note, however, that after this episode, the girls were careful to attend to middle values in their

Summary and Discussion

Over the course of roughly twenty minutes, Molly and Sarah explored three different strategies for describing heart rates from their PAD data and ultimately attended to something approximating central tendency. First, they began by arbitrarily picking boundary points, reading the values associated with them, and then checking to make sure those values confirmed their existing ideas about what activities required more exertion. They also made broad inferences based on individual points, as was the case when discussing the minimums for Stan's activity readings.

Molly and Sarah continued to misread the data when they computed the total number of heartbeats from the total number of cases. Although they had the initial sense that the values they were using in their computations were for some other purpose, they did not consider that issue seriously until the results conflicted with Sarah's conceptualization of herself as being very athletic. That was an important moment in the girls' analysis of the data as it prompted them to reconsider what was being represented in their plots, and ultimately attempt a new strategy.

Following that, they constructed a new plot representation and began to see that the first method they used to make inferences about the data, in which single boundary points were considered, had some limitations. It was Molly's turn to reconcile her thinking about her exertion during the two activities. Just as Sarah did, she raised specific issues that she needed to resolve in order to accept the results. Sarah was ultimately able to guide Molly's thinking by helping her to see that areas with more points would be more important to consider. By the end of that episode, Molly began to see the data in that way as well. In effect, the girls started to go from a point-based perspective of the data (in which the points were represented by boundary values) to an aggregate-one, in which an emergent property (central tendency) was discovered. The latter was considered to be more useful for describing the distributions that they were examining.

From this case and our understanding of what transpired, we believe that three influences led to this eventual realization. First, expectations were violated. Schank (1999) notes how "expectation failures" prompt interest in a problem or subject and we believe that some version of this is what both Sarah and Molly each encountered during vignettes two and three during this investigation. What is of note is that the expectations that each of the girls had was in relation to their physical and athletic selves. Sarah had certain expectations because of her status as an official 'athlete' in the group of students and based on ideas that had emerged earlier in the week about athletes. That prompted the girls to rethink what quantities were being represented. Molly had formed an image of herself in relation to the activities she had participated in and had expectations regarding her heart rate violated. Those violations prompted the girls to devise entirely new strategies for examining and describing distributions.

These expectations and the subsequent violations of those expectations are productive ones we expect many students to have when using PAD technologies because the values that were being referenced real, recently-lived experiences. To some extent, learning about one's body and how it reacts has an intrinsic appeal, but the real weight behind using personally obtained activity data is that it sets the stage for students to feel confident about claims about their effort or abilities. They are experts on their bodies and efforts. This results in some

---

subsequent analyses. They proceeded to using means as evidence for their interpretations on day 5.

investment on the part of students into the meanings of the values that they are computing and a need to reconcile their ideas and their computations.

The second influence relates to the data visualization tool that the students were using. *TinkerPlots* enables students to create their own visualizations and each visualization offers different affordances and interpretations. While the expectation violations were an important motivator for the discussions that Sarah and Molly had, they could only have come about because of the specific plot representations and the capabilities of the software. Although *TinkerPlots* has a great deal of power in this respect, it alone was not sufficient for supporting the learning that took place for with these two participants. There were occasions in which incorrect ideas emerged and seemed reasonable. For example, the total number of cases happened to be in the range of a reasonable heart rate measurement. Those numbers could have been accepted and the incorrect interpretations maintained had the girls not had intimate knowledge about their own activity efforts or their own general sense of athleticism. For instance, if the data that the girls had inspected were of some other, unknown individuals, they might have stuck with the incorrect computations of total number of cases in vignette two and not given their results a second thought.

Finally, we believe that ambiguity and uncertainty regarding results played an important role. This particular investigation had an uncertain answer and the data were not straightforward. The students had predicted that HORSE would demand more from their hearts, but after the investigation, the students were evenly split. It turned out that some of the students had higher average heart rates than others in each of the two activities. Because the answer was not so obvious upon inspection of the data, there was a genuine interest among the students to find out what the answers would be. In contrast, we observed a lack of comparable mathematical engagement from the students in the video recordings of the final day of the workshop. Granted, this was the final day and the students could have been eager to finish. However, on the final day, the students completed analysis of data collected for the investigations they themselves had designed. We anticipated that the final day would have been the most interesting and satisfying for the students, as they were completing their own projects. In the video of Sarah and Molly of day 5, they raced through their own investigation after seeing that their original hypotheses were correct. The data displays clearly showed that the elliptical machines, in which one is standing and using their arms, required greater activity from the heart than the stationary bicycles where one can sit and rest their arms at their sides. The fact that the Frisbee/HORSE data lent itself to so many interpretations seemed to be both a motivational and conceptual affordance. Those data created a rich enough environment for students' explorations to be worthwhile and required the students to refine their thinking in order to come up with defensible answers.

Conclusion

Our goal in this paper was to identify a new class of data collection technologies that could potentially be used for students' data investigations. Ideally, they would be used in such a way that students engage in authentic practices related to data collection and analysis. We discussed some of the currently available technologies and then described how we explored their use in a designed learning context that centralized data about students' physical activities. In support of these ideas, we presented a descriptive case to show what engagement with physical activity data could look like, particularly with secondary school students.

From a feasibility perspective, we believe a case has been made that further exploration of PAD devices and physical activity data as an object of study can be fruitful. Relatively

speaking, the cost of commercial PAD technologies is comparable or cheaper than some calculators or probeware device suites. These prices will only continue to drop, and newer devices that include more features or are easier to use will appear in the coming years. Beyond their ability to generate a lot of personally relevant data for investigations, PAD devices have the added advantage of being designed intentionally for rough use. For children and technology, this is always an important consideration.

In reflecting on our efforts to design a learning experience, we would like to note that what we had done as a design effort was, relatively speaking, quite modest. We simply gathered a few students, planned a few seed activities, and devised ways of quickly accessing data stored in XML format and moving it into an existing data visualization tool. Granted, there was more nuance in our design, facilitation, and research activities than those steps alone. But relative to large-scale curriculum or platform development efforts, our use of and introduction of students to PAD technology could certainly be replicated and enhanced by others. If, in this single design iteration, we were able to observe some substantive conversation about how to read data and shifts in strategies for describing large sets of data, then an extension of this work could yield new ways for students to learn about and use data. We could perhaps find additional synergies by leveraging reform-based instruction, such as the *Investigations in Number, Data, and Space* elementary math curriculum (Russell, Tierney, Mokros, & Economopoulos, 2004) which spirals students through data description activities or the *I, Bio* middle school project-based science curriculum (Kanter, 2010) that engages students in a multi-week investigation involving their individual energy expenditures through analysis of their own physical activities.

Ultimately, our parting point is that the work described here is simply a first step toward conceptualizing students' experiences with data in new ways. We do not expect that all students will behave as Sarah or Molly did in the case above, but we now know that the potential is certainly there. In the coming years, we expect to continue exploring further the potential of PAD devices as a new form of learning technology. We hope that this article inspires at least a few others to do the same.

References

Bakker, A., & Gravemeijer, K. (2006). An historical phenomenology of mean and median. *Educational Studies in Mathematics, 62*(2), 149-168.

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.

Bers, M. U. (2008). *Blocks to robots: Learning with technology in the early childhood classroom*. New York, NY: Teachers College Press.

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.

Cai, J. (1998). Exporing students' conceptual understanding of the averaging algorithm *School Science and Mathematics, 98*, 93-98.

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.

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

Dyanstream Innovations Inc. (2003). *SpeedMax White Paper* (No. D00000187). Cochrana, AB, Canada.

Finzer, W. (2005). Fathom Dynamic Data Software. Emeryville, CA: Key Curriculum Press.

Hancock, C., Kaput, J. J., & Goldsmith, L. T. (1992). Authentic inquiry with data: Critical barriers to classroom implementation. *Educational Psychologist, 27*(3), 337-364.

Hug, B., & McNeill, K. L. (2008). Use of First-hand and Second-hand Data in Science: Does data type influence classroom conversations? *International Journal of Science Education, 30*(13), 1725-1751.

Jacobbe, T. (2008). Elementary school teachers' understanding of the mean and median. In C. Batanero, G. Burrill, C. Reading & A. Rossman (Eds.), *Proceedings of the ICMI Study 18 and 2008 IASE Round Table Conference.* Monterey, Mexico.

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.

Karvonen, J., Chwalbinska-Moneta, J., & Saynajakangas, S. (1984). Comparison of heart rates measured by ECG and microcomputer. *Physican and Sportsmedicince, 12*, 65-69.

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., & Pollatsek, A. (2002). Data analysis as the search for signals in noisy processes. *Journal for Research in Mathematics Education, 33*(4), 259-289.

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.

Krajcik, J. S., Czerniak, C. M., & Berger, C. (1999). *Teaching Children Science: A Project-Based Approach*: McGraw Hill College Press.

Leger, L., & Thivierege, M. (1988). Heart rate monitors: Validity, stability and functionality. *Physican and Sportsmedicince, 16*, 143-151.

Lehrer, R., Giles, N. D., & Schauble, L. (2002). Children's work with data. In R. Lehrer & L. Schauble (Eds.), *Investigating real data in the classroom: Expanding children's understanding of math and science* (pp. 1-26). New York: Teachers College Press.

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., & Romberg, T. A. (1996). Exploring children's data modeling. *Cognition & Instruction, 14*(1), 69-108.

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.

Marx, R. W., Blumenfeld, P., Kracjik, J., Fishman, B. J., Soloway, E., Geier, R., et al. (2004). Inquiry-based science in the middle grades: Assessment of learning in urban systemic reforms. *Journal of Research in Science Teaching, 41*(10), 1063-1080.

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

Metcalf, S. J., & Tinker, R. (2004). Probeware and handhelds in elementary and middle school science. *Journal of Science Education and Technology, 13*(1), 43-49.

Mevarech, Z. (1983). A deep structure model of students' statistical misconceptions. *Educational Studies in Mathematics, 14*(4), 415-429.

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

Mokros, J., & Tinker, R. (1987). The impact of microcomputer-based labs on children's ability to interpret graphs. *Journal of Research in Science Teaching, 24*(4), 369-383.

Nemirovsky, R. (1994). On ways of symbolizing: The case of Laura and the velocity sign. *Journal of Mathematical Behavior, 13*(4), 389-422.

Nemirovsky, R. (2002). On guessing the essential thing. In K. Gravemeijer, R. Lehrer & L. Verschaffel (Eds.), *Symbolizing, modeling, and tool use in mathematics education* (pp. 233-255). Dordrecht, Netherlands: Kluwer Academic Publishers.

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

Papert, S. (1980). *Mindstorms : children, computers, and powerful ideas*. [New York?]: Basic Books.

Pollatsek, A., Lima, S., & Well, A. D. (1981). Concept or computation: Students' understanding of the mean. *Educational Studies in Mathematics, 12*(2), 191-204.

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., 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.

Resnick, M., Martin, F., Sargent, R., & Silverman, B. (1996). Programmable Bricks: Toys to Think with. *IBM Systems Journal, 35*(Nos. 3 & 4).

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.

Russell, S. J., Tierney, C., Mokros, J., & Economopoulos, K. (2004). *Investigations in number, data, and space.* Glenview, IL: Scott Foresman.

Schank, R. C. (1999). *Dynamic memory revisited* ([2nd]. ed.). Cambridge ; New York: Cambridge University Press.

Sipitakiat, A., Blikstein, P., & Cavallo, D. P. (2004). GoGo board: Augmenting programmable bricks for economically challenged audiences. In Y. Kafai, W. Sandoval, N. Enyedy, A. Nixon & F. Herrera (Eds.), *Sixth International Conference of the Learning Sciences* (pp. 481-488). Santa Monica, CA: Lawrence Erlbaum Associates.

Strauss, S., & Bichler, E. (1988). The development of children's concepts of the arithmetic average. *Journal for Research in Mathematics Education, 19*(1), 64-80.

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

Watson, J. M., & Moritz, J. B. (2000). The longitudinal development of understanding of average. *Mathematical Thinking and Learning, 2*(1&2), 11-50.

Wolf, G. (2010). The data-driven life. *New York Times Magazine.* Retrieved from http://www.nytimes.com/2010/05/02/magazine/02self-measurement-t.html?_r=1&ref=magazine

Wu, H.-K., & Kracjik, J. S. (2006). Inscription practices in two inquiry-based classrooms: A case study of seventh graders' use of data tables and graphs. *Journal of Research in Science Teaching, 43*(1), 63-95.