

Making Sense of Data:

Context Matters

Gain insight into the ways that students reason about measurement units and use data to draw conclusions.

Jennifer N. Lovett and Hollylynn S. Lee

“

d

“Data are not just numbers, they are numbers with a context” (Cobb and Moore 1997, p. 801). Take, for example, these data:

32, 20, 38, 35, 36, 28, 68, and 33.

What possible contexts could students generate from these numbers? Could these values represent the pounds lost on a weight loss program, the ages of new mothers, or the length of women’s hair? How would you rea-

son about the data values, given each of these contexts?

If the context is weight loss, the data values would represent the pounds lost. All the values seem plausible and reflect the facts that most people lost around 35 pounds but that some lost more than others. In the context involves the age





of new mothers, one data value (68) is potentially unreasonable. Was a data entry mistake made or did a woman adopt a child (since biologically it is extremely unlikely for a woman to give birth at age 68). If the context is length of women's hair, we should stop and ask ourselves what unit of measurement makes sense for these data. Given our tendency in the United States to use customary units, we may first assume that the numbers reflect a measurement in inches. It is hoped that we could think about these values and realize that a different unit of measurement is more likely, such as centimeters. In all these cases, the way we reason about the seven numbers is different if we understand the context of the data.

To draw conclusions from data, we need to understand the context of the data and the measurement units. In 2007, Franklin and colleagues presented the *Guidelines for Assessment and Instruction in Statistics Education (GAISE) Report* to provide teachers with a framework to assist in developing students' investigation of statistical questions in an investigative cycle. Although the role of context and the use of measurement units for data are important components, they continue to be underemphasized in many curricula (Franklin et al. 2007).

As teachers, are we providing opportunities for students to engage in this type of reasoning? Too often, the answer is "no." The Common Core State Standards for Mathematics (CCSSM) recommends using tasks in which students consider the context and in which measurement units of data align with a sixth-grade standard that encourages "describing the nature

JACK HOLLINGSWORTH/THINKSTOCK

Fig. 1 One student's TinkerPlots data card showed a variety of data values.

● case 1 of 25 ◀▶

Attribute	Value	Unit	Formula
A	16		○
B	10		○
C	7		○
D	7		○
E	2		○
F	0		○
G	0		○
H	2		○
I	5:00		○
J	8		○
K	0		○
L	27		○
M	7		○
N	8		○
O	11:30		○
P	6		○

of the attribute under an investigation, including how it was measured and its units of measurement” (CCSSI 2010, p. 45). Although others (e.g., Groth and Bargagliotti 2012) have provided more recent guidance for teachers on how the GAISE framework can be used to complement the CCSSM’s Standards, more attention is still

needed concerning the role of context and measurement units that can support students’ statistical reasoning. This article will examine how data collected from a personal information survey provides students such an opportunity to reason about data by considering the units of measurement and reasonable values within a context.

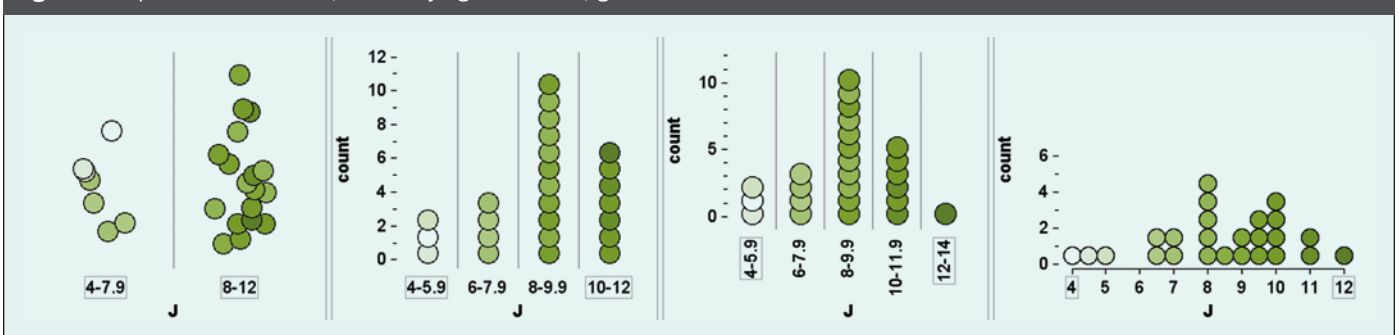
THE DATA TASK

Jennifer Lovett taught the lesson to several sixth-grade classes during their regular sixty-minute mathematics class in a middle school in the southeastern region of the United States. The students’ work presented in this article came from one class of 25 students (15 boys, 10 girls). Hollylynne S. Lee was an observer during the class session and took field notes of the students’ and teacher’s work and interactions. These students had previous experience in writing statistical questions, constructing graphical representations, and interpreting graphical representations throughout their statistics unit. However, they had never used TinkerPlots™ (Konold and Miller 2011), software for exploring data.

The lesson was based on a task by Garfield and Ben-Zvi (2008). The original task was designed to help introductory statistics students develop an understanding that different statistical questions produce different types of variables. In this task, students completed a personal information survey in class, and then a question number from the survey was taped to their backs. Students asked their peers to respond to the question; from those responses, the students had to figure out which question was taped to their back.

This task was modified with the goal for students to explore survey questions, types of variables that questions produce, and measurement

Fig. 2 Four plots for attribute J, with varying bin widths, gave different information.



units; they were also to reason about expected data values. To begin, students completed the personal information survey (see the **activity sheet**) containing 16 questions. Survey questions were chosen because the answers would produce different types of data values and measurement units, such as whole numbers, decimals, and time values. Some questions also required students to answer a categorical question numerically (e.g., “What is your gender?” was answered using 0 for male and 1 for female). Data from the survey were used to create a data set in TinkerPlots in which the 16 attribute names were labeled A, B, C, and so on, rather than names such as “shoe size” or “bed time” (see **fig. 1**). The order of the attribute list was also randomized so as not to match the order of the questions on the survey.

TinkerPlots was incorporated into this lesson because it allowed students to quickly make graphical representations with data so that they could focus on interpreting the data and drawing conclusions. The software allowed students to drag and drop attribute (variable) names from a data card or data table onto an axis of a plot. However, instead of generating the appropriate plot, as with other technology tools, students constructed their own representations by placing an attribute name on an axis and then separating and stacking the data. When a user would drag a quantitative attribute to the x -axis, the data were divided into two intervals, known as bins, by default. **Figure 2** shows four plots with different bin widths and how the representation changed as the data were separated and stacked.

TinkerPlots also allowed users to dynamically link representations to make connections between representations. For example, when a data point was selected in one plot, that case was highlighted in any other visible plot, as well as that case card

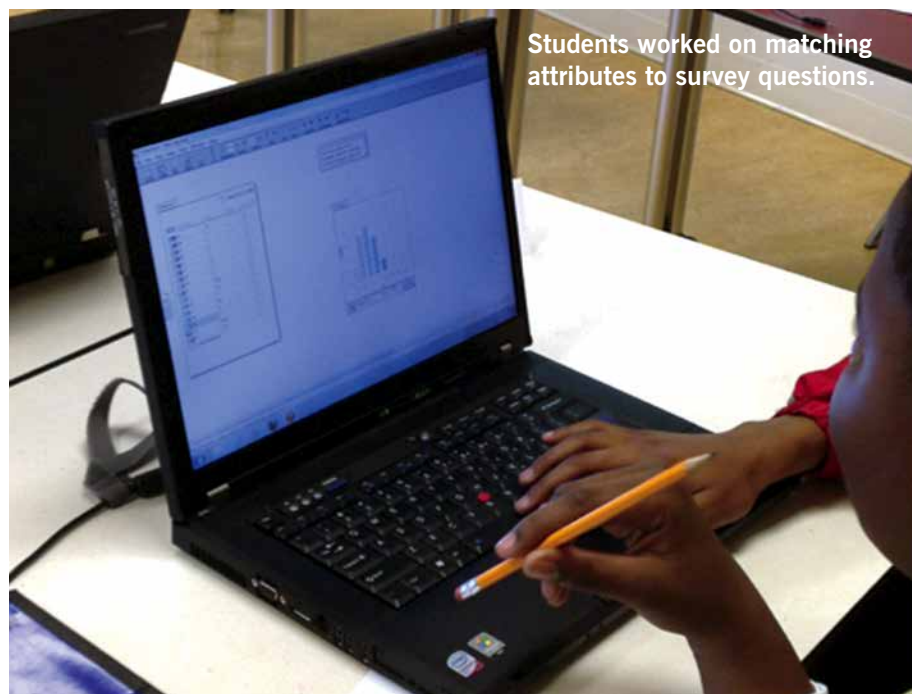
The teacher asked students what was unusual about the information on the cards. “The question numbers were not on the cards,” they said.

shown in the data cards. All measures computed on a data set were likewise updated if a data point was changed through dragging (using the drag tool in the plot window). For example, Hudson (2012) used dynamic linking to have students explore the relationship between data values and the mean as a measure of center.

Several days after completing the survey, student pairs worked on a laptop with the TinkerPlots program. To get acquainted with TinkerPlots, students first opened a fictitious data set and were shown the organization of data in the data cards; how to create plots of data, including different-size bin widths; and a dot plot. While examining the cards, the teacher asked the students what was unusual about the information. Students responded

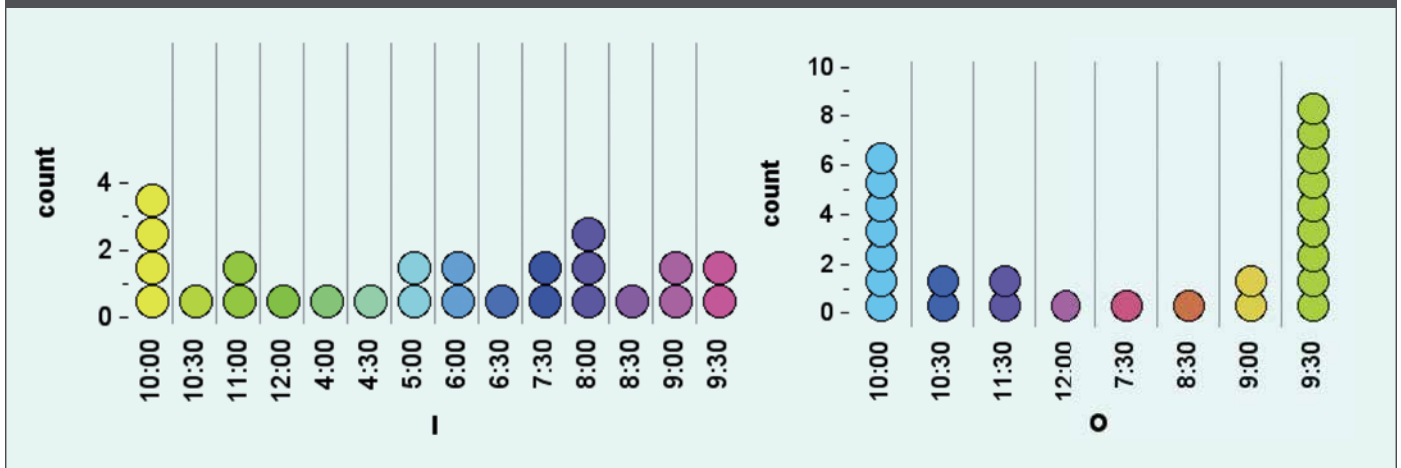
that the question numbers were not identified on the cards. This helped establish that students should not expect to see question numbers as variable names that matched certain questions from the survey.

Students then opened the file with their survey data (25 cases of 16 attributes). They also had a copy of the survey with the 16 questions. Each pair of students was assigned two attributes to examine, asked to make a conjecture of which survey questions the two attributes most likely came from, and write a justification to support their reasoning. Students had the freedom to use TinkerPlots in whatever ways they needed to reason which question was the source of the data. Around the classroom were 16 posters, one for each question from the survey.



JENNIFER N. LOVETT AND HOLLYNNE S. LEE

Fig. 3 Ciara and Gabriella compared attributes I and O.



Students were required to write down evidence they used to support their claim. Therefore, all students were engaged in constructing a claim, identifying evidence to support it, and making their claim public to the class (Standard for Mathematical Practice 3; CCSSI 2010). As the class started finishing the task and identifying the question sources of their attributes, the teacher issued a challenge to the class to now pick a question from the survey and identify the attribute that corresponded to it.

STUDENTS' REASONING

To approach this task, students had to familiarize themselves with the data in TinkerPlots and the survey they had taken earlier in the week. Students read through the questions on the survey provided to them or on the posters hanging around the room and explored the data cards. They made the connections that the 16 different attributes had to come from the 16 questions on the survey that they took and that the number of cases in the dataset ($n = 25$) matched the number of students in the class. As students engaged with the task to determine the question that corresponded to their assigned attributes, three different approaches emerged.

Examining Data and Properties of Data Values

The most common approach used by students was first, to examine a graph of an attribute, then identify possible survey questions that could produce similar data. For example, Ciara and Gabriella examined attribute O, made a plot, and stacked it to view the data (see fig. 3). To describe the data in the graph, Ciara and Gabriella wrote, "All of the answers are time measurements." They then looked over the survey and crossed out any question that did not produce time measurements. They identified only two questions on the survey that were answered in terms of time, "What time did you go to bed last night?" and "What time did you wake up Saturday morning?" The students could not decide which question matched attribute O, so they decided to look for the other attribute that had time measurements, which was not one of their assigned attributes. By using the data cards, they found that attribute I had time measurements, as well, and then graphed those data (see fig. 3).

Ciara and Gabriella discussed the data points from both graphs in the context of both questions. Although the data appear numerical, they were being treated as a categorical attribute

in the software, with each "time" category appearing in the data having its own bin. These bins were not ordered as one may have expected, from low times (e.g., 4:00) to high times (e.g., 11:00). Thus, students had to search through the x -axis to find various times. They discussed the range of times represented, whether these were likely or unlikely times to wake up and go to bed, and how Saturday morning played a role in the context of the problem. After several discussions, Ciara and Gabriella decided it was more likely that attribute I matched "What time their classmates woke up on Saturday morning?" and that attribute O must be "What time they went to bed last night?" Their justification was that "there are a lot of early answers [in I] and kids our age don't go to bed that early but we get up that early."

Another example of this approach was Kim and Keisha investigating attribute J. They created a plot for attribute J and changed the bin width to view the data in ordered subgroups (see fig. 4). This attribute was a quantitative (or numerical) measurement that contained both whole numbers and decimal values, but these were not evident in the binned plot that the students had made. By examining the data cards, Kim and Keisha noticed

that this attribute had some data points that were decimal values, specifically, 4.5, 7.5, 8.5, and 9.5. They examined the survey and identified two possible questions for attribute J, “What is your shoe size?” and “How many hours of sleep did you get last night? (Round to the nearest 1/2 hour).” Kim and Keisha decided that data from attribute J matched hours of sleep because “some people don’t get a lot of sleep and some people do. It seems that most kids get 8 to 9 hours of sleep because it says 8–9.9 has the most dots.”

This approach allowed students to narrow a focus to possible sources of data to a few questions from the survey. However, students had to use the context of the question, and sometimes examine additional attributes not assigned to them, to form an argument about their proposed matches.

Examining Questions for Expected Data Values

A second approach, taken by Nassir and Ebba, involved examining the questions and identifying possible attributes that produced data fitting the context. On the survey, Nassir and Ebba listed possible attributes for each question, staying organized and eliminating options when they believed they had found a match. They graphed multiple plots on their screen to view multiple attributes at the same time. They listed several possible attributes and could not make a decision between the attributes. However, with the question “In what month were you born? (January = 1, February = 2, . . . December = 12)” they were able to identify the attribute as D because it was plotted on their screen (see fig. 5). Their justification was that “all the answers lie between 1 and 12, which is the same number of months in a year.”

Xander and Carlos also used this approach when trying to identify the attribute matching the question

“What is your gender? (1 = Male; 0 = Female).” First, the students had to understand that the categorical answers of male and female were coded with numeric values of a 0 or a 1. This allowed them to narrow the choices to attributes C and F since they both had “answers of 0s and 1s.” Xander and Carlos plotted both attributes next to each other to make the decision of which one matched the gender of students in their class (see fig. 6). To determine the attri-

bute, Xander looked around the room and counted the number of male and female students. He counted more male students than female students, so he identified attribute C to match the question “What is your gender?”

This approach allowed students to reason about the attributes that matched questions from the survey. Some survey questions followed this approach more than others, which helped students to be confident in their claim. However, this approach

Fig. 4 Kim and Keisha scrutinized attribute J.



Fig. 5 Nassir and Ebba plotted data for attribute D.

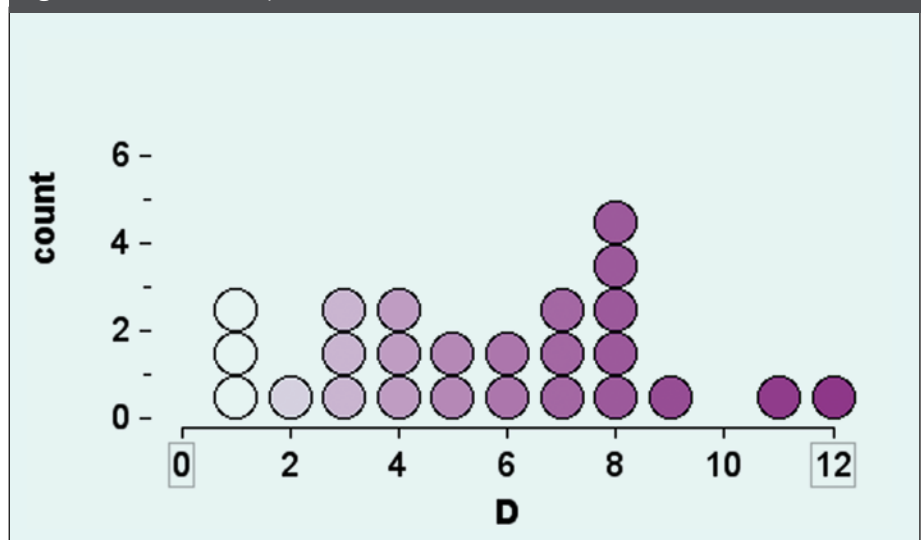


Fig. 6 Xander and Carlos compared plots of attributes C and F to decide which one denoted gender.

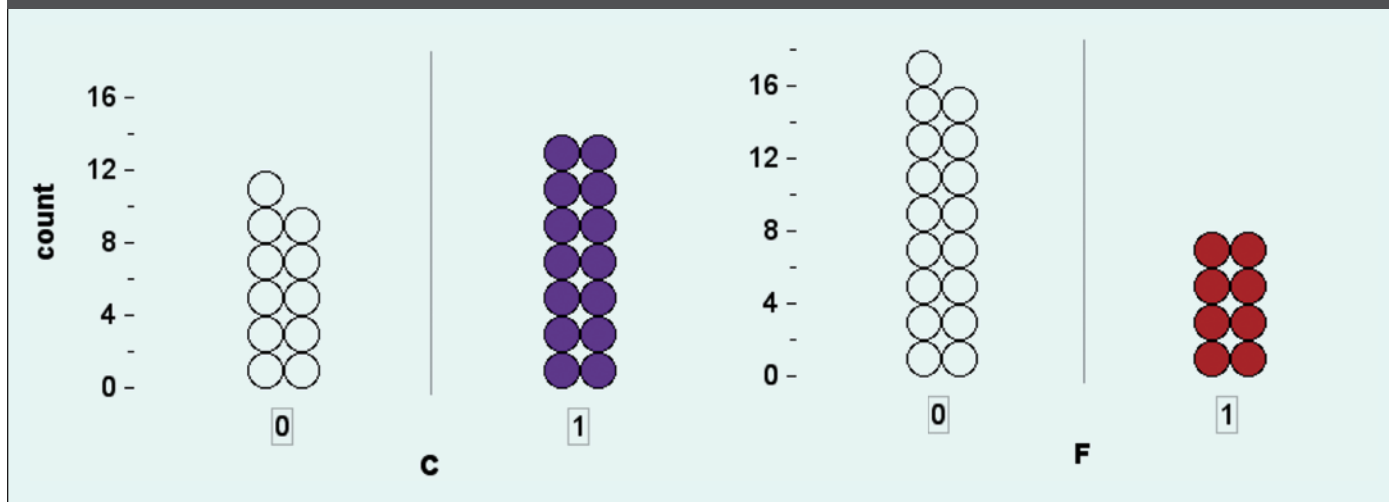
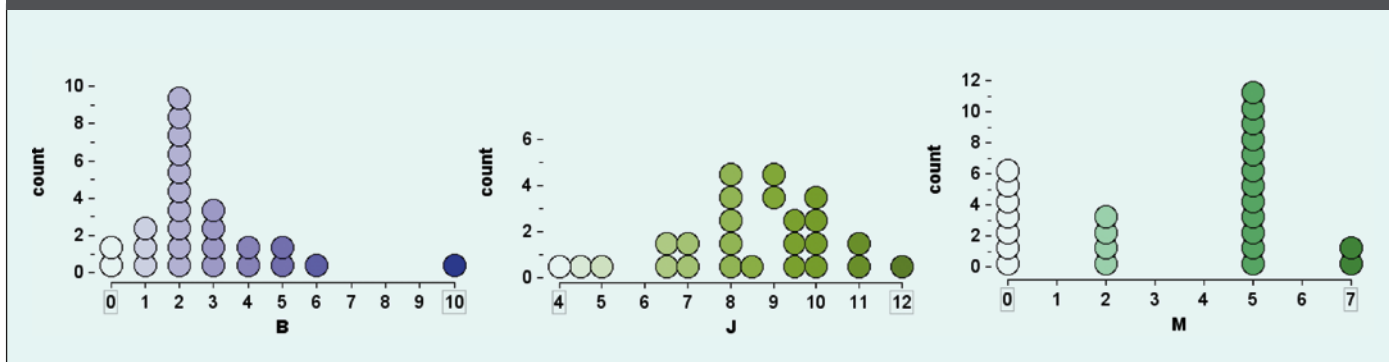


Fig. 7 Graphs of attributes B, J, and M illustrated the lunch at school question.



resulted in questions having many possible matching attributes, and students were unable to form an argument and draw a conclusion (e.g., “How many movies did you watch in the theater last month?” or “How many siblings do you have?”).

Examining an Individual's Data

Instead of taking a graphical approach to examining the attributes or the questions, like many students in the class, Jamal and Teresa took a third approach and decided to explore the data cards in TinkerPlots (see **fig. 1**). They used the arrows at the upper right of the data cards to examine several cases in the “stack” of data cards for data values for attributes A and J. Teresa located a card, examined the values for each attribute, and was convinced she had

found her card. The teacher challenged Teresa, asking her how she knew it was her card. To support her claim, Teresa showed the teacher several convincing data values, specifically, the values of attributes she believed matched her responses to questions about the number of letters in her first name, day and month she was born, and shoe size. The teacher then challenged both students, Jamal and Teresa, to locate Jamal's case card. They believed they were successful in finding Jamal's card. Teresa and Jamal examined other cards and were able to make claims about the source for 10 out of the 16 attributes before class ended.

This approach taken by Teresa and Jamal, looking at the attributes and questions in the collective to find their individual data, was unanticipated.

This type of reasoning seemed to be an effective way to reason to solve this task. However, this approach to the task would not have been possible if the data had not come from students' responses to the survey.

CLASS DISCUSSION

Toward the end of class, the teacher brought the class back together and asked students to explain their claim, justify their choice using the data as evidence, and critique one another's reasoning. In a few cases, disagreements ensued about which attribute corresponded to which question, such as occurred with the question “How many times did you buy lunch at school last week?” The students provided three different possible matching attributes: B, J, and M (see **fig. 7**). For each claim,

the students presented their arguments and then the teacher polled the class to see if the class reached a conclusion. Most of the time, students were able to reach a conclusion. However, for some questions, the class did not reach a consensus. Students were not given the question-attribute correspondences. This was done purposefully to help students develop the statistical habit that justifying their claim with evidence was important, rather than emphasizing whether they were right or wrong.

IMPACT OF TECHNOLOGY

TinkerPlots provided students an opportunity to engage in three different approaches to the task. It allowed students to quickly construct not only graphical representations of the attributes they were assigned but also other attributes that they

Students were able to construct graphical representations easily and quickly; therefore, they could spend the majority of the task reasoning and drawing conclusions.

wanted to explore. Graphing data on paper would likely inhibit such an approach. Because students were able to construct graphical representations easily and quickly, they could spend the majority of the task reasoning and drawing conclusions about the context for each attribute in the data.

The original version of this task presented students with data for only one attribute, and the students had to draw a conclusion from that limited amount of information. However, incorporating TinkerPlots for this

task gave students access to the entire data set, allowing them to participate in a more open investigation of all the attributes and to explore other attributes, so that they could reason and draw conclusions about the two they were assigned (e.g., exploring attributes I and O to determine which question corresponded to I). Many pairs of students chose to continue to explore and draw conclusions about other attributes once they had made claims about the first two they were assigned.

Equip Students to Make Strong Financial Decisions

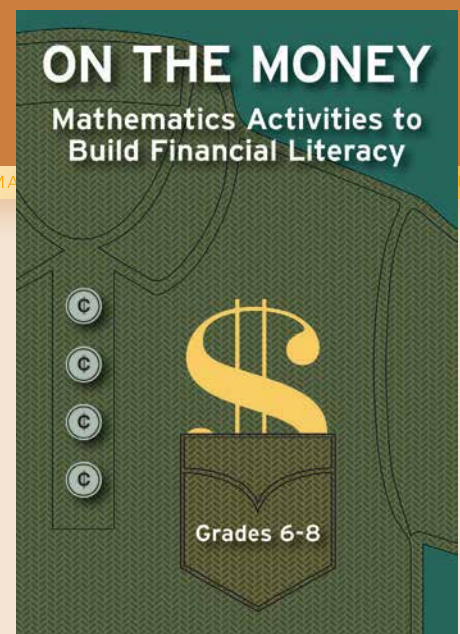
MATH IS ALL AROUND US | MATH IS ALL AROUND US | MATH IS ALL AROUND US | MATH IS ALL AROUND US

NEW | On the Money: Math Activities to Build Financial Literacy, Grades 6-8

BY JENNIFER M. BAY-WILLIAMS, SARAH B. BUSH, SUSAN A. PETERS, AND MAGGIE B. MCGATHA

More than half of today's teens wish they knew more about how to manage their money. Students who develop financial literacy are equipped to make better financial decisions—about budgeting, saving, buying on credit, investing, and a host of other topics. Math is essential to money management and sound financial decision making, and activities in this book draw on and extend core concepts related to ratios and proportions, expressions and equations, functions, and statistics, while reinforcing critical mathematical practices and habits of mind. The authors show how the activities align with the Common Core State Standards and the Jump\$tart Financial Literacy Standards.

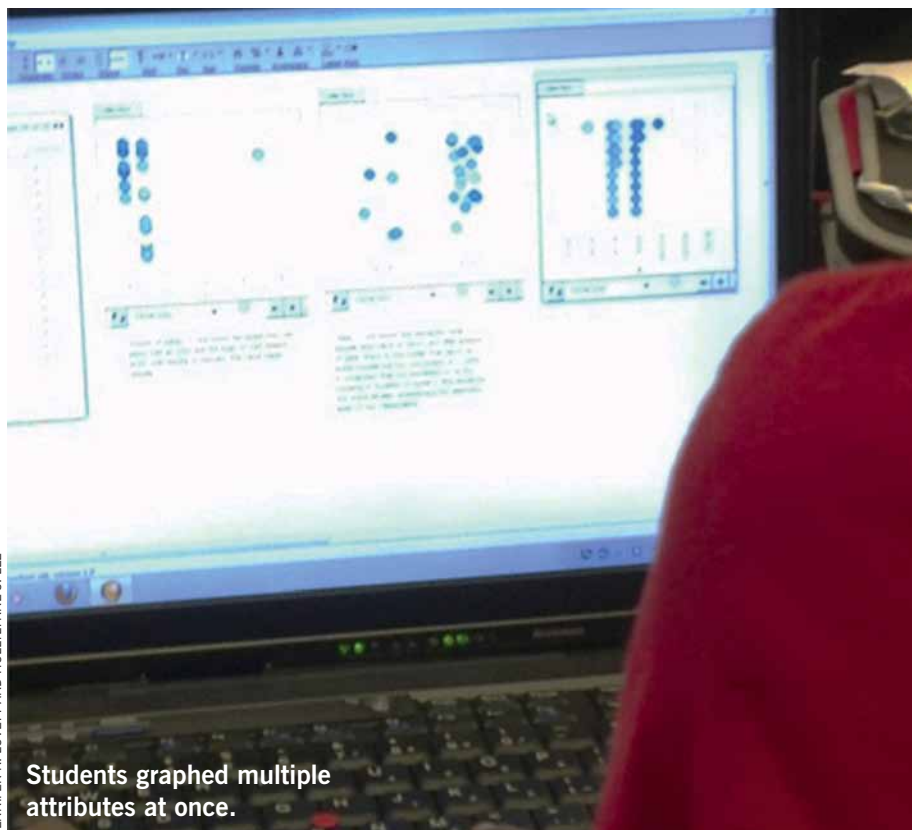
©2015, Stock #14534



NATIONAL COUNCIL OF
TEACHERS OF MATHEMATICS

Visit nctm.org/store for tables of content and sample pages.

For more information or to place an order, call (800) 235-7566 or visit nctm.org/store.



Students graphed multiple attributes at once.

THE IMPORTANCE OF CONTEXT

Our experience with this task emphasized the importance of providing students opportunities to explicitly reason about the context of data, including anticipation of reasonable values. It allowed students to consider multiple variables, construct an argument, and critique the arguments of others. Incorporating TinkerPlots also allowed students to focus on reasoning and decreased the time of constructing graphical displays.

This task is just one example of how teachers can engage their students in reasoning about measurement units and expected values of data. When students engage in these types of tasks, they are developing the statistical habits of making an argument and using data to support their claims. The Common Core recognizes these statistical habits as a sixth-grade standard; however, students need to develop such reasoning throughout middle school and high school. The

next time that you use precollected data in your classroom, take the time to challenge your students to reason about the context because without that understanding, data are just numbers.

REFERENCES

- Cobb, George W., and David S. Moore. 1997. "Mathematics, Statistics, and Teaching." *The American Mathematical Monthly* 104 (9): 801–23.
- Common Core State Standards Initiative (CCSSI). 2010. Common Core State Standards for Mathematics. Washington, DC: National Governors Association Center for Best Practices and the Council of Chief State School Officers. http://www.corestandards.org/wp-content/uploads/Math_Standards.pdf
- Franklin, Christine, Gary Kader, Denise Mewborn, Jerry Moreno, Roxy Peck, Mike Perry, and Richard Schaeffer. 2007. *Guidelines for Assessment and Instruction in Statistics Education*

- (GAISE) Report: A Pre-K–12 Curriculum Framework. Alexandria, VA: American Statistical Association.
- Garfield, Joan, and Dani Ben-Zvi. 2008. *Developing Students' Statistical Reasoning: Connecting Research and Teaching Practice*. New York: Springer.
- Groth, Randall E., and Anna Bargagliotti. 2012. "GAISEing into the Common Core of Statistics." *Mathematics Teaching in the Middle School* 18 (August): 39–45.
- Hudson, Rick A. 2012/2013. "Finding Balance at the Elusive Mean." *Mathematics Teaching in the Middle School* 18 (December/January): 300–306.
- Konold, Cliff, and Craig D. Miller. 2011. *TinkerPlots: Dynamic Data Exploration*. Emeryville, CA: Key Curriculum Press.

Any thoughts on this article? Send an email to mtms@nctm.org.—Ed.

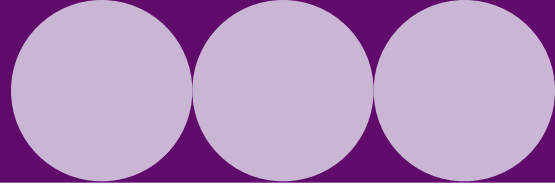


Jennifer N. Lovett, jnickel@ncsu.edu, is a doctoral candidate at North Carolina State University, Raleigh. **Hollylynn S. Lee**, hollylynn@ncsu.edu, is a professor in the Science, Technology, Engineering, and Mathematics Education Department at North Carolina State

University. Both authors are interested in students' reasoning about statistical ideas when using technology tools and how secondary mathematics teachers effectively use such tools with students.



The **activity sheet** on p. 347 is also available online at www.nctm.org. More4U content is available to NCTM members only.



Name _____

PERSONAL INFORMATION SURVEY

1. How many pets do you have in your family? _____
2. What time did you go to bed last night? _____
3. How many hours of sleep did you get last night? (Round to the nearest 1/2 hour.) _____
4. How many letters are in your first and last name? _____
5. How many books have you read this school year? _____
6. What is your gender? (1 = Male; 0 = Female) _____
7. How many siblings do you have? Include half-siblings and stepbrothers and sisters. _____
8. How many movies did you watch in the theater last month? _____
9. How many times did you buy your lunch at school last week? _____
10. What time did you wake up last Saturday morning? _____
11. How many glasses of milk do you drink daily? _____
12. Do you wear glasses? (1 = Yes; 0 = No) _____
13. On what day of the month were you born? _____
14. How many letters are in your first name? _____
15. In what month were you born? (January = 1; February = 2; . . . December = 12) _____
16. What is your shoe size? _____