Nesrin Cengiz and Theresa J. Grant

Children Generate Their Own Representations

"How many teeth have *you* lost?"

Imagine second graders collecting data from their peers about how many teeth each child is missing and then creating their own data representations that mean something to them. Instead of showing her students how to create a bar graph of the data, this teacher asks them to display the information in some way that helps them make sense of their data. Students work in small groups, discuss their ideas with peers, and create their own data representations. Then they share their data with the entire class, discussing differences in their representations and interpretations.

This classroom scene stands in stark contrast to conventional instruction on data analysis, which focuses on teaching students how to create and read particular representations (e.g., pictographs, bar graphs) and on determining the data's mean, median, or mode—but not necessarily contextually connecting these measures (Russell 2006). Conventional instructional goals may encourage students to think about representation features, such as title, axis, and key, and to use algorithms to compute the measures of center; however, little focus may be brought to bear on the data themselves.

Why might children benefit from a different approach to data analysis? Having students generate their own data representations and interpretations better prepares them to make sense of an increasingly data-driven society and thus become "informed citizens and intelligent consumers" (NCTM 2000, p. 48). Furthermore, research suggests that the more active a role that students have in constructing their own knowledge, the more useful that knowledge is. For example, when students are simply *shown* steps for drawing representations, rather than creating their own, they often misinterpret the information in those representations (Economopoulos and Wright 1998; Konold and Higgins 2003; Russell 2006).

We illustrate this perspective by sharing our observations of elementary school students who learned to make sense of data via this new approach. This article describes ways that students represent data when they are encouraged to create their own displays and discuss the issues they grapple with during this process. Additionally, we share ideas on how teachers can use student-generated representations to extend student thinking about representing data in ways that illuminate the data's important features.

Student-Generated Representations

When youngsters are encouraged to create their own data representations, they use a variety of forms. In this section, we describe the most prevalent representations that we have observed in elementary classrooms, along with an indication of how the tasks influenced students. For the lower elementary grades that we observed, instruction was guided by the unit *How Many Pockets? How Many Teeth?* (Economopoulos and Wright 1998), which focuses on helping students learn to analyze numerical data, and more specifically on assisting students with finding ways to count and keep track of data, to create their own representations of data, and to use those representations as a way of communicating information. In the upper elementary grades, students had previous experiences with some aspects of data analysis, such as drawing bar graphs and pie graphs, and were engaged in *The Shape of the Data: Statistics* unit (Russell et al. 1998), which focuses on recording, analyzing, representing, and comparing data sets about real-life situations. We invite you to analyze some examples of students' representations in the paragraphs below.

Nesrin Cengiz, nesrinc@umd.umich.edu, and Theresa J. Grant, terry.grant@wmich.edu, are mathematics educators at the University of Michigan–Dearborn and Western Michigan University, respectively. They are interested in how teachers help students learn mathematics with understanding through reasoning about mathematics.

Edited by **Cindy Langrall**, *Langrall@ilstu.edu, a professor in the mathematics department at Illinois State University in Normal; and* **Dorothy White**, *dywhite@uga.edu, an associate professor in the Department of Mathematics and Science Education at the University of Georgia in Athens. "Research, Reflection, Practice" describes research and demonstrates its importance to practicing classroom teachers. Readers are encouraged to send manuscripts appropriate for this section by accessing* **tcm.msubmit.net.** *Manuscripts should be no more than 2500 words.*

Cube towers

How Many Pockets? How Many Teeth? contains lessons in which students physically represent their own data point by creating a cube tower for a given situation, such as the number of teeth they have lost. Students are prompted to line up their towers on the edge of the board, allowing them to see the lowest and highest values and the number of instances of the same value. They are not shown how to represent the data on paper in any particular way but are encouraged to create representations that make sense to them. We observed a common approach: drawing the entire collection of cube towers. **Figures 1** and **2** show representations that are similar to listing data values. For example, the cube towers (see **fig. 1**) correspond with the following list: 5, 7, 5, 2, 5, 12, 4, 12, 13, 1, 7, 0, 14, 13, 0, 5, 7.

Line plots

A line plot is typically introduced as a quick way to sketch numerical data. In the classrooms that we observed, students were encouraged through whole-group discussions to create informal line plots from cube towers. Students first built cube towers to represent the number of teeth each had lost, wrote that number on a sticky note, and attached the note to the cube tower. For example, a class of seventeen students made seventeen cube towers. Seventeen corresponding sticky notes represented the lost-teeth data from all class members.

Students next arranged the sticky notes—rather than the cube towers. This method of organizing the data spawned displays such as Heidi's (see **fig. 3**), in which the stack of three squares in the center represents the fact that three students had all lost exactly five teeth.

Bar graphs

As mentioned, many upper elementary students had been taught to create bar graphs earlier in the school year. It is therefore not surprising that some of the representations created by these students were variations of a traditional bar graph. **Figures 4a**, **b**, and **c** are representations created by three different groups as they sought to answer the question, "How much taller are sixth graders than second graders?" Some students drew two separate bar graphs, which they placed at opposite ends of their paper (see **fig. 4a**). Others combined the data sets in various ways (see **figs. 4b** and **c**).

Before you continue, reflect on these questions: Which representations might be more useful in making sense of data and answering questions

Figure 1

Todd's drawing of the entire collection of cube towers is a typical approach.

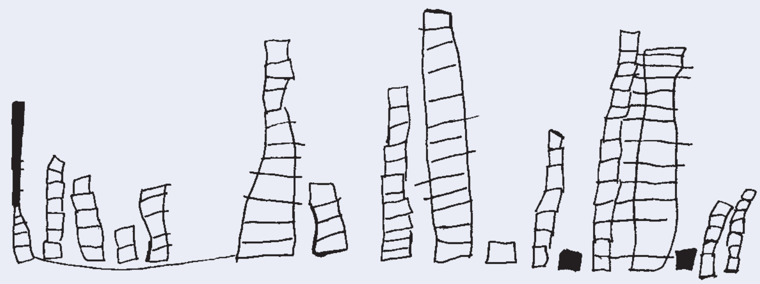


Figure 2

Jane's illustration shows representations that are similar to listing data values.

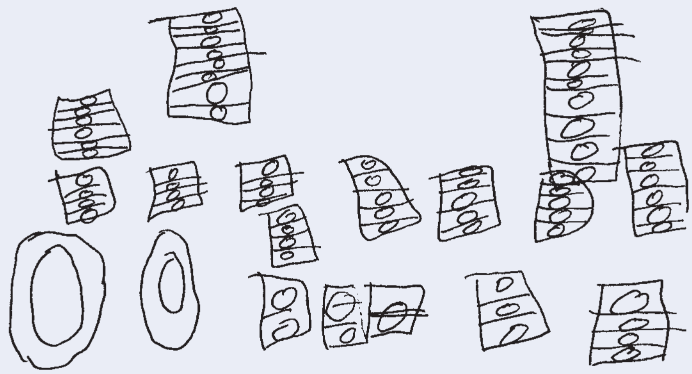


Figure 3

The stack of three squares in the center of Heidi's drawing represents the fact that three students all lost exactly five teeth.

about data? Which aspects of these representations would you want your students to focus on during class discussions? What kinds of issues might students face? What understanding do you want students to develop as they construct and share their own data representations?

Issues That Arise

When students are encouraged to generate their own data representations, they must make decisions that impact how their data story is told: how to present and organize data, keep track of it, and represent gaps. In the next section, we use student-generated data displays to discuss these closely related issues that we observed in classrooms.

Presenting and organizing data

Student-generated representations clearly look different from the standard pictographs and bar graphs—which we have become accustomed to—and may initially be somewhat disconcerting to teachers. For example, young children often present data in drawings with little regard for scale. Consider Todd's representation of cube towers (see **fig. 1**) depicting the number of missing teeth. Although the heights of Todd's cubes are fairly consistent, particularly for a young child, the resulting display provides false information. For example, what appears to be the largest tower (in the center of the display) is partitioned to show thirteen cubes, whereas the greatest number of lost teeth is actually fourteen (represented by the next tallest stack to the right). Similarly, the last tower in the display represents seven lost teeth, but this tower is considerably smaller than the second tower in the display, which also represents seven teeth. Interestingly, issues related to scale have been shown to persist into the upper elementary and middle school grades as students engage in coordinate graphing involving the scaling of axes (Watson 2006).

At first glance, Jane's cube tower representation (see **fig. 2)** seems no different from Todd's in that they both have issues with scale and both appear to be disorganized. However, closer inspection reveals that Jane has organized her data: The bottom left corner shows the two students who have no missing teeth, followed by three students who have lost two teeth, and so on, moving from left to right, row by row, from the bottom of the page to the top. Jane not only put her data points in order but also accounted for the two students who have not lost any teeth. Each cube in a tower represents a lost tooth, so the physical representation is restrictive in the sense that it does not provide a means for showing "zero teeth." Notice how Jane chose a different shape—circular rather than rectangular—for the zero values. This points to another issue that we observed: keeping track of all the data values.

Keeping track of data

Lower elementary students commonly forget either to represent all data values in their diagrams or to insert extra data values that do not belong in the original data set. Keeping track of the data was an issue for upper elementary students as well. They usually represented each data value in their diagrams but not necessarily with the correct frequency. **Figure 4a**, for example, shows two second graders with a height of fifty-six inches; in fact, only one such student existed. **Figure 4c**, on the other hand, indicates one second grader who was forty-seven inches tall; there were actually two.

Although some of these inaccuracies might be viewed as careless errors, the case of representing values of zero poses a particular challenge for young students, as mentioned above. In **figure 2**, Jane literally represented the number of missing teeth by drawing each cube tower; that is, each rectangle (cube) represents a tooth. Thus, she needed to create a "no tooth" icon to be able to indicate that two children had not lost any teeth. In contrast, Heidi represented the *frequency* of lost teeth (see **fig. 3**), where each rectangle stands for a child, and the number of lost teeth is recorded above each stack. Heidi's informal line plot allowed her to represent any number of lost teeth, even zero. Furthermore, some students were challenged to distinguish the case of five students losing zero teeth from that of nobody losing five teeth—an issue that has been shown to persist in upper grades (Konold and Higgins 2003). This brings us to a third related issue.

Representing data gaps

Students must understand when and why it is important to represent gaps in data; this may not be as trivial an issue as one might think. Consider that physical models, such as cube towers, focus on data that *exist*, rather than data values that do not occur. Thus, when students in the classrooms that we observed displayed their cube towers, they simply lined them up, one next to another. When the towers were organized from least to greatest (see **fig. 5**), gaps in the data were discernable. In this display, the large change in the heights of the towers identify gaps; for example, the jump from the seven-tower to the twelve-tower indicates that no students lost eight, nine, ten, or eleven teeth. In contrast, the one-cube difference between consecutive data values—such as from four to seven, where the cube towers look like a staircase—denotes no gap between these data points. That is, students lost four, five, six, and seven teeth. However, students might not easily engage in this kind of analysis of the display and thus would not readily recognize gaps in this type of representation.

The conventional line plot, on the other hand, has a location for every data value between the minimum and the maximum data points, regardless of whether or not these data values are observed. So, gaps in the data are obvious at a glance. When students first encounter the concept of a line plot, however, they commonly focus only on *observed* data and thus create representations like Heidi's (see **fig. 3**). As students have more experiences with this kind of representation, they begin to create informal line plots with gaps (see **fig. 6**) and, eventually, more conventional line plots (such as **fig. 7**) with a scaled axis. Unlike **figure 3**, the latter two figures reveal the data's clusters, peaks, and valleys. Such features that define the shape of the data are important to the story portrayed in a data display.

Extending Students' Thinking

The two main tasks discussed thus far require students to specifically examine different ways of representing the same data set and to find meaningful ways to represent two sets of data on one graph. These tasks not only elicit the issues we have discussed but also provide the means for helping students confront those issues. Encouraging students to compare and contrast one another's drawings and to base data descriptions and interpretations on these drawings were effective instructional strategies in the classrooms that we observed.

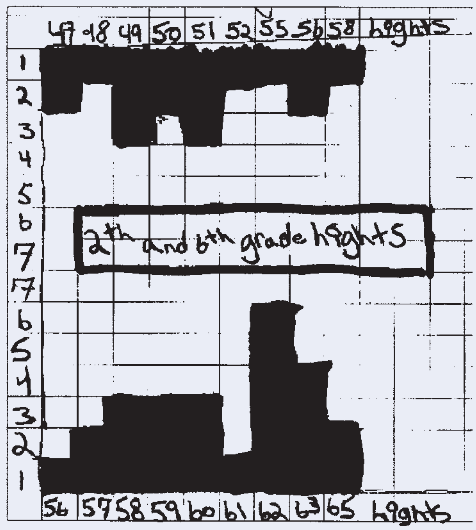
Connections between representations

The young elementary students we observed were exposed to both cube towers and line plots as they learned to collect and analyze data. The shift from cube tower representations to line plots was not trivial for these students, and thus required great attention from teachers (Russell 2006). An effective strategy for facilitating this shift has students verify the displayed information by matching data points in different representations of the same data set and discussing the similarities and differences

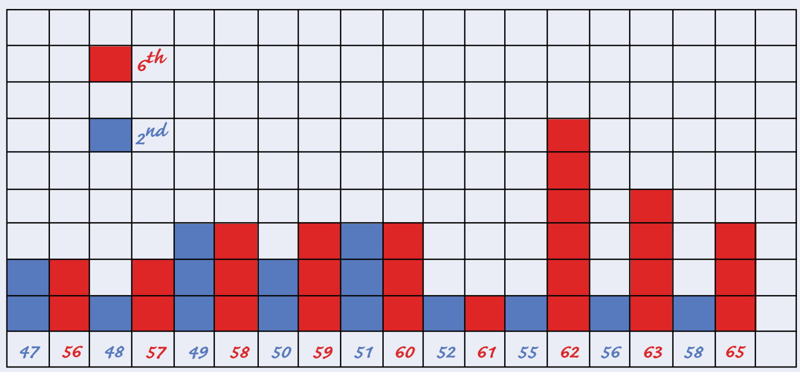
Figure 4

Three different groups sought to answer the question, "How much taller are sixth graders than second graders?"

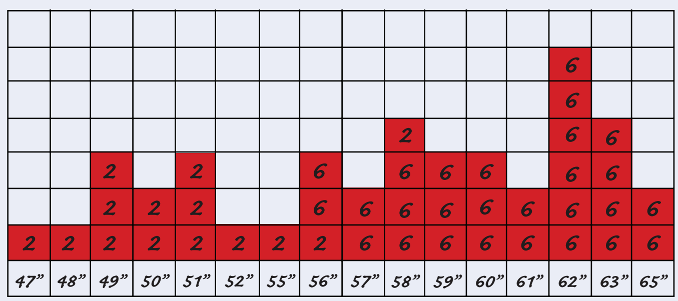
(a) Joe and Hunter placed separate bar graphs at opposite ends of their paper.



(b) Other students combined the data sets. Sally and Jessica alternated data values from two different grade levels.



(c) Melissa and Nick created one bar graph with the data sets combined, labeling each data value to indicate whether it came from a second or sixth grader.



in the various representations (Economopoulos and Wright 1998). For example, in whole-class discussions, teachers prompted students to consider how a tower of seven cubes constructed by a student who lost seven teeth was represented on the line plot by a single sticky note labeled with the number 7.

To make sense of the overall shape of the data, students must recognize the existence of gaps in the data and understand the reasons for representing these gaps in their data displays. Several strategies can help students confront this issue. Research suggests that one source of difficulty with data gaps may stem from the tendency to represent data as they are gathered, rather than organizing and reflecting on the data set before making a plan for creating a representation (Konold and Higgins 2003). Thus, it might be beneficial to explicitly suggest that students take time to look at their data set and make a plan for their representation before starting to draw.

Another effective strategy is to have students discuss the similarities and differences in various representations, particularly when some students highlight gaps in the data and others do not. Finally, encourage students to consider the purpose for representing the data (i.e., what we hope to learn from the display) because different representations can reveal different aspects of data and might be more or less useful in answering the questions at hand.

Two data sets, one representation

In addition to comparing and contrasting different representations of a single data set, upper elementary students created representations in order to compare multiple data sets. In one task, they compared a bar graph of their heights with two bar graphs depicting the heights of students in two other classes. In a subsequent task, students created a single representation of two data sets. These tasks gave students an opportunity to extend their understanding of the issues by creating representations in the context of a new challenge: dealing with two data sets at one time. Once again, whole-class settings were critical for providing a forum to discuss these issues. Consider the following whole-group discussion between a teacher (T) and her students (A–J) using student work from the classrooms that we observed.

T: [On the basis of your representations] what can we say about the heights of second graders and your [heights]? [*See* ***figs. 4a***, ***b***, *and* ***c***.]
A: We are much taller than the second graders.
T: What makes you say that?
A: The shortest person in our class is almost as high as their tallest. Look, the highest second grader is fifty-eight inches, and the shortest sixth grader is fifty-six inches.
T: What do the rest of you think about A's claim?
B: I agree with her. Most of the second graders are shorter than fifty-six inches, and we are as tall as, or taller than, fifty-six inches. This makes us taller.

Figure 5

Ordered cube towers represent the number of lost teeth.

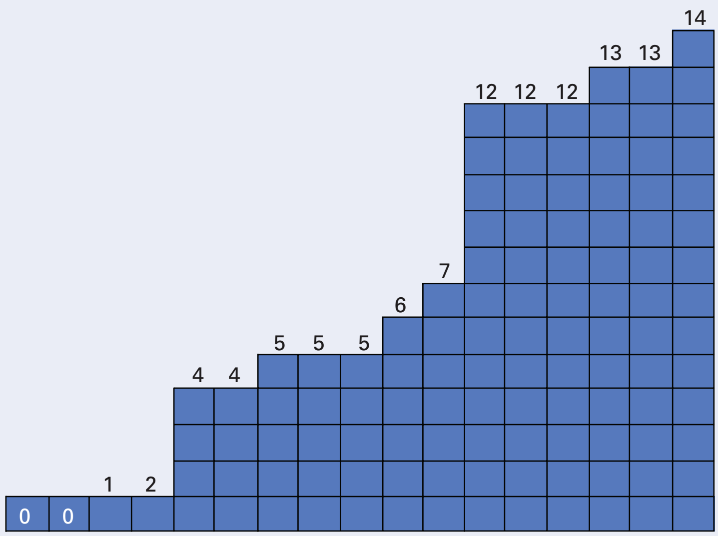


Figure 6

A line plot representation is typically a quick, informal way to sketch numerical data.

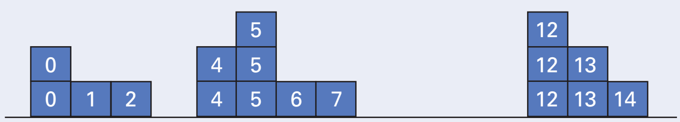
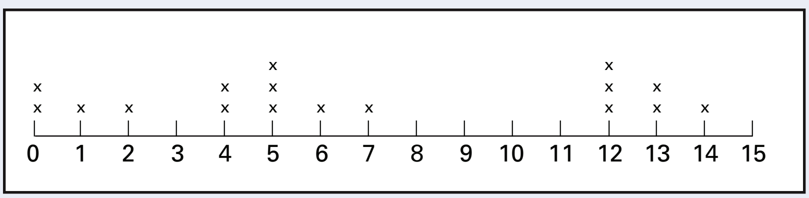


Figure 7

A scaled line plot represents the number of missing teeth.



T: What are other things that you notice about your heights and second graders' heights by looking at these three representations?
C: I see that I am the shortest in our class. I am fifty-six inches tall, and there is a second grader who is taller than I am!
D: Looking at Joe and Hunter's graph [*see* ***fig. 4a***], I can see clumps where most heights are.
T: How about the other two? Can you see where the clumps are in them also?
E: It is harder to see the clumps in Sally and Jessica's [*see* ***fig. 4b***].
T: What makes it harder?
F: Well, they started with the shortest second grader and the shortest sixth grader, and then they kept alternating until they reached the tallest. This makes it harder to see the clumps.
T: How about Melissa and Nick's representation? [*See* ***fig. 4c***.]
G: They labeled the heights according to grade levels, and I can easily see the heights that are in common—like there are students from both classes that are fifty-eight inches tall.
H: It's not as easy to see the common heights in Sally and Jessica's or in Joe and Hunter's.
I: I think Joe and Hunter's feels like two separate graphs. They're not connected, so I don't see how it is different from putting those two graphs on different papers.
T: Are there other things that you notice about the data?
J: There are no people that are fifty-three or fifty-four inches tall, but it's kind of hard to see this.
B: You could easily fix this on Melissa and Nick's graph [*see* ***fig. 4c***] by putting some space between 52 and 55….

Several purposes overlap in this kind of discussion. First, it is critical to direct students to focus on what we can learn from the data by asking questions such as, "What kinds of things do you notice about the height of the second graders and the sixth graders?" After all, the purpose of collecting, representing, and analyzing data is to learn something about the world around us.

Second, by comparing and contrasting student-generated representations, youngsters can focus on how particular features of different representations illuminate or obscure certain aspects of the data. This, then, provides students with a rationale, for example, for showing gaps in the data; that is, comparing representations helps illuminate interesting data features.

What Teachers Can Do

Analyze the activities you currently use for data analysis:

- Do students have opportunities to create their own representations and to discuss the features of those representations? (If not, try some of the tasks we describe. Help students identify an interesting question or topic to explore.)
- Do students always have a purpose for analyzing data? For instance, the data answer a question or tell us more about something that interests us.

With younger students, focus on a question involving a single data set:

- How many teeth have all the second graders lost?
- How many pockets are in our classroom?

With older students, consider a question requiring a comparison:

- Which fourth-grade class is taller?
- Can fifth graders jump farther than sixth graders?

At all levels, allow students to create and use their own representations to describe and analyze data. Consider the issues raised in this article as you facilitate whole-group discussions about students' representations. Add to this list as you observe your own students' efforts at constructing data displays.

Conclusions and Suggestions

Teaching data analysis by encouraging students to generate their own data representations is a challenging undertaking that requires teachers to be knowledgeable about the types of representations that students create on their own, the issues that students face during the process of creating those representations, and how to facilitate discussions to help students make sense of their representations. Because many teachers have learned to represent data by following step-by-step procedures, they may have difficulty considering a different way of teaching this topic. What, then, makes it worthwhile to have students generate their own data representations?

When we show students how to generate a particular type of representation by providing them with a step-by-step procedure for creating it, we essentially deprive them of opportunities for learning *why* certain graphing conventions are followed. For example, when teachers demonstrate the method to create a bar graph, we make assumptions about how to order and group the data. In our experience, many elementary students do not initially recognize the need to organize or group data. Therefore, allowing them to generate and analyze their own representations provides the opportunity to arrive at these assumptions in their own time and in a meaningful way and thus develop a solid understanding of the role of representations in analyzing and interpreting data. Furthermore, allowing students to generate their own representations helps them recognize that more than one way exists to solve problems and represent situations. Students realize that their own thinking is important and valued. In short, the classroom environment supports teaching that is built on student thinking, where students learn data analysis with understanding.

By allowing students to generate their own data representations, you will learn more about your students' thinking and be better equipped to capitalize on that thinking to further their understanding and use of data.

References

Economopoulos, Karen, and Tracey Wright. *How Many Pockets? How Many Teeth?* 1998. Investigations in Number, Data, and Space series. Menlo Park, CA: Dale Seymour Publications.

Konold, Clifford, and Traci L. Higgins. "Reasoning about Data." In *A Research Companion to "Principles and Standards for School Mathematics,"* edited by Jeremy Kilpatrick, W. Gary Martin, and Deborah Schifter, pp. 193–215. Reston, VA: National Council of Teachers of Mathematics, 2003.

National Council of Teachers of Mathematics (NCTM). *Principles and Standards for School Mathematics*. Reston, VA: NCTM, 2000.

Russell, Susan J."What Does It Mean That '5 Has a Lot'? From the World to Data and Back." In *Thinking and Reasoning with Data and Chance*, edited by Gail F. Burrill and Portia C. Elliott, pp. 17–29. Reston, VA: National Council of Teachers of Mathematics, 2006.

Russell, Susan J., Rebecca B. Corwin, Andee Rubin, and Joan Akers. *The Shape of the Data: Statistics.* 1998. Investigations in Number, Data, and Space series. Menlo Park, CA: Dale Seymour Publications.

Watson, Jane M. *Statistical Literacy at School: Growth and Goals*. Mahwah, NJ: Lawrence Erlbaum Associates, 2006.

This work was supported by National Science Foundation grant ESI-0333879. Any opinions expressed are those of the authors and do not necessarily represent the view of the National Science Foundation.