

From Gal, I. & Garfield, J. B. (editors). The Assessment Challenge in Statistics Education. IOS Press, 1997 (on behalf of the ISI). Pages 55-63. ISBN 90 5199 333 1. Copyright holder: International Statistical Institute. Permission granted by ISI to the IASE to make the chapters from this book freely available on the Internet. The References section is a separate chapter and is available on the Internet at <http://www.stat.auckland.ac.nz/~iase/publications/assessbkref>.

## 5. A Framework for Assessing Knowledge and Learning in Statistics (K-8)

Susan N. Friel  
George W. Bright  
Dargan Frierson  
Gary D. Kader

### Purpose

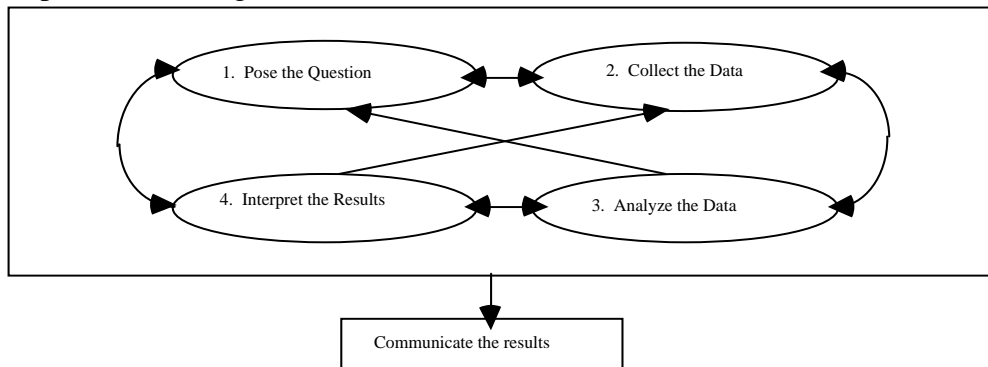
This chapter provides an overview framework for thinking about what teachers and students should know and be able to do with respect to learning statistics at the K-8 levels. Given the number of concepts to be considered and our limited knowledge about the complexities of learning these concepts, we focus on the understanding of graphical representations, examine examples of “good tasks” that may be used to assess graph knowledge, and reflect on what we have learned about the complexities of assessing students’ graph knowledge when using these tasks.

### INTRODUCTION

What do teachers and students need to know and be able to do with respect to statistics in the elementary and middle grades? Our answer centers on the development of “data sense”, which includes being comfortable with posing questions, collecting and analyzing data, and interpreting the results in ways that respond to the original question asked. It also includes comfort and competence in reading, listening to, and evaluating reports based in statistics, such as those found in newspapers, magazines, television, and other forms of popular press. That is, data sense encompasses not only understanding the graphs and statistics that are presented but also evaluating the statistical investigation process used to generate that information from which the graphs and statistics are constructed.

One critical part of data sense is understanding that a statistical investigation is really a process. A statistical investigation typically involves four components (1) posing the question, (2) collecting data, (3) analyzing data, and (4) interpreting the results, in some order (Graham, 1987). Kader and Perry (1994) suggest a fifth stage of a statistical investigation: communication

of results. The resulting model gives structure to our understanding of the type of reasoning used in statistical problem solving.



**Figure 1: The Process of Statistical Investigation**

As with any *process*, there is inherent difficulty in attempting to capture the *quality* of its use by students. Having students do projects that involve the use of the statistical investigation process may be one of the better ways to assess capability, yet such assessment is fraught with difficulties. How do we determine what students know? How do we know that they know it? Part of the answer lies in a better understanding of the components of the process and the related statistical concepts.

This has led us to consider what it means to understand and use graphical representations as a key part of what it means to know and be able to do statistics. The current literature tells us very little about how such knowledge develops. There is some anecdotal and written evidence obtained through developmental research (Gravemeijer, 1994) associated with three different curriculum development projects which helps to frame a set of issues related to understanding graphical representations. Too, graphicacy (Wainer, 1980) in the curriculum, particularly at the elementary and early childhood levels has become a focus of curriculum analysis (Lappan, et al., 1996; Russell & Corwin, 1989).

We have begun developmental research to examine upper elementary and middle grades students' learning of key concepts related to the use and interpretation of graphical representations. We have looked at how such understanding changes over time and with instructional intervention provided by knowledgeable teachers in order to develop a structure for understanding the mathematics that is involved. In the remainder of this chapter, we focus on assessment strategies for the "analysis of data" and "interpretation of results" components of the statistical investigation process that relate specifically to the understanding of different ways to represent data. The focus on data representation in general and graph knowledge more specifically is due to the fact that this is such an important part of being able to use statistics in the real world. Very little seems to be known about students' understanding in this area. Hopefully, learning more about students' understanding of data representation may also allow us to address more general assessment issues for other aspects of statistics.

## ISSUES RELATED TO GRAPH KNOWLEDGE

Understanding the process of data reduction and the structure of graphs are factors that influence graph knowledge. The transition from tabular and graphical representations which

## 5. A Framework for Assessing

display raw data to those which present grouped data or other aggregate summary representations is called data reduction. The purpose of data reduction is to identify appropriate representations of the data which remove as much detail from the data as is possible while providing sufficient information to address the specific question at hand.

Graphical representations of numerical data reflect different levels of data reduction. A representation may display the original raw data or grouped data. For example, line plots and stem plots present the original data, while box plots and histograms present grouped data. Most graphical representations used in the early grades (e.g., picture graphs, bar graphs) involve either the original data or tallied data from which the original observations may be obtained. Students in upper grades more often use graphical representations of grouped data (histograms, box plots) from which it is not possible to return to the data in its original form.

In addition to misunderstandings that may emerge with data reduction, the structure of graphical representations of data may also impact understanding. For example, graphical representations utilize one axis or two axes or, in some cases, not have an axis. For graphical representations that use both axes, the axes may have different meanings. In some simple graphs, the vertical axis may display the value for each observation while the vertical axis for more typical bar graphs and histograms provides the frequency of occurrence of each observation (or group of observations) displayed on the horizontal axis. Confusion may develop if the different functions of the x- and y-axes across these graphs are not explicitly recognized.

Another factor that may influence graph knowledge is the ways students are asked to read graphs to gather information. Curcio (1987 and Chapter 10 in this book) conducted a study of graph comprehension assessing fourth and seventh grade students' understanding of four traditional "school" graphs: pictographs, bar graphs, circle or pie graphs, and line graphs. She identified three components to graph comprehension: reading the data ("How many students have 12 letters in their names?"), reading between the data ("How many students have more than 12 letters in their names?"), and reading beyond the data ("If a new student joined our class, how many letters would you predict that student would have in her name?")

### SOME TASKS TO ASSESS GRAPH KNOWLEDGE

Informally, teachers can frequently cite instances of confusion that exist when students work with different graphs. Instructional and assessment strategies seem to be needed to help students focus on the characteristics of the graphs, on transitions from one data representation to another, and on different levels of interpretation tied to graph comprehension. Appropriate questions for both instruction and assessment may be developed through application of Curcio's scheme of reading the data, reading between the data, and reading beyond the data. Transitions among data representations may be structured using the transformational strategy of developing bar graphs from work with line plots and histograms from work with stem plots.

There are a number of different kinds of questions used in both written tests and interview settings that we have been exploring with both upper elementary and middle grades students and practicing teachers which provide possible directions for assessment that focus on earlier-noted concerns. With each assessment, we have sought to embed tasks within understandable contexts and to design questions that would call attention to the process of statistical investigation. That is, we have chosen contexts that seem to relate to students' everyday lives. Examples of written

## S. Friel, G. Bright, D. Frierson, & G. Kader

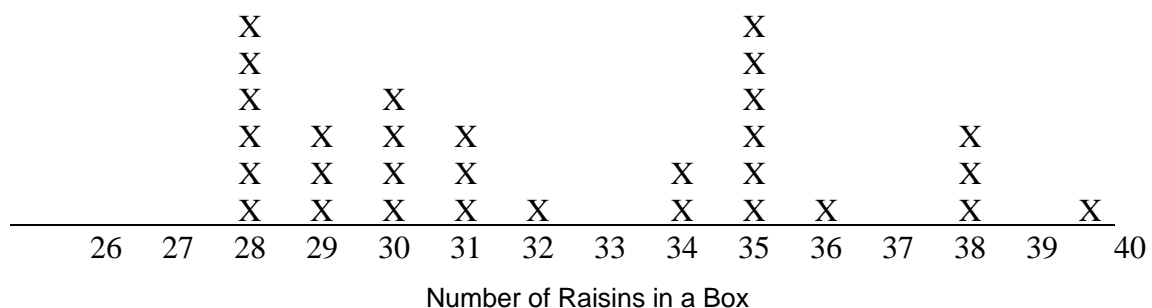
problems that we have used as both pre- and post-test instruments with students address the use of line plots and bar graphs and stem plots and histograms (Friel & Bright, 1995).

### Assessing First Steps in Data Reduction

One problem undertaken by student involves the context of raisins (as shown in Figure 2).

---

*Students brought several different foods to school for snacks. One snack that lots of them like is raisins. They decided they wanted to find out just how many raisins are in half-ounce boxes of raisins. They wondered if there was the same numbers of raisins in every box. The next day for snacks they each brought a small box of raisins. They opened their boxes and counted the number of raisins in each of their boxes. Students are presented with a line plot showing the information the class found:*



*Are there the same number of raisins in each box? How can you tell?*

---

**Figure 2: Counts of Raisins**

In our research, we grouped students' responses into categories based on the ways they talked about data and the features of the graph, as shown below:

*Properties of the graph (considering both range of data and frequency)*

- No, because the x's are not all on one number.
- No, because the X's show how many boxes had that many of raisins. Like 28 had 6 and 29 had 3.
- No. If there were the same number in each box there would be X's all above the same number.

*Literally "reading" the data from the graph.*

- No there aren't the same number of raisins in each box, I found my answer by looking at the data, 6 boxes have 28, 3 have 29, 4 have 30, 3 have 31, 1 has 32, 2 have 34, 6 have 35, 1 has 36, 3 have 38, and 1 has 40.

*Properties related to the context or to the data.*

- No, because they weigh the boxes until they equal 1/2 ounce. They don't count the raisins.
- No, because some raisins can be smaller and that means you can have more.

*Range of the data (considering only range and does not include frequency)*

- Because it says the number of raisins goes from 26 to 40.

*Frequency of occurrence/height of bars*

- No, the X's have different numbers, so there are different numbers of X's in each box.
- No, because some do not have as much X's and some have more.

## 5. A Framework for Assessing

*Other (includes incomplete, unclear, incorrect, or not statistically-reasoned responses)*

- No there are not. They all have different amounts.

In this problem, a limited number of students (roughly 28% of sixth graders) were able to reason using information about the data values themselves (from the axis) and the frequencies of occurrence of these data values (the X's). The number of students who seemed to focus on the frequency or number of X's as the data values indicates that there may well be confusion even when using line plots about the role of data values and frequencies. We have found such confusions exist with students' reading of bar graphs; we attribute some of these confusions to having to read the frequency using the vertical axis. For a line plot, this is not the case.

### Assessing Next Steps in Data Reduction

Another problem is an investigation undertaken by a middle grades class of students (as shown in Figure 3).

---

*Students were interested in how they used their time. They brainstormed a list of ways such as sleeping, eating, after school sports, and so on. Jim reminded them that some of their time is used just traveling back and forth to school. Some of the students thought this shouldn't count because it really wasn't much time at all. Others disagreed. The class wondered, "What is the typical time it takes to travel to school?"*

Students taking the test are told that these students spent time discussing how they would collect their data. For the first look at this problem, they each decided that they would time how long it took them to reach school the next day. They used stop watches (which the teacher had) or timed their trips using their own watches. Once the data were collected, the students made a stem-and-leaf plot (stem plot) for the data number of minutes it took them to travel to school.

#### Minutes to Travel to School

0	3 3 5 7 8 9
1	0 2 3 5 6 6 8 9
2	0 1 3 3 3 5 5 8 8
3	0 5
4	5

2 | 5 means 25 minutes

1. How many students are in the class? How can you tell?
2. How many students took less than 15 minutes to travel to school? How can you tell?
3. Write down the three shortest travel times students took to get to school.
4. Write down the three longest travel times students took to get to school.
5. What is the typical time it takes for students to travel to school? Explain your answer.
6. Make a histogram to show the information about travel time that is displayed on the stem-and-leaf plot.

---

**Figure 3: Travel times to school**

In our research, we found that, after instruction, students were able to correctly answer all four of questions 1-4 which required them to either 'read the data' or 'read between the data' on the given stem plot. Their answers to question 5, which required them to 'read beyond the data',

## S. Friel, G. Bright, D. Frierson, & G. Kader

varied somewhat but did show a range of responses correctly utilizing the statistical measures of center they studied. Students were less likely to compute measures of center as part of their responses. Some samples of various types of responses follow:

### *Responses that identified the mode in the data:*

- The typical time it takes for the students to travel to school is 23 minutes, because there are more students in that class that takes them 23 minutes to travel to school.

### *Responses that identified a cluster of times:*

- 10 to 28 because they have most of the data on it. And that is where it is clumped.
- I think between 3 to 20 minutes. Because 15 people are between 3 to 20 and that is more than half.

### *Responses that provided a tally or range of numbers that occurred most frequently:*

- 23-28 min. I know because these numbers are repeated more.
- 3 & 5, you see how many of the same number you can find. [Note: counted digits in leaves]
- Probably 3, 16, 23, 15 & 28 because more people have them 5 numbers.

In the question noted, responses that identified the mode and responses that identified clusters of data are both appropriate; however, how do we evaluate these two responses? Is one “more appropriate” than the other; do we want students to move beyond the use of the mode as a tool in this case to using clusters of data as a way of describing what’s typical? If so, what is a “good sized” cluster to be highlighting?

## WHAT OTHER KINDS OF TASKS COULD OR SHOULD BE USED IN ASSESSING GRAPH KNOWLEDGE?

Not surprisingly, we can learn a great deal by trying to understand how students explain their thinking about graphs. Although students often give “correct answers” to questions about graphs (as evidenced by the fact that 97% of the students answered the “raisin question” in Figure 2 correctly), the reasoning that supports these answers is often faulty. Students seem to be interpreting graphs in ways that are inconsistent with clear understanding of the underlying mathematics.

Too, a large number of students provide vague or incomplete responses that seem to say, “The answer is this because the graph says so.” This level of clarity may reflect the usual emphasis in mathematics on “getting an answer” with little emphasis on follow-up questioning that would create an expectation for clear explanations of why an answer is given. Students may simply not be experienced enough at explaining to be able to do it in writing. On the other hand, students’ difficulty at explaining may result from our own lack of clarity about how we expect students to be able to talk about graphs and our corresponding lack of clear criteria for what we expect students to say. For example, for question 5 in the “travel time” scenario (Figure 3), it is not clear how we ought to respond to the students who tallied occurrences of numbers. What do such responses tell us about where the students are in their thinking about the concept of “typical” and about what we might want to do to develop this concept over time? In and of itself, we cannot tell if this response reflects an initial approach to answering the question, the result of a full investigation by this student, or something else. It is also interesting that for this question few students used the median or the mean, although these ways for summarizing data had been

## 5. A Framework for Assessing

included in the instruction and would have been an appropriate way to answer this question. What does their inattention to appropriate statistical concepts tell us about their thinking?

Based on even this little information about students' thinking, it is possible to propose other kinds of questions that could be asked to help both us and the students clarify reasoning strategies further or to help focus on both the data values and their frequencies as components involved in reading a graph. The first two examples below deal with how to use data to create a graph. The third deals with detailed interpretation of one small part of a given representation.

1. Someone has opened 5 boxes of raisins and each box has the same number of raisins. How might a line plot that shows these data look? Why?
2. Someone has opened 5 boxes of raisins and two of the boxes have the same number of raisins. Of the remaining 3 boxes, each has a different number of raisins. How might the line plot that shows these data look? Why?
3. In the line plot in Figure 2, what do the four X's above 31 mean? Explain.

We do not know a great deal about how to use forced choice items as a way to assess students' graph knowledge. However, we do have models for ways to think about doing this. One model involves students being asked to choose among representations as they read through a number of short problem statements; either students simply choose the appropriate graph (Mathematical Sciences Education Board, 1993) or they choose the appropriate graph and identify certain of the graph elements as well. These examples of forced choice items also require that students justify their reasoning and, in some situations, respond to a short answer interpretive question as well.

Currently, very little research has been done on the usefulness of such items a part of an overall strategy for assessing knowledge. Such problems reflect an attempt to address the problem solving process of reversibility (Rachlin, 1992) in which we move from students working from a problem to data collection to creation of representations; here we have students working from a problem, bypassing data collection and construction of representations, to interpretive reading of an incomplete representation constructed by someone else.

Another model, the use Mystery Graphs (Russell & Corwin, 1989), involves giving students an incompletely labeled graph showing data from a specific context. Students describe what context is being described by the graph. One example is a graph that shows the weights of a number of different lions in zoos across the United States; as part of the data set, there are a few cubs included. This type of problem is not a "fixed choice" problem yet the responses are contained within a realm of prediction about what is possible, given the data as they are displayed in the selected representation.

### IMPLICATIONS

Among the reasons for encountering difficulties in assessing knowledge of statistics is the need to clarify what it is that we want students and teachers to know and be able to do with respect to statistics. We currently know very little about the development of many statistical concepts over time and through instructional interventions. Our work in the area of data representations has provided a model of one way to build both understanding based on

## **S. Friel, G. Bright, D. Frierson, & G. Kader**

consideration of the development of graph knowledge and of strategies for assessment that may be used to support and evaluate this development.

Instructional and assessment questions need to provide opportunities for students' thinking to be revealed. Just asking for answers can be misleading for teachers as they try to understand what students know. Sometimes this thinking can be revealed by asking students to explain their answers. But we also need to develop questions (such as the three examples in the previous section) that focus on detail that students might not otherwise talk about or think about. Questions like these have the potential to reveal students' levels of concept attainment as they talk about their thinking. The process of answering these questions may help students expand and refine their thinking. That is, these questions become learning events as well as assessment events.

We need to use a variety of questions so that we can determine whether students are consistent in their thinking. In this way we will be better able to "triangulate" their thinking. At the same time, we are the first to admit that it is often difficult to know what questions to ask to probe students' thinking. It often takes a lot of reflection on our part to even begin to understand what students might be thinking. It is only in hindsight that sometimes we can think what we should have asked to probe a response.

Part of what we need to know about a student's response is whether the basic understandings (e.g., reading the data) are in place. We don't want to leave the impression that there is a tightly sequenced list of prerequisite skills that students need to master in order to answer questions about graphs, but we don't want to confuse lack of understanding about the question with lack of these prerequisite skills.

The use of graphs and other kinds of representations needs to be viewed as part of the process of statistical investigation and not as an end in itself. Some will argue that such study needs to emerge during the process of investigating reasonably "big" problems that engage students in this process; it is not a question of whether we "teach histograms now" but rather that we wait for a need for various representations to emerge out of "big" problems and to be taught within "big" problem contexts. Our work supports consideration of the process, but does address graph knowledge more explicitly as part of the development of the overall process. Part of the reason for choosing such a direction revolves around the need to understand what we don't understand about ways in which the use and reading of graphs may be misunderstood or misinterpreted by students.

In statistics, data reduction is an essential part of analysis of the data; different graphs emphasize different degrees of data reduction. Past instruction and assessment (e.g., often found in school mathematics textbooks) has not demonstrated an awareness of the connections between the process of data reduction and the choice of graphs. Indeed, there appear to be natural transitions between some representations (i.e., line plot and bar graph or stem plot and histogram) that support building the use of multiple representations in a way that may facilitate understanding.

The use of technology as a tool (e.g., computer database, graphing software, or graphing calculators) offers real potential for students in helping to promote exploration of different representations of data as well as the structure of these representations. There has been very little work done in this arena; we don't know a lot about ways that such data display software or calculators support students' understandings. We have seen students use such tools without much thought; they make graphs that are not appropriate for the data or they do not consider exploring the impact of using different graphs or experimenting with scaling. Specifically, we wonder



## 5. A Framework for Assessing

about the fact that when graphing calculators are used, the graphs have no labels on axes or titles displayed with the representations. Also we wonder about the array of implicit “definitions” about the nature of certain kinds of graphs provided by different computer software programs (e.g., *Cricket Graph III*, *Statistics Workshop*, *MacStat*, and *Data Insights* ). Each program defines the parameters involved in making histograms in very different ways. In addition, each program provides different options with respect to scaling axes and using relative frequencies as an alternative to counts.

There is a need to monitor learners’ changes in thinking as they move among ungrouped and grouped data representations. Once we have some knowledge of learners’ thinking, are we clear about what attributes of statistical thinking we want to promote and about ways to promote these attributes? How do we rank responses following an instructional unit? For example, what if we want students to begin to understand what the data tells them and also to understand when they are adding their own lens without support from the data? What instructional and evaluation strategies do we use? Do we begin to push students to ask questions about their observations or conclusions such as “Do the data tell you this? Are you making a judgment based on personal experience and reasoning rather than on the data? Do you have the information you need to make a judgment using only the data?”

We need to understand what students understand prior to and following instruction *and* be clearer about how we will judge their responses in light of what we think reflects sound statistical thinking.