THE JOURNAL OF THE LEARNING SCIENCES, *11*(2&3), 163–185

Guest Editor's Introduction

Appendix: The Object and the Context: What Our Data Are and Where They Come From

Kay McClain
Department of Teaching and Learning
Vanderbilt University

OVERVIEW OF THE STATISTICS PROJECT

The data discussed in the articles in this issue are taken from a classroom teaching experiment conducted in the fall semester of 1997 with a group of 29 American seventh-grade students (age 12). During the 12 weeks of the teaching experiment, the research team[1] assumed total responsibility for the class sessions, including teaching.[2] The primary goal for the teaching experiment was to investigate ways to proactively support middle-school students' ability to reason about data while developing statistical understandings related to exploratory data analysis. An integral aspect of that understanding entailed students coming to view data sets as distributions. In that process, they would structure and organize data sets multiplicatively as they created data-based arguments that were grounded in their analysis. The research team's interest was motivated by current debates about the role of statistics in school curricula (cf. Burrill, 1996; Burrill & Romberg, in press; Cobb, 1997; Gal, 2000; Lajoie, in press; Lajoie & Romberg, in press; Lipson & Jones, 1996; National Council of Teachers of Mathematics, 1989, 1991, 2000; Shaughnessy, 1992, 1996). The guiding image that emerged as the research team read and synthesized

Correspondence and requests for reprints should be sent to Kay McClain, 5330 Old Harding Road, Franklin, TN 37064. E-mail: mcclain@vanderbilt.edu

[1]The research team was composed of Paul Cobb, Kay McClain, Koeno Gravemeijer, Maggie McGatha, Lynn Hodge, Jose Cortina, Carla Richards, and Cliff Konold.

[2]Kay McClain had primary responsibility for the teaching during the classroom teaching experiment. However, on numerous occasions Paul Cobb assisted. In addition, additional members of the research team circulated around the room to monitor students' activity during small-group work. In these instances, the team then discussed which solutions to highlight in subsequent whole class discussions.

the literature was that of students engaging in instructional activities that they both developed and critiqued data-based arguments (cf. Wilensky, 1997).

Before conducting the classroom teaching experiment, the research team spent time reading the literature and gathering data on students' ways of reasoning statistically to determine appropriate starting points for the instructional sequence. As part of this process, the research team conducted whole-class performance assessments with seventh-graders in the school that the team intended to work during the following year. Analysis of these tasks revealed that the students' prior instruction in statistics had focused on procedures for calculating measures of central tendency and conventions for creating graphs. In solving data-based tasks, students would determine the measure to be found or graph to be created without reasoning about its utility with respect to the problem or question at hand. The research team described the students' activity as attempting "to do something with numbers" to satisfy a school requirement (for detailed analysis, see McGatha, Cobb, & McClain, 1998). The research team found similar results in analysis of individual clinical interviews conducted with all students in the project classroom before the teaching experiment began.

As a result of these analyses, the initial goal of the teaching experiment was that students would act on data. This would be followed by coming to view data sets as distributions. The research team therefore envisioned that measures of center would be viewed as characteristics of these distributions and graphs as ways of structuring and organizing distributions. The teacher's goal would be to capitalize on students' ways of reasoning about data to advance the mathematical agenda. The crucial norm that would be necessary to support this classroom environment, therefore, would be that of explaining and justifying solutions in the context of the problem being explored. This is a radically different approach to statistics than is typically introduced in United States' middle schools. It highlights the importance of curricula that provide opportunities for students to engage in genuine problem solving that supports the development of central mathematical concepts.

Design of the Instructional Sequence

As the research team began to design the instructional sequence to be used in the seventh-grade classroom, the team attempted to identify the "big ideas" in statistics. The plan was to develop a single, coherent sequence and thus tie together the separate, loosely related topics that typically characterize United States' middle-school statistics curricula. In doing so, the research team came to focus on the notion of distribution. This enabled measures of center, relative frequency, and other features such as "skewness" and "spread-outness" to be treated as characteristics of distributions. It also allowed various conventional graphs such as histograms and box-and-whisker plots to be viewed as different ways of structuring distributions. The instructional goal was, therefore, to support students' gradual development of a single, multifaceted notion, that of distribution, rather than a collection of topics to be taught as separate components of a curriculum unit. As the team formulated hypotheses about how the students might reason about distributions, one of the primary goals was that students would think about data sets as entities that have properties in their own right rather than as collections of points (e.g., Hancock, Kaput, & Goldsmith, 1992; Konold, Pollatsek, Well, & Gagnon, in press; Mokros & Russell, 1995). The research team conjectured that if students began to think about data in this way, they could then investigate ways of structuring data sets that would help them identify trends and patterns.

When reasoning about data sets, students typically do so either additively or multiplicatively. For instance, in reasoning additively, students might talk in terms of part–whole relations (i.e., 15 of the cars were traveling faster than the speed limit and 30 were traveling slower). When reasoning multiplicatively, students would reason proportionally (i.e., one third of the cars were traveling faster than the speed limit). In the latter case, they would be reasoning about qualitative proportions of the data. The research team viewed multiplicative reasoning as not only more sophisticated but also as the basis for many concepts addressed at the middle-school level. The research team therefore conjectured that it would be important to support students' ability to reason multiplicatively about data.

As the research team began mapping out the instructional sequence, we were guided by the premise that the integration of computer tools was critical in supporting the mathematical goals. Students would need efficient ways to organize, structure, describe, and compare large data sets. This could best be facilitated by the use of computer tools for data analysis. However, the research team tried to avoid creating tools for analysis that would offer either too much or too little support. This quandary is captured in the current debate about the role of technologies in supporting students' understandings of data and data analysis. This debate often is cast in terms of what has been termed as expressive and exploratory computer models (cf. Doerr, 1995). In the expressive model, students are expected to recreate conventional graphs with only an occasional nudge from the teacher. In the exploratory model, students work with computer software that presents a range of conventional graphs with the expectation that the students will develop mature mathematical understandings of their meanings as they use them. The approach that the research team took when designing computer-based tools for data analysis offers a middle ground between the two approaches. It introduces particular tools and ways of structuring data that are designed to fit with students' current ways of understanding, while simultaneously building toward conventional graphs (Gravemeijer, Cobb, Bowers, & Whitenack, 2000). As such, the computer tools we designed were intended to support students' emerging mathematical notions while simultaneously providing them with tools for data analysis.

Computer-Based Tools for Analysis

The instructional sequence developed in the course of the seventh-grade teaching experiment involved two computer-based minitools. In the initial phase of the sequence, which lasted for almost 6 weeks, the students used the first minitool to explore sets of data. This minitool was explicitly designed for this instructional phase and provided a means for students to manipulate, order, partition, and otherwise organize small sets of data in a relatively routine way. Part of the rationale in designing this tool was to support students' ability to actually engage in analyzing data as opposed to performing meaningless calculations. When data were entered into the tool, each individual data value was shown as a bar, the length of which signified the numerical value of the data. This can be seen in the data on the longevity of two brands of batteries shown in Figure 1.

A data set therefore was shown as a set of parallel bars of varying lengths that were aligned with an axis. The bars could be either pink or green, allowing for comparisons of two data sets. The minitool's use in the classroom made it possible for students to act on data in a relatively direct way. This would not have been possible had we used commercially available software packages for data analysis whose options typically include only a selection of conventional graphs. The first computer minitool also contained a red value bar that could be dragged along the horizontal axis to partition data sets or to estimate the mean or to mark the median.

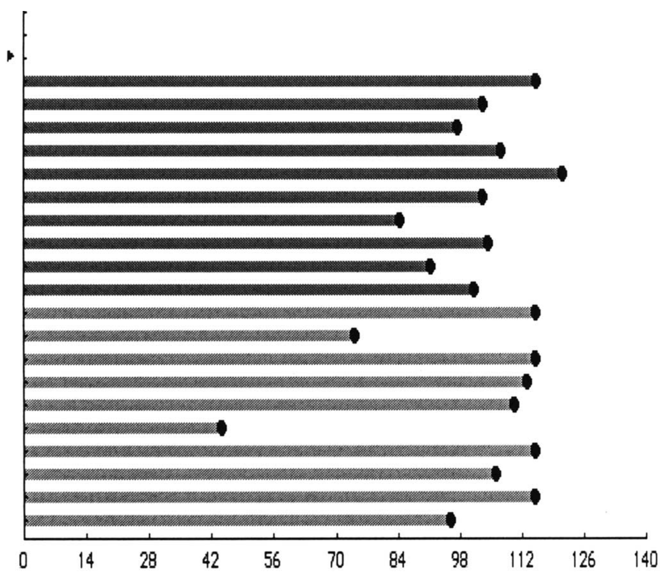

FIGURE 1 Batteries display in first minitool.

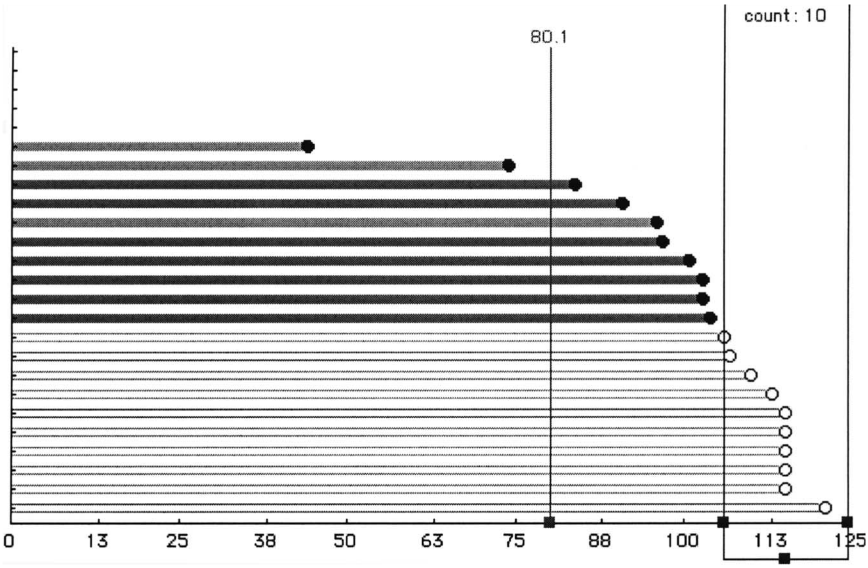

FIGURE 2 First minitool with value bar and range tool shown.

As the bar was moved along the axis, a number appeared at the top of the bar indicating its location on the axis, as shown in Figure 2 (see red value bar at 80.1).

In addition, there was a range tool that could be used to determine the number of data points within a fixed range. Students used this feature by dragging two connected blue bars to specific locations on the axis as shown in Figure 2. The bars could be moved separately or in tandem. As the bars moved, a count appeared between the bars indicating the number of data values captured by the bars. The count adjusted as the bars were moved.

The second computer minitool can be viewed as an immediate successor of the first. As such, the endpoints of the bars that each signified a single data point in the first minitool were, in effect, collapsed down onto the axis so that a data set was now shown as collection of dots located on an axis. This can be seen in Figure 3 (e.g., an axis plot).

The tool offered a range of ways to structure data. It is important that this palate of options did not correspond to a variety of conventional graphs as is typically the case with commercially available software packages. Instead, the various options were designed after we identified in the literature the various ways that students structure data when they are given the opportunity to develop their own approaches while conducting genuine analysis (cf. Hancock, in press). Two of the options can be viewed as precursors to standard ways of structuring and inscribing data. These are organizing the data into four equal groups so that each group contains one fourth of the data (precursor to the box-and-whisker plot) and organizing data into groups of a fixed interval width so that each interval spans the same range

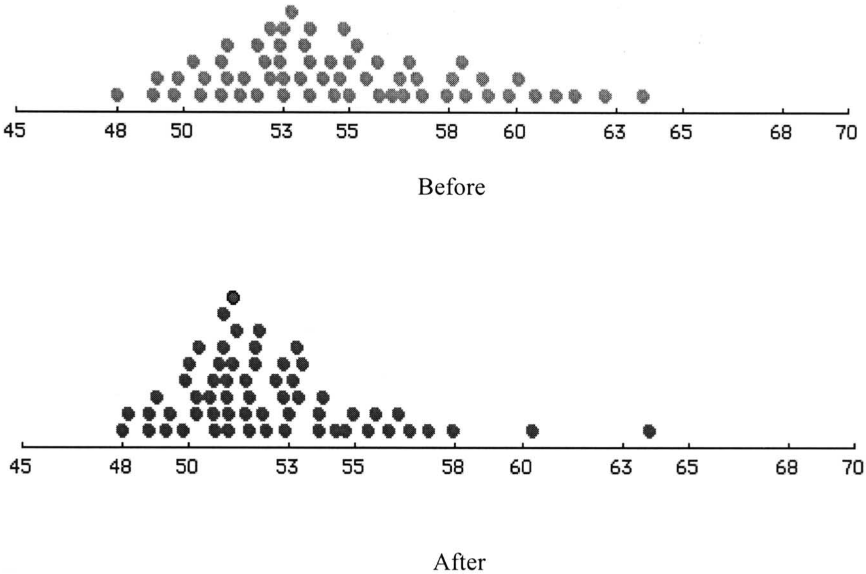

FIGURE 3 Data as displayed in second minitool.

on the axis (precursor to the histogram). However, the three other options available to students do not correspond to graphs typically taught in school. These involve structuring the data by (a) creating your own groups, (b) partitioning the data into groups of a fixed size, and (c) partitioning the data into two equal groups. The first and least sophisticated of these options simply involved dragging one or more bars to chosen locations on the axis to partition the data set into groups of points. The number of points in each group was shown on the screen and adjusted automatically as the bars were dragged along the axis. The key point is that this tool was designed to fit with students' ways of reasoning while simultaneously taking important statistical ideas seriously.

Instructional Activities

As the research team worked to outline the sequence, we reasoned that students would need to encounter situations in which they had to develop arguments based on the reasons for which the data were generated. They therefore would need to develop ways to analyze and describe the data to substantiate their recommendations. The research team anticipated that this would best be achieved by developing a sequence of instructional tasks that involved either describing a data set or analyzing two or more data sets to make a decision or a judgment. The students typically engaged in these types of tasks to make a recommendation to someone about a practical course of action that should be followed.

An important aspect of the instructional sequence involved talking through the data creation process with the students. In situations where students did not actually collect the data themselves, the research team found it very important for the students to think about the types of decisions that are made when collecting data to answer a question. During this process, the students made conjectures and offered suggestions about the information that would be needed in make a reasoned decision. Against this background, they discussed the steps that they might take to collect the data. In the process, they delineated the design specifications for a study that would yield the measures of the attributes necessary to make an informed decision. These discussions proved critical in grounding the students' data analysis activity in the context of a recommendation that had real consequences.

Classroom Structure

The classroom teaching experiment consisted of 37 class sessions. The class met on Monday, Thursday, and Friday for 45 min and on Tuesday for 90 min. The typical format of these sessions involved an initial introduction to the task including the data creation discussion. This was followed by students working in pairs or small groups at the computers to analyze the data. As part of this process they were asked to make a judgment or recommendation about a procedure or course of action based on the results of their analysis. In addition, as the teaching experiment progressed, students were asked to create an inscription to support their argument. These inscriptions were supposed to summarize the data without reproducing the data sets in their entirety. The goal was to have the students focus on features and characteristics of the distributions that would be critical in making a reasoned decision. The final phase of the activity structure involved the students and teacher engaged in a whole class discussion during which time the students shared the results of their analysis. These discussions were deliberately facilitated by the teacher in concert with other members of the research team so that issues of significance might arise from the students' diverse ways of analyzing the data.

CLASSROOM EPISODES

Batteries Episode

The first of the two episodes featured in this issue occurred during the ninth class session. It was the fourth session during which the students worked with the first computer minitool. Before this task, students had been engaged in a task involving temperatures and were trying to determine the accuracy of reporting a "daily high temperature" for a city if the temperature was recorded at only one location. This task was followed by the introduction of the first minitool. Initially, students were shown paper-and-pencil graphs of data on the braking distances of 10 each of two different makes of cars. After working in pairs to determine which was the safer make of car, students were shown the same data inscribed in the first minitool. They then were given the opportunity to revise or refine their arguments.

The batteries data constituted the second task posed by using the first minitool (see Figure 1). The data were introduced to the students as being the result of tests conducted on 10 each of two brands of batteries. Before introducing the data, I (as the classroom teacher) engaged the students in a discussion about the data creation process. I began by posing a problem in which I asked the students how they might test batteries to determine which brand lasts longer. They offered suggestions such as placing the batteries in flashlights and seeing how long they burn, or placing the batteries in toys and seeing how long the toys are able to move. Students mentioned the importance of using the same device for each test and discussed accurate ways of timing. After this discussion, the data were introduced to the students. Students were asked to use the computer minitool to analyze the data to determine which brand of battery they would recommend. Students then spent the remainder of the class period working in pairs at the computer on their analyses.

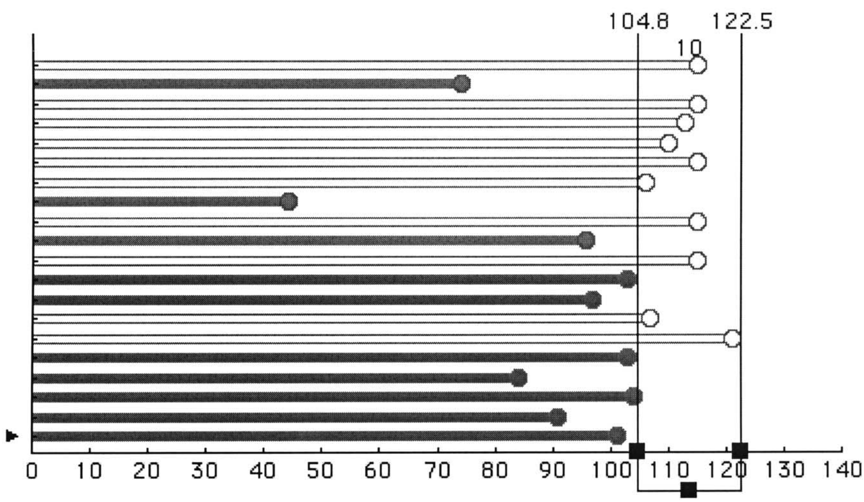

FIGURE 4 Batteries display with the range tool feature activated.

As the students worked, I and other members of the research team circulated around the room to monitor the students' activity. Our goal was to determine the diversity of ways in which the students were approaching the task. These ways of analyzing the data would be the basis of the whole class discussion on the following day. As a result, the goal was to gather information, not correct students' approaches. Students spent the remainder of the period working on their analyses.

I began the whole class discussion on the following day by asking Ceasara to explain how she and her partner solved the task. I used a computer projection system to project the data sets onto the white board. Ceasara began by asking me to activate the range tool feature on the minitool and use it to capture the top 10 batteries as shown in Figure 4.

Transcript of *Batteries* episode:

1. Kay: OK, Ceasara.
2. Ceasara: Could you put the range on?
3. Kay: What do you want me to change it to [moves the value bar]?
4. Ceasara: Oh, not that, I mean like…[inaudible].
5. Kay: I'm sorry I'm not understanding…
6. Ss: The blue thing [referring to the blue bars of the range tool].
7. Kay: Ah, count within the range. Sorry. Didn't hear you. Big voice.
8. Ceasara: OK. You go to the longest lasting battery on the pink. OK, and then, narrow it down to the top 10. Count of 10.
9. Kay: Count of 10; OK, I was wondering if I was supposed to do that. OK [places blue range tool bars in the data as shown in Figure 4].
10. Ceasara: And I was saying see like there's 7 green that last longer…
11. Kay: OK, the greens are the *Always Ready*, so let's make sure we keep up with which set is which, OK?
12. Ceasara: OK, the *Always Ready* is more consistent with the seven right there, and then like seven of the *Tough* ones, they're like the … further back … I was just saying 'cuz like all 7, 7 out of 3—7 out of 10 of the greens were the longest, and there was…
13. Kyle: Good point.
14. Jamie: I understand.
15. Kay: You understand? OK, Jamie, I'm not sure I do. So could you say it for me?
16. Jamie: She's understanding, I mean she's saying that out of 10 of the batteries that lasted the longest, 7 of them are green, and that's the most number, so the *Always Ready* batteries are better, because more of those batteries lasted longest.
17. Kay: Jason? Jessica, no.

18. Jason: Ah, see, still, the pink ones, the *Tough Cell*, has more higher ones, like even though it does have more in the end? There's a bunch of close ones in the pink right next, almost in that area. And so then if you put all those in, you'd have 7.
19. Kay: So you're saying if I open this out a little bit (referring to the two bars on the range tool). Well, maybe, Ceasara, you can explain to us why you chose 10. That would be really helpful.
20. Ceasara: All right, there was 10 of the *Always Ready*, and there was 10 of the *Tough*. So that's 20, and half of 20 is 10, so that's how I chose it.
21. Kay: But why would it be helpful for us to know about the top 10? Why did you choose that? Why did you choose 10 instead of 12?
22. Ceasara: Because I was trying to go with the half.
23. Kay: Ah. OK. Blake?

At this point, Blake introduced a different way to think about structuring the data that involved using the value bar feature. He referred to it as the "representative value" and asked me to place it in the data at 80 hr as shown in Figure 5.

24. Blake: Can you put the representative value up there please?
25. Kay: I sure will.

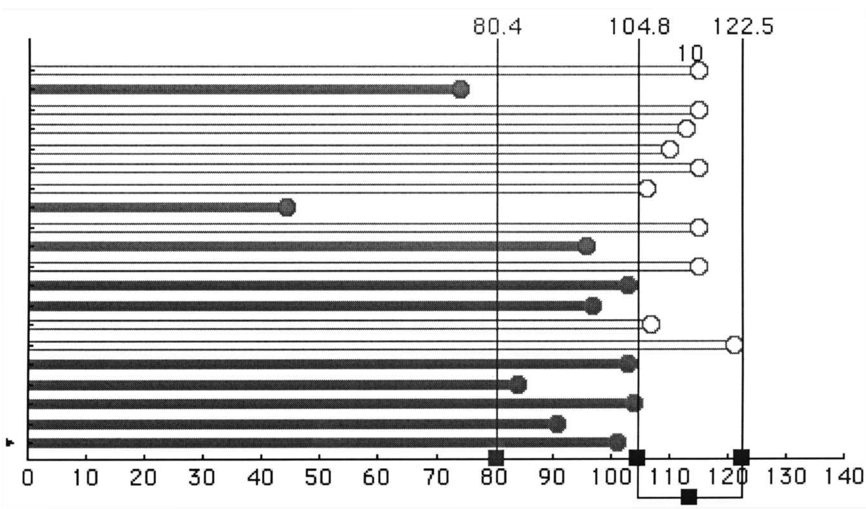

FIGURE 5 Batteries data with range tool and value bar activated.

26. Blake: Will you put it on 80?
27. Kay: I don't know if I can get it [attempts to place the value bar on exactly 80 hr].
28. Blake: That's close.
29. Kay: Is that close enough (drags red value bar into data as shown in Figure 5)?
30. Blake: Now, see, there's still green ones behind 80, but all of the *Tough Cell* is above 80. So I'd rather have a consistent battery that I know that'll get me over 80 hours than one that just try to guess.
31. Kay: Why? Why were you picking 80?
32 Blake: Well, because most of *Tough Cell* batteries were all over 80.
33. Kay: Ah. OK. So it's like a lower limit for you. OK. Questions for Blake? Yes, Jamie?
34. Jamie: Um, why wouldn't the *Always Ready* batteries be consistent?
35. Blake: Well, because all your *Tough Cell* is above 80, but you still have 2 that are behind 80 in the *Always Ready*.
36. Jamie: I know, but that's only 3 out of 10.
37. Blake: No, but see, they only did, what, 10 batteries? So the 2 or 3 will add up. They'll add up to more bad batteries and all that.
38. Kay: Oh, I see; as you get more and more batteries, it's going to get more, more bad ones if that's representative. OK, is that…Jamie?
39. Jamie: So why wouldn't that happen with the *Tough Cell* batteries?
40. Blake: Well, because the way that those 10 batteries show on the chart that they're all over 80 that means that it seems to me that they would have a better quality.
41. Kay: Kyle. [To Jessica] Do you have something to add or ask?
42. Jessica: I was just going to say that well, even though 7 of the 10 longest lasting batteries are the *Always Ready* ones, the 2 lowest are also *Always Ready* and if you were using those batteries for something important then you might end up with one of the bad batteries and could [inaudible]
43. Kay: Yeah, it could. OK, everybody understand what Jessica said?
44. S: No, I didn't…
45. Kay: No? Ceasara?
46. Ceasara: I didn't hear it.
47. Kay: Sequoria, did you understand her? Could you explain what she said?
48. Sequoria: She said that if, even though that the highest 7 were *Always Ready* batteries, the lowest ones were always the *Always Ready* batteries. And if you had something important to do then

you could end up with the ones that were the lowest…you know, it'd jeopardize whatever you were gonna do.

49. Kay: Nice. OK, questions? Any other comments? Yes, Marcus? It is about this, I didn't just mean a random comment about anything.

AIDS Episode

The second episode is taken from the 32nd class session. At this point in the teaching experiment, the students were using the second of the two computer minitools to analyze univariate data sets (see Figure 2). The second minitool had been introduced to the students during the 24th class session. Analysis involving the second minitool typically entailed the students making a comparison between two sets of data about a course of action or a choice of product or treatment. The initial data sets introduced to the students contained equal numbers of data points. In these instances, direct additive comparisons were sufficient for analysis. However, the research team was interested in initiating shifts toward multiplicative ways of reasoning and therefore introduced data sets containing unequal numbers of data points. The *AIDS* task was, in fact, the first instance of a task in which the two data sets contained different numbers of data values (see Figure 6).

In this particular instance, the differences were large enough (46 data values vs. 186 data values) so that multiplicative reasoning might emerge as a tool to solve a problem in the course of analysis. The research team hoped to build on these issues in the course of discussing students' solutions.

By this point in the instructional sequence, students were proficient with the computer minitool. Ways of structuring the data to support arguments involved the following:

1. Partitioning the data at a cut point and reasoning about the number of data points on either side.

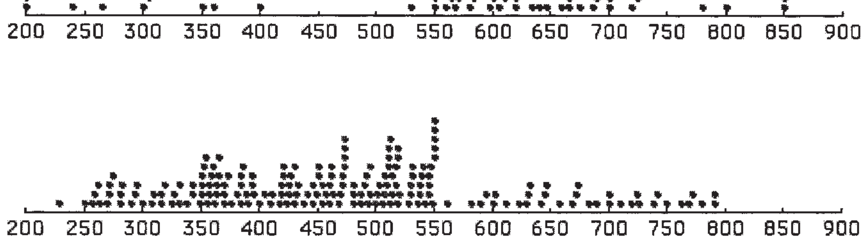

FIGURE 6 *AIDS* data as displayed in the second minitool.

2. Using the four-equal-groups option and reasoning about comparative proportions of the two data sets.
3. Using the fixed-interval-widths option and comparing corresponding cells across data sets.

Students were asked to create informal inscriptions from these ways of structuring the data to accompany their arguments. Because students usually agreed on the course of action from their analysis, discussion shifted toward focusing on the adequacy of reports and their accompanying inscriptions.

In the course of introducing the AIDS data, it became apparent that the students were quite knowledgeable about AIDS. They engaged in a lengthy discussion of health-related issues and discussed the importance of drug treatments that could boost the autoimmune system. During the data creation discussion, the students were able to delineate features of the study on which they would want to collect data and talked about accounting for variation across the samples. After this discussion, I introduced the data as coming from patients enrolled in two treatment programs—a traditional treatment protocol and a new, experimental protocol. The experimental treatment had fewer patients because many of the people in the traditional treatment did not want to risk their health on untried medicines (see Figure 6).

The students were asked to analyze the data and make a report to a chief medical officer at the hospital in which they would recommend one of the two treatment protocols. Students spent the remainder of the class session at the computers, analyzing the data and writing their reports. Ways of using the computer minitool that emerged as they worked included using the create-your-own-groups feature to partition the data as shown in Figure 7 and the four-equal-groups feature as shown in Figure 8.

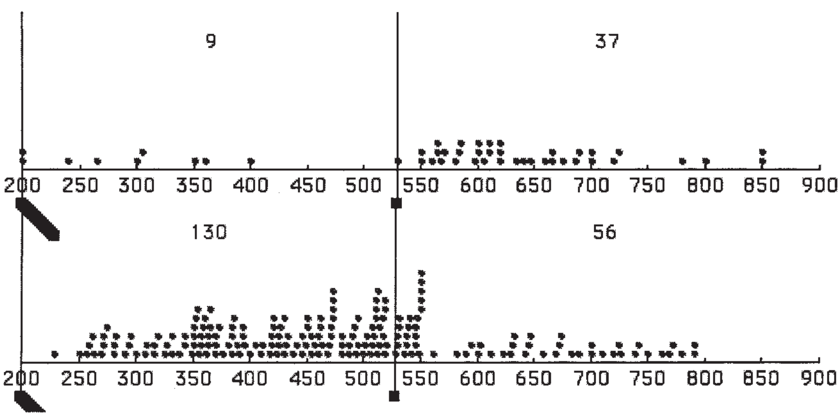

FIGURE 7 Data partitioned by using the create-your-own-groups feature.

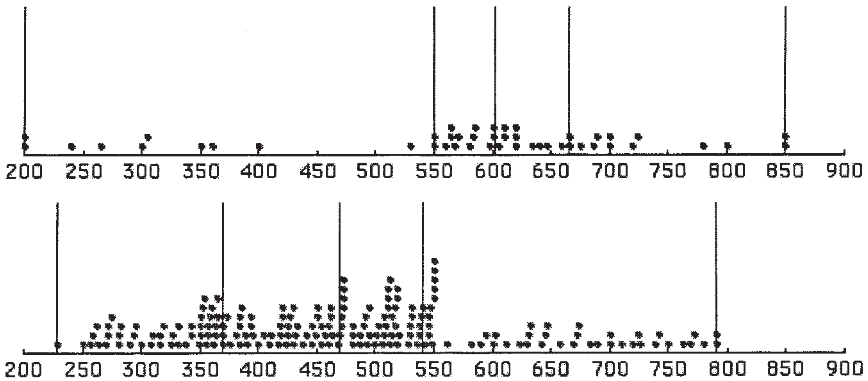

FIGURE 8 Data partitioned by using the four-equal-groups feature.

Because the students' analyses required the entire class session, the research team was able to collect their reports and use them in deciding how best to structure the subsequent whole class discussion. In these deliberations, the research team determined that it would be beneficial to reproduce the students' reports and inscriptions on chart paper so that they could be displayed for whole class discussion.

AIDS Subepisode 1. I began the whole class discussion of the *AIDS* task by explaining the task to the class.

50. Kay: What I want to do today is I want to take a look at some of the things that you did and let's talk about them. I want to see if as a group I want us to look at them and decide if we think that they are an adequate way to represent this data and if we actually understand what these folks are doing. So start with one?

At this point, the first report was put on the board for critique. This first inscription was based on attention to where the majority of the data points were clustered in each of the two data sets by noting where one might find the "highest numbers" of data points as shown in Figure 9.

AIDS Subepisode 2. This subepisode begins with Jamie giving her interpretation of the adequacy of the first report.

51. Kay: Jamie, one more time, big voice.
52. Jamie: I think it's a pretty adequate way of showing information because you can see where the range is starting and ending and you can see where the majority of the numbers are.

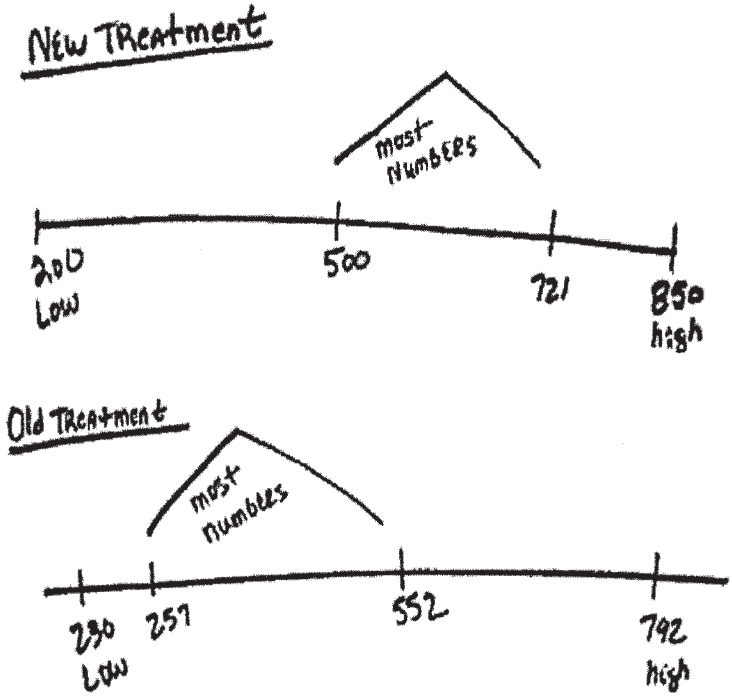

FIGURE 9 Inscription 1 of *AIDS* data.

53. Kay: OK, comments about what Jamie said, or other comments about this?
54. Derrick: I didn't hear what she said.
55. Kay: You didn't hear what she said. Jamie, you just have to really, you have a really quiet voice and it's a big room.
56. Jamie: OK, I think that it's an adequate way of showing the information because you can see where the range is started and ended and you can see where the majority of the numbers were.
57. Derrick: What do you mean by the majority of the numbers?
58. Kay: What do you mean he, Derrick, doesn't know what you mean by the majority of the numbers.
59. Jamie: Where most of the numbers were.
60. Derrick: Where most of the numbers are…
61. Kay: Sheena, can you help?
62. Sheena: Like, when she talks about like when she says, like when you say where the majority of the numbers were, where the, where the point is, like, you see where it goes up?

63. Kay: I do see where it goes up.
64. Sheena: Yeah, like right in there, that's where the majority of it is.
65. Kay: OK.
66. Derrick: The highest range of the numbers?
67. Sheena: Yeah.
68. Kay: The highest range?
69. Sheena: Oh, no.
70. Vallory: No.
71. Kay: Vallory?
72. Vallory: However many people were tested, that's where most of those people fitted in, in between that range. With those little…
73. Paul: You mean this range here?
74. Sheena, Jamie, and Vallory: Yes.

AIDS Subepisode 3. The next two solutions were then posted on the board. They were similar in argument, although one group had written out the information whereas the other had made a table of values, as shown in Figure 10. Both groups partitioned the data by creating cut points, one at 550 T-cells and the other at 525 T-cells. These two inscriptions can be seen in Figure 10.

75. Kay: OK, these folks kind of wrote out what they were talking about. This one right here they said, "The new drug was better than

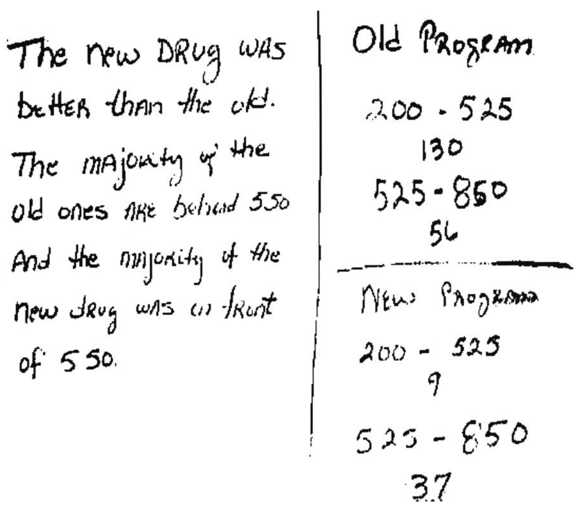

FIGURE 10 Inscriptions 2 and 3 of the *AIDS* data.

the old. The majority of the old ones are behind 550, and the majority of the new drug was in front of 550."

76. Paul: First of all, do you know what that person did, what they were talking about and how they were thinking about it?
77. Kay: Vallory?
78. Vallory: Why did they…550? I don't know? Why is 550 so important? Because the median is really 500—no it's not, but it's not 550.
79. Kay: Kiri?
80. Kiri: Because 550 is in the middle of the whole thing, like the whole, the whole scale; 550 is in the middle. It might not be the middle of the data, but it's the middle of whole scale.
81. Kay: Oh, So it's like the middle of the range, not necessarily the middle of the…Megan?
82. Megan: 'Cause, me, when I looked at it, 550 had, it's kind of where…
83. Kay: Wait just a minute. Derrick, you are going to have to sit up and put that hood down.
84. Kyle: He doesn't have his hood off.
85. Kay: Derrick.
86. Marcus: He's sleeping.
87. Kay: Sit up. I'm sorry, can you start over.
88. Megan: When I looked at the computer, they were like lined up straight on 550. And then starting to get above it, or all the dots for the old treatment, and the new treatment had like…
89. Kay: Oh, so it was like a natural break.
90. Paul: When I was going around talking to some groups, and I think it was the group over in the corner here said something similar. What I heard them say is the same thing. They were trying to find the place sort of between the hills. So if they looked at where this kind of hill started and where this kind of hill finished, they said it's around 550. And so that's why they took that break.
91. Kay: OK, and this group did a similar thing because they said they looked at how many of the T-cells counts were between 200 and 525, and they looked at how many were between 525 and above. So they used 525, and these people used 550. So they were doing like you were talking about Megan, looking at where the hills started to change on the graph. Questions or comments about these two ways? Marissa?
92. Marissa: I would think the second one would be more confusing because it has, since the old program has more numbers than the new program.

93. Kay: Oh. So it looks like that there's more. They had 56 that were above 525, and they only had 37?
94. Marissa: So it's like, I guess what I'm trying to say is it's harder to compare them.
95. Kay: What about what Marissa said? She just said there were more people in the old program so if you actually looked at the actual numbers of people, you find out that they had 56 that were in this upper range, which is where we want to be, and these only had 37. So somebody might say the old program was better because there were more. Jamie?
96. Jamie: This is related. You know like, like a, like a scale drawing where it's not the same size, but it's the same thing. If you look at it, so you have to…
97. Kay: I know what you're trying to say.

As the discussion continued, Kyle suggested that the class construct a type of graphical diagram of the second report that would be "more helpful" in making comparisons across the data sets. The result can be seen in Inscription 4 in Figure 11.

AIDS Subepisode 4. This subepisode begins with a discussion of the diagram that was produced from the data in the previous reports as shown in Figure 11.

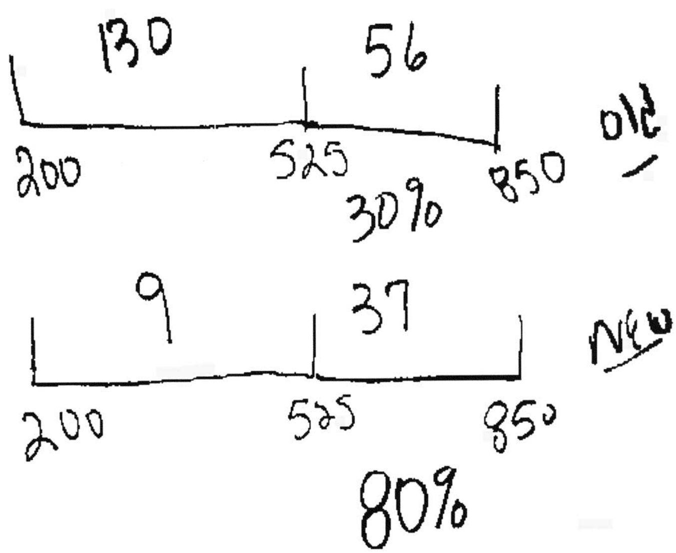

FIGURE 11 Inscription 4 of the *AIDS* data.

98. Kay: So Kyle said from this information he was able to make this diagram and that by looking at information he could also have the diagram. And that that was helpful to him.
99. Paul: I've got a question for everybody. Couldn't you just argue, hey, this shows really convincingly that the old treatment was better, right? Because there were 56 of them, 56 scores above 525, 56 people with T-cell counts above 525, and here there's only 37 above, so the old one just had to be better, there's more people. I mean, there's 19 more people in there, so that's the better one, surely.
100. Kay: Blake?
101. Blake: But then there's more people with the old program than there is with the new program.
102. Kay: Jason.
103. Jason: Then you see that there's 37 more than half over 525 and 56 is not more than half of 130…more of them on the bottom than on the top.
104. Kay: OK.
105. Paul: Can somebody help me out? Can somebody paraphrase what Blake and Jason are talking about?
106. Kay: What was Jason just saying? Somebody help us out with what Jason was just saying.
107. Derrick: I couldn't hear him.
108. Kay: You couldn't hear him?
109. Brian: If you would pay attention.
110. Derrick: OK, I couldn't hear him.
111. Kay: OK, guys come on.
112. Jason: OK, you see how 37 is more than half of 9 and 37 together? But 56 is not more than half of 30, 130, and 56 put together. There's more on the bottom one than on the top one.
113. Kay: OK, who can help me out with that, who can say that a different way so that I might could understand that? Will, can you say it a different way?
114. Will: Well, in that situation it wouldn't matter how many people were in there because see like…
115. Kay: Big voice, Will.
116. Will: On the bottom one you have, see what Jason was saying there's more than there is below 525 and so that means that that one is better because the top one it doesn't even have close to half of what the one below 525 is on that one. So that means that if, if that was the same amount of people it had like, if they both had the same amount of people and, but, and they had the numbers and ev-

erything, and this one, the bottom one was a however much more that of…

117. Paul: And Will…
118. Derrick: No, he's trying to say that…
119. Kay: All right, Derrick, you go ahead and then we'll get … Marcus is on. Hang on Will. Big voice now, Derrick.
120. Derrick: He's saying that, that half, like 37 is over half of 526, 3, 5, 6, uh, uh, 56. And uh, like if it was more people than it would probably be higher than, than 6 like…
121. Kay: I see.

AIDS Subepisode 5. The last solution to be shared was one based on a four-equal-groups structuring of the data. The resulting inscription is shown in Figure 12.

122. Kay: All right, so same job as before, guys. Take a look at this one. This is the old treatment and this is the new treatment. Your first job is to see if you can understand what these folks did, and then decide if you think it's adequate. OK, let's take a look. I've got two people, three people…

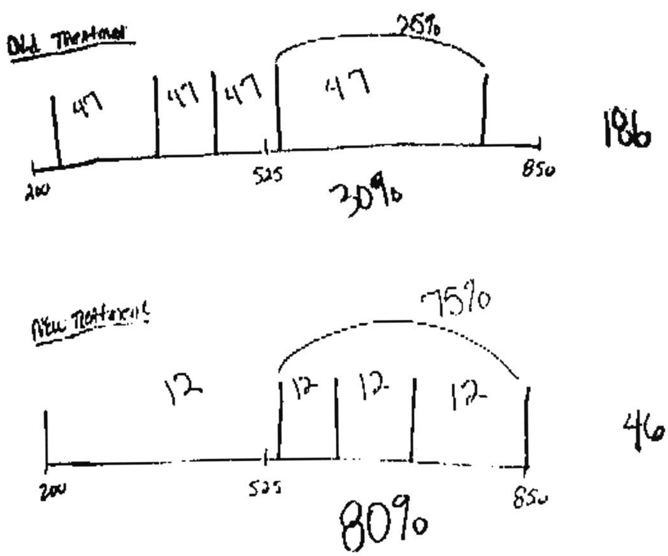

FIGURE 12 Inscription 5 of the *AIDS* data.

123. Marcus: Four people.
124. Kay: Blake, you want to take a shot at it?
125. Blake: Yeah, um…
126. Kay: Big voice, Blake.
127. Blake: I think it's adequate because, uh, for one, like you know when people use the range and stuff but the numbers were the same on this. So that's gonna be like a comparison. They split into four groups.
128. Kay: So it's helpful that the ranges are the same.
129. Blake: And then with the four equal groups, you can tell where the differences in the four groups.
130. Kay: Can you do that by looking at this? Can you? So what do you see when you look at this, Blake?
131. Blake: That the new treatment worked better than the old treatment.
132. Kay: What, I just couldn't hear you.
133. Blake: That the new treatment was better than the old treatment.
134. Kay: And what are you basing that comment on?
135. Blake: Because the three lines for the equal groups were all, what is that—525?
136. Kay: Yeah.
137. Blake: Above 525 compared to only one of them was over on the top.
138. Kay: Ah. Marcus, something to add?
139. Marcus: Yeah, last time like, it was like kind of crooked, and this time it's easier to see because it's right under each other. So then you have to sort of move it in. Like, I'm not trying to copy what Blake said. But it is good though. Anyway, so it might have helped if they put the numbers in the groups so you have a better idea of what you're seeing.
140. Kay: Put the numbers in here?
141. Marcus: In say the fourth group would be a 12 in there, and just write the number 12 in there so you know…
142. Kay: All right, now guys, if this is, if they divided it into four equal groups, which is what Blake said, then is there some way for us to know how many are in each of those groups?
143. Derrick: No.
144. Kay: Could we do that right now? Derrick says no.
145. Brian: If we knew how many…
146. Paul: There's 40, on the bottom one, there were 46 and there was 186 on the top one.
147. Derrick: Divide it into half.

148. Blake: It doesn't really matter.
149. Derrick: Divide it into half.
150. Blake: It doesn't really matter cause you can already, it, even if it was all up against one you that that's still gonna be better because of where…
151. Kay: Wait a minute guys, this is important. Blake, go ahead.
152. Blake: Well, it doesn't really matter where the all the data is because you know from where the groups are what, what treatment is better or where the data stands on both treatments.
153. Kay: So Blake says it doesn't really matter exactly how many, we just know where they are and that's important. McKenna?

REFERENCES

Burrill, G. (1996). Data analysis in the United States: Curriculum projects. In B. Phillips (Ed.), *Papers on statistical education* (pp. 15–26). Hawthorn, Australia: Swinburne.

Burrill, G., & Romberg, T. A. (1987). Statistics and probability for the middle grades: Examples from mathematics in context. In S. P. Lajoie (Ed.), *Reflections on statistics: Learning, teaching, and assessment in K–12* (pp. 33–62). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Cobb, G. W. (1998). Mere literacy is not enough. In L. A. Steen (Ed.), *Why numbers count: Quantitative literacy for tomorrow's America* (pp. 75–91). New York: College Examination Board.

Doerr, H. M. (1995, April). *An integrated approach to mathematical modeling: A classroom study*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco.

Gal, I. (2000). Assessing statistical knowledge as it relates to students' interpretation of data. In S. P. Lajoie (Ed.), *Reflections on statistics: Learning, teaching, and assessment in K–12* (pp. 275–298). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Gravemeijer, K., Cobb, P., Bowers, J., & Whitenack, J. (2000). Symbolizing, modeling, and instructional design. In P. Cobb, E. Yackel, & K. McClain (Eds.), *Symbolizing and communicating in mathematics classrooms: Perspectives on discourse, tools, and instructional design* (pp. 225–274). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Hancock, C. (1995). The medium and the curriculum: Reflection on transparent tools and tacit mathematics. In A. diSessa, C. Hoyles, R. Noss, & L. Edwards (Eds.), *Computers and exploratory learning* (pp. 221–240). Heidelberg, Germany: Springer-Verlag.

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

Konold, C., Pollatsek, A., Well, A., & Gagnon, A. (1987). Students analyzing data: Research of critical barriers. In J. B. Garfield & G. Burrill (Eds.), *Research on the role of teachnology in teaching and learning statistics: 1996 Proceedings of the IASE Round Table Conference* (pp. 151–167). Voorburg, The Netherlands: International Statistical Institute.

Lajoie, S. P. (1997). Reflections on a statistics agenda for K–12. In S. P. Lajoie (Ed.), *Reflections on statistics: Learning, teaching, and assessment in K–12* (pp. 299–316). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Lajoie, S. P., & Romberg, T. A. (1997). Identifying an agenda for statistics instruction and assessment in K–12. In S. P. Lajoie (Ed.), *Reflections on statistics: Learning, teaching, and assessment in K–12* (pp. i–vii). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Lipson, K., & Jones, P. (1996). Statistics: Towards the 21st century. In B. Phillips (Ed.), *Papers on statistical education* (pp. 67–73). Hawthorn, Australia: Swinburne.

McGatha, M., Cobb, P., & McClain, K. (1998, April). *An analysis of students' statistical understandings*. Paper presented at the annual meeting of the American Education Research Association, San Diego, CA.

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

National Council of Teachers of Mathematics. (1989). *Curriculum and evaluation standards for school mathematics*. Reston, VA: Author.

National Council of Teachers of Mathematics. (1991). *Professional standards for teaching mathematics*. Reston, VA: Author.

National Council of Teachers of Mathematics. (2000). *Standards 2000*. Reston, VA: Author.

Shaughnessy, J. M. (1992). Research on probability and statistics: Reflections and directions. In D. Grouws (Ed.), *Handbook of research on the teaching and learning of mathematics* (pp. 465–494). New York: Macmillan.

Shaughnessy, J. M. (1996). Emerging research issues in the teaching and learning of probability and statistics. In B. Phillips (Ed.), *Papers on statistical education* (pp. 39–48). Hawthorn, Australia: Swinburne.

Wilensky, U. (1997). What is normal anyway? Therapy for epistemological anxiety. *Educational Studies in Mathematics, 33,* 171–202.