THE EVOLUTION OF TEACHERS' UNDERSTANDINGS OF DISTRIBUTION

Kay McClain
Arizona State University, United States of America
kay.mcclain@yahoo.com

This paper provides an analysis of the evolution of the statistical understandings related to exploratory data analysis of a cohort of middle-school mathematics teachers. The analysis is grounded in a design experiment in the context of teacher development where the teachers' understandings of statistical data analysis, in particular distribution, were the mathematical endpoint. Activities from an instructional sequence designed to support ways to reason statistically about data were the basis of the engagement. Analyses of the episodes in this paper document the teachers' learning that occurred.

INTRODUCTION

In his seminal chapter in Steen's (1990) *On the Shoulders of Giants*, Moore (1990) makes a compelling argument for the teaching of statistics: "If the purpose of education is to develop broad intellectual skills, statistics merits an essential place in teaching and learning" (p. 134). An advocate of statistics instruction throughout schooling, Moore notes that he is not calling for "detailed instruction in specific statistical methods for their own sake" but argues that statistical thinking, broadly defined, "should be part of the mental equipment of every educated person" (p. 134-135). Moore's call for attention to statistical ways of reasoning is framed in arguments that are premised on meaningful explorations of data instead of an exclusive emphasis on procedures. Cobb (1991; 1992) supports this stance by noting that almost any course in statistics can be improved by "more emphasis on data and concepts at the expense of less theory and fewer recipes."

However, in order for changes to occur in the teaching of statistics in K-12 classrooms in the United States of America (USA), teachers must become knowledgeable about the content. The purpose of this paper is to contribute to these efforts by providing an analysis of the development of one group of teachers' understandings of statistical data analysis. In particular, the analysis in this paper focuses on a collaborative effort conducted between the author and a cohort of seventeen middle-school teachers. The instructional activities utilized were taken from an instructional sequence designed to support the development of sophisticated ways of reasoning statistically about univariate data. The overarching goal of the sequence is a focus on reasoning about data in terms of distributions.

INSTRUCTIONAL SEQUENCE

The goal in developing the instructional sequence that was the focus of the teacher collaboration was to create a coherent sequence that would tie together the separate, loosely related topics that typically characterize USA middle-school statistics curricula. The statistical notion that emerged as central to instruction from a synthesis of the literature was that of distribution. Moore (1990) points to distribution as an "important part of learning to look at data" (p. 106). In the case of univariate data sets, for example, this focus on distribution enables one to treat measures of center, spreadout-ness, skewness, and relative frequency as characteristics of the way the data are distributed. In addition, it allows conventional graphs such as histograms and box-and-whiskers plots to be viewed as different ways of displaying distributions. The development efforts were also guided by the premise that the integration of computer tools was critical in supporting the mathematical goals. The instructional sequence, in fact, involves two computer tools for data analysis. In the initial phase of the sequence, the first computer tool serves as a means to manipulate, order, partition, and otherwise organize small sets of data in a relatively routine way. When data are entered into the tool, each individual data value is shown as a bar, the length of which signified the numerical value of the data point (see Figure 1). A data set was therefore shown as a set of parallel bars of varying lengths that were aligned with an axis. The first computer tool also contained a value bar that could be dragged

© ICMI/IASE 2008. In C. Batanero, G. Burrill, C. Reading & A. Rossman (Eds.), *Joint ICMI/IASE Study: Teaching Statistics in School Mathematics. Challenges for Teaching and Teacher Education. Proceedings of the ICMI Study 18 and 2008 IASE Round Table Conference.*

along the axis to partition data sets or to estimate the mean or to mark the median. In addition, there was a tool that could be used to determine the number of data points within a fixed range. The second computer tool 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 tool were, in effect, collapsed down onto the axis so that a data set was now shown as a collection of dots located on an axis (i.e., an axis plot as shown in Figure 2).

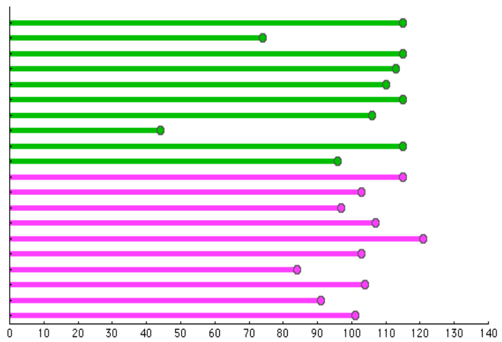

Figure 1. First computer-based tool

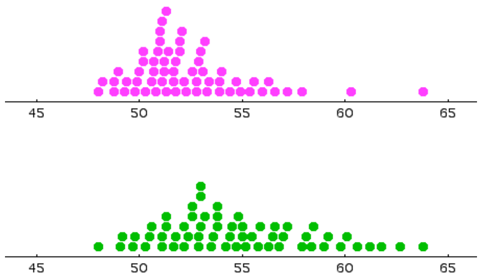

Figure 2. Second computer tool

The second computer tool offered a range of ways to structure data. The options included: (1) making your own groups, (2) partitioning the data into groups of a fixed size, (3) partitioning the data into equal interval widths, (4) partitioning the data into two equal groups, and (4) partitioning the data into four equal groups.

ANALYSIS OF EPISODES

In the first work session conducted with the teachers, I presented a series of data tasks designed to assess the teachers' current understandings of the utility of various graphs and their ability to develop data based arguments from these representations. The tasks involved asking the teachers to find ways to represent a set of data in order to make an argument. As an example, the teachers were given data on the number of hours that thirty seventh-grade students watch television in a week and asked to create a display that could be shared with parents. Their activity on this task was focused on creating the correct form of the graph, including appropriate titles and labels, and colorful displays. They argued about which type of graph should be used with this data. They believed there was a mapping between the type of data and the correct graph for that data.

In the discussion that followed, all of the explanations offered by the teachers focused on the result of calculating measures of center. Further, they argued among themselves about which average (e.g., mean, median or mode) was most useful. These arguments were based on rules they had formulated for each measure. As an example, they used statements such as, "when you have an outlier, you can't use the mean", and "you can only use the mode if it's in the middle of the range." I was not surprised by their discussions as much instruction in statistical data analysis at the middle grades in the USA is focused on learning to correctly calculate measures of center and to create graphs from a set of prescriptive procedures. Their activity fit with this notion of statistics.

Batteries Task

The next task investigated by the teacher cohort was an analysis of data on the longevity of two brands of batteries. The results of tests of ten each of two different brands of batteries were provided as shown in Figure 1. The intent of the task was to problematize comparison of the means. The data sets were distributed such that the more consistent brand of battery had the lower mean. A focus on the variability, or distribution, would therefore be important in the analysis.

As the teachers began their analysis, many of them first calculated the mean. Although most teachers wanted to the know the value of the mean, they subsequently discounted it as "not useful," arguing that the two means were "so close." They then shifted to less conventional ways of structuring the data. In particular, most of the teachers created cut points in the data and reasoned about the number or proportion of data values above or below the cut point. As an example, Van created cut points at 100 hours, 110 hours and 120 hours, noting the number of batteries from each brand above the cut point. In each case he found more Brand A batteries above the cut point. This was justification for his selecting Brand A. However, Regis countered Van's argument by noting that Brand A has "a lot lower lows." Regis went on to state that the two lowest Brand A batteries were "too low and unacceptable." This, he reasoned, was an argument in favor of Brand B. The conversation continued when Mick noted, "if you throw out the bottom two in each group then they are really close." Julian then shifted the discussion to the "top ten of the twenty" by stating that seven of the top ten were Brand A. This discussion segued into probabilistic statements such as "the chance of a good battery" is greater with Brand A.

It is important to note that in their discussions, the teachers were focused on comparing and contrasting the procedures they employed. Their conversations did not build from each other's analyses but resembled a 'show and tell' of ways of structuring the data. In these conversations, procedure took precedent over analysis. In addition, the teachers focused on particular data values within the set (e.g. those above 100 hours, the two low values) instead of the complete distribution. For this reason, their activity could be characterized at this point as reasoning about collections of data points.

Throughout the discussion of the teachers' arguments, I worked to support the need for warrants for each of the claims. As an example, as Van discussed the results of his cut points, I questioned him about his choice of 100 hours, 110 hours and 120 hours. He responded that, "those were the numbers on the axis." For me, this was not a justifiable warrant for his claim. The warrant needed to be grounded in the data. However, none of the teachers questioned his response. In a similar manner, when I asked Julian why the top ten batteries would be significant, she responded that she "used the median." Again, I viewed this as procedural and not grounded in the analysis.

Braking Distance Task

The next task introduced to the teacher cohort involved analyzing data on the braking distances of ten each of two makes of cars (see Figure 3). Similar to the batteries task, this task was also intended to elicit a focus on the distribution of values instead of the use of the means. I introduced the task by first talking through the data creation process with the teachers. In doing so the teachers clarified how the data might be collected to allow for reasonable conclusions to result.

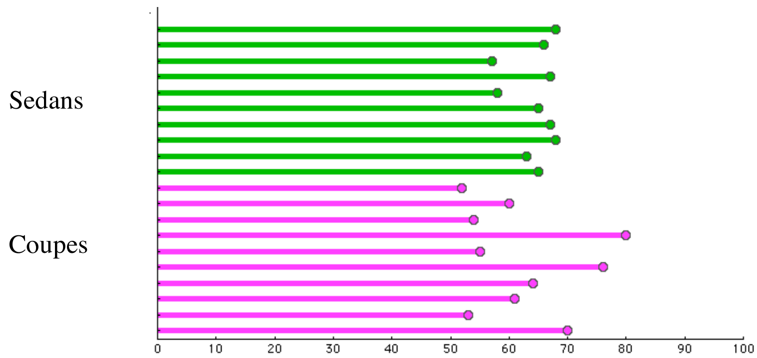

Figure 3. Braking distance data as inscribed in first computer tool.

I asked the teachers to decide which brand of car they thought was safer, based on these data. As the teachers began their analyses, many of them again initially calculated the mean of each set of data. They subsequently judged that measure to be inadequate for making the decision and proceeded to find ways to structure the data using the computer tool that supported their efforts at analysis. In this process, they placed vertical lines in the data to create cut points and to capture the range of each set of data. As an example, Mary Jean noted that, "A full forty percent of the coupes stopped in less than 60 feet, and if you go to 62 [feet] it goes to sixty [percent], and there are only twenty percent of the sedans below even 62 [feet]." But Gayle disagreed noting that, "Those two that took a long time to stop are significant." She continued by stating that, "all the sedans stop in around sixty to seventy feet and it might even be better (pointing to the two data values that were less than sixty feet)." Alice followed by arguing that "all of the sedans took over 58 feet to stop" whereas "forty percent of the coupes were able to stop in less than 58 feet."

It is important to note that the discussions of the various solutions were based on what the teachers judged to be important about braking not about the ways of structuring the data. This was a shift from their prior activity on the batteries task where their discussions were focused on the details of their procedures instead of the results of their analysis. An important distinction can therefore be made in the ways the teachers reasoned about these two tasks. Although both discussions focused on reasoning around partitions, in the latter task the teachers were able to justify their way of structuring the data with respect to the analysis. This is significant in that it indicates that although the process of partitioning was emerging as normative, the teachers were still teasing out significant aspects of this activity. As part of this process, the importance of justifying the location of the cut point was still being negotiated within the group. For example, creating cut points and reasoning about percentages or proportions of the data set above or below the cut point was accepted without justification as a way to structure the data. Questions arose not over the method (e.g., creating cut points) but over warrants for the claims. As an example, Gayle's disagreement with Mary Jean's argument was not based on the manner in which Mary Jean had structured the data but on the choice of cut point and the resulting conclusion she reached as a result of her particular cut point. This way of reasoning had first emerged on the batteries task and was now characteristic of the arguments that were presented. As a result, what was becoming established in the course of public discourse was partitioning data sets and reasoning about the proportions or relative frequencies formed. Arguments then were formulated to justify claims made from such partitions—not to justify the act of partitioning and reasoning about proportions. This is an important distinction in that it indicates that the first normative way of reasoning or mathematical practice within the teacher cohort was that of partitioning data sets and reasoning about the relative frequencies.

Speed Trap Task

A shift to the second computer tool began with the introduction of the speed trap task. This task was based on data collected on the speeds of two sets of sixty cars. The first data set was collected on a busy highway on a Friday afternoon. Speeds were recorded on the first sixty cars to pass the data collection point. The second set of data was collected at the same location on a subsequent Friday afternoon after a speed trap had been put in place for a week. The goal of the speed trap (e.g., issuing a large number of speeding tickets by ticketing anyone who exceeds the speed limit by even 1 mile per hour) was to slow the traffic on a highway where numerous accidents typically occur. The task was to determine if the speed trap was effective in slowing traffic (see Figure 3 for the speed trap data displayed in the second computer tool). The intent of the task was to provide a setting in which the distributions of data values would provide an opportunity for attention to 'shape of the distribution' to emerge.

As the teachers worked on the data, most of them created cut points at the speed limit and reasoned about the number of drivers exceeding the speed limit both before and after the speed trap. One teacher focused on the shape of the two data sets and noted that at first it "looked like a Volkswagen Beetle", and then it "flattened out like a large Town Car." I found this particularly significant because it was the first occasion where a teacher found a way to describe the shape of the distribution.

In analyzing the form of the teachers' arguments, it appeared that the teachers were reasoning about the data as aggregate (cf. Konold, et al., in preparation). In particular, the perceptual unit in their analysis was the entire distribution of values. They reasoned about the relative number of cases in various parts of the distribution (e.g., exceeding the speed limit) and did so in terms of percentages and/or proportions, while reasoning across the two data sets. For this reason, they were concerned with the relative density of the data in certain intervals across the two data sets. This can be characterized as a concern for the amount of data clustered within an interval across the data sets (e.g., number of cars speeding both before and after the speed trap). In the course of making their arguments, the teachers were able to coordinate the differences in the frequencies (e.g., analogous to the *y* values) across the *x*-axis in a multiplicative sense (cf. Thompson, 1994). This type of reasoning was typical of the solutions developed by the teachers and indicates that the second normative way of reasoning involved a concern for relative density across data sets where the teachers viewed data as aggregate.

AIDS Task

The final collection of tasks in the instructional sequence involved data sets with unequal numbers of data points. In the first task from this collection, data were presented on two sets of AIDS patients enrolled in different treatment protocols—a traditional treatment program with 186 patients and an experimental treatment program with 46 patients. T-cell counts were reported on all 232 patients (see Figure 4). The task was to determine which treatment protocol was better at producing high T-cell counts. As the teachers worked on their analysis, they initially noted that the clump, cluster, or hill of the data shifted between the two groups. In particular, they characterized the shift by creating cut points around a T-cell count of 525 and reasoning about the percentage of patients in each group with T-cell counts above the cut point. They noted that the cluster of T-cell counts in the traditional treatment program was below the cut point whereas the cluster of T-cell counts in the experimental was above. In addition, they could use the four-equal-groups inscription to further tease out the differences in how the data were distributed. As an example, Diane reasoned that, “seventy-five percent of the patients in the experimental treatment group are in the same range as only 25% of the patients in the traditional treatment.”

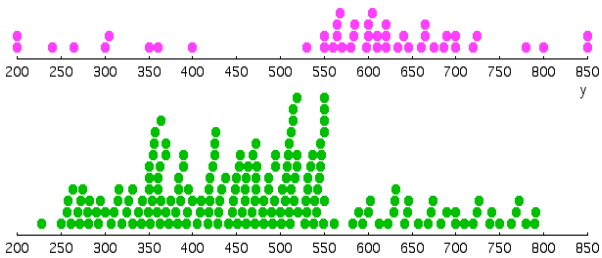

Figure 4. AIDS data displayed in the second computer tool.

Although the teachers all agreed that the experimental treatment was more effective, they engaged in a lengthy discussion concerning the adequacy of the various arguments. In the course of their discussion, they focused on features of the different data displays that would strengthen their argument. Here again, the norms for argumentation did not include discussion of the differing ways of structuring the data but of the resulting argument that could be made. In the course of ongoing conversations, the teachers worked to find ways to characterize the difference in density of data across the axis. In other words, they were attempting to coordinate the differences in the relative densities of the data sets as they clarified their arguments.

Thompson (personal communication, October 2004) notes that the ability to scan the axis from left to right and read the frequency as the rate at which the total accumulates over the *x*-axis is what is entailed in seeing distributions as density functions. This concern for relative density is a step towards what Thompson views as an endpoint in reasoning about distributions. Although data for this analysis do not permit claims about the teachers’ ability to view the data sets in such a sophisticated manner, the results of their analysis do indicate that they were coordinating the relative densities as they worked to find ways to describe the shifts in the data. The third normative way of solving tasks that emerged was therefore that of structuring the data multiplicatively to describe shifts and changes in the distributions by coordinating the relative densities.

CONCLUSION

Research is needed that more fully explicates the process of teacher development that is focused on statistics. This paper does so by addressing both theoretical and pragmatic concerns. Theoretically, it involves reflecting on the activity of planning and conducting professional development interventions to develop conjectures about what might be involved for teachers to pursue an instructional agenda that aims at developing statistical understanding of significant ideas. Pragmatically, it supports teachers in this endeavor by engaging them in similar learning processes. In this way, the field can build theory in a grounded, systematic way.

REFERENCES

Cobb, G. (1991). Teaching statistics: More data, less lecturing. *Amstat News*, December 1991, p. 1, 4.

Cobb. G. (1992). Teaching statistics. In L. Steen (Ed.), *Heeding the call for change: Suggestions for curricular action*. MAA Notes 22. Washington, D.C.: Mathematical Association of America.

Konold, C., Higgins, T., Russell, S. J., & Khalil, K. (2004). Data seen through different lenses. Unpublished manuscript. Amherst, MA: University of Massachusetts.

Moore, D. (1990). Uncertainty. In L. Steen (Ed.), *On the shoulders of giants*, (pp. 95–138). Washington, D.C.: National Academy Press.

Steen, L. (1990). *On the shoulders of giants*. Washington, D.C.: National Academy Press.

Thompson, P. W. (1994). The development of the concept of speed and its relationship to concepts of speed. In G. Harel and J. Confrey (Eds.) *The Development of Multiplicative Reasoning in the Learning of Mathematics*. (pp. 179-234). NY: SUNY Press.