SUPPORTING TEACHER CHANGE: A CASE FROM STATISTICS

Kay McClain

Vanderbilt University

This paper provides an analysis of a teacher development experiment (cf. Simon, 2000) designed to support teachers' understandings of statistical data analysis. The experiment was conducted as part of collaborative efforts between the author and a cohort of middle-school mathematics teachers during the 2000-2001 academic year. Analysis of the episodes in this paper document the evolution of the teachers' understandings as they participated in activities from an instructional sequence designed to support conceptual understanding of statistical data analysis. In this process, I highlight the mathematical issues that emerged as the teachers worked to further their own understandings.

INTRODUCTION

The purpose of this paper is to provide an analysis of the development of one group of teachers' understandings of statistical data analysis. The analysis builds from the literature on students' understandings by taking prior research in classrooms as a basis for conjectures about means of supporting teachers' development. In particular, the analysis in this paper will focus on a collaborative effort conducted between the author of this paper and a cohort of middle-school teachers. The collaboration occurred during the 2000-2001 academic year. The teachers participated with the author in monthly work sessions designed to support their understandings of effective ways to teach statistical data analysis in the middle grades (ages 12-14). Fundamental to this effort was attention to the development of the teachers' content knowledge. The instructional activities utilized during the course of the collaboration were taken from a classroom teaching experiment conducted with a group of seventh-grade students during the fall semester of 1997 (for a detailed analysis of the classroom teaching experiment see Cobb, 1999; McClain & Cobb, 2001). The intent of the instructional sequence is to support middle-school students' development of sophisticated ways to reason statistically about univariate data. The overarching goal is that they come to reason about data in terms of distributions. Inherent in this understanding is a focus on multiplicative ways of structuring and organizing data.

The intent of the teacher collaboration was then to take the seventh-grade instructional sequence as a *means of support* for the learning of the teacher cohort. This support included tasks from the instructional sequence, computer-based tools for analysis that accompanied the sequence, the teachers' varied inscriptions and solutions to tasks from the sequence, and norms for argumentation that were negotiated during the work sessions.

In the following sections of this paper, I begin by describing the theoretical framework that guided the analysis. I then provide a description of the method of analysis and the data corpus. I follow by outlining the instructional sequence that was the basis of the teacher collaboration. Against this background, I provide an analysis of episodes from the work sessions intended to document the teachers' developing understandings of statistical data analysis.

THEORETICAL FRAMEWORK

The analysis reported in this paper was guided by the emergent perspective (cf. Cobb & Yackel, 1996). The emergent perspective involves coordinating constructivist analyses of individual activities and meanings with an analysis of the communal mathematical practices in which they occur. This framework was developed out of attempts to coordinate individual students' mathematical development with social processes in order to account for learning in the social context of the classroom. It therefore places the students' and teacher's activity in social context by explicitly coordinating sociological and psychological perspectives. The psychological perspective is constructivist and treats mathematical development as a process of self-organization in which the learner reorganizes his or her activity in an attempt to achieve purposes or goals. The sociological perspective is interactionist and views communication as a process of mutual adaptation wherein individuals negotiate mathematical meaning. From this perspective, learning is characterized as the personal reconstruction of societal means and models through negotiation in interaction. Together, the two perspectives treat mathematical learning as both a process of active individual construction and a process of enculturation into the mathematical practices of wider society. Individual and collective processes are viewed as reflexively related in that one does not exist without the other. Together, these two aspects provide a means for accounting for the teachers' activity in the social context of the work sessions.

METHOD OF ANALYSIS

The particular lens that guided my analysis of the data was a focus on the normative ways of solving tasks, or what Cobb and Yackel (1996) have defined as *mathematical practices*. These practices focus on the collective mathematical learning of the classroom community and thus enable one to talk explicitly about collective mathematical learning (cf. Cobb, Stephan, McClain, & Gravemeijer, 2001). This analytical lens therefore enabled me to document the collective mathematical development of the teacher cohort over the course of the year.

In order to conduct an analysis focused on the learning of the community, it is important to account for the diverse ways in which the teachers participate in communal practices. For this reason, the participation of the teachers in discussions where their mathematical activity is the focus then becomes data for analysis. The diversity in reasoning also serves as a primary means of support of the collective mathematical learning of the teacher cohort. As a result, "an analysis of classroom mathematical practices characterizes changes in collective mathematical activity while taking into account the diversity in individual [teachers]' reasoning" (Cobb, 1999, p. 10). An analysis focused on the emergence of classroom mathematical practices is therefore a conceptual tool that reflects particular goals (Cobb, et al., 2001).

DATA

Data for this study consist of videorecordings of each monthly work session and of the weeklong summer work sessions. In addition to the videotape there is a set of field notes taken by a research assistant, copies of the teachers' work, copies of their students' work on the same tasks, and audiotape of interviews conducted with each teacher. This comprehensive data corpus allowed for the longitudinal analysis of the emergence of the mathematical practices by testing and refining conjectures against both the activity of the cohort and of the individual teachers within the cohort. This was done in a manner described by Cobb and Whitenack (1996) and is consistent with Glaser and Strauss' (1967) constant comparative method.

INSTRUCTIONAL SEQUENCE

In developing the instructional sequence for the seventh-grade classroom teaching experiment, the goal of the research team was to develop a coherent sequence that would tie together the separate, loosely related topics that typically characterize American middle-school statistics curricula. The notion that emerged as central from the synthesis of the literature was that of distribution. In the case of univariate data sets, for example, this enabled the research team to treat measures of center, spreadout-ness, skewness, and relative frequency as characteristics of the way the data are distributed. In addition, it allowed the research team to view various conventional graphs such as histograms and box-and-whiskers plots as different ways of structuring distributions. The instructional goal was therefore to support the development of a single, multi-faceted notion, that of distribution, rather than a collection of topics to be taught as separate components of a curriculum unit. A distinction that was made during this process which later proved to be important is that between reasoning additively and reasoning multiplicatively about data (cf. Harel & Confrey, 1994; Thompson, 1994; Thompson & Saldanha, 2000). Multiplicative reasoning is inherent in the proficient use of a number of conventional inscriptions such as histograms and box-and-whiskers plots.

As the research team[1] began mapping out the instructional sequence, it was guided by the premise that the integration of computer tools was critical in supporting the mathematical goals. The instructional sequence developed in the course of the seventh-grade teaching experiment in fact involved two computer tools. In the initial phase of the sequence, the students used the first computer tool to explore sets of data. This tool 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. 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 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 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.

---

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

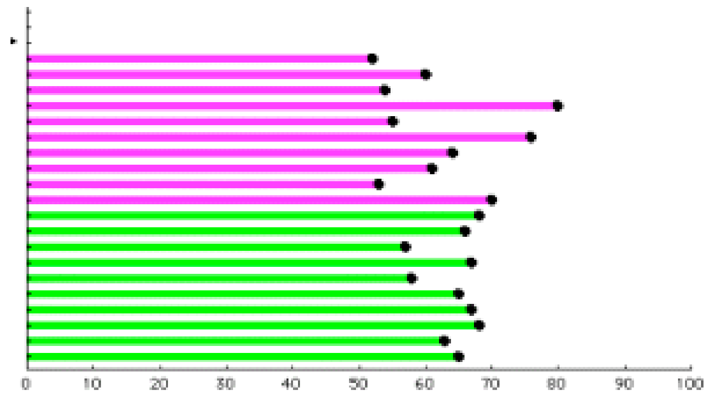

Figure 1. Data displayed in first computer tool.

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 computer 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).

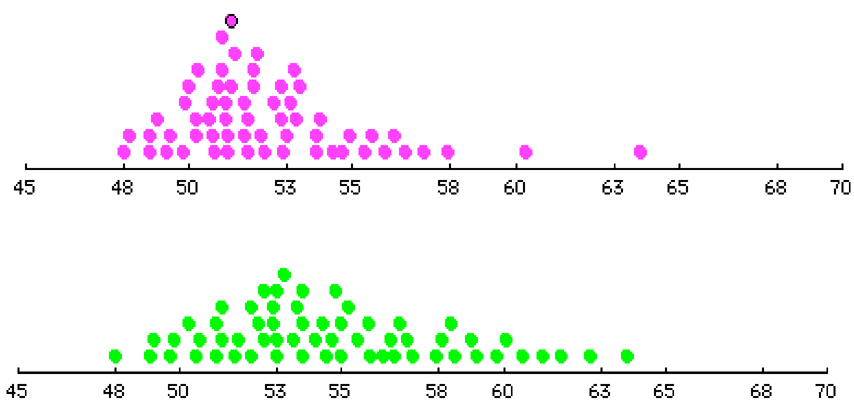

Figure 2. Data displayed in second computer tool.

The 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. The key point to note is that this tool was designed to fit with students' ways of reasoning while simultaneously taking important statistical ideas seriously.

As the research team worked to outline the instructional sequence for the seventh-grade classroom, it 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 would then need to develop ways to analyze and describe the data in order 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 in order to make a decision or a judgment. The students typically engaged in these types of tasks in order to make a recommendation to someone about a practical course of action that should be followed.

RESULTS OF ANALYSIS

The initial activities of the teacher cohort involved the teachers analyzing data on the braking distances of ten each of two makes of cars, a coupe and a sedan. The teachers were given printouts of the data inscribed in the first computer tool as shown in Figure 1 and asked to decide which make of car they thought was safer, based on this data. My decision to use printouts of the data was based on my own experience in working with students on these tasks. I had noticed that when students were asked to make initial conjectures based on informal analysis of the printouts, their activity on the computer tool seemed more focused. They used the features on the tool to substantiate their preliminary analysis instead of to explore the structures that resulted from the use of the features. I was also curious to see if the tools we had designed offered the teachers the means of analyzing data that fit with their initial, informal ways of analyzing the data.

As the teachers began their analyses, they proceeded to find ways to structure the data 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. 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. 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 conclusion she reached as a result of her particular cut point. This was typical of the discussions of all arguments that were presented on tasks using the first computer tool. As a result, what became constituted in the course of public discourse was partitioning data sets and reasoning about the proportions formed. Arguments then had to be 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 that became constituted within the cohort was that of partitioning data sets and reasoning about resulting proportions.

A shift to the second computer tool began with the introduction of the speed trap task. The 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. 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 2 for the speed trap data displayed in the second computer tool).

As the teachers worked on their analysis, most of them drew a vertical line across the two data sets to create a cut point at the speed limit and reasoned about the number of drivers exceeding the speed limit both before and after the speed trap. They used a range of strategies to describe the partitions including ratios and percentages. As they worked, their arguments indicated that they 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. For this reason, they were concerned with the relative density of the data in certain intervals. In particular, they were concerned about the amount of data clustered within an interval across the data sets (e.g. number of cars exceeding the speed limit both before and after the speed trap).

As an example, Regis created cut points at 50 miles per hour (mph), 53 mph, and 55 mph. He then examined the data to look for shifts within those intervals. He argued that before the speed trap, 25 drivers were traveling in excess of 55 mph. After the speed trap, that number was reduced to 10. He then argued, “that’s 15 less and since the sample was 60, 15 out of 60 is 25%. So 25% fewer drivers were speeding.” It appeared that in the course of making this argument, Regis was 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.

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 was 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 3). 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."

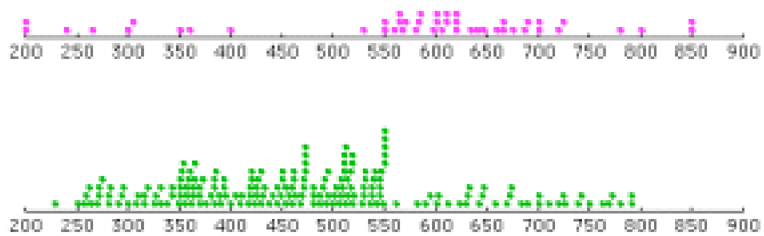

Figure 3. AIDS data displayed in the second computer tool.

Thompson (personal communication, October, 2002) 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. This density perspective is consistent with what Khalil and Konold (2002) found in their analysis of expert data analysts.

Although this analysis does 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 frequencies 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.

CONCLUSION

The resulting shifts that emerged in the normative ways of reasoning indicate a mathematical progression over the course of the year. In particular, the first practice to emerge was that of partitioning data sets and reasoning about resulting proportions. The second practice entailed a concern for relative density across data sets where the teachers viewed data as aggregate. The third and final practice involved the ability to view the data in two data sets distributed on the x-axis and simultaneously coordinate the relative density of the distributions when structured multiplicatively. These normative ways of reasoning or mathematical practices can be thought of as the realized learning route (cf. Simon, 1995) of the community. As a result, the practices that emerged in the course of interaction document the learning of the teacher cohort by characterizing changes in collective mathematical activity while highlighting the diversity in individual teachers' reasoning.

References

Cobb, P. (1999). Individual and collective Mathematical learning: The case of statistical data analysis. *Mathematical Thinking and Learning*, 5-44.

Cobb, P., Stephan, M., McClain, K., & Gravemeijer, K. (2001). Participating in classroom mathematical practices. *Journal for the Learning Sciences, 10*(1&2), 113-164.

Cobb, P., & Whitenack, J.W. (1996) A method for conducting longitudinal analyses of classroom videorecordings and transcripts. *Educational Studies in Mathematics, 30*, 458-477.

Cobb, P. & Yackel, E. (1996). Constructivist, emergent, and sociocultural perspectives in the context of developmental research. *Educational Psychology, 31,* 175-190.

Glaser, B. & Strauss, A. (1967). *The discovery of grounded theory: Strategies for qualitative research*. Chicago: Aldine.

Harel, G. & Confrey, J. (1994) *The development of multiplicative reasoning in the learning of mathematics*. NY: SUNY Press.

Khalil, E. & Konold, C. (2002) *How experts and novices visually inspect graphs*. Poster presented at the annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education, Athens, GA.

Konold, C., Khalil, E., Higgins, T., & Russell, S.J. (in preparation), *Data seen through different lenses*.

McClain, K., & Cobb, P. (2001). Supporting students' ability to reason about data. *Educational Studies in Mathematics, 45*(1-3), 103-129.

Simon, M. A. (2000). Research on mathematics teacher development: The teacher development experiment. In A. Kelly & R. Lesh (Eds.), *Handbook of research design in mathematics and science education*. Hillsdale, NJ: LEA.

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.

Thompson, P. & Saldanha, L. (2000). *To understand post counting numbers and operations*. White paper prepared for the National Council of Teachers of Mathematics.