Developing Experts' Points of View on Local-Global Approaches to Data and Data Representations

Dani Ben-Zvi
dani.ben-zvi@weizmann.ac.il

Abraham Arcavi
abraham.arcavi@weizmann.ac.il

The Weizmann Institute of Science, Department of Science Teaching
P.O. Box 26
Rehovot, 76100, Israel

Introduction

The purpose of this study is to analyze the learning processes of seventh grade students working through a carefully designed Statistics Curriculum (*SC*) in a computerized environment. The study has two parts. In part I, we follow at a very fine level of detail, the ways in which two 13-year-old students began to make sense of data and data representations, as well as adopt and exercise the habits and points of view that are common among experts. In particular, we focus on the ways they started to develop global views (and tools to support them) of data and data representations on the basis of their previous knowledge and different kinds of local observations. In part II we analyze an assessment task administered to 80 seventh grade students two weeks after the end of the *SC*. The goal is to provide a quantitative indication of the extent and scope of the phenomena identified in the first part. In the light of the analyses, we propose a description of what it may mean to learn EDA, and draw educational implications.

The Statistics Curriculum

The Statistics Curriculum (*SC*) Project (described in detail in Ben-Zvi & Friedlander, 1997 and Ben-Zvi & Arcavi, 1998), was developed in order to introduce junior high school students to statistical thinking and the "art and culture" of EDA -- Exploratory Data Analysis (in the spirit of Shaughnessy, Garfield & Greer, 1996). The design of the Project was based on the creation of "small scenarios" in which students can experience some of the processes involved in the experts' practice of data based enquiry. The *SC* stresses: (a) student active participation in organization, description, interpretation, representation and analysis of data situations (on topics close to the students' world), with a considerable use of visual displays as analytical tools (in the spirit of Garfield, 1995); (b) peer collaboration and classroom discussions; and (c) incorporation of technological tools for simple use of various data representations and transformations of them (as described in Biehler, 1993 and Ben-Zvi, 2000). The scope of the curriculum is 20-30 classroom hours, and it includes a "research project" to be carried out by the students (mostly out of the classroom) with the teacher's guidance.

Theoretical Perspectives

In establishing the setting and in conducting the analysis, we relied on both cognitive and socio-cultural perspectives. The cognitive perspective led us to focus on the development and change in students' conceptions and the evolution of their statistical reasoning on the basis of what they knew. The socio-cultural perspective, guided us to focus on learning of a complex domain, such as EDA, as the adoption of the points of view of a community of experts, in addition to skills and procedures (Resnick, 1988). Thus, we looked at learning as an "*enculturation*" process, and the design of the

SC is seen as an attempt to bring "*the practice of knowing mathematics* (statistics) *in school closer to what it means to know mathematics* (statistics) *within the discipline*" (Lampert, 1990, p. 29).

Wild and Pfannkuch (1999) suggest a description of the common practice of statistics experts involved in data based enquiry, which includes processes, perspectives and points of view. In this framework, we focus on a key perspective of experts' view of data: an inclination to see *propensities* (in the sense of Konold, Pollatsek & Well, 1997, p. 151) in data, and a flexible and dynamic shift between local observations and global observations (Ben-Zvi & Arcavi, in press). *Local understanding* of data representation involves focusing on an individual value (or a few of them) within a group of data. *Global understanding* refers to the ability to search for, recognize, describe and explain general patterns in a set of data (e.g., change over time, trends).

Methodology

The students were 80 seventh graders (13-year-old) of mixed ability in a progressive experimental school, taught by three skillful and experienced teachers, who were aware of the spirit and goals of the *SC*. The two students of the case study were above-average ability students, and had good cooperative working praxis. They agreed to participate in this study, which took place within their regular classroom periods and included being videotaped and interviewed, and furnishing their notebooks for analysis. The students were videotaped at almost all stages (20 hours of tapes) and their notebooks were collected. We interviewed the teachers as an additional source of data. In this paper, we concentrate first on the work of the two students investigating real sports data during the first four class periods, and then on the responses of all the students in a final assessment task administered two weeks after the end of the *SC*.

The analysis of the videotapes was based on interpretive microanalysis (see, for example, Meira, 1991, pp. 62-3): a qualitative detailed analysis of the protocols, taking into account verbal, gestural and symbolic actions within the situations in which they occurred. The goal of such an analysis is to infer and trace the development of cognitive structures and the socio-cultural processes of understanding and learning. In order to validate the analysis, it underwent two stages, one within the researchers' team and one with fellow researchers, who had no involvement with the data or the SC Project (triangulation in the sense of Schoenfeld, 1994). In both stages, the researchers discussed, presented, advanced and/or rejected hypothesis and interpretations, and inferences about the students' cognitive structures.

The Case Study

We analyze how the views of the two students slowly changed and evolved from a novice to an expert perspective. We focus on how their use of local information promote the development of global points of view of data sets in different representations (table, graph). This process includes descriptions of how they developed an understanding of:

- what to focus on: from irrelevant towards relevant information;
- how to read and make sense of local information in tables and in graphs;
- how to look at differences between adjacent local entries;
- how to look globally by struggling to learn the ideas underlying global perspectives, the language to describe those ideas, and the difficulties in applying both;
- how to construct global views through handling special local entries (outliers); and
- how to flexibly transfer between local and global views through re-scaling of graphs.

On the basis of the detailed analysis (Ben-Zvi & Arcavi, in press), we argue that learning to look globally at data can be a complex and devious process from a local-pointwise view towards a flexible combination of local and global views of data and data representations. The local pointwise view sometimes restrained the students from "seeing globally", but on other occasions it served as a basis upon which the students started to see globally. Moreover, in a certain context, even looking globally may indicate different meanings for a novice than for an expert. Most of the learning took place through dialogues between the students themselves and in conversations with the teacher. *Appropriation* (in the sense of Moschkovich, 1989) by the teacher (in order to support learning) or by the students (in order to change the sense they make of what they do) seems to be a central mechanism of "enculturation". This study shows that this mechanism is especially salient when students learn the dispositions of the subject matter rather than its skills and procedures.

Part II: The Final Assessment Task

In the second part of this study, we analyze responses to a final assessment task, in which 80 students were asked to examine in pairs real immigration data using spreadsheets and a worksheet with minimal instructions. The goal is to offer a *quantitative* indication of the extent and scope of the phenomena identified in the first part. The analysis provides an indication of students' accumulated knowledge after the end of the *SC*. They were asked to independently formulate research questions and hypotheses, handle data representations and statistical measures, and interpret the results of their investigation. In addition, students' actions were assessed in terms of their local - global views of data.

After a short introduction of the investigation context, the students were asked to formulate research questions and related hypotheses. Their initial responses (without looking at data) were categorized to three types of research questions: global (58%), irrelevant (22%), and local (20%). When prompted by two data entries (number of immigrants in the first and last years) more students (68%) were able to formulate global research questions. Students' responses reflected an understanding of key concepts, such as trends and relationship, although the use of them was often incomplete or vague.

Most of the students were capable of choosing an appropriate graph to display the data (e.g., 62% used time plots). Many chose to display the data in multiple graphical methods, in order to arrive at a more convincing visual result. Furthermore, when investigating trends in the data, 30% of the students changed the design of their graphs, by scaling, adding a regression line, or using colors to emphasize affluent immigration periods. A vast majority (88%) was capable of interpreting their graphs globally by describing trends, cycles, or suggesting various groupings of the data. In addition, 90% of the students were able to present and interpret a frequency distribution of the data (in table and graph).

Concluding Comments

We propose that in learning environments such as the *SC*, the students' learning involves the following: (a) prior knowledge is engaged in multifaceted and sometimes unexpected ways – possibly hindering progress in some instances, but making the basis for construction of new knowledge in others; (b) during the learning process, many questions either make little sense, or, alternatively, are interpreted and answered differently from the original intention; (c) students' work is inevitably based on partial understandings, which grow and evolve towards more complete meanings; and (d) most of the learning takes place through dialogues between the students themselves and in conversations with the teacher.

Students' remarks at the first activities of the *SC* indicated that questions like *"what do you learn from a table/graph?"* may encounter an initial inability to focus attention on relevant (even informal) views of the data representation. Novices seemed to be either at a loss (when asked this kind of questions), or their perceptions of what is relevant are very different from the experts' view. Furthermore, their views of data representations were mostly local at the beginning, and only later slowly changed and evolved towards an expert perspective that includes global points of view of data sets in different representations.

In the final assessment activity, students were capable to make sense of the language and intention of the tasks. Moreover, most of them were able to formulate global research questions and hypotheses, interpret data (in tables and graphs) in a global way attending to general patterns, rather than local features of the data, or intertwining both perspectives. We suggest that the two students observed in the case study, were not an exception among their peers in terms of their local-global view of data.

Finally, this study also confirmed that even if students initially do not make more than partial sense of their tasks, through the support of appropriate teacher guidance, class discussions, peer work and interactions, and ongoing cycles of experiences with realistic problem situations, students slowly build meanings and develop experts' points of view on local-global approaches to data and data representations.

References

Ben-Zvi, D. (2000). Towards Understanding the Role of Technological Tools in Statistical Learning, *Mathematical Thinking and Learning*, **2** (1&2), 127-155.

Ben-Zvi, D., & Arcavi, A. (1998). Towards a characterization and understanding of students' learning in an interactive statistics environment, in L. Pereira-Mendoza (Ed.), *Proceedings of the Fifth International Conference on Teaching of Statistics*, **II**, 647-653. Voorburg, The Netherlands: International Statistical Institute.

Ben-Zvi, D., & Arcavi, A. (In press). Junior high school students' construction of meanings for data representations, *Educational Studies in Mathematics*.

Ben-Zvi, D., & Friedlander, A. (1997). Statistical thinking in a technological environment, in J. B. Garfield & G. Burrill (Eds.), *Research on the Role of Technology in Teaching and Learning Statistics*, 45-55. Voorburg, The Netherlands: International Statistical Institute.

Biehler, R. (1993). Software tools and mathematics education: The case of statistics, in C. Keitel & K. Ruthven (Eds.), *Learning from Computers: Mathematics Education and Technology*, 68-100. Berlin: Springer.

Garfield, J. (1995). How students learn statistics, *International Statistical Review*, **63** (1), 25-34.

Konold, C., Pollatsek, A., & Well, A. (1997). Students analyzing data: research of critical barriers, in J. Garfield & G. Burrill (Eds.), *Research on the Role of Technology in Teaching and Learning Statistics*, 151-168. Voorburg, The Netherlands: International Statistical Institute.

Lampert. M. (1990). When the problem is not the question and the solution is not the answer: mathematical knowing and teaching, *American Educational Research Journal*, **27**, 29-63.

Meira, L. R. (1991). *Explorations of Mathematical Sense-Making: An Activity-Oriented View of Children's Use and Design of Material Displays*, an unpublished Ph.D. dissertation, Berkeley, CA: University of California.

Moschkovich, J. D. (1989). *Constructing a Problem Space through Appropriation: A Case Study of Guided Computer Exploration of Linear Functions*, an unpublished manuscript, available from the author.

Resnick, L. (1988). Treating mathematics as an ill-structured discipline, in R. Charles & E. Silver (Eds.), *The Teaching and Assessing of Mathematical Problem Solving*, 32-60. Reston, VA: National Council of Teachers of Mathematics.

Schoenfeld, A. H. (1994). Some notes on the enterprise (research in collegiate mathematics education, that is), *Conference Board of the Mathematical Sciences Issues in Mathematics Education*, **4**, 1-19.

Shaughnessy, J. M., Garfield, J., & Greer, B. (1996). Data Handling, in A. J. Bishop, K. Clements, C. Keitel, J. Kilpatrick & C. Laborde (Eds.), *International Handbook of Mathematics Education*, **1**, 205-237. Dordrecht, Netherlands: Kluwer.

Wild, C. J., & Pfannkuch, M (1999). Statistical thinking in empirical enquiry, *International Statistical Review*, **67** (3), 223-265.