TEACHING STATISTICS IN ELEMENTARY AND HIGH SCHOOL AND TEACHER TRAINING

Cileda de Queiroz e Silva Coutinho
Pontifícia Universidade Católica de São Paulo, Brazil
cileda@pucsp.br

The focus of the reflections presented here is the teacher's point of view on the teaching and learning of statistical concepts. This paper reports the research efforts carried out recently by our group with converging results. We identified that teachers have difficulty in teaching topics related to statistics, particularly when some analysis of the data is required. Teacher discourse shows they favor the philosophy of Exploratory Data Analysis (EDA), but in practice, they restrict their work, according to the results of our study, to a more technical approach that emphasizes the use of algorithms. This research suggests the need for initial and continued training in statistics for mathematics teachers.

INTRODUCTION–AN OVERVIEW OF THE RESEARCH

Teaching and learning statistical concepts are the focus of several research studies conducted by the community of mathematics and statistics educators. Some of the most important results indicate a strong relationship between the theory and practice of teachers and, therefore, a strong relationship between those practices and the ones established by the pupils involved in the learning process. In this paper we will discuss some of the main results obtained by the research carried out in our research group. The objective was to determine what teachers thought about teaching statistics and their knowledge of the subject and to compare those results to identify patterns that would explain the behavior of teachers concerning this content and consequences for the classroom. This research is related to the issues of teachers' attitudes, knowledge, conceptions and beliefs in relation to statistics education, more specifically to: Which research instruments and strategies are useful for determining the knowledge of statistics possessed by statistics teachers? The main theoretical framework for the research we have analyzed here was the development of statistical thinking and reasoning, facilitating progress in statistical literacy for both teachers and learners. In our team, we have been working on the levels of literacy proposed by Shamos (1995) and assumed by Gal (2002). The progression of these levels range from cultural to functional and then, from functional to the scientific.

The first level, 'cultural', refers to a grasp of basic terminology commonly used in the media to convey information about science. For example, we use this level when we want to know how the teachers in our study use data published in the media to work on the comprehension of graphs. The second level, 'functional', requires that people are able to converse, read and write coherently, in addition to the abilities required for the cultural level. They must use scientific terms to communicate, perhaps in a non-technical but meaningful context. Looking at the teachers in their classrooms, we work at this level as we study their discourse when they are planning to improve the levels of statistical reasoning by teaching students how to analyze graphs. The most advanced level, scientific, requires "some understanding of the overall scientific enterprise, coupled with an understanding of scientific and investigative processes" (Gal, 2002, p.2).

The research carried out by this group uses the principles of Exploratory Data Analysis, and, therefore, the main focus was to first diagnose whether the in-service elementary and high school teachers would favor these principles or not.

Several theories that try to model the learning process for elementary and high school education affirm that the student should be the protagonist, responsible for the construction of his/her knowledge, and that the teacher should leave behind his/her role as a mere content transmitter to become a tutor. It is intended to present the students with problems in contexts that are familiar to them. In this sense, our team carries out research on curriculum development and textbooks for elementary and secondary schools.

© 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.*

CURRICULUM FOR ELEMENTARY AND HIGH SCHOOL EDUCATION

Since 1998, Brazil has published the National Curriculum Parameters for basic education (elementary through high school). This document was elaborated with the purpose of establishing a common national educational reference for all Brazilian regions while respecting the regional, cultural and political diversities that exist in our country. The volume written for mathematics includes, for the first time, guidance for statistics, probability, and combinatorial analysis. This content should, therefore, be taught starting from elementary school. Among the proposed objectives, it is important to emphasize those that suggest students should question reality, formulate problems and use logical thinking, creativity, intuition, and capacity for critical analysis, in addition to selecting procedures and verifying whether they are appropriate or not. We find here the basics for the development of statistical reasoning and the first steps toward statistical literacy. Particularly for 5th and 6th graders (11 and 12 year-olds), we find the following guidelines for work with statistical content:

> As far as statistics is concerned, the purpose is to allow the students to build skills in collecting, organizing and conveying data using tables, charts and representations that appear frequently in their daily life. Besides that, they should also calculate a few statistical measures such as mean, median and mode with the purpose of offering new elements for interpreting statistical data. (Brazil, 1997, p. 52)

For high school students, the National Curriculum Parameters, published in 2000, also include content on statistics, combinatorial analysis and probability:

> The capacity to describe and analyze a large set of data, infer and predict based on a sample of the population, apply the ideas of probability and combination to natural and daily phenomena are applications of mathematics in issues of the real world that have had a large growth or that have become very complex. This shows how important a careful approach of the contents of counting, statistics and probability will be in high school, widening the interface between learning mathematics and the other sciences and subjects. (Brazil, 2000, p.44-45)

In 2006, the Brazilian Ministry of Education published the Curriculum Guidelines for High School to complement the National Curriculum Parameters. According to this document, the content choice should allow students to the build their knowledge of mathematics through an investigative process. The document affirms that the study of statistics enables learning how to formulate questions that may be answered by means of data collection and their subsequent organization and representation. For that, the students are expected to improve the abilities they have acquired in the previous school grades using all the available technologies. It is worth pointing out in the document for the high school syllabus the need for working with measures of central tendency of a distribution and measures of dispersion, which should have been addressed in a more intuitive way in elementary school.

The National Curriculum Parameters determine what teachers should cover in the classroom, but it is the textbook that effectively determines the syllabus. For Lajolo (1996),

> (..) a textbook is the book that will be used in classes and courses, which has probably been written, edited, sold, bought, and designed for systematic school use. Its importance increases more in countries like Brazil, where an extremely precarious educational scenario leads to a situation where textbooks end up determining contents, conditioning teaching strategies, marking, therefore, in a decisive way, what is taught and how it is taught. (p. 3)

A recent study on teaching approaches for statistics in elementary and high school textbooks (Friolani, 2007) concluded that the three book series chosen by the author allow the students to develop only the cultural level of statistical literacy (the lowest), while the desirable would be the functional (intermediate) or even the scientific level. Friolani analyzed textbooks for elementary school using the Praxeological Organization proposed by Chevallard (1996) and the levels of statistical literacy proposed by Shamos, already quoted in this paper as a theoretical framework. For Friolani, the tasks contained in the analyzed educational book series are not effective, since they ask students to simply interpret data already registered in tables and graphs, exploring in particular direct reading without researching, collecting, organizing or presenting data, which are all necessary for preparing students for decision-making. They do not fulfill the requirements of the National Curriculum Parameters and of the results of the research in the area of mathematics education. The results obtained by Friolani confirm the results obtained previously by Morais (2006), who also analyzed three book series different from Friolani's choice in her Master's degree dissertation.

TEACHER DISCOURSE ON TEACHING STATISTICS

In this section, we have recorded the main results obtained by the participants in our research group, named PEA-ESTAT (Statistics Teaching and Learning Processes), during teacher training. We have concentrated on issues that concern statistics but also deal with the probabilistic and combinatory conceptions held by teachers (Gonçalves, 2004; Costa, 2003, respectively).

First, we will quote the work of Morais (2006) who, besides analyzing textbooks, gave a questionnaire to elementary and high school teachers in order to identify their conceptions about statistics. The analysis was carried out with the use of the software called CHIC (Hierarchical, Implicative and Cohesive Classification). The data collected suggested that around 90% of the teachers who answered the questionnaire worked on the cultural level of statistical literacy, since they did not approach activities with continuous data, did not explore the analysis of the results, did not study the variation existing in the data, and therefore, they did not give students the opportunity to make decisions based on those analyses.

Those results were confirmed by Bigattão (2007), whose purpose was to investigate 5th-8th grade mathematics teachers' conceptions about the components and the formation of statistical and probabilistic thinking. The results indicate that the vast majority of the teachers questioned did not fulfill the requirements of the cultural level of statistical literacy with the same difficulties identified by Morais (2006). Among the 23 teachers who answered the questionnaire in Bigattão's research, seven wrote that they worked with statistical contents more frequently in the 8th grade. The reasons given by some teachers for not teaching this content were: the textbooks do not include this content (eight teachers); this content is too complex for the students at this level of schooling (seven teachers); they themselves have not mastered this content. Working on Bigattão's results, we can suppose that besides their discourse that shows favorable attitudes toward the teaching of statistical content, in the daily routine of the classroom teachers do not teach this content.

In the same project as Bigattão, Cardoso (2007) applied an activity to 29 high school teachers organized into three stages: a data set, two distributions represented by tables, and two represented graphically. The teachers were asked to analyze the data through the association between mean and standard deviation and between median and quartiles. The problem asked the teachers to calculate the mean and standard deviation and to calculate the median and quartiles. After calculating these summaries, the teachers were asked to explain the meaning of the results obtained. At the end of the first part of the diagnostic instrument (the task involving mean and standard deviation), a discussion was established with the teachers on these values and their meanings. This discussion showed the extreme difficulty the teachers had in answering and giving a clear critical analysis. In other words, the teachers interviewed could not explicitly make any kind of oral or written analysis to justify the results found.

In a data analysis using the software CHIC, Cardoso shows, among other results, that the teachers who provide a good analysis of the values of mean and standard deviation, also give a good analysis of the values of median and quartiles. Moreover, they affirm, in the last part of the questionnaire, that the median is a very difficult topic to teach 8th grade students, which contradicts some of their previous discourse about the teaching of descriptive statistics.

Another important issue that was raised in Cardoso's research was the fact that, at some point during the resolution of the proposed activity, the teachers incorrectly analyzed concepts like the mean and median. This fact was diagnosed through the observation that some teachers attributed symmetry to all data distributions; for them, all distributions were symmetric (leading to the confusion between median and mean). This can be explained by the almost exclusive use of symmetric distributions in textbook examples both for elementary and high school as well as in college textbooks.

Another example of erroneous interpretation of the attribution of symmetry to distribution is found in the research by Silva (2007), who conducted training for in-service teacher on descriptive statistics during the first semester of 2005. The purpose in this research was to study the reasoning about variability of elementary and high school teachers. In the activity the teachers were asked to interpret and analyze a data set built from research that they had organized. They analyzed the average for the variable ‘age’ as a value that represented the majority, excluding any need to complement that analysis through the study of standard deviation. When discussing whether the analysis was correct or not, other strategies appeared to complement the information supplied by the mean, such as finding the maximum, minimum and mode, but the need for the standard deviation was not perceived.

The task of interpreting the standard deviation allowed for the identification of the difficulty in understanding that it is a measure of variation around the mean. The prevailing idea referred to a measure of the differences between the observations. This can be explained by the fact that the educational books analyzed by Silva (books recommended by the teachers who took part in the project and that were used in their classrooms) featured standard deviation as a measure of sample homogeneity or data regularity.

Coutinho and Miguel (2007) carried out another study in this project and gave a questionnaire to 33 high school teachers. The instrument contained statements with which the teachers were asked to agree, partially agree or disagree. The statements included points that the teachers should consider in lesson planning when working with descriptive statistical concepts for any educational level (See Table 1).

Table 1. Distribution of answers by the 33 teachers consulted, according to their agreement with the item. (D: disagreement; A: agreement; PA partial agreement)

In planning the development of lessons with the aim of working with descriptive statistics (at any school level), the teacher must:	D	PA	A	Total
a) Generate learning situations that are contextualized in themes in which the students have an interest;	1	8	24	33
b) Use graphical representations that, among other aspects, may facilitate the perception of the variability in a group of observed data;	8	6	19	33
c) Apply, preferably, position measures to which students attribute meaning in an easier way;	5	14	13	33
d) Refrain from using (because it is not necessary) formulas that require complex mathematical manipulation;	7	22	4	33
e) Consider the need for the students to collect data and how the data can be produced;	1	6	26	33
f) Consider the need for familiarity with basic terms and ideas of descriptive statistics such as mean, mode, median, quartiles, standard deviation and average deviation;	2	7	23	33
g) Consider the need for familiarity with basic terms and ideas of graphical and tabular presentation such as relative and absolute frequency, accumulated frequency, bars, columns, sector, histogram, ogive, etc.;	1	12	19	33
h) Consider the comprehension of basic notions of probability;	1	11	20	33
i) Consider the descriptive study for drawing inferences or statistical conclusions.	3	14	14	33

In their arguments, the teachers showed they favored working with statistics in their classes within the Exploratory Data Analysis principles and allowing students to fulfill the role of the real learner instead of reproducing techniques previously presented. However, the reading of the questionnaires showed that the teachers' knowledge of descriptive statistics was limited to the collection of data and the respective construction of tables (with absolute, relative and accumulated frequency) and graphs (bar, column and sector). Those who referred to measures of central tendency and dispersion were unanimous in affirming that they taught how to apply formulas and that the difficulty faced by students is in the calculations. Some even mentioned that the work could be made easier with calculators and computers, but that the concepts would become meaningless (term used by teachers). We observed a trend towards a technicist approach, which is compatible with the results identified by the use of analysis of the similarities of the variables that we used in this data set (carried out with CHIC).

One part of the questionnaire asked teachers to solve a problem, and it was noted that none of them correctly solved the tasks proposed. One teacher calculated the mean, variance and standard deviation, in such a way that the last two statistics showed absurd values, incompatible with the data. The absence of a mathematical solution may be interpreted as just an omission, but the comments recorded indicate that the reason was the lack of knowledge of the content. Some teachers explicitly declared that they did not teach the median and quartiles because they are too difficult for the student to understand the formulas and others because they themselves did not know. On the other hand, we can see that in some other countries, like France, that this subject is taught to 12 and 13-year-old students.

Among the 33 teachers, only five disagreed totally with the use of positional measures (such as median and quartiles), while the others are almost equally divided between total and partial agreement on that issue. The question raised is how the teachers could agree with teaching something they do not master.

A FEW CONSIDERATIONS

The results of studies carried out by our team, focusing on elementary and high school teachers, agree concerning the identification of difficulties or even the lack of statistical knowledge in the teachers and allow us to focus on the need for initial and continuing training in statistics. Although in their declarations, the teachers seem to favor the philosophy of Exploratory Data Analysis, in practice, they restrict themselves to a more technicist approach, centered on the use of algorithms. Research efforts carried out with elementary and high school teachers show that the training should cover a range of themes starting with the most basic descriptive statistic concepts, the representations of a distribution of frequencies and measures that characterize some data. Since we have chosen the point of view of the philosophy of the Exploratory Data Analysis, the most important point should not be the calculation of the construction of a representation but the analysis of the data through the use of the several available tools: tables, charts and measures. One should strive to bring teachers up to a scientific level of statistical literacy so that they may have autonomy to create efficient learning situations for the literacy of their students on all educational levels.

REFERENCES

Bigattão, P. A. Jr. (2007). *Concepção do professor de matemática sobre o ensino da estatística.* (Mathematics' teachers conceptions about the teaching of statistics). Unpublished Master's thesis . PUC-SP. Online: www.pucsp.br/pos/edmat.

Brasil, Ministério da Educação. Secretaria de Educação Fundamental. (1998). *Parâmetros curriculares nacionais: Matemática* (National curricular parameters: Mathematics). Brasília: Ministério da Educação/Secretaria de Educação Fundamental.

Brasil, Ministério da Educação. Secretaria de Educação Média e Tecnológica. (2000). *Parâmetros curriculares nacionais (ensino médio)–Parte III: Ciências da natureza, matemática e suas tecnologias.* (National curricular parameters (high school)–Part III: Sciences of the nature, mathematics and its technologies). Brasília: Ministério da Educação.

Brasil, Ministério da Educação. Secretaria de Educação Média e Tecnológica. (2006). *Orientações curriculares nacionais para o ensino médio* (Curricular national orientation for high school). Brasília: Ministério da Educação.

Cardoso, R. (2007). *O professor de matemática e a análise exploratória de dados no ensino médio.*(The mathematics' teachers and the exploratory data analysis in high school). Unpublished Master's thesis. PUC-SP. Online: www.pucsp.br/pos/edmat.

Chevallard, Y. (1996). La fonction professorale: Esquisse d'un modèle didactique. (The teacher's function: Outline of a didactic model.) *Actes de l 'Ecole et Universite d'Été de Didactique des Mathematiques* (pp. 83–122). Clermont- Ferrand: IREM .

Costa, C. A. (2003). As concepções dos professores de matemática sobre o uso da modelagem no desenvolvimento do raciocínio combinatório no ensino fundamental. (The mathematics teachers' conceptions about model building in the development of the combinatory reasoning in basic school). Unpublished Master's thesis. PUC-SP. Online: www.pucsp.br/pos/edmat.

Coutinho, C. Q. S., & Miguel, M. I. R. (2007). *Análise exploratória de dados: Um estudo diagnóstico sobre concepções de professores* (Exploratory data analysis: A diagnostic study on teachers' conceptions). Online: www.anped.org.br.

Friolani, L. C. (2007). O pensamento estocástico nos livros didáticos do ensino fundamental. (Stocastic thinking in textbooks for basic school.) Unpublished Master's thesis. PUC-SP. Online: www.pucsp.br/pos/edmat.

Gal, I. (2002). Adults' statistical literacy: Meanings, components, responsibilities. *International Statistical Review*, *70*(1),1-50.

Gonçalves, M. C. (2004). Concepções de professores e o ensino de probabilidade na escola básica (Teachers' conceptions and the teaching of probability in basic school). Unpublished Master's thesis. PUC-SP. Online: www.pucsp.br/pos/edmat.

Lajolo, M. (1996). Livro didático: Um (quase) manual de usuário. (Textbooks: An (almost) user's manual). *Aberto*, *69*(26), 3-7.

Morais, T. M. R, (2006). *Um estudo sobre o pensamento estatístico: Componentes e habilidades.* (A study about statistical thinking: Components and abilities). Unpublished Master's thesis. PUC-SP. Online: www.pucsp.br/pos/edmat.

Silva, C. B. (2007). Pensamento estatístico e raciocínio sobre variação: Um estudo com professores de matemática (Statistical thinking and reasoning about variation: A study with mathematics' teachers). Unpublished Master's thesis. PUC-SP. Online: www.pucsp.br/pos/edmat.

Shamos, M.H. (1995). *The myth of scientific literacy.* New Brunswick, NJ: Rutgers University Press.