STATISTICAL GRAPHS PRODUCED BY PROSPECTIVE TEACHERS IN COMPARING TWO DISTRIBUTIONS

Carmen Batanero*, Pedro Arteaga*, Blanca Ruiz**
*Universidad de Granada
**Instituto Tecnológico y de Estudios Superiores, Monterrey, México

We analyse the graphs produced by 93 prospective primary school teachers in an open statistical project where they had to compare two statistical variables. We classify the graphs according its semiotic complexity and analyse the teachers' errors in selecting and building the graphs as well as their capacity for interpreting the graphs and getting a conclusion on the research question. Although about two thirds of participant produced a graph with enough semiotic complexity to get an adequate conclusion, half the graphs were either inadequate to the problem or incorrect. Only one third of participants were able to get a conclusion in relation to the research question.

Keywords: Statistical graphs, semiotic complexity, prospective teachers, assessment, competence.

INTRODUCTION

Graphical language is essential in organising and analysing data, since it is a tool for *transnumeration*, a basic component in statistical reasoning (Wild & Pfannkuch, 1999). Building and interpreting statistical graphs is also an important part of statistical literacy which is the union of two related competences: interpreting and critically evaluating statistically based information from a wide range of sources and formulating and communicating a reasoned opinion on such information. (Gal, 2002). Because recent curricular guidelines in Spain introduce statistics graph since the first year of primary school level and therefore, this research was oriented to assess prospective primary school teachers' graphical competence in order to use this information in improving the training of these teachers.

Understanding statistical graphs

In spite of its relevance, didactic research warn us that competence related to statistical graphs is not reached in compulsory education, since students make errors in scales (Li & Shen, 1992) or in building specific graphs (Pereira Mendoza & Mellor, 1990; Lee & Meletiou, 2003; Bakker, Biehler & Konold, 2004). Other authors define levels in graph understanding (Curcio, 1989; Gerber, Boulton-Lewis & Bruce, 1995; Friel, Curcio & Bright, 2001) that vary from a complete misunderstanding of the graph, going through reading isolated elements or being able to compare elements to the ability to predict or expand to data that are not included in the graph. More recently, these levels were expanded to take into account the critical evaluation of information, once the student completely reads the graph (Aoyama, 2007):

1. *Rational/literal level.* Students correctly read the graph, interpolate, detect the tendencies and predict. They use the graph features to answer the question posed but they do neither criticise the information nor provide alternative explanations.
2. *Critical level*: Students read the graph, understand the context and evaluate the information reliability; but they are unable to think in alternative hypotheses that explain the disparity between a graph and a conclusion.
3. *Hypothetical level*: Students read the graphs, interpret and evaluate the information, and are able to create their own hypotheses and models.

Graphical Competence in Prospective Teachers

Recent research by Espinel, Bruno & Plasencia (2008) also highlight the scarce graphical competence in future primary school teachers, who make errors when building histograms or frequency polygons, or lack coherence between their building of a graph and their evaluation of tasks carried out by fictitious future students. When comparing the statistical literacy and reasoning of Spanish prospective teachers and American university students even when the tasks were hard for both groups, results were much poorer in the Spanish teachers, in particular when predicting the shape of a graph or reading histograms. Monteiro and Ainley (2007) studied the competence of Brazilian prospective teachers and found many of these teachers did not possess enough mathematical knowledge to read graphs taken from daily press. A possible explanation of all these difficulties is that the simplicity of graphical language is only apparent, since any graph is in fact a mathematical model. In producing a graph we summarize the data, going from the individual observations to the values of a statistical variable and the frequencies of these values. That is, we introduce the frequency distribution, a complex object that refers to the aggregate (population or sample) instead of referring to each particular individual and this object can be not grasped by the students.

THE STUDY

As stated in the introduction, the main goal in our research was to assess the graphical competence of prospective primary school teachers. A secondary aim was to classify the graphs produced by these teachers as regards its complexity. More specifically we analyse the graphs produced by 93 prospective teachers when working in an open statistical project with the aim of providing information useful to teacher educators. These students had studied descriptive statistics (graphs, tables, averages, spread) the previous academic year (their first year of University) as well as in secondary school level. The data were collected along a classroom practice (Godino, Batanero, Roa & Wilhelmi, 2008) that was carried out in a Mathematics Education course (second year of University) directed to prospective teachers in the Faculty of Education, University of Granada. In this practice (2 hours long) we proposed prospective teachers a data analysis project. At the end of the session, participants were given a sheet with the data obtained in the classroom and were asked to individually produce a data analysis written report to answer the question set in the project. Participants were free to use any statistical graph or summary and work with computers if they wished. They were given a week to complete the reports that were collected and analysed.

The statistical project: "Check your intuitions about chance"

This project is part of a didactical unit designed to introduce the "information handling, chance and probability" content included in the upper level of primary education. Some aims are: a) showing the usefulness of statistics to check conjectures and analyse experimental data; b) checking intuitions about randomness and realising these intuitions are sometimes misleading. The sequence of activities in the project was as follows.

1. *Presenting the problem, initial instructions and collective discussion.* We started a discussion about intuitions and proposed that the future teachers carry out an experiment to decide whether they have good intuitions or not. The experiment consists of trying to write down apparent random results of flipping a coin 20 times (without really throwing the coin, just inventing the results) in such a way that other people would think the coin was flipped at random.

2. *Individual experiments and collecting data.* The future teachers tried the experiment themselves and invented an apparently random sequence (simulated throwing). They recorded their sequences using H for head and T for tail. Afterwards the future teachers were asked to flip a fair coin 20 times and write the results on the same recording sheet (real throwing).

3. *Classroom discussion, new questions and activities.* After the experiments were performed we started a discussion of possible strategies to compare the simulated and real random sequences. A first suggestion was to compare the number of heads and tails in the two sequences since we expect the average number of heads in a random sequence of 20 tosses to be about 10. The lecturer posed questions like: If the sequence is random, should we get exactly 10 heads and 10 tails? What if we get 11 heads and 9 tails? Do you think in this case the

sequence is not random? These questions introduced the idea of comparing the number of tails and heads in the real and simulated experiments for the whole class and then studying the similarities and differences.

4. At the end of the session the future teachers were given a copy of the data set for the whole group of students. This data set contained two statistical variables: number of heads for each of real and simulated sequences and for each student; n cases with these 2 variables each. As prospective teachers were divided in 3 groups, n varied (30-40 cases in each group). They were asked to complete the analysis at home and produce a report with a conclusion about the group intuitions concerning randomness. Students were able to use any statistical method or graph and should include the statistical analysis in the report.

RESULTS AND DISCUSSION

Once the students' written reports were collected, we made a qualitative analysis of these reports. By means of an inductive procedure we classified into different categories the graphs produced as a part of the analysis, the interpretations of graphs and the conclusions about the group intuitions. The classification of graphs took into account the type of graph, number of variables represented in the graph, and underlying mathematical objects as well as some theoretical ideas that we summarise below.

Font, Godino and D'Amore (2007) generalize the notion of *representation*, by taking from Eco the idea of semiotic function "*there is a semiotic function when an expression and a content are put in correspondence*" (Eco, 1979, p.83) and by taking into account an ontology of objects that intervene in mathematical practices: *problems, actions, concepts-definition, language properties and arguments*, any of which could be used as either expression or content in a semiotic function. In our project we propose a *problem* (comparing two distributions to decide about the intuitions in the set of students) and analyse the students' practices when solving the problem. More specifically we study the graphs produced by the students; these graphs involve a series of *actions, concepts-definitions* and *properties* that vary in different graphs. Consequently the semiotic functions underlying the building and interpretation of graphs, including putting in relation the graphs with the initial question by an argument also vary. We therefore should not consider the different graphs as equivalent representations of a same mathematical concept (the data distribution) but as different configurations of interrelated objects that interact with that distribution. Five students only computed some statistical summaries (mean, median or range) and did not produce graphs; we are not taking into account these students in our report. Using the ideas above we performed a semiotic analysis of the different graphs produced by the other 88 students and defined different levels of semiotic complexity as follow:

L1. Representing only his/her individual results. Some students produced a graph to represent the data they obtained in his/her particular experiment, without considering their classmates' data. These graphs (e.g. a bar chart) represent the frequencies of heads and tails in the 20 throwing. Students in this level tried to answer the project question for only his /her own case (tried to assess whether his/her intuition was good); part of these students manifested a wrong conception of chance, in assuming a good intuition would imply that the simulated sequence would be identical to the real sequence in some characteristic, for example the number of heads. Since they represented the frequency of results in the individual experiment, in fact these students showed an intuitive idea of statistical variable and distribution; although they only considered the Bernoulli variable "result of throwing a coin" with two possible values: "1= head", 0= tail" and 20 repetitions of the experiment, instead of considering a Binomial distribution "number of heads in the 20 throwing" that have a wider range of values (1-20 with average equal to 10) and r repetitions of the experiments (r= number of students in the classroom).

L2. Representing the individual values for the number of heads. These students did neither group the similar values of the number of heads in the real nor in the simulated sequences. Instead, they represented the value (or values) obtained by each student in the classroom in the order the data were collected, so they did neither compute the frequency of the different values nor explicitly used the idea of distribution. The order of data in the X-axis was artificial, since it only indicated the arbitrary order in which the students were located in the classroom. In this category we got horizontal and vertical bar graphs, line graphs of one or the two variables that, even when did not solve the problem of comparison, at least showed the data variability. Other students produced graphs such as pie chart, or stocked bar charts, that were clearly inappropriate, since they did not allow visualizing the data variability.

L3. Producing graphs separate for each distribution. The student produced a frequency table for each of the two variables and from it constructed a graph or else directly represented the graph with each of the different values of the variable with its frequency. This mean that the students went from the data set to the statistical variable “number of heads in each sequence” and its distribution and used the ideas of frequencies and distribution. The order in the X-axis was the natural order in the real line. In case the students did not use the same scale in both graphs or used different graphs for the two distributions the comparison was harder. Examples of correct graphs in this category were bar graphs and frequency polygons. Students also produced incorrect graphs in this category such as histograms with incorrect representation of intervals, bar graphs with axes exchanged (confusing the independent and dependent variable in the frequency distribution), representing the frequencies and variable values in an attached bar graph or representing variables that were not related.

L4. Producing a joint graph for the two distributions. The students formed the distributions for the two variables and represented them in a joint graph, which facilitated the comparison; the graph was more complex, since it represented two different variables. We found the following variety of correct graphs: attached bar chart; representing some common statistics (e.g. the mean or the mode) for the two variables in the same graph; line graphs or dot plots in the same framework. Example of incorrect graphs in this category were graphs presenting statistics that were not comparable (e.g. mean and variance in the same graphs) or the same statistics for variables that cannot be compared.

In Figure 1 we present an example of graphs produced in each category. Even when within each of these categories we observe a variety of graphs and configurations of mathematical objects it is evident a qualitative gap between each of the different levels. In Table 1 we present the distribution of students according the semiotic complexity of the graph, it correctness, the interpretation of the graph and the conclusion about intuitions.

Table 1. Results

	Correctness of the graph			Interpretation of graph			Conclusion on the intuitions			Total in the level
	1	2	3	1	2	3	1	2	3	
L1. Representing only the student data	1		1	1	1				2	2
L2. Representing individual results	10	1	4	4	10	1		3	12	15
L3. Separate graphs	15	17	14	15	15	16	1	12	33	46
L4. Joint graphs	14	6	5	9	11	5	1	7	17	25
Total	40	24	24	29	37	22	2	22	64	88

(1) Correct; (2) Partially correct; (3) Incorrect or no interpretation / conclusion

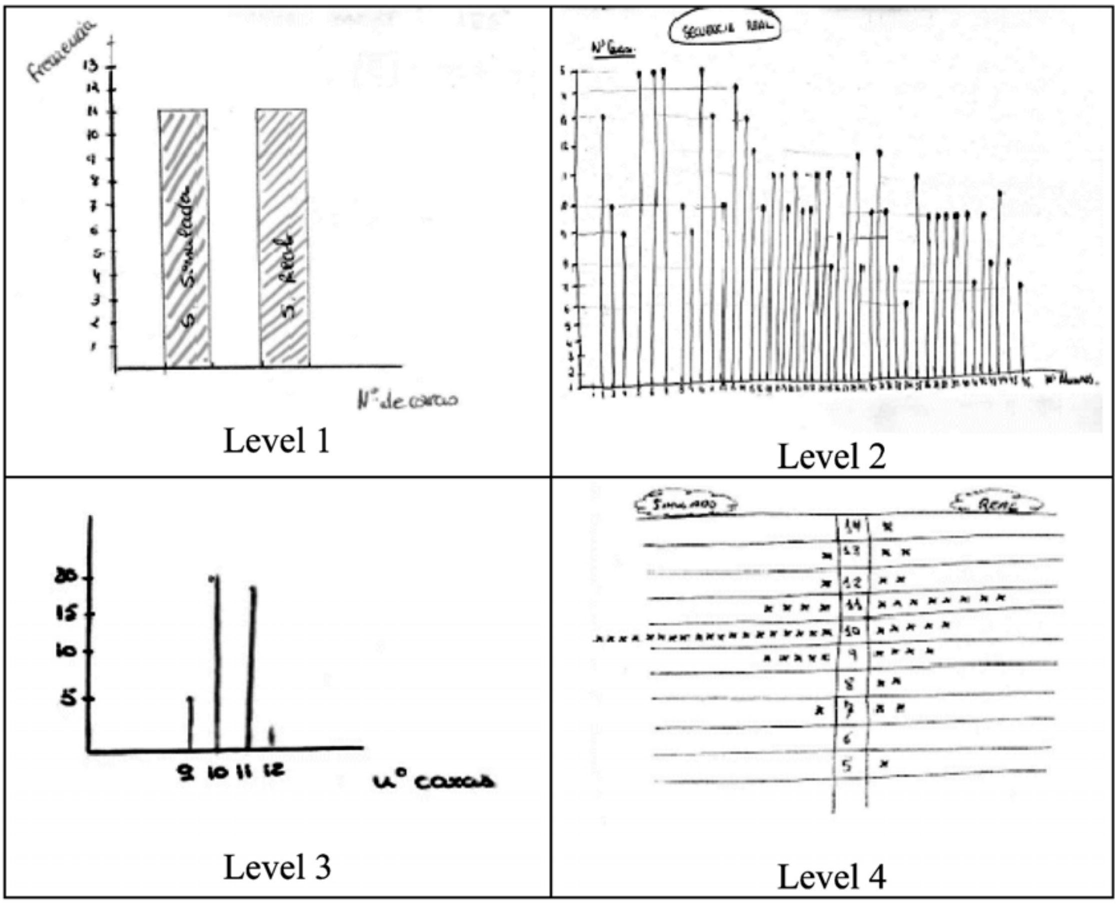

Figure 1. Examples of graphs in each different level of semiotic complexity

From a total of 93 students 88 (94,6%) produced some graphs when analysing the data, even if the instructions given to the student did not explicitly require that they constructed a graph. This fact suggests that students felt the need of building a graph and reached, by a transnumeration process some information that was not available in the raw data. Most students (52,2%) produced separate graphs for each variable (level 3), that were generally correct o partly correct (correct graph with different scales or different graph in each sample; not centring the rectangles in the histogram, or missing labels).

14 students in this level constructed a non-meaningful graph since they represented the product of values by frequencies, exchanged the frequencies and values of variables in the axes thus confusing the independent and dependent variable in the frequency distribution. 28,4% students worked at level 4, and produced only a joint graph for the two variables, although 6 of these graphs were partly correct and 5 incorrect (same reasons than those described in level 3). Few students only analysed their own data (level 1) and only 17% of participants studied the value got

by each student without forming the distribution. Consequently the concept of distribution seemed natural for the majority of students who used it to solve the task, although the instructions did not require this explicitly.

In general, these prospective teachers interpreted correctly or partially correctly the graphs in all the levels, reaching the Curcio's (1989) intermediate level (reading between the data) and the difficulty of interpretation of graphs increased with its semiotic complexity. However, an important part of students in our levels 3 and 4, even when they built correct graphs did not reached the "reading between the data" level, because either they did not interpret the graph either made only a partial interpretation. As regards the Aoyama's (2007) levels, the majority of prospective teachers only read the graphs produced at a rational/literal level, without being able of read the graphs at a critical or a hypothetical level. The teachers performed a mathematical comparison of the graphs but did not get a conclusion about the intuitions in the classroom (e.g. they correctly compared averages but did not comment what were the implications in relation to the students' intuitions). Only two students in the group reached the hypothetical level in reading the graphs, as they got the correct conclusion about group's intuition. These two students realised that the group have correct intuitions about the average number of heads but poor intuitions about the spread. Students were supposed to get this conclusion from comparing the averages and range in the variables in the simulated and real sequences distributions. At higher level statistical tests could also be used to support this conclusion that have been observed in previous research about people perception of randomness. 22 participants got a partial conclusion that the intuition as regards averages was good, as they were able to perceive difference or similitude in the averages, but they did not considered the results obtained in comparing spread of the variable (number of heads) in the two sequences. These students also work at the Aoyama's (2007) hypothetical level, although they did not considered spread in comparing the two distributions. Those working at levels 1 and 2 got few partly correct conclusions and none correct conclusion, so that these levels of complexity in the graph were not adequate to get a complete conclusion.

CONCLUSIONS

In the project posed the prospective teachers went through the different steps in the statistics method as described by Wild and Pfannkunch (1999) in their PPCAI cycle: setting a problem, refining the research questions, collecting and analysing data and obtaining some conclusions. They also practiced the process of modelling, since, beyond working with the statistics and random variables, they should interpret the results of working with the mathematical model in the problem context (whether the students' intuitions was good or not). This last step (relating the result with the research question) was the most difficult for the students, who lacked familiarity with statistical projects and modelling activities. Since these activities are today recommended in the teaching of statistics since primary school level in Spain and are particularly adequate to carry out group and individual work as recommended in the Higher European Education Space we suggest they are particularly suitable for the training of teachers. Our research also suggest that building and interpreting graphs is a complex activity and confirm some of the difficulties described by Espinel, Bruno and Plasencia (2008) in the future teachers, in spite that they should transmit graphical language to their students and use it as a tool in their professional life. Improving the teaching of statistics in schools should start from the education of teachers that should take into account statistical graphs.

Acknowledgment. Project SEJ2007-60110 and grant FPU AP2007-03222 .

REFERENCES

Aoyama, K. (2007). Investigating a hierarchy of students' interpretations of graphs. *International Electronic Journal of Mathematics Education 2* (3). On line: http://www.iejme/.

Bakker, A., Biehler, R., & Konold, C. (2004). Should young students learn about box plots? In G. Burrill & M. Camden (Eds.), *Curricular Development in Statistics Education. Proceedings of the International Association for Statistical Education (IASE) Roundtable* (pp. 163-173). Auckland: IASE. On Line: www.stat.auckland.ac.nz/~iase/publications/.

Eco, U. (1979). *Tratado de semiótica general* (General semiotics). Barcelona: Lumen.

Curcio, F. R. (1989). *Developing graph comprehension*. Reston, VA: N.C.T.M.

Espinel, C., Bruno, A., & Plasencia, I. (2008). Statistical graphs in the training of teachers. In C. Batanero, G. Burrill, C. Reading & A. Rossman (2008). *Proceedings of the Joint ICMI /IASE Study Teaching Statistics in School Mathematics. Challenges for Teaching and Teacher Education.* Monterrey, México: ICMI &IASE. CD ROM.

Friel, S., Curcio, F.,& Bright, G. (2001). Making sense of graphs: critical factors influencing comprehension and instructional implications. *Journal for Research in Mathematics Education 32*(2), 124-158.

Gal, I. (2002). Adult's statistical literacy: Meaning, components, responsibilities. *International Statistical Review 70*(1), 1-25.

Gerber, R., Boulton-Lewis, G & Bruce, C. (1995). Children's understanding of graphic representation of quantitative data. *Learning and Instruction 5*, 70-100.

Godino, J. D., Batanero, C., Roa, R.,& Wilhelmi, M. R. (2008). Assessing and developing pedagogical content and statistical knowledge of primary school teachers through project work. InC. Batanero, G. Burrill, C. Reading & A. Rossman (2008). *Proceedings of the Joint ICMI /IASE Study eaching Statistics in School Mathematics. Challenges for Teaching and Teacher Education.* Monterrey, México: ICMI e IASE. CD ROM.

Font, J. D., Godino, J. D.,& D'Amore, B. (2007). An ontosemiotic approach to representations in mathematics education. *For the Learning of Mathematics, 27* (2), 3-9.

Lee, C.,& Meletiou, M. (2003). Some difficulties of learning histograms in introductory statistics. *Joint Statistical Meetings- Section on Statistical Education.* On line: http://www.statlit.org/PDF/2003LeeASA.pdf

Li, D. Y.,& Shen, S. M. (1992). Students'weaknesses in statistical projects. *Teaching Statistics 14* (1), 2-8.

MEC (2006). *Real Decreto 1513/2006, de 7 de diciembre, por el que se establecen las enseñanzas mínimas correspondientes a la Educación Primaria* (Royal Decrete establishing the minimum teaching contents for Primary Education).

Monteiro, C., & Ainley, J. (2007). Investigating the interpretation of media graphs among student teachers. *International Electronic Journal of Mathematics Education* 2 (3), 188-207. On line: http://www.iejme/.

Pereira-Mendoza, L.,& Mellor, J. (1990). Student's concepts of bar graph: Some preliminary findings. In D. Vere-Jones (Ed.), *Proceedings of the Third International Conference on Teaching Statistics* Voorburg: International Statistical Institute.

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