

SEMIOTIC COMPLEXITY LEVELS AND ACTIVITIES RELATED TO STATISTICAL GRAPHS IN CHILEAN PRIMARY EDUCATION TEXTBOOKS

Danilo Díaz-Levicoy, Pedro Arteaga, Carmen Batanero, María Magdalena Gea

University of Granada, Spain

In this work we analyze the activities proposed and the semiotic complexity levels in the statistical graphs included in two series of Chilean primary education textbooks using content analysis. The most common activities related to these graphs in these textbooks are performing computations, building graphs and using the graphs as examples. The most common complexity level was level 3 (representing a data distribution). These results also show (with smaller frequency) a varied type of activity and higher semiotic levels with increasing age of students.

INTRODUCTION

The scientific and technological development of modern society requires an educational process that helps to promote the understanding of information, which is becoming more abundant and specialized every day.

Statistical graphs play an important role in the communication and dissemination of this information (Monroy, 2007). According to Arteaga, Batanero, Cañadas, and Contreras (2011) these graphs are widely used in different situations of daily life. Carvalho (2011) remarks that the presence of graphs, tables and statistical summaries in the media is aimed at summarizing information and facilitating its understanding; however, these representations are often used to get some benefit out of particular situations, and include biased information. Consequently it is important that the citizen knows how to read and interpret graphs, to ensure his/her adequate integration into our society (Arteaga, Batanero, & Contreras, 2011).

This need is reflected in the relevance given to statistical literacy, which Gal (2002) defined as the union of two competences:

- a) People's ability to interpret and critically evaluate statistical information, data-related arguments, or stochastic phenomena, which they may encounter in diverse contexts; ... and, when relevant;
- (b) their ability to discuss or communicate their reactions to such statistical information, such as their understanding of the meaning of the information, their opinions about the implications of this information (p. 2-3).

Del Pino and Estrella (2012) compare statistical literacy to a civil right, since people should be able to use elementary arguments, language and tools to act in an informed and critical way in the knowledgeable society.

These demands for statistical culture have been considered in various curricular documents, e.g., The GAISE project (Franklin, Kader, Mewborn, Moreno, Peck, Perry, & Scheaffer, 2007), the Common Core State Standards Initiative (CCSSI,

2010), the Spanish regulations of the Ministry of Education, Culture and Sports (MECD, 2014) and the Chilean regulations of the Ministry of Education (MINEDUC, 2012). These documents emphasize the relevance of statistical graphs, which represent data coming from familiar situations and contexts where children participate actively. Thus, as shown in the Chilean curricular guidelines (MINEDUC, 2012), statistical graphs are approached throughout primary education from the first level with increasing sophistication and with constant work (Díaz-Levicoy, 2014).

This background led us to start a research project intended to analyze the way in which the statistical graphs are introduced in the Chilean primary education textbooks, identify the main variables that may affect their difficulty and use this information to help teachers to organize the teaching of the topic. Below we present the study background and some preliminary results.

BACKGROUND

Statistical graphs

Statistical graphs have been analyzed from various points of view. Bertin (1967) remarked that a statistical graph is a complex semiotic system, since its understanding requires the local interpretation of each element, as well as a global interpretation of the graph. Moreover, Arteaga (2011), and Batanero, Arteaga, and Ruiz (2010) analyzed the semiotic activity involved in the construction of graphs, that varies in complexity, according to the mathematical objects needed in this process. Taking into account this activity, the authors described four semiotic levels in statistical graphs that will be used in our analysis (with examples reproduced in Figure 1):

- *Representing isolated data.* At this level, the graph only represents some isolated data (either individual data or a few data), without considering the whole set from which the data were extracted. In these graphs, the author does not need to consider the ideas of variable or distribution, since no comprehensive analysis of a whole data set is performed. In the example reproduced in Figure 1, only a dot is represented in a scatter plot.
- *Representing a data set without building a distribution.* At this level a list of data or a data set is represented in the same order in which the data were collected. There is no grouping of similar values of the variable or computation of frequencies. Consequently, although in this graph the idea of variable is used, the distribution is absent. Moreover, often the order of the data in the graph is not a numerical order, but is artificial. In Figure 1 the speed in each kilometre is represented with no order (of these speeds).
- *Representing a data distribution.* These graphs include the representation of a distribution, with values and frequencies for each value; the order of the variable values in the graph axes (if used) is the ordinary numerical order. The example in Figure 1 shows the frequencies of people going to the movies in the morning,

afternoon and night. So there is a distribution for a qualitative variable (time to go to the movies).

- *Representing several distributions on the same graph.* At the highest level more than one distribution is represented on the same graph. This often requires a decision on a common scale that makes the graph understandable. In the example in Figure 1, the distribution of the type of food taken by children in two different classes (two distributions) is compared on the same graph.

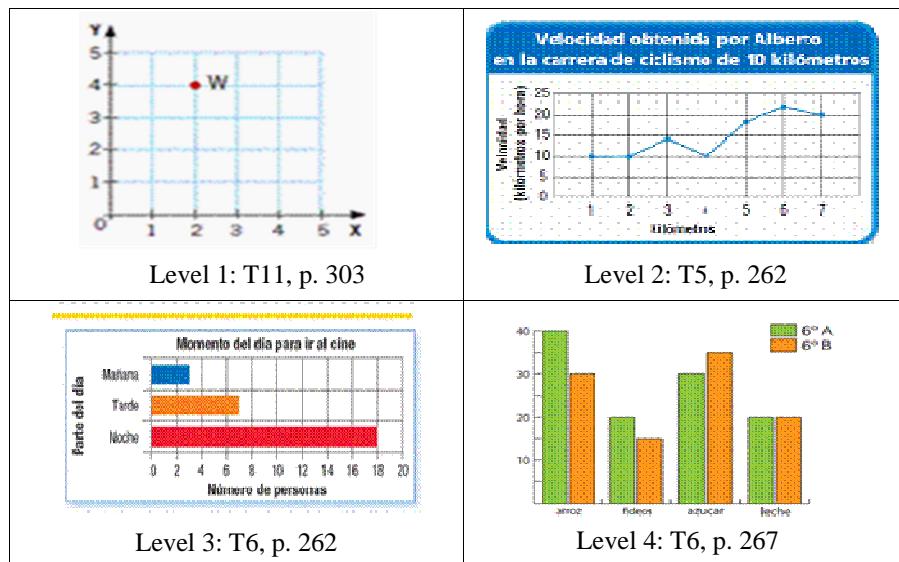


Figure 1. Different levels of semiotic complexity in the graphs included in the textbooks (T_i , $1 \leq i \leq 12$).

Textbooks

There is increasing research on textbooks in different areas of learning and mathematics is not an exception. Despite this research tradition, the analyses of statistics and probability textbooks are scarce, and the analysis of statistical graphs in the textbooks even more so.

A first study related to this topic is that by May (2009), who studied the statistical graphs in 25 university textbooks directed towards psychologists. The most common graphs in these books were the line graph, histogram, density curve and bar graph; most activities were directed towards the construction of graphs and the reading level in Curcio's classification (Curcio, 1989) was reading the data. According to these authors these activities promote statistical literacy but not statistical reasoning.

Gómez, Ortiz, Batanero, and Contreras (2013) studied the language of probability in two series of Spanish primary education textbooks. Their results suggest an extensive

use of tabular and graphical representations (bar graph, sectors, pictogram and histogram); the authors did not analyse the variables that characterize these graphs.

Gea, Batanero, Arteaga, Cañadas, and Contreras (2014) analysed the presentation of correlation and regression in 8 Spanish high school textbooks. They concluded that the main graph used is the scatter plots around which many exercises are set and in which the students should analyze the sign and strength of linear relationships. The authors also found some bar charts, three-dimensional histograms and bubble graphs in these textbooks.

Mateus (2014) analyzed 5 Colombian textbooks (primary and secondary school levels). These textbooks showed a predominance of contexts related to the students' life; the activities included some basic construction, and elements of reading graphs; the authors also identified some errors in the graphs.

Díaz-Levicoy (2014) studied the statistical graphs on a sample of 18 Spanish Primary Education textbooks. He found that the most common representation was the bar graph; usually the reading level in Curcio (1989) classification was "reading into the data"; the semiotic level in the graph was "representation of a data distribution" and the most common activity was building a graph or performing computations from data represented in the graph. Our research is intended to complement these studies; in particular to compare the analysis of Chilean textbooks with our previous study of Spanish textbooks for the same educational level (Díaz-Levicoy, 2014).

METHOD

In this paper we follow a qualitative methodology, based on content analysis (López, 2002). The sample consists of 12 primary education textbooks that were selected by intentional sampling, based on a controlled selection and according to certain characteristics that have been previously defined. More precisely, we selected textbooks that were widely used in the schools in the academic year 2013-2014 and edited according to the curriculum guidelines for these courses. The list of the textbooks used in this research is included as an appendix.

We considered all the textbook sections that included any statistical graphs (exercises, examples, definitions, problems) or that involved any activity related to statistical graphs. Each paragraph where one or more statistical graphs intervene was an analysis unit. A total of 421 different paragraphs were identified and analysed; for each of them the categories in the relevant variables were coded; in this case we considered the variables "semiotic complexity levels" and "type of activity". All the data were recorded in SPSS and the data file was analysed to obtain some conclusions.

RESULTS

Semiotic complexity levels

In Table 1 we classify the activities analyzed according to the semiotic level of the graph involved in the same (taking the higher level, in case the activity involves two

distributions). The predominance of level 3 is clear (representing a distribution in the graph). Consequently in the majority of activities the children have to work with a data distribution (a 66% of the activities) that is, with graphs representing one or more distributions (if we take into account levels 3 and 4). There are also 26.1% of activities in level 2, where the data set is represented with no computation of frequencies or building of the distribution.

School level (grade)							
Level	1 (n=43)	2 (n=65)	3 (n=95)	4 (n=38)	5 (n=87)	6 (n=93)	Total (n=421)
1					2.3	1.1	0.7
2	44.2	33.8	30.5	15.8	26.4	11.8	26.1
3	55.8	66.2	69.5	84.2	58.6	66.7	66
4					12.6	20.4	7.1

Table 1: Percentage of activities in the texts analyzed according to their semiotic complexity level.

When we analyse this variable by school year we observe that in the first four grades we only found levels 2 and 3 statistical graphs, with constant predominance of level 3, which gradually increase in these four levels. In our opinion it would be reasonable to include more level 2 graphs and even some level 1 graphs for such young children. In grade 5 and 6, we found level 4 graphs, as well as sporadic level 1 activities (just 3 activities, which correspond to scatter plots). These results agree with those of Díaz-Levicoy (2014) in his analysis of Spanish textbooks; although in the Spanish textbooks he found statistical graphs of level 4 from the second level of primary education.

Type of activities

A second variable analysed was the type of activity that children are requested in relation with the statistical graphs. These activities were classified in the following categories that were defined taking into account previous researches:

- *Reading the graph.* In this activity a graph is given and children are asked to read the information displayed in the graph. The activity may involve different reading levels according to Curcio (1989): reading the data; reading between the data; reading beyond the data or reading behind the data. We did not differentiate this variable in this paper.
- *Calculation.* In this case children are given a graph and are asked to perform some comparisons (for example, finding the mode) and / or simple calculations (for example, transforming a frequency to percentage), using the information in a statistical graph; this activity also involves reading the graph and some additional computations.

- *Building a graph.* In this activity children are asked to build a statistical graph using data presented in statistical tables or lists.
- *Complete a graph.* Similar to the previous activity; but in this case the children finish the construction of a statistical graph according to a pattern given (the first steps are presented to the children).
- *Translating.* In this activity the children should build a new a graph with the information provided in another, or translate the graph to a table.
- *Example.* This is the section of the textbook that uses a graph to clarify ideas and / or concepts.
- *Invent a problem.* In this type of activity the children should create a problem that makes sense for data presented in a statistical graph, i.e., they should generate a context where the data are relevant and coherent.
- *Comparing and justifying.* Students have to select a graph according to the nature of the data, pointing out the advantages and disadvantages of the specific graph selected, and indicate the best form of presenting the information (in a table or graph).

More than one type of activity can be asked from a statistical graph (e.g., some activity ask to build a graph and later calculate from the graph); in these cases to perform this analysis we considered these activities separately.

In Table 2 we show the distribution of the types of activities that we found after the analysis of these textbooks. We observe that the most frequent activities are computing (54.2%), building (19.2%), examples (15.2%), comparing and justifying (9.7%) and reading (7.6%).

Reading, building, examples and comparing and justifying are included in all the school levels. The activity “completing graphs” is presented in the first levels to guide the students when they are learning to build statistical graphs. The percentage of computation activities is high and increases with school level. We remark that this type of activity does not help students to develop their statistical reasoning or literacy; it is only a mathematical activity where the statistical graphs are only a tool to present data. Our suggestion is that this type of activity should be less frequent and changed by more interpretative activities.

Our results in this variable agree with those of Díaz-Levicoy (2014), where the activities of reading, building and example predominate. An important difference is that computation activities were only 8.8% in the Spanish textbooks while reading activities amounted to 40% and therefore there were more interpretative activities in these textbooks. Another difference is that the examples are presented for the first grade in Chilean texts, while in Spanish it is done from the second grade; finally, in the Spanish books some activities asked to describe the variables included in the graph; while we do not identify these activities in the Chilean text books.

School level (grade)							
Type of activity	1 (n=43)	2 (n=65)	3 (n=95)	4 (n=38)	5 (n=87)	6 (n=93)	Total (n=421)
Reading	9.3	12.3	9.5	2.6	6.9	4.3	7.6
Calculation	39.5	47.7	42.1	63.2	59.8	68.8	54.2
Building	20.9	20	30.5	10.5	10.3	18.3	19.2
Completing	9.3	6.2	2.1				2.4
Translating	9.3		1.1			1.1	1.4
Example	14	18.5	15.8	21.1	13.8	11.8	15.2
Inv. Problems			2.1	2.6	1.1		1
Comparing and justifying	9.3	1.5	7.4	23.7	11.5	10.8	9.7

Table 2: Percentages of tasks in primary education textbooks according to the requested activity.

CONCLUSIONS

Textbooks are a teaching and learning resource with an important tradition within the classroom, because they provide support to teachers and students throughout the instructional process. It is therefore necessary to study how the textbooks present mathematical and statistical topics as a first step to suggest possible improvement in their content and to check that they follow the curricular guidelines.

Statistical graphs are a new topic in Chilean primary education and for this reason our analysis is particularly needed. A first conclusion is the large number of activities (421 as compared with 215 activities in the same number of Spanish textbooks for these curricular levels). Although this number (double proportion of graph activities in the Chilean books as compared with the Spanish books) provides the teacher with rich material, it is also important to guide the teacher with criteria to select those that will be used in the classroom, since it is clear that it is impossible for the children to complete such a large number of tasks.

The two common semiotic complexity levels in these books are "representation of a data distribution" and "representation of a dataset without producing the distribution (90% of all the activities). Although these two levels are appropriate for the upper levels of primary education, and even level 4 (representing two or more distributions in the same graph), we recommend to include some more simple level 1 situations for the grades 1 and 2 as they appear in the Spanish texts.

We also recommend diminishing the emphasis on computation and reinforcing other activities relevant for learning. For example, the activity of "completing a graph" should be given more emphasis at all primary school levels, because this activity

helps students to become familiar with new graphs as they are introduced in the curriculum. The "reading" activity should have a higher presence in the early levels, starting with simple "reading the data", and progressing through the levels of "reading between the data", "reading beyond the data" and "reading behind the data" in Curcio (1989) framework throughout these school years. In the same way, the activities of "inventing problems" and "translating graphs" should have greater presence as they are more cognitively challenging for students and reinforce their statistical reasoning.

We hope our results are useful for teachers and teacher educators, who have the responsibility to make statistical literacy a reality for everyone and to develop statistical sense (Batanero, Díaz, Contreras, & Roa, 2013) in children and teachers.

ACKNOWLEDGEMENT

Project EDU2013-41141-P (MEC), Group FQM126 (Junta de Andalucía) and Grant CONICYT PFCHA 72150306.

REFERENCES

- Arteaga, P. (2011). *Evaluación de conocimientos sobre gráficos estadísticos y conocimientos didácticos de futuros profesores* (Doctoral dissertation). Universidad de Granada, Granada, España.
- Arteaga, P., Batanero, C., Cañadas, G., & Contreras, J. M. (2011). Las tablas y gráficos estadísticos como objetos culturales. *Números. Revista de Didáctica de las Matemáticas*, 76, 55–67.
- Arteaga, P., Batanero, C., & Contreras, J. M. (2011). Gráficos estadísticos en la educación primaria y la formación de profesores. *Indivisa* (12), 123-135.
- Batanero, C., Arteaga, P., & Ruiz, B. (2010). Análisis de la complejidad semiótica de los gráficos producidos por futuros profesores de educación primaria en una tarea de comparación de dos variables estadísticas. *Enseñanza de las Ciencias*, 28(1), 141-154.
- Batanero, C., Díaz, C., Contreras, J. M., & Roa, R. (2013). El sentido estadístico y su desarrollo. *Números*, 83, 7-18.
- Bertin, J. (1967). *Semiologie graphique*. Paris: Gauthier-Villars.
- Carvalho, J. I. F. (2011). *Média aritmética nos livros didáticos dos anos finais do ensino fundamental* (Master dissertation). Universidade Federal de Pernambuco, Recife, Brazil.
- CCSSI (2010). *Common Core State Standards for Mathematics*. Washington, DC: National Governors Association for Best Practices and the Council of Chief State School Officers.
- Curcio, F. R. (1989). *Developing graph comprehension*. Reston, VA: NCTM.

- Del Pino, G., & Estrella, S. (2012). Educación estadística: Relaciones con la matemática. *Pensamiento Educativo. Revista de Investigación Educativa Latinoamericana*, 49(1), 53-64.
- Díaz-Levicoy, D. (2014). *Un estudio empírico de los gráficos estadísticos en libros de texto de Educación Primaria española* (Trabajo Fin de Máster). Universidad de Granada. Granada, España.
- Franklin, C., Kader, G., Mewborn, D., Moreno, J., Peck, R., Perry, M., & Scheaffer, R. (2007). *Guidelines for assessment and instruction in statistics education (GAISE) report: A Pre-K- 12 curriculum framework*. Alexandria, VA: American Statistical Association.
- Gal, I. (2002). Adult's statistical literacy: Meaning, components, responsibilities. *International Statistical Review*, 70(1), 1-25.
- Gea, M. M., Batanero, C., Arteaga, P., Cañadas, G. R., & Contreras, J. M. (2014). Análisis del lenguaje sobre la correlación y regresión en libros de texto de bachillerato. *SUMA*, 76, 37-45.
- Gómez, E., Ortiz, J. J., Batanero, C., & Contreras, J. M. (2013). El lenguaje de probabilidad en los libros de texto de Educación Primaria. *UNIÓN. Revista Iberoamericana de Educación Matemática*, 35, 75-91.
- López, F. (2002). El análisis de contenido como método de investigación. *XXI. Revista de Educación*, 4, 167-180.
- May, R. (2009). *La representación gráfica en estadística a nivel superior: un análisis de libros de texto en psicología y educación* (Master dissertation). Universidad Autónoma de Yucatán. Mérida de Yucatán, México.
- Mateus, L. (2014). Estudio de gráficos estadísticos usados en una muestra de libros de matemáticas para la educación básica y media en Bogotá. In L. Andrade (Ed.), *Memorias del I Encuentro Colombiano de Educación Estocástica* (pp. 274-280). Bogotá: Asociación Colombiana de Educación Estocástica.
- MECD (2014). *Real Decreto 126/2014, de 28 de febrero, por el que se establece el currículo básico de la Educación Primaria*. Madrid: Author.
- MINEDUC (2012). *Matemática educación básica. Bases curriculares*. Santiago: Author.
- Monroy, R. (2007). Categorización de la comprensión de gráficas estadísticas en estudiantes de secundaria (12-15). *Revista Electrónica de Investigación en Educación en Ciencia*, 2(2), 29-38.

APPENDIX: TEXTBOOKS ANALYSED

- T1. Salazar, R., & Sprovera, M. (2014). *Matemática 1º Básico. Texto del Estudiante*. Santiago: Fe y Alegría.
- T2. Ho Kheong, F., Ramakrishnan, C., Pui Wah, B. L., & Choo, M. (2014). *Mi Matemática. Texto del Estudiante 2º*. Santiago: Marshall Cavendish.
- T3. Charles, R., Caldwell, J., Cavanagh, M., Chancellor, D., Copley, J., Crown, W., Fennell, F., Ramirez, A., Sammons, K., Schielack, J., Tate, W., & Van de Walle, J. (2014). *Matemática 3º Educación Básica. Texto del estudiante*. Santiago: Pearson.
- T4. Andrews, A., Dixon, J., Norwood, K., Roby, T., Scheer, J., Bennett, J., Luckie, L., Newman, V., Scarcella, R., & Wright, D. (2014a). *Matemática 4º Básico. Texto del estudiante*. Santiago: Galileo.
- T5. Andrews, A., Dixon, J., Norwood, K., Roby, T., Scheer, J., Bennett, J., Luckie, L., Newman, V., Scarcella, R., & Wright, D. (2014b). *Matemática 5º Básico. Texto del estudiante*. Santiago: Galileo.
- T6. Andrews, A., Dixon, J., Norwood, K., Roby, T., Scheer, J., Bennett, J., Luckie, L., Newman, V., Scarcella, R., & Wright, D. (2014c). *Matemática 6º Básico. Texto del estudiante*. Santiago: Galileo.
- T7. Baeza, A., López, F., Sandoval, M., & Urrea, A. (2013). *Matemática 1º Básico. Tomo II*. Santiago: Santillana.
- T8. Baeza, A., Blajtrach, P., Kükenshöner, C., & Sandoval, M. (2013). *Matemática 2º Básico. Tomo II*. Santiago: Santillana.
- T9. Véliz, C. (2013). *Matemática 3º Básico. Tomo II*. Santiago: Santillana.
- T10. Batarce, Y., Cáceres, B., & Kükenshöner, C. (2013). *Matemática 4º Básico. Tomo II*. Santiago: Santillana.
- T11. Ávila, J., Fuenzalida, C., Jiménez, M., & Ramírez, P. (2013). *Matemática 5º Básico. Tomo II*. Santiago: Santillana.
- T12. Ávila, J., Castro, C., Merino, R., & Ramírez, P. (2013). *Matemática 6º Básico. Tomo II*. Santiago: Santillana.