ACTA SCIENTIAE

ISSN: 2178-7727
DOI: 10.17648/acta.scientiae.5884

Chilean Primary School Children's Understanding of Statistical Graphs

Pedro Arteaga[a]
Danilo Díaz-Levicoy[b]
Carmen Batanero[a]

[a]Universidad de Granada, Facultad de Ciencias de la Educación, Departamento de Didáctica de la Matemática, Granada, España
[b]Universidad Católica del Maule, Facultad de Ciencias Básicas, Departamento de Matemática, Física y Estadística, Talca, Región del Maule, Chile

Received for publication on 28 May. 2020. Accepted after review on 7 Sep. 2020
Designated editor: Claudia Lisete Oliveira Groenwald

ABSTRACT

Context: Statistical graphs are pervasive in the media and professional settings and therefore its understanding is a relevant component of statistical literacy. **Objectives***:* The aim of this research was to assess the understanding of statistical graphs achieved by Chilean children when finishing their primary education and how much of this understanding is maintained one year later. **Design:** To achieve this aim, we built a comprehensive questionnaire which takes into account a previous analysis of the Chilean curricular guidelines and textbooks for primary education. The main variables characterizing graphical competence in the literature, and recommendations from experts in statistics education were also considered in the construction of the questionnaire. **Participants and setting:** The sample is intentional and is made of 745 6th and 7th grade students in Chile. **Data selection and analysis:** The responses to this questionnaire are analysed quantitatively and qualitatively. **Results:** We inform of the task difficulty according to the different variables in the questionnaire and the average reading level of statistical graphs achieved in the grades. We also inform of the main difficulties related to the different graphs and activities proposed in the questionnaire. **Conclusions**: We conclude that hat the students learnt by the end of the 6th grade is remembered one year later, dot plot and pie charts were the most difficult graphs, building of graphs was difficult when the frequencies were not provided and few children attained the upper level of reading graphs.

Keywords: reading levels and understanding, primary education, statistical graphs.

Comprensión de gráficos estadísticos por niños chilenos de Educación Primaria

RESUMEN

Contexto: Los gráficos estadísticos permean los medios y contextos profesionales y por tanto su comprensión es una componente relevante de la cultura estadística. **Objetivo**: La finalidad de este trabajo fue evaluar la comprensión de los gráficos estadísticos que alcanzan los niños chilenos al finalizar su educación primaria y lo que recuerdan un año después. **Diseño**: Para

Corresponding author: Danilo Diaz-Levicoy. Email: dddiaz01@hotmail.com

alcanzarlo, se construyó un cuestionario comprensivo que tiene en cuenta un análisis previo de las directrices curriculares y libros de texto chilenos de educación primaria. Se tienen en cuenta en la construcción del cuestionario las principales variables que caracterizan la competencia gráfica en la investigación y las recomendaciones de expertos. **Entorno y participantes**: La muestra es intencional y estuvo formada por 745 niños de los cursos 6º y 7º en Chile. **Recopilación y análisis de datos**: Se analizan las respuestas en forma cualitativa y cuantitativa. **Resultados:** Se describe la diferencia de competencia y el nivel medio de lectura de gráficos estadísticos en los dos cursos. También informados de las principales dificultades relacionadas con los gráficos y las actividades propuestas. **Conclusiones**: Concluimos que lo aprendido al final de 6º curso sobre los gráficos se recuerda un año más tarde, que los gráficos de punto y de sectores son los más difíciles, la construcción de gráficos es difícil cuando no se dan las frecuencias y pocos niños alcanzan el nivel más avanzado de lectura de gráficos.

Palabras clave: niveles de lectura y comprensión, escuela primaria, gráficos estadísticos.

Compreensão de gráficos estatísticos por crianças chilenas do ensino fundamental

RESUMO

Contexto: Os gráficos estatísticos permeiam a mídia e os contextos profissionais e, portanto, seu entendimento é um componente relevante da cultura estatística. **Objectivos:** O objetivo deste trabalho foi avaliar a compreensão dos gráficos estatísticos que as crianças chilenas alcançam quando terminam o ensino fundamental e o que lembram um ano depois. **Design:** Para isso, foi fornecido um questionário, que leva em consideração uma análise prévia das diretrizes curriculares chilenas e dos livros didáticos da educação primária. As principais variáveis que caracterizam a competência gráfica na pesquisa e as recomendações de especialistas são levadas em consideração na construção do questionário. **Entorno e participantes:** : A amostra é intencional e consistiu em 745 crianças do 6º e 7º ano no Chile. **Coleta e análise de dados:** As respostas são analisadas quantitativamente e qualitativamente. **Resultados:** Se estuda a diferença de competência e o nível médio de leitura de gráficos estatísticos nos dois cursos. Também informamos das principais dificuldades relacionadas aos gráficos e às atividades propostas. **Conclusões:** Concluímos que o que foi aprendido no final do 6º curso de gráficos é lembrado um ano depois, que os gráficos de ponto e setor são os mais difíceis, a construção de gráficos é difícil quando as frequências não são dadas e poucas crianças atingem o nível leitura de gráficos mais avançada

Palavras-chave: níveis de leitura e compreensão, ensino fundamental, gráficos estatísticos

INTRODUCTION

The need for statistical literacy to assure active citizen participation in public decision processes is today reinforced by the amount of statistical information on the media (Gal, 2002; Engel, 2017). Often this information is presented in statistical graphs that the citizen has to interpret in order to understand the related reports and /or make decisions concerning different relevant aspects of his or her life (Ridgway, 2016). Consequently, a basic competence to understand and work with elementary statistical graphs is one essential aspect of critical statistical literacy (Aoyama, 2007; Chick & Pierce, 2012; Ridgway, Nicholson, Sutherland, & Hedger, 2019).

Statistical graphs have been taught at secondary school level for years, but it is only in the past decades when this content was included in the primary school curricula around the world (e.g. CCSSI, 2010; MEC, 2006; NCTM, 2000). In Chile, primary school children are requested to collect and record data to answer statistical questions about themselves and their environment, using tables and graphs from the 1st grade of primary school. More specifically, in the current curricular guidelines (MINEDUC, 2012), there is an explicit recommendation to work with the following types of graphs: pictograms (Grades 1-4), bar graphs (Grades 2-5), dot plots (Grades 3 and 6), line graphs (Grade 5), stem and leaf diagrams (Grades 5 and 6) and pie charts (Grade 6).

Including statistical graphs in the primary school curricula is a first step to educate future statistically literate citizens. Notwithstanding, this relevant educational change should go along with research into the children's competence to accomplish the types of tasks suggested in these guidelines. As suggested by Freedman and Shah (2002), a fundamental research question is the characterization of both the graph comprehension processes and the characteristics of tasks that help the students develop this understanding.

Research focusing on children's ability to read and build statistical graphs and on the level of complexity of the graphical tasks that can be proposed to the children is recently increasing. However, this research has centered on isolated types of graphs and has been carried out in countries different from Chile; furthermore, we found no comprehensive questionnaire that takes into account the curricular content in Chile and the main variables characterizing graphical competence. In order to fill this gap, the aim of this research was to assess the global understanding of statistical graphs achieved by Chilean children when finishing their primary education and how much of this knowledge is maintained one year later when students enter the middle school. A second goal was to build a questionnaire based on the study of literature and in a previous analysis of the content related to statistical graphs in the most frequently used Chilean textbooks (Díaz-Levicoy, Batanero, Arteaga, & Gea, 2016). Below we summarize the research background and method, and then describe the construction of the questionnaire and its psychometric features. Finally, we present a summary of the main results concerning the effect of the main variables included in the questionnaire on the tasks' difficulties, the children´s average reading levels in the reading tasks and their main difficulties in building statistical graphs.

BACKGROUND

Different authors have defined graphical understanding which includes the ability to build, read and interpret graphs, and to translate graphs to a different representation. We base on the work by Friel, Curcio and Bright (2001), who defined graphical understanding as the "graph readers' abilities to derive meaning from graphs created by others or by themselves" (p. 132). For these authors, graph understanding includes the following competences:

- Recognizing the graph structural elements (axes, scales, labels, specific elements) and their relationships. In our questionnaire this competence is taken into account by including questions involving the reading of these different elements of the graph.
- Assessing the impact of each structural element on the presentation of information in a graph (for example, being able to predict the change of the graph when changing the scale in an axis). We consider this ability in a task, where the students need to select between two line graphs with different scales.
- Translating the relationships reflected in the graph to the data represented in the same and vice versa. We included some tasks where students have to translate from a graph to a table and others where they have to build a graph.
- Recognizing when a graph is more useful than another, depending on the problem and data represented. Again, one task requires selecting between different graphs, the most adequate for a given table.

Reading Levels

Different authors have defined levels in reading the graphs to recognize that questions of different difficulty can be proposed for the same graph. We considered the reading levels introduced by Curcio (1989), Friel et al. (2001) and Shaughnessy, Garfield and Greer (1996):

Reading the data, where only the literal reading of a graph element is requested; for example, asking a child to read the frequency for a given value of the variable.

Reading between the data. Besides the literal reading of the graph, in this level the child has to perform comparison of various data represented in the graph or complete some arithmetical calculations with the data. For example, the child is asked to obtain the mean or mode in a distribution represented in a graph.

Reading beyond the data. This level involves obtaining or predicting information that is not directly represented in the graph, which requires the interpolation or extrapolation of the information displayed. For example, from a graphical representation of data corresponding to six consecutive years, the child is asked to predict the value of the variable represented in the following year.

Reading behind the data. A child attains this level when he or she is able to make a critical valuation of the graph, of the way it has been constructed or can discuss a statement related to the graph content. An example is discussing the truth or falsity of an assertion, using the information represented in the graph.

Graph semiotic complexity

Each element of the graph and the graph itself represent a different object and the child has to interpret this representation. According to Font, Godino and D'Amore (2007) symbols or words used in mathematics to represent abstract objects can be conceived as semiotic functions in terms of Eco (1976), where an expression (the word or the symbol) and content (the abstract object) are related by a link, which is usually conventionally established. When producing a graph, the author performs a series of actions (such as deciding the particular type of graph or fixing the scale), and uses concepts and properties that vary in different graphs. Consequently, the author implicitly establishes a series of semiotic functions that can be more or less difficult to interpret by the child, depending on the particular type of graph. Taking into account the above ideas, Batanero, Arteaga and Ruiz (2010) defined different levels in the graph semiotic complexity, as follow:

- *C1. Graph representing only individual results.* When the graph only represents isolated data values, for example the student's personal data. As a consequence, to interpret these graphs there is no need to use the ideas of statistical variable or distribution.
- *C2. Representation of a data list*. Sometimes the data in a list are represented one by one, without an attempt to order the data or to combine identical values. Consequently, to interpret these graphs students do not need the idea of frequency or distribution.
- *C3. Representation of one distribution*. In these graphs, the data have been ordered, and the frequency for the different values of the variable are obtained. To interpret the graph, the student needs to understand the idea of distribution.
- *C4. Joint graph for two or more distributions.* In this case two or more distributions are represented in the same graph, and therefore the interpretation is more complex.

Children's graphical competence

Our work is also based on previous research which describes school children's competence to deal with statistical graphs. This research is scarce, but is receiving increasing attention; the main examples of this type of research are summarized below:

Sharma (2006) performed interviews to 29 students aged 14-16 years in relation to statistical tables and graphs. The author suggests the students could read and compare data presented in bar graphs and classified the children's responses in non-statistical, partial statistical and statistical responses. Pagan, Leite, Magina and Cazorla (2008) proposed the reading of two bar graphs to 399 5th and 8th grade students in Brazil and informed that 84% of them correctly performed the tasks involving the reading level R1 and 43% those of reading level R2. Canché (2009) applied a multiple-choice questionnaire, with bar graphs to 206 6th grade students in Mexico; the percentage of correct answers ranged between 19.4% and 82% in the task requesting R1and R2 levels questions, and between 22.3% and 79.6% in the level R3 tasks.

Fernandes and Morais (2011) analyzed the responses by 108 9th grade students to activities related to the reading of a bar graph, a pie chart and a line graph. In the bar graph 90% of students reached the level R1 and 23% the level R2. In the pie chart, 96% of the sample succeeded at the level R1, 31% at R2 and 23% at R3. In the line graph 19% of children obtained the level R1, 14% the level R2 and 43% the level R3. Evangelista (2014) studied the performance of 60 5th grade students in Brazil when working with bar and line graphs using activities of reading levels R1 and R2. On average, the students correctly answered 59% of activities related to bar charts and 43% the tasks related to line graphs.

Fernandes and Morais (2011) analyzed the graphs produced by 108 9th degree children to represent ages by gender and life span of 21 different animal species. This second task was easier (61% of correct graphs), while in the first task the percentage of success was only 35%. Most errors consisted of selecting an inadequate graph, omitting titles or labels, and using non-proportional scales. Evangelista, Oliveira and Ribeiro (2014) asked 46 5th grade children to build two statistical graphs, obtaining 88.1% of correct graphs, although none of the children added a title to the graph and only a few of them used a label to describe the variables on the axis. Only 19.6% used a correct scale.

None of this research has dealt with Chilean children and used isolated activities that did not take into account at the same time the different types of graphs, activities, reading levels and semiotic complexity of the tasks proposed to the children. To fill this gap, we built a comprehensive questionnaire and analyzed the responses to the same by 745 Chilean primary school children. The aim of this paper is to describe the questionnaire and present a summary of the main results that describe the children's graphical understanding when finishing primary school, and when entering the middle school, the following year. Further details of the children's reading levels in some tasks of the questionnaire have been published in Díaz-Levicoy, Batanero, Arteaga and Gea (2019) and of children's difficulties in translating pictograms in Batanero, Díaz-Levicoy and Arteaga (2018).

METHOD

A total of 745 (380 6th grade children (11-12-year-olds) and 365 7th grade children (12-13-year-olds)) took part in the sample. These students were recruited from 13 public or private schools in different cities of Chile (Osorno, Castro, Queilen, Puerto Octay, La Union, Viña del Mar and Ñuñoa), to achieve a reasonable representativeness of the sample. The centers collaborated in the research, with permission from the educational authorities, and the questionnaire was given to the children as a classroom activity in the presence of the classroom teacher and one of the authors[1].

Questionnaire design and construction

To build a questionnaire that took into account the main activities related to graphs that Chilean children are given in the school, we previously performed an analysis of Chilean primary textbooks education (three complete series of books for grades 1rst to 6th), including the series which is provided free of cost to students in the public schools in the country (Díaz-Levicoy et al., 2016). In all these books we studied the different types of graphs included, as well as three main variables that we used in the questionnaire to define the children's graphical understanding: a) the graph semiotic complexity, b) the type of task proposed in relation to the graph; and c) the reading level required to solve the activity proposed. Taking into account the frequency of the different types of graphs, activities, semiotic complexity and reading levels, in Tables 1 and 2 we present the combination of categories decided for these variables in each of the tasks of the questionnaire.

Table 1
Task Distribution by Type of Graph and Complexity Level

Type of graph	Graph complexity (data represented)		
	C2: Data list	**C3: One distribution**	**C4: Two or more distributions**
Bar graph	Task 1 and 2	Task 7	Task 3
Pictogram	Task 4	Task 5	-
Line graph	Task 6	-	Task 8
Stem and leaf	-	Task 9	-
Pie chart	-	Task 10	-
Dot plot	-	Task 11	-

[1] The research, where ethical aspects have been respected, did not go through an ethics committee, as it was not a requirement. Participants were informed of the purpose of the study and that their identity would not be disclosed. Entrance to the educational centres was requested through the directors, heads of pedagogical units or mathematics teachers, obtaining authorization in all cases. Consequently, we discard Acta Scientiae from any responsibility arising for data collection and analysis described in the paper, according to the Brazil Ministry of Health Resolution nº510, April 07, 2016.

Table 2
Task Distribution by Activity and Reading Level

Activity requested to the child	Reading level		
	R1	R2	R3 - R4
Reading the graph	Task 6	Task 9	Task 2, 5, 11
Building a graph	Task 1	Task 7	-
Translating from graph to table	Task 3	Task 4	-
Selecting a graph for some given data	-	-	Task 8, 10

Starting from these tables, we selected three different possible tasks for each combination of categories (33 tasks in total), among many other possible tasks taken either from the primary school Chilean textbooks or from previous research. The tasks were translated to Spanish when needed, the related questions and graphs modified if requested to fit Tables 1 and 2 and the context was adapted to situations familiar for Chilean children. The set of 33 preliminary tasks was sent to 10 expert statistical education researchers, from different countries in Latin-America, Spain and Portugal, who were fluent in Spanish, had performed a doctoral dissertation in statistical education and worked in mathematics education at the time the tasks were sent. These expert researchers completed a questionnaire in which, for each of the 33 preliminary tasks, they were asked to score (0 to 5 points) the task adequacy to represent the given combination of variables; for example, for task 4, they should score the adequacy of each of the three possible examples to evaluate the children´s competence to translate a pictogram representing a data list to a table, requiring a maximum reading level R2. Additionally, the experts were requested to judge the clarity and simplicity (for the children) of the task and to provide suggestions to improve the task.

Once the questionnaires filled by the experts who collaborated in our work were collected, we computed the average score obtained with the expert data in each task. For each possible combination of categories in Table 1 and 2, we selected the task with the maximum mean score and minimum standard deviation. All the tasks finally included in the instrument got an average score of 4 or higher and some of them were reworded taking into account the experts' suggestions. Finally, and before starting the collection of data, we tried the questionnaire with a small sample of children that did not take part in the final sample to estimate the time needed to complete all the tasks. We also asked these children if they understood all the tasks, and to point to any word or expression difficult for them; as a consequence of this pilot trial, still some small changes in the wording of the tasks were performed. The translation to English of the final questionnaire is presented in the Appendix. The Cronbach reliability coefficient of the questionnaire in the sample was Alfa=0.773, which is reasonably high, given the variety of different tasks included in the questionnaire.

RESULTS

Task Difficulty

An indicator of a questionnaire quality is the distribution of the task's difficulty, which reflects the proportion of correct responses to the same. In our study, we took into account both the correct and partly correct responses (responses that are basically correct with minor errors, such as, for example, forgetting the title of the graph when building a completely correct graph). We then assigned 2 points to correct responses and 1 point to partly correct responses in each of the questions included in each task. Then we divided the score obtained by each student by the maximum possible score in the task and computed the mean of all the students in each school grade (See Table 3, where the different tasks are ordered by difficulty).

We also performed the t-test of difference in averages in both groups in each task and marked with an asterisk (*) those tasks where the difference in average was significant (p-value in the t-test lower than .05) and with two asterisks (**) those tasks with a highly significant difference in the groups (p-value in the t-test lower than .05). In this table we observe that the average score is usually lower in the older students (7th grade), although the difference is usually non-significant, except in a few tasks. This means that what the students learnt by the end of the 6th grade is remembered one year later. The exceptions are task 2 (extrapolating a value in a bar graph) and task 9, reading the stem and leaf.

Table 3
Task Average Difficulty by Group and Task Characteristics

Task	Average difficulty		Characteristics of the task			
	Grade 6 (n=380)	Grade 7 (n=365)	Type of graph	Complexity level	Reading level	Activity
T11	0.41	0.39	Dot plot	C3	R4	Reading
T3	0.42	0.40	Bar graph	C4	R1	Translating
T10	0.42	0.43	Pie chart	C3	R4	Selecting
T7	0.42	0.43	Bar graph	C3	R2	Building
T2**	0.46	0.39	Bar graph	C2	R3	Reading
T8	0.5	0.5	Line graph	C4	R4	Selecting
T1*	0.56	0.51	Bar graph	C2	R1	Building
T5	0.57	0.56	Pictogram	C3	R4	Reading
T9**	0.64	0.52	Stem and leaf	C3	R2	Reading
T6	0.75	0.75	Line graph	C2	R1	Reading
T4	0.91	0.90	Pictogram	C2	R2	Translating

*significant (p-value <.05); **highly significant (p-value <.01

We observe a reasonable difficulty in all the tasks with a range of variation in difficulty from .4 to .91; half the tasks had difficulty .5 or higher, which reflects a reasonable understanding of statistical graphs in the sample of children. Taking into account these results the most difficult tasks were the following:

- *Task 11* (mean difficulty= .41 and .39 in the groups). In this task the child is asked for a critical reading of a dot plot, where the information in the graph should be used to refute or confirm an assertion about the scores in two different subjects. About 60% of the students recognised that the assertion was false, that is, they were able to read the graph and related the data to the sentence given in the task. However, not all of the children provided a sound argument to refute the sentence (they did not achieve the reading level R4), although about 50% of the children reached the reading level R2, because they were able to read the graph and perform comparisons of parts of the graph.
- *Task 3* (mean difficulty= .42 and .40). In this task the students should translate a double bar graph (semiotic complexity C4) to a table and only 36% of them built a correct table, due to problems in either reading the graph or in building the table.

The easiest tasks, where the majority of students succeeded were:

- *Task 4* (mean difficulty = .91 and .90). Translating a pictogram (complexity level C2 and reading level R2) to a table. In this task, 75.4% of students built a correct table and 18.5% a partly correct table. Besides, 92.6% of them reached the maximum reading level in the task. When comparing with task 3, the influence of the complexity level of the graph that is much lower in task 4, is clear.
- *Task 6* (mean difficulty= .75). Reading a line graph of semiotic complexity C2 and reading level R1. In this task there was 94.4% correct responses in the reading of the title, 93.3% in the direct reading of the graph and only 31.4% in the description of the variables represented in the graph, probably because the students' lack of argumentative competence.

In Table 4 we present the average difficulty of tasks for the different categories in each variable included in the questionnaire design that have been computed by averaging the different tasks of a same type. For example, when we compute the average difficulty in tasks 1, 2, 3 and 7 we obtain the average difficulty in bar graphs. Since the differences between the two grades in Table 3 were small and mainly non statistically significant we have not split the groups in Table 4. In this table the relative difficulty of the different graphs and activities have been ordered by increasing value of difficulty.

Table 4

Average Difficulties for the Different Variables included in the Questionnaire

Type of graph		Activity		Semiotic complexity		Reading level	
Dot plot	0.40	Selecting	0.47	C2	0.66	R1	0.57
Pie chart	0.43	Building	0.49	C3	0.48	R2	0.64
Bar graph	0.45	Reading	0.55	C4	0.46	R3-4	0.47
Stem and leaf	0.59	Translating	0.66				
Line graph	0.63						
Pictogram	0.66						

Relative difficulty of different graphs.

Dot plots and pie charts represented more challenges for the children, probably because these graphs are less frequently used in their textbooks and because they are introduced at higher grades, so that the students had less practice with them.

In the pie chart (Task 10), a frequent error was confusing frequencies and percentages:

> I select graph A, since it represents, the data on the table, where 100 children play football, 40 tennis and 60 basketball (S38).

As regards the dot plot (Task 11), some students did not use the data on the dot plot to answer the proposed question, as happens in the following response:

> I agree, as mathematics is very difficult (S6).

The children's understanding of stem and leaf (Task 9) was reasonable, in considering that this graph was only worked in grades 5th and 6th. Common errors in reading this graph were not considering the leaves, interpreting the leaves corresponding to the same stem as the decimal part of a number, or answering with the most frequent value of leaves:

> The most frequent score is 40 (S50).
>
> 4.111169 (S166).
>
> 6 (S266).

Finally, line graphs and pictograms were easy for these children and the difficulty of bar graph depended on the activity.

Relative difficulty of the various activities.

The most difficult activity was selecting an appropriate graph for a given data set, again because this activity is scarce in the textbooks and requires an understanding of the features of the graphs, and a high reading level. For this reason, there was a high percentage of no response in Tasks 8 (15.3%) and 10 (25.2%). Translating a graph to a table had moderate difficulty. The failure in the activity was usually due to errors in reading the graph, building a different graph instead of a table, building a different type of table or not computing the frequencies (See example in Figure 1).

Figure 1

Using iconic representation of students, instead of frequencies to build a table in Task 3 (S63).

Finally, building and reading graphs had moderate difficulty and the children's most common difficulties in these tasks are discussed in sections 4.2 and 4.3.

Effect of semiotic complexity and reading level

As predicted, the task was more difficult when the semiotic complexity of the graph increased. There was not such a clear pattern in the reading levels, because of the interaction between these levels and the other variables and because not all the tasks took into account all the possible reading levels.

Children's Reading Ability

We included five tasks with questions that served to evaluate the reading level these children were able to reach. We considered that the child reached the *Level R1. Reading the data,* when he or she correctly made a simple reading of data in the graph, but failed in making some computation or comparisons of these data. For example, in Task 5, S55 correctly counts the number of icons for the science fiction row, but fails in realizing that each icon counts for 15 books:

> True, because in the inventory row for science books there are two books (S55).

A child was assigned the reading *Level R2, reading between the data,* when he or she correctly read the graph and, in addition performed some computations or comparisons with the same. For example, in Task 8, S17 selected the correct graph and pointed to the increasing tendency in Pedro's data; however, he failed in observing that Gabriel's data also rose and, moreover, he did not mention the effect produced in the graph by the change of scale.

> The right-handed graph because Pedro is increasing (S17).

Level R3. Reading beyond the data is achieved when, in addition to reading data and comparing data, a child is able to extrapolate a piece of data not represented in the graph. A response at this level in Task 2 is that by S25, who read all the data values in the line graph, observed the increasing tendency in the data series, computed the average daily increase in the number of calls and added this average to the last number in the series to estimate the numbers of calls for Sunday:

> 374,4. In the graph the number of calls was increasing and then I got the average that I added to the number of calls on Saturday and obtained this number (S25).

Finally, *level R4, reading behind the data,* involves not only the correct reading of the graph as well as the performing of the calculations needed to respond to the task, but being able to question an argument based on the data. For example, in Task 11, S42 read the graph and computed the maximum, as well the number of children with scores over 6 or equal and higher than 6. In addition, he combined all this data to offer a correct argument:

> I do not agree, since there are the same number of students with score 6,0 or higher; besides the worst mathematics score is 5,9 while in language the worst score is

5,8 [;]there are also more children with scores higher than 6 in mathematics than in language and the best score is 6.3 higher than the maximum in language (6,2) (S42).

In Table 5, we present the percentage of children achieving at least each reading level in the five tasks where several reading levels were possible. Few children were able to provide a response at the upper reading levels R3 (in task 2) and R4 in the other four tasks. This time the percentage of 7th grade students achieving these levels is higher than that of 6th graders, because of the higher maturity of the older children. In most tasks, the percentage of students´ reasoning only at a literal reading level R1 of the data is higher in the smaller students and ranges between 50.7 and 98.7%, which is compatible with results from Pagan et al. (84%), Canché (19.2-82%), Fernandes and Morais (19-96%).

Table 5

Percentage of Children Achieving at least the Different Reading Levels in Each Task

	Task 2		Task 5		Task 8		Task 10		Task 11	
Reading Level	6th n=380	7th n=365	6th n=380	7th n=365	6th n=380	7th n=365	6th n=380	7th n=365	6th n=380	7th n=365
R1	73.4	60.5	98.2	98.7	70.8	68.2	82.4	78.6	55.1	50.7
R2	56.3	43.8	62.7	59.5	65	63.8	32.9	39.7	48.5	47.4
R3-R4	8.4	9.6	6.6	6.3	11.6	15.9	7.9	11.2	1.1	1.4

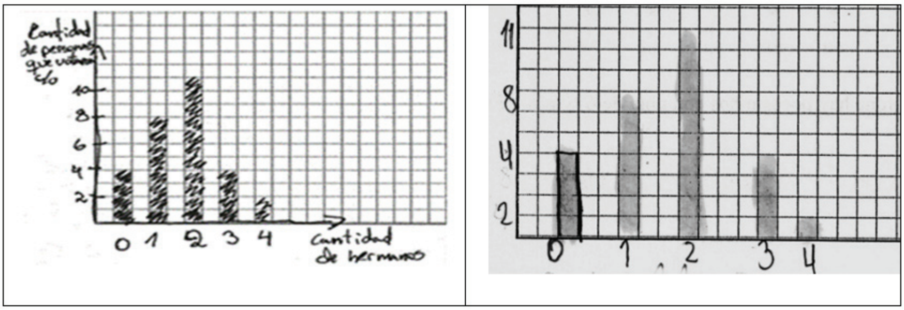
The percentage of children reaching at least level R2 in task 10 was small, due to confusion between absolute frequencies and percentages. In the other tasks these percentages (43.8-63.8%) are also compatible with those of Pagan et al. (43%), Canché (19.2-82%) and higher than those from Fernandes and Morais (14-31%). More details about the children’s reading levels in these tasks have been published in Díaz-Levicoy et al. (2019).

Children Competence in Building Graphs

In two tasks the students were given some data and were asked to build a bar graph using a grill that was provided to them, in order to facilitate the construction of the vertical scale and the location of modalities on the horizontal axis. The difference between both tasks is that in Task 1 the distribution of frequencies is given to the children, while in Task 7 they only receive a data list and should classify the numerical values and obtain the frequency distribution of the variable. Graphs were considered partly correct when the graph only contained some minor mistakes, such as for example, forgetting the title of the graph (an example is provided on the left hand side of Figure 2) omitting the labels for the modalities or the scale on the axes, using bars of different width, or not separating the bars in the graph (see example on the right hand side of Figure 2).

Figure 2

Examples of partially correct graphs



A graph was considered incorrect when the scale did not include the whole range of variation of the variable, the students built a different graph (e.g. a line graph) or the scale was not proportional. Some examples of non-proportional scales are given in Figure 3, in the second example the students place the labels on the Y axes in the order the data are given in Task 2, and the length of the bars do not correspond to the value of these labels. Both in Fernandes and Morais (2011) and Evangelista et al. (2014) the students tended to forget the title or labels and some of them used non proportional scales.

Figure 3

Examples of incorrect graphs

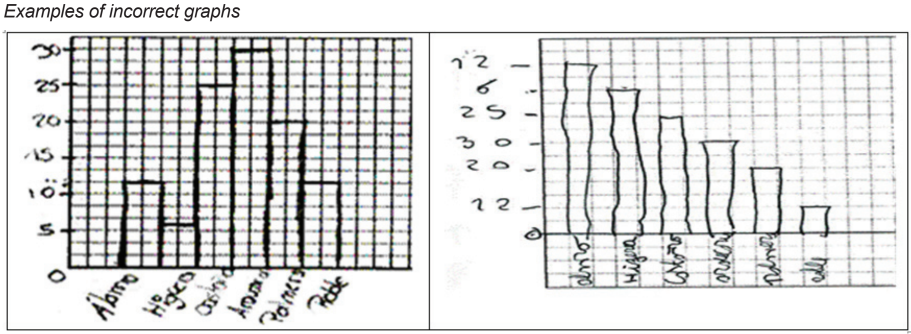


Table 6
Percentage of Students According to the Correction in the Building of a Bar Graph

	Task 1: frequencies provided		Task 7: no frequencies provided	
Graph correctness	Grade 6 (n=380)	Grade 7 (n=365)	Grade 6 (n=380)	Grade 7 (n=365)
Correct	5.8	6.3	4.7	7.9
Partly correct	72.9	65.5	40.3	44.1
Incorrect	6.1	5.5	31.3	17.8
No graph	15.3	22.7	23.7	30.1

The percentages of students in each group that provided no graph, incorrect, partly correct or correct graph are given in Table 6. We observe that most students succeeded in constructing a correct or partly correct graph in Task 1, where the frequency distribution was given to the children. This percentage is higher to that of correct graphs in Fernandes and Morais (2011) (61%) and smaller than that of Evangelista et al. (2014) (81%). However, these authors did not discriminate between correct and partly correct graphs and between providing or not the frequency distribution. In our research the success is much lower in Task 7 where students should form the frequency distribution, which was difficult in both groups.

IMPLICATIONS FOR RESEARCH AND PRACTICE

In this paper we provided a complete picture of the statistical graph understanding in a wide sample of Chilean students finishing their primary education and, in another sample, entering the middle school. Although other previous research provided isolated pieces of information concerning the variables considered in our questionnaire, only our previous work (Díaz-Levicoy et al., 2019) took into account Chilean children and none of that research used an instrument that completely reflected the curricular and textbooks content in primary education in Chile.

Our results concerning the reading levels that the children can achieve by the end of primary education do not differ much from previous research in other countries, such as for example, that by Canché (2009), Curcio (1987), Evangelista (2014), Fernandes and Morais (2011), and Pagan et al. (2008), but however does complement this research with a bigger sample of children in a different country.

We also provide new information to research on children's competence to build graphs, where the most frequent errors in this task reproduced those described in Fernandes and Morais (2011) and Evangelista et al. (2014). Moreover, we compared the children´s performance depending on whether the frequency distribution is given to the children or they should compute this distribution themselves.

In addition, we used a greater variety of graphs than that proposed in previous research, such as for example, the stem and leaf or the dot plot, for which there is not much information about the children´s understanding. Furthermore, we studied other activities such as translating the graph to a table or selecting a graph that adequately represent a given data set. All these variables, as well as the reading levels and the graph semiotic activities were considered in the textbooks, as was showed in our previous analysis of the same (Díaz-Levicoy et al., 2016) and therefore should be taken into account in the assessment of children's learning. This was the purpose of our questionnaire that was built using a sound methodological procedure and considering all the variables characterizing the activities proposed to children in the Chilean primary school curriculum.

The relative difficulty which for the children implied the different types of graphs and activities, reading levels and graph semiotic complexity should be taken into account when planning the introduction of graphs and related tasks throughout the primary school curriculum. Since the graph content in the Chilean curriculum is similar to those of many other countries, both the questionnaire and the results described in the paper can be useful to teachers and researchers, and in addition to the educational authorities responsible for revising the primary school statistical curriculum.

ACKNOWLEDGEMENT

Project PID2019-105601GB-I00 (MICIN) and Group FQM126 (Junta de Andalucía).

AUTHORS' CONTRIBUTIONS STATEMENTS

The three authors conceived the presented idea, developed the theory and built the questionnaire, collected the questionnaire and codified the data. All authors actively participated in the discussion of the results, reviewed and approved the final version of the work.

DATA AVAILABILITY STATEMENT

The data supporting the study will be available by D.D.L on reasonable requests to develop other possible studies in collaboration with the authors.

REFERENCES

Aoyama, K. (2007). Investigating a hierarchy of students' interpretations of graphs. *International Electronic Journal of Mathematics Education, 2*(3), 298-318.

Batanero, C., Arteaga, P. y 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-Levicoy, D., & Arteaga, P. (2018). Evaluación del nivel de lectura y la traducción de pictogramas por estudiantes chilenos de educación básica *Avances de Investigación en Educación Matemática, 14*(1), 49-64. https://doi.org/10.35763/aiem.v0i14.231
Canché, L. (2009). *La comprensión gráfica de los alumnos del nivel primaria* Master Dissertation. Universidad Autónoma de Yucatán, México.
Chick, H. L., & Pierce, R. (2012). Teaching for statistical literacy: Utilising affordances in real-world data. *International Journal of Science and Mathematics Education, 10*(2), 339-362.
CCSSI (2010). *Common Core State Standards for Mathematics*. National Governors Association Center for Best Practices and the Council of Chief State School Officers.
Curcio, F. R. (1989). *Developing graph comprehension.* NCTM.
Díaz-Levicoy, D., Batanero, C., Arteaga, P., & Gea, M. M. (2016). Gráficos estadísticos en libros de texto de primaria: Un estudio comparativo entre España y Chile *BOLEMA. Boletim de Educação Matemática, 30*(55), 713-737.
Díaz-Levicoy, D., Batanero, C., Arteaga, P., & Gea, M. M. (2019). Chilean children's reading levels of statistical graphs. *International Electronic Journal of Mathematics Education*, *14*(3), 689-700. https://doi.org/10.29333/iejme/5786
Eco, U. (1976). *A theory of semiotics.* Indiana University Press.
Engel, J. (2017). Statistical literacy for active citizenship: A call for data science education. *Statistics Education Research Journal, 16*(1), 44-49.
Evangelista, B. (2014). *Aprendendo a representar escalas em gráficos: um estudo de intervenção.* Master Dissertation. Universidade Federal de Pernambuco, Brasil. Retrieved from https://repositorio.ufpe.br/handle/123456789/13049
Evangelista, B., Oliveira, F. S., & Ribeiro, P. M. (2014). Analisando a construção de gráficos de alunos do 5º ano do ensino fundamental In CONEDU (Ed.), *Anais I Congresso Nacional de Educação* (pp. 1-5). Campina Grande, Brazil: Realize. Retrieved from https://www.editorarealize.com.br/revistas/conedu/anais.php
Fernandes, J. A. & Morais, P. C. (2011). Leitura e interpretação de gráficos estatísticos por alunos do 9º ano de escolaridade. *Educação Matemática Pesquisa, 13*(1), 95-115.
Freedman, E. G., & Shah, P. (2002). Toward a model of knowledge-based graph comprehension. In M. Hegarty, B. Meyer, & H. Narayanan (Eds.), *Proceedings of the International Conference on Theory and Application of Diagrams* (pp. 18-30). Springer.
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. https://doi.org/10.2307/749671
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.
Gal, I. (2002). Adult´s statistical literacy: Meaning, components, responsibilities. *International Statistical Review, 70*(1), 1-25.

MEC (2006). *Real Decreto 1513/2006, de 7 de diciembre, por el que se establecen las enseñanzas mínimas correspondientes a la educación primaria.* Autor.
MINEDUC (2012). *Matemática educación básica. Bases*. Unidad de Currículum y Evaluación.
NCTM (2000). *Principles and standards for school mathematics*. National Council of Teachers of Mathematics.
Pagan, A., Leite, A. P., Magina, S., & Cazorla, I. (2008). A leitura e interpretação de gráficos e tabelas no ensino fundamental e. In V. Gitirana, F. Bellemain, & V. Andrade (Eds.), *Anais do 2º Simpósio Internacional de Pesquisa em Educação Matemática* (pp. 1-10). Universidad Federal de Pernambuco.
Ridgway, J. (2016). Implications of the data revolution for statistics education. *International Statistical Review, 84*(3), 528-549. https://doi.org/10.1111/insr.12110
Ridgway, J., Nicholson, J., Sutherland, S., & Hedger, S. (2019). Critical statistical literacy and interactive data visualisations. In J. Evans, S. Ruane, & H. Southall (Eds.), *Data in society: challenging statistics in an age of globalisation* (pp. 349-357). Policy press.
Sharma, S. V. (2006). High school students interpreting tables and graphs: Implications for research. *International Journal of Science and Mathematics Education, 4*(2), 241-268.
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* (pp. 205-237). Kluwer. https://doi.org/10.1007/978-94-009-1465-0.

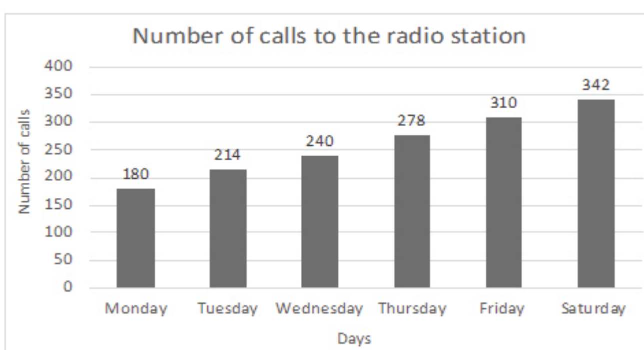
APPENDIX. QUESTIONNAIRE

Task 1. In the following table height in meters of some tree species are shown. Build a bar chart to represent this data (children were provided with a grid pattern)

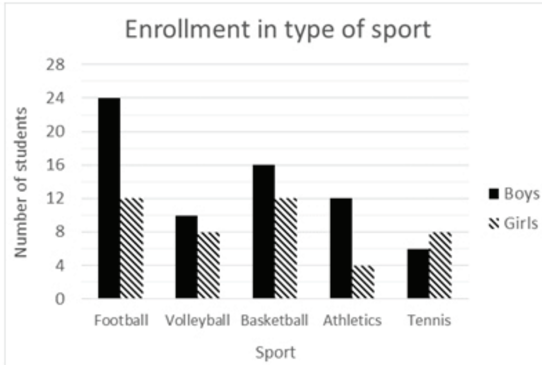
Specie	Poplar	Fig tree	Araucaria	Chestnut	Palm	Oak
Height	12	6	30	25	20	12

Task 2. Every day a radio station records the calls from the listeners. The first six days data are represented in the graph.

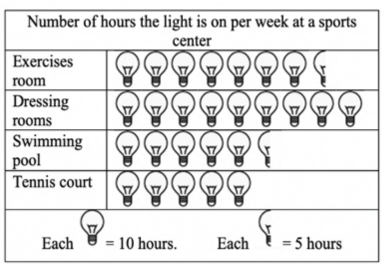
How many calls do you expect on the next day (Sunday)? Explain why you gave this number:



Task 3. The graph shows the number of students enrolled in each sport practiced in a school. Each student only practices one sport. Represent these data in a frequency table



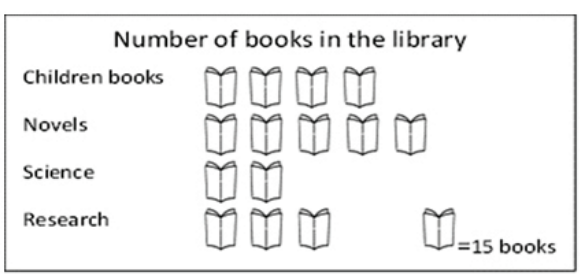
Task 4. Complete the following table with the information shown in the graph



Number of hours the light is turned on per week at a sports center

Place	Number of hours
Exercise room	
Dressing room	
Swimming pool	
Tennis court	
Total	

Task 5. The school librarian performed an inventory in the library and displayed the results in this graph. A child said that there are only two science books in the library. Another child suggested that there are 60 children´s books. Are the children right? Why or why not?



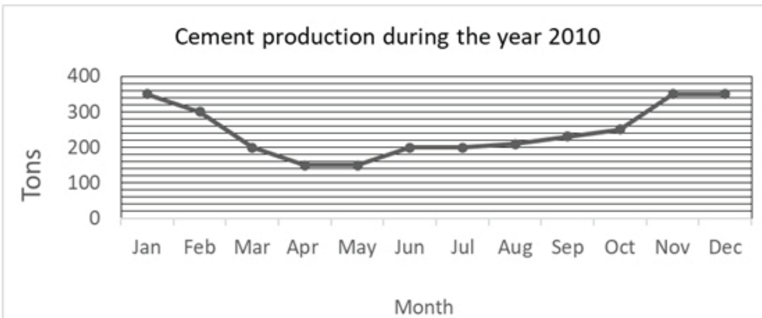
Task 6. Look at the following line graph and then answer.

1. What is the title of the graph?

2. Which are the variables represented?

3. How many tons of cement were produced during the month of February?

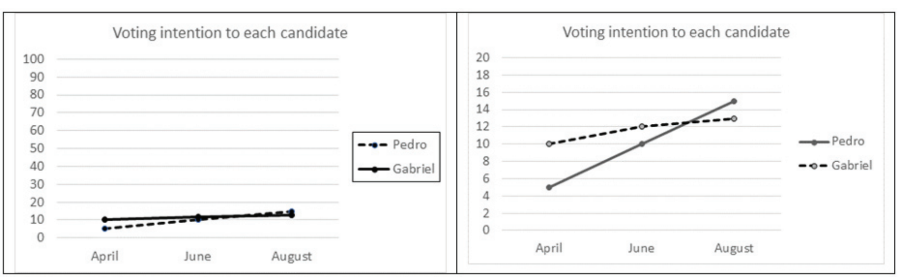
4. In what month (s) was the production 200 tons of cement?



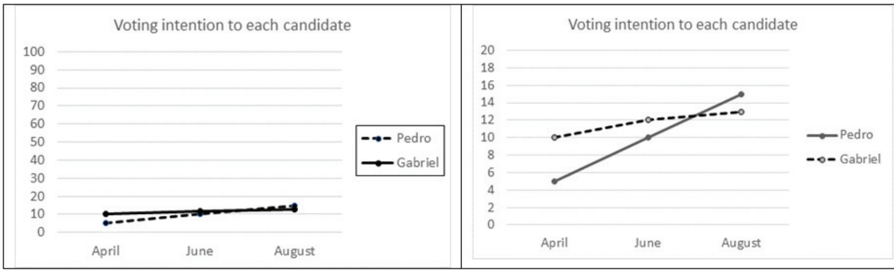
Task 7. Below the results of a survey with the question, “How many brothers do you have?” are given:

0, 1, 3, 2, 2, 3, 2, 4, 1, 2, 1, 2, 0, 2, 3, 1, 1, 0, 2, 4, 0, 1, 2, 3, 1, 1, 2, 2, 2

Build a bar graph with the information provided. (The children were provided with a grid pattern)



Task 8. These two graphs show the same results from a voting survey. If you were Pedro, which graph would you prefer? Why?



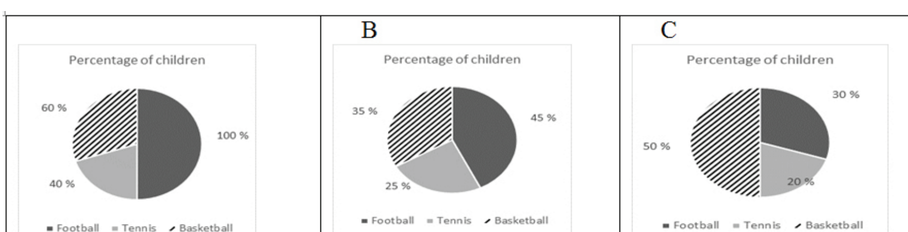
Task 9. The following graph shows the scores (between 2 and 7) obtained by 6th grade students in a maths test (with a decimal number). According to this information, respond:

1. What is the most frequent score?
2. How many students obtained a score greater than or equal to 5?

Stem	Leaves
2,	2 2
3,	0 0 0 9
4,	1 1 1 1 6 9
5,	0 3 5 7 8
6,	0 1 5 6 6
7,	0

Task 10. The following table shows the number of children in school practicing various sports. Which graph corresponds to the information given in the table? Why?

Sport	Number of children
Football	100
Tennis	40
Basketball	60



Task 11. In the following graphs we represent the language and mathematics scores of some children. Maria suggests that Mathematics is more difficult, as there are fewer children with 6 or more points. Do you agree with Maria? Why or why not?