Zetetiké

DOI: 10.20396/zet.v28i0.8656854

Comprehension of statistical graphs and tables by primary school teachers-in-training

Comprensión de gráficas y tablas estadísticas de estudiantes para profesor de educación primaria

Daniel Eudave Muñoz[1]

Ana Cecilia Macías Esparza[2]

Margarita CarvajalCiprés[3]

María Guadalupe Muñoz Macías[4]

Abstract

Understanding statistical graphs and tables is essential for elementary school teachers' performance, as they are part of the basic education curriculum content; moreover, they are used for the elaboration of educational planning and diagnosis. For this reason, teachers' initial training must consider these subjects. An exploratory-descriptive research, which was made in two stages, is reported: in the first stage, 240 students of the B.A. in Elementary Education at a Teacher Education School in Mexico took a test to analyze a statistical table and a graph; in the second stage, a series of task-focused interviews were applied to a sample of nine students in order to recognize their level of comprehension and difficulties. The results demonstrate different levels of comprehension of the graph and the table, as well as the basic underlying statistical concepts. The findings suggest that students have not worked enough with statistical graphs and tables.

Keywords: teachers' initial training; graphs and tables comprehension; statistical sense; fundamental statistical ideas.

Resumen

La comprensión de gráficas y tablas estadísticas es fundamental para el desempeño de los profesores de educación primaria, por ser un contenido curricular de la educación básica, y por su utilización en la realización de diagnósticos y planeaciones educativas. Por tanto, la formación inicial de los docentes debe contemplar estos temas. Se reporta una investigación exploratoria-descriptiva realizada en dos fases: en la primera se aplicó una prueba para analizar una gráfica y una tabla estadística, a los 240 alumnos de la Licenciatura en Educación Primaria, en una Escuela Normal, en México; en la segunda, se realizaron entrevistas centradas en tareas a una muestra de nueve alumnos, para reconocer sus comprensiones y dificultades. Los resultados muestran diferentes niveles de comprensión de la gráfica y la tabla, así como de los conceptos estadísticos fundamentales subyacentes. Los hallazgos sugieren que los estudiantes no han trabajado lo suficiente con gráficas y tablas estadísticas.

Submetido em: 27/09/2019 – **Aceito em**: 16/01/2020 – **Publicado em**: 14/02/2020

[1] Doctor en Educación, profesor-investigador de la Universidad Autónoma de Aguascalientes, México. Email: deudave@correo.uaa.mx.

[2] Doctora en Investigación Educativa, profesora-investigadora de la Universidad Autónoma de Aguascalientes, México. Email: acmacias@hotmail.com.

[3] Maestra en Educación, profesora-investigadora de la Universidad Autónoma de Aguascalientes, México. Email: mcarvaja@correo.uaa.mx.

[4] Maestra en Ciencias de la Educación. Email: lupi-89@hotmail.com.

DOI: 10.20396/zet.v28i0.8656854

Palabras clave: Formación inicial de profesores; Comprensión de gráficas y tablas; Sentido estadístico; Ideas estadísticas fundamentales.

Background

In recent decades, statistics has been incorporated as part of the basic educational content in many countries, due to the need to provide all citizens with a statistical culture that allows them to understand the large amount of information present in multiple forms of print and electronic media, on all kinds of topics and for different purposes (Batanero, 2002). This has as a natural consequence the implementation of strategies for training and upgrading the skill sets of teachers who have to meet this educational demand, as well as the development of multiple teaching proposals in order to achieve the best results. In the case of Mexico, it was the 1972 curriculum for primary education (ages 6 to 11), that introduced for the first time statistical issues such as data recording, the notion of frequencies and bar graphs, and basic elements of probability. In the subsequent curricular reforms to primary education, carried out in 1980, 1993, 2009, 2011 and 2017, the emphasis on statistics and probability have been maintained, and from the 1993 plan on, there has been more emphasis on the work of obtaining and analyzing statistical data and its graphical and tabular representations. In the 2019-2020 school year, there is a transition between the 2011 and 2017 curriculums, with the former being implemented from the third to sixth grade, and the latter in the first and second grade (SEP, 2019).

In the 2011 mathematics curriculum for primary education (SEP 2011), the thematic axis Information Management and the sub-axis Analysis and Representation of Data are included, with subject matter for development from the third grade up to the sixth grade of primary school. In this thematic axis various activities are contemplated, such as: representation and interpretation of tables or pictograms of quantitative or qualitative data collected in the environment; reading information contained in bar graphs; reading explicit or implicit information contained in different mediums and addressed to a particular audience, resolution of problems in which it is necessary to extract information from tables or bar graphs; identification and analysis of the utility of the most frequent datum in a set of data (mode); analysis of conventions for the construction of bar graphs and the calculation of the average (mean), and analysis of its relevance with respect to the mode. Finally, the reading of data contained in tables and pie charts is included, in order to answer various questions; the use of the three measures of central tendency, the mean, median and mode, in the resolution of problems. The 2017 curriculum (SEP, 2017) follows a similar logic to the 2011 curriculum, emphasizing the development of activities related to everyday tasks, focusing on data obtention and the construction and analysis of tables and graphs, as well as measures of central tendency. Although there are slight variations in the grade level at which some content from both curricula is presented, the expected learning curve does not vary considerably, the emphasis being on the construction and reading of simple tables and graphs, as well as a first approach to the measures of central tendency.

DOI: 10.20396/zet.v28i0.8656854

As for the training of primary school teachers in Mexico, this is done in the Normal Schools. There is a single curriculum for the training of primary school teachers throughout the country. At the start of the 2019-2020 school year, there is also a transition in the B.A. in Elementary Education, from the 2012 curriculum (SEP, 2012) to the 2018 curriculum (SEP, 2018). In Mexico, the previous studies required to obtain an undergraduate degree are nine years of basic education - primary and secondary - and three years of upper secondary education. In both curricula, courses for the training of mathematics teaching are included, focused on Arithmetic (one in the 2012 curriculum and two in the 2018 curriculum), Algebra, Geometry, and one more on Statistics and Probability (called Statistical Information Processing in the 2012 curriculum and Probability and Statistics in the 2018 curriculum). The purpose, focus and contents of these last two courses are very similar. Following is a brief description of the Statistical Information Processing course, from the 2012 curriculum, as it was the one taken by the students who participated in this research. The Statistical Information Processing course consists of four thematic units: 1) Descriptive Statistics (including topics such as frequency distribution tables and graphical representations; measures of central tendency; measures of position; measures of dispersion; study of populations with bivariate data); 2) Probability and Sampling; 3) Statistical Inference; 4) Linkage to the Information Management axis (including analysis of the concepts of the Information Management and Statistics axis in primary education, its importance and challenges; development of teaching strategies) (SEP, 2012 b). This course, apart from offering didactic training, also aims to ensure that the future teacher understands and applies the basic concepts and procedures of probability and of descriptive and inferential statistics, in the collection, organization, presentation and analysis of data for knowledge and problem solving in an educational context; in the same way, they are expected to apply this knowledge to the research they have to carry out in order to obtain their undergraduate degrees. Although the time lag between the 2011 primary school curriculum and the B.A. in Elementary Education curriculum implemented since 2012 is evident, in the case of statistical content there has not been much change in the last twenty-five years, so we consider the statistical training of primary school teachers to be consistent with the purposes of the educational level at which they will work.

Graphs and tables as synthetic resources for statistical dissemination

A basic proficiency required to achieve a statistical culture is the capacity to read and interpret graphs and statistical tables, as these formats are privileged resources when it comes to grouping and synthesizing large quantities of information in an efficient and visually attractive manner. They are not only widely used by the mass media, they are also an important part of the dissemination of official statistics and of investigative reports in a large number of fields of knowledge (Arteaga, Batanero, Cañadas & Contreras, 2011; Estrella, 2014). A statistically educated citizen should have the capacity to organize data, construct graphs and tables and work with different types of data representation, likewise, they should understand concepts, vocabulary and the symbols with which they are represented, including

DOI: 10.20396/zet.v28i0.8656854

an understanding of the notion of probability and the measurement of uncertainty, in short, they should possess statistical literacy (Ben-Zvi & Garfield, 2004, p. 7). In the professional field, a capacity to obtain, process and analyze information from official sources is also required, and users are confronted with resources and tasks that can be very complex (Gal & Murray, 2011), as in the case of teachers, who face different needs, from the preparation of records of monthly or semiannual grades and other indicators of student performance, to the handling of different statistics offered by official bodies, such as the Secretary of Public Education (SEP) and the National Institute of Statistics and Geography (INEGI).

However, various studies conducted with teachers of different education levels and teachers-in-training demonstrate that their performance is limited when facing tasks such as the construction and interpretation of graphs and tables, and lay bare a series of difficulties that show incorrect or underdeveloped understanding when it comes to the basic criteria for the preparation and reading of these kind of statistical records and of the statistical notions they represent (Batanero, Arteaga & Ruiz, 2009; Jacobbe & Horton, 2010; Juárez & Inzunsa, 2014; Estrella 2014; Estrella, Olfos & Mena-Lorca, 2015; Arteaga, Batanero, Contreras & Cañadas, 2015; Arteaga, Díaz-Levicoy & Batanero, 2018; Arredondo, García & López, 2019). Although there are common difficulties when it comes to understanding graphs or tables, there are certain characteristics unique to each type of record that should be analyzed separately.

The complexity of a graph lies in the totality of its elements, its nature and composition, which we can synthesize as follows (adapted from Kosslyn, 1985, cited by Arteaga et al., 2018):

- Structure: whether it has a geometric form (such as pie charts, the use of bars or rectangles) or consists of Cartesian axes.
- Pictorial content: for example, the rectangles of a histogram, points in a scatter graph, and dispersion, lines, etc.
- Labels: headers, footers, denomination of variables and the axes, range of values of the variables.
- Mathematical content: numerical sets used, the sections of a pie chart, etc.
- Statistical content: absolute frequencies, relative frequencies, variation, dispersion, central tendency, etc.

The difficulties in the reading and construction of statistical graphs have to do with not recognizing all or some of the elements they are composed of and the multiple meanings they contain (such as the nature of the data that is represented and the relationships that are established between them). Arteaga et al. (2015) y Arteaga et al. (2018) describe four levels of semiotic complexity that occur in the construction of statistical graphs:

DOI: 10.20396/zet.v28i0.8656854

L1: *Representation of individual data.* The students make graphs that represent only a single datum or a portion thereof, without taking into account the distribution of all measured values.

L2: *Representation of all individual data of one or several variables, without summarizing their distribution.* When students, for example, represent all the data of a set of values involving the heights of a group of people, but without taking into account their order, and therefore without having any clear notion of frequency or distribution.

L3: *Representing the distribution of each variable with different graphs.* At this level, students are capable of representing data from the calculation of frequencies, representing it in an orderly manner and taking into account its distribution, but without including more than one variable in their representation.

L4: *Representation of various distributions in the same graph.* Students who are at this level can represent the distribution of data from two or more variables, for example representing in a single graph the frequencies of the heights of a group of people, separating them by their gender.

These levels show us the gradual appropriation of the different elements, codes and conventions of representation in statistical graphs, but also the appropriation of notions that can be generated from a data set, such as frequency, variation, distribution, central tendency, etc. These levels, as we will discuss later, also explain the processes of reading and analyzing graphs and tables.

We can find similar conditions with respect to statistical tables:

> Despite the central role of tables in scientific practice, as well as their wide use in science classes and scientific texts, there is evidence that the interpretation of tables is not an easy task, and the acquisition of the skills required to interpret tables is not a transparent process (Estrella, 2014, p. 16).

This is explained in part by the wide variety of types of tables, which can range from simple lists of values corresponding to a single variable, to complex arrangements that show aggregate information drawn from multiple variables. Generically speaking, we can say that a table is:

> A rectangular arrangement with a structure comprising a set of rows and columns, allowing the presentation of data corresponding to one or more variables (characteristics of the phenomenon being studied) in an ordered and summary fashion, in order to allow visualization of data behavior and facilitate understanding of the information that can be extracted (Estrella, 2014, p. 6).

In a table, a row generally corresponds to a class of cases and a column corresponds to a variable. Just like graphs, tables come with a series of conventions for their construction, and although some elements or criteria may vary, the most common are as follows (Estrella, 2014): title, body of data (a rectangular block consisting of rows, columns and cells that

DOI: 10.20396/zet.v28i0.8656854

contain numerical information), side header or first column, top header (named for the content of the columns), totals (in the last row and/or last column). Other elements of the format of a table are also important: the arrangement or spatial location of the elements; groupings (types of data ordering following a certain criteria); the number of figures of the numbers represented, the type of statistical measures (absolute or relative frequencies, central tendency, etc.), nonverbal and non-numeric codes that facilitate the display of table elements, such as: shape, size, symbolism, colour, shading, use of bold, use of lines and spaces, different fonts, icons (Estrella, 2014). This author points out that, just like graphs, tables represent a series of difficulties for students, as much to do with their construction as their reading, and that said difficulties have to do with cognitive and graphic processes, or rather, semiotics. Also at play in these processes of comprehension are other processes that not all students come to master, such as data registration and organization, categorization, classification and combinatorics. However, unlike graphs, tables require a greater numerical mastery.

It must be taken into account that graphs and statistical tables are at once cultural objects and complex semiotic objects (Arteaga, et al., 2011), constrained by a series of conventions and technical criteria that define their elements and structure (the signifier), with the purpose of communicating certain messages (the signified). The uses required by different groups of professionals condition the characteristics of the representational systems, in this case graphs and tables as cultural objects. Thus, for example, the needs of schools are of a different nature to those required by educational planning and evaluation systems at the state or national level, or to the statistics of the health sector, although they have certain basic principles in common (Gal & Murray, 2011). On the other hand, the complex nature of these semiotic objects is that in a graph or table there is no biunivocal relationship between the signifier and the signified, and the elements that compose them have multiple meanings, that combined are amplified and enhanced (Batanero, et al., 2009).

This research aims to describe the knowledge that primary school teachers-in-training exercise when interpreting graphs and statistical tables, taking a special interest in their comprehension and difficulties. From this starting point it is expected in the future to promote curricular and didactic actions that will be useful for the initial training of teachers of basic education.

Methodology

The study consisted of an exploratory-descriptive research to identify the comprehension and difficulties that primary school teachers-in-training exhibit when confronted with statistical analysis tasks presented by means of a graph and a frequency table. Two complementary methodological approaches were carried out: the first was the application of a written test; secondly, task-focused interviews were conducted. The knowledge test was applied to all students of the B.A. in Elementary Education in a teacher training school in Aguascalientes, Mexico. In total 240 students participated: 74 from the

DOI: 10.20396/zet.v28i0.8656854

first semester; 71 from the second semester; 54 from the fifth semester; 41 from the seventh semester.

The interview went deeper into the comprehension and difficulties of a sample of nine students. For these interviews two types of students were chosen: those who demonstrated a greater understanding of the task performed during the written test, and those who experienced more difficulties while completing the tasks. Three students were selected from each semester (third, fifth and seventh) without considering those from the first semester, as they had not yet had enough contact with the training activities of the curriculum. In order to conduct the interviews, the same questions from the written test were taken as a starting point, but with the difference that the interviewers could suggest some additional activities, and the students could explain in more detail how they had arrived at their answers. The interviews were carried out in the facilities of the institution, in provisional spaces where there were no interruptions (classrooms or cubicles of the tutors). The interviews were carried out by three members of the research team, and they were recorded and subsequently transcribed for analysis. For the purpose of triangulation, the categorization of the responses from each student was carried out by at least two of the participating researchers.

Tools

Before designing the knowledge test, an analysis of the expected proficiencies in the 2012 Curriculum of the B.A. in Elementary Education (SEP, 2012 a) proceeded. With this analysis, the importance for future teachers of knowledge and comprehension of the statistical information present in graphs and tables was detected, along with the type of representations and most pertinent information to face the teaching and school planning and didactic tasks. A questionnaire made up of two tasks was designed, the first consisting of a series of questions based on a line chart, and the second consisting of questions based on a frequency table. Both the chart and the table were obtained from an official educational statistics report from the state of Aguascalientes, prepared by the government agency responsible for the provision of basic education (IEA, 2012). The answers were open, and for their analysis categories were made and three people reviewed and ranked the responses. Prior to the application of the questionnaire to teachers-in-training, a pilot of the tool was carried out with pedagogy students from a public university in the state of Aguascalientes. Based on the pilot, some corrections were made to the tool.

The statistical data graph. The graph includes three variables: new students (continuous variable: number of students enrolled); school level (dichotomous variable: primary and secondary) and school year (variable with 9 categories: from the 2004-2005 school year to the 2012-2013 school year). (See Figure 1).

DOI: 10.20396/zet.v28i0.8656854

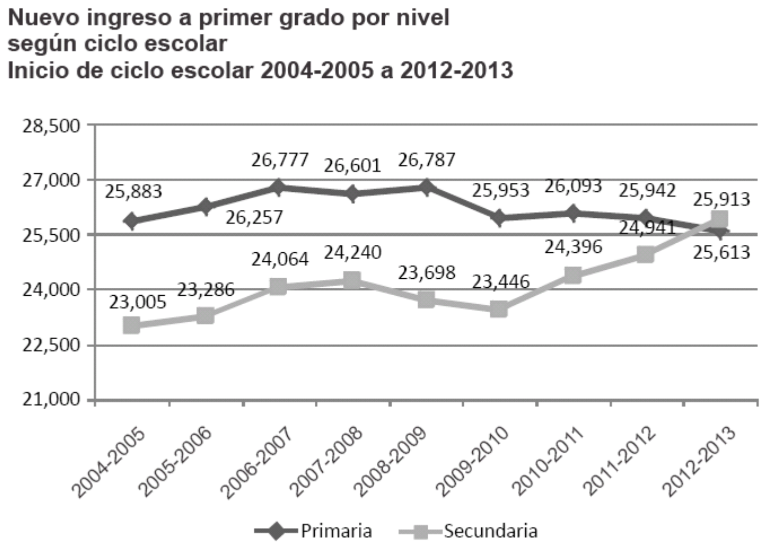

Figure 1. Graph presented in the test and in the interview.
Source: IEA (2012) *The education figures. Start of the 2012-2013 school year. Basic education statistics.*

The line graph represents the evolution of enrollment in primary and secondary school in the state of Aguascalientes, Mexico, during nine school years. The lines explicitly show the ups and downs in enrollment at both educational levels, and the corresponding figures for each year allow an accurate assessment of said evolution. The five questions that were presented to the students were aimed at investigating their recognition of the information presented in the graph, of the variables involved, of their recognition of the trends of the enrollment variable over the years, for both primary and secondary school, as well as an analysis of the enrollment variable year by year (implying a comparison between educational levels).

The frequency table. As for the other task, a table of basic education statistics from the state of Aguascalientes was adapted (IEA, 2012), in which the data corresponding to the initial enrollment of primary school students in the 2012-2013 school year is presented. In this table, composed of 8 columns and 14 lines, three variables are included: school grade (from 1st to 6th), age of the students (from 5 to 15 years), and initial student enrollment (number of students enrolled). One aspect that makes the table more complex is the variation that the age variable shows with respect to the school grade, the modal value corresponding to the appropriate age for each grade level. (See Figure 2).

Initial enrollment of students in basic primary education
Start of 2012-2013 school year

Primary							
Age	**1°**	**2°**	**3°**	**4°**	**5°**	**6°**	**Total**
5	8,106	1					8,107

DOI: 10.20396/zet.v28i0.8656854

6	17,512	7,654					25,166
7	946	17,153	7,117				25,216
8	160	1,530	17,169	7,185			26,044
9	49	285	1,791	16,534	7,122		25,781
10	20	89	437	1,948	17,229	6,881	26,604
11	7	29	97	457	2,161	16,763	19,514
12	2	9	34	120	538	2,328	3,031
13	1	2	12	19	99	521	654
14		1	1	3	21	98	124
15				2	1	12	15
Total	26,803	26,753	26,658	26,268	27,171	26,603	160,265

Figure 2. Table presented in the test and in the interview.

Source: IEA (2012) The education figures. Start of the 2012-2013 school year. Basic education statistics.

For this table, three questions were presented, two aimed at conducting a comparative analysis between the enrollment of the different school grades (analysis by column to identify the school grades with higher and lower enrollment), and one intended to identify which school grade contained the greatest number of overage students (with a greater age than the ideal for each grade), which implies an analysis of the trends of the age variable in each school grade, and a comparison between the different grades.

Results

Responses to the written test

While about half of the students who completed the test were able to account for the information contained in the graph and the table, a better performance was expected, considering that the construction and reading of these types of representations is taught from primary school on and is part of the training of future teachers. In general, we find trends in the responses similar to the results reported by other authors (Batanero, et al., 2009; Jacobbe & Horton, 2010; Arteaga et al., 2015; Estrella et al., 2015; Arredondo et al., 2019).

***Responses to the statistical graph*.** The first question sought to identify the information contained in the graph (What information does the graph present?). We have that 55.4% of the respondents managed to identify the elements of the graph (title, educational phenomenon described, outstanding features), while 35% only partially identified the content of the table. A small portion (9.2%) didn’t recognize the meaning of the graph, even though it came with its title and its elements are explicit. Only one person did not answer this question.

Upon being asked to identify the variables described in the graph, only 12.1% of respondents recognized the three variables involved (new students; school level; school year), which implies a recognition of the values of the phenomena that are represented in the graph. 32.9% recognized two variables and 26.3% only mentioned one variable. 6.3% mentioned another phenomenon related to the theme of the graph, but without mentioning any of the

Zetetiké
DOI: 10.20396/zet.v28i0.8656854

variables involved. 17.9% mentioned the differences or variations between the data, but without explicitly mentioning the variables. 4.6% did not answer.

Before the prompt asking for their comments on the primary level, 56.7% were able to correctly describe the trends of new enrollment in primary school during the period considered, apparently relying on the line that represents this trend (See Figure 1), as could be corroborated in the interviews, which implies a visual analysis of the increases and decreases in the enrollment of new students over time, but without necessarily involving a fine numerical comparison (which only 5.4% of respondents accomplished). It should be noted that there were some students who incorrectly described the trend, although they were few (5.8%). Interestingly, 27.9% of the students mentioned other things associated with the phenomenon represented in the table, but which didn't have anything to do with the variables involved in the graph, as if the subject of the graph were a trigger causing them to mention associated situations, such as population growth, the construction of a greater number of schools in the area, etc. 4.2% did not answer this question.

In a similar manner to what is identified in the description of the trends of new primary school enrollment, in the case of secondary school enrollment, the response rates of students were very similar, predominately a correct description but without detailed analysis of the corresponding figures (59.6%), that is, only alluding to the variations that can be seen directly in the line that represents the values of new students, making statements such as "the number of students increases", "enrollment goes up over the years". A very small portion, 3.8%, made a detailed description supported by numerical variations. The percentage of students who made an incorrect description (5.4%) is also similar, as with those who mentioned something that had nothing to do with the variables in the graph (27.5%). 3.8% did not answer.

Analyzing the enrollment trends of new entries to primary or secondary school, separately, and supported by a graphic representation (a line where it is relatively easy to identify ups and downs) can be considered a level 3 task, according to the typology of Arteaga et al. (2015). Comparing the enrollment data of new entries to primary and secondary school over time (fifth task), implies comparing two distributions (level 4). Comparing the evolution of the figures from both educational levels over time, numerically identifying the variations year after year and between levels, was a task that few students could accomplish (9.2%). 22.9% made a comparison of the trends between the two levels, but without making reference to the quantities, which implies having a correct notion of the meaning of the graph, but staying at a very intuitive or visual level. Although 38.3% could make some comment on the trends shown by the graph, they could not really make a comparison between the data sets. For 22.1%, the theme of the graph led them to comment on other related themes, but not ones linked directly to the information in the graph. 7.5%, amounting to 18 students, did not answer.

We can suppose that the majority of the students surveyed had no experience reading statistical graphs (at least the type of graphs presented to them), which surely implies that this

DOI: 10.20396/zet.v28i0.8656854

was an infrequently covered topic in their studies during basic and secondary education, and even at a higher level, despite being a topic included in the curricula.

***Responses to the statistical table*.** As noted by Estrella (2014) and Estrella et al. (2015), reading a frequency table may be more difficult for students than reading a line graph or bar graph, because it involves making a numerical comparison without any kind of visual aid, even more so if the table contains a considerable amount of information. This is the case with the table used in this test, since the number of students who gave an adequate response was lower, even in the first two tasks, which only involved making a comparison between the figures in the line of totals, but here the key issue was to know precisely in what part of the table to find the required information.

Only 15.8% of respondents could identify the grade in which there was the greatest increase in enrollment (the fifth grade), and give an argument based on the values in the table to describe or explain this increase. Another 27.5% gave a correct response, but did not give a consistent justification, or gave no justification at all. It should be noted that a little more than half (52.9%), answered without considering all the figures in the table, pointing out some other grade as the one in which there was the largest increase in enrollment (perhaps only making a comparison between two school grades). 3.8% did not respond. These responses make us assume that the majority of the students were capable of identifying individual data, but without recognizing their distribution (which is equivalent to level 1 in the typology of Arteaga et al., 2015).

A similar situation occurred in the case of the task in which respondents were asked to identify the grade in which there was the greatest decrease in students. This task required identifying the greatest difference between grades (which implies subtracting the totals of the grades, two by two). Between the 1st and 2nd grades, 2nd and 3rd, 3rd and 4th, and 5th and 6th, a decrease in enrollment is clearly visible, but between which grades is there the greatest decrease? Only 10.4% of respondents gave a correct answer (between 5th and 6th grade). 15.4% noted that the greatest decrease occurred in the 6th grade, but without justifying their response using the data from the table, and 9.6% without giving any justification. In this task as well, a little more half (55.8%), gave a response supported perhaps by a hasty analysis of the table (one that was for that reason incorrect) or by some assumption based on experience, but not on the information provided by the table. 8.8% of students did not answer this question, the highest percentage throughout the test.

To answer the third and last question related to the table of statistical data, consisting of identifying the school grade in which there was the greatest number of overage students (older than the expected age for that grade), it would first have been necessary to identify the ideal age corresponding to each grade, and then from this to identify the students exceeding that age. To answer this question, two criteria could be taken. The first criterion is pragmatic; considering that children enter primary school at 6 years of age, the ideal age for each grade would be 7 for the second grade, 8 for the third, and so on successively until arriving at an age of 11 years for the sixth grade. The other criterion is of a statistical nature, and is

DOI: 10.20396/zet.v28i0.8656854

reflected in the mode of each grade (corresponding to 6 years in the first grade, 7 in the second, and so on successively). If we take either of the two previous criteria as a reference point, it is in the 6th grade in which the greatest number of overage students appear. A little more than half of respondents noted that the overage phenomenon was greatest in the 6th grade (53.3%). 3.8% of the students surveyed said that this phenomenon occurred in the 5th grade, giving numerical support to their response, and 14.6% gave the same response, but without giving a reason (the 5th grade takes second place in respect to this occurrence, which could have led them to favour this response). 24.2% mentioned some other grade in their response, and 4.2% did not answer.

As we will see below, in the interviews the students gave similar answers to those they gave during the test, but also showing us other factors that allow us to appreciate their comprehension and difficulties.

Analysis of the interviews

In the interviews, the selected students were first presented with the graph, and then with the table. They were asked to review each of these representations, making a general assessment of them based on the question: what is the graph or table about? To continue the interview, the questions from the written test were resumed, with the possibility of asking additional questions to expand upon the responses of the students, this being at the discretion of each interviewer.

The analysis of the students' responses when carrying out the task of reading and interpreting the graph and the table led us to consider two parallel realities: a) the recognition and understanding of the elements of the graph and the table; b) the recognition and understanding of the variables involved and their nature and their referents as displayed in the parameters of the graph and the table, as well as the fundamental statistical ideas of variation and distribution underlying the values of each variable. The assessment of the graph and the table is conditioned in part by the students' understanding of the variability of the statistical data and the conceptual resources they may have to explain it (such as the notions of frequency, minimum and maximum values, range, mode). But we also have that the statistical concepts, being mediated by the forms of representation, are conditioned by the resources that make them explicit (the range of values in each of the axes, the shape of the lines, in the case of the graph; or in the minimum and maximum values, the frequencies and the mode, displayed in each column, in the case of the table).

We appreciate that the students interviewed displayed different levels of comprehension. In some cases, their limited knowledge of the characteristics of the graph and the table and the conventions of their construction was clear. But in other cases, a static idea of the phenomena represented is also evident, a kind of negation of the entire range of possible values, even when they are represented in the graph and table. This inability to see the multiple values of the variables (especially quantitative variables), makes us assume a limited understanding of the notion of variation, combined with an ignorance of the criteria

DOI: 10.20396/zet.v28i0.8656854

and conventions of construction of this type of representation. This is what Arteaga et al. (2015) identified as representing (or recognizing) only individual results, characteristic of level 1 in his typology. On the other hand, students who experienced less difficulties due to their greater comprehension, in addition to knowledge of all or at least most of the constituent elements of the graph and the table, also demonstrated a greater understanding of the notions of variation and distribution, since in their explanations they made reference to all the values of each variable and even their trends. What developed first in these students: an understanding of the characteristics of the graph and the table, or the notions of variation and distribution? Because of the type of answers, they provided, we suspect a dynamic relationship between the signifier and the signified, which over time and with the support of different experiences, has led them to their current understanding.

How was this manifested in the responses of the students? First, we were struck by the fact that the lowest performing students began their analysis by referring to an outstanding aspect, or rather, an element that seemed outstanding to them, either a visual element (for example the top line of a graph, which is also the one that stands out for its colour) or a numeric one (for example, the highest amount in the graph or the table, or a number located in a highly visible place, such as the row or column of totals). This could be a strategy to start the analysis of a representation, however, in the case of the lowest performing students, the greater part of their analysis was limited to this outstanding element, ignoring or giving less importance to the rest of the elements of the graph or table, and therefore resulting in an incomplete or biased reading. Only paying attention to an outstanding element and not looking any further (or doing so in a very limited manner) entails other kinds of conceptions, some that became explicit during the interviews, others more implicit, for example: giving more importance to the pictograph than to the numbers; analyzing only isolated data, without recognizing the importance of the data sets represented, and, therefore, their properties, standing out among them variation and composition as a specific distribution; an incorrect or partial understanding of the statistical notions involved, reflected by the use of colloquial terms when the use of specialized terminology would be more appropriate (Park, Park, Lee y Lee, 2016, demonstrate how difficult it can be to achieve a comprehensive mastery of statistical terminology); and finally, making reference to phenomena related to the subject of the graph and the table, but not directly represented in them, or, similarly, making an idiosyncratic reading as described by Arredondo et al. (2019), consisting of giving opinions based on experience or personal perspectives, without really reading into the values and their trends.

By contrast, those who performed the best demonstrated a more complete understanding of the table: they recognized all the variables represented in the graph and the table (naming them correctly according to the headings of each representation), the range of values of each variable and their main trends (either of each variable separately or of more than one together); although they made use of pictographic resources, their analysis was also based on numerical values, they tried to read most of the numbers and elements of the graph and the table, which permitted them to identify the variation in the data and make a

DOI: 10.20396/zet.v28i0.8656854

description of the distribution, even if it was approximate or intuitive; they used less colloquial expressions and demonstrated a greater mastery of specialized terminology and its meanings; they tried to complete their analysis with elements of the graph and the table without resorting to external factors, and if they did, they were phenomena directly related to the variables represented.

We identified three students with the lowest performance, who could be located in level 1 of the Arteaga et al. (2015) typology, and two that corresponded to level 4. The remaining five can be placed between levels 2 and 3 (See Table 1). It is worth clarifying that this classification is approximate, because some students demonstrated a level of knowledge that would place them in intermediate positions between the levels, and some achieved a better performance in the graph than in the table.

Table 1. Classification of interviewed students based on their level of performance.

Semester	Students (with pseudonym)		
3rd	Carmen. L2	Irma. L1	Graciela. L1
5th	Gerardo. L3	Eduardo. L2	Soledad. L2
7th	Arturo. L4	Eréndira. L3 (L4)	Pilar. L1

Source:own preparation.

The fact that the majority of the students located in L1 are from the 3rd semester stands out, and that the students with the highest level of comprehension are from the most advanced semesters, although there are exceptions, which shows us that progress toward a complete understanding of graphs and tables and the statistical concepts involved may require more time than is allocated to the subject of Statistical Information Processing, and that the development of this understanding is fueled by additional activities, possibly developed in other courses, such as those related to educational evaluation and planning.

Conclusions

It is essential for future primary school teachers to have the capacity to read and analyze graphs and statistical tables. This proficiency will guarantee an adequate performance as they guide the learning of their students, and will enable them to obtain, understand and use different sources of official statistical information on which to base their school planning activities.

However, the limitations in their comprehension of the graph and the table and the difficulties identified in at least half of the students show us that the Statistical Information Processing course has fallen short of fulfilling its objective to at least offer statistical literacy. Likewise, it reveals that, in other subjects, such as Basic Tools for Educational Research and Assessment for Learning, apparently limited use is made of statistics. We believe that the curriculum of the B.A. in Elementary Education establishes a favorable framework for the

DOI: 10.20396/zet.v28i0.8656854

transversal development of competence in statistical analysis, but for it to be achieved, teamwork on the part of the teachers and a revaluation of the usual practices are required.

It should be considered, as well, that the academic background of teachers-in-training is diverse, as are their cultural and economic conditions and possibilities, and that not all of them, throughout their basic training, have had the opportunity to face situations where statistical information is present, such as print media, research reports and Internet access. In these cases, special support will be required to help correct these possible deficiencies. The development of the proficiency required to capably interpret graphs and statistical tables requires lots of practice and familiarity with different types of representation, and this is precisely the deficiency reported by various authors (Batanero et al., 2009; Jacobbe & Horton, 2010; Estrella et al., 2015), therefore, it is necessary to open the largest possible number of analysis spaces for these types of representations in the curriculum, especially for those students with more deficiencies and difficulties.

Insufficient attention paid to the statistical training of primary school teachers may have various implications, the most direct and the one with the greatest long-term effects having to do with their teaching activity in the classroom. The promotion of a statistical culture for all children must begin with training teachers and updating the skill sets of teachers already in service. This will be another challenge involved with the process of curricular transition that is currently taking place in the Mexican school system, and one of those foreseen in the immediate future.

Acknowledgments:

This research was conducted with funds from the Educational Research Program of the Autonomous University of Aguascalientes, Mexico.

References:

Arredondo, E.H., García, J.I. & López, C. (2019). Niveles de lectura de estudiantes de licenciatura: el caso de una tabla y una gráfica de líneas. *Matemática, Educación e Internet*, 19 (2)

Arteaga, P., Batanero, C., Cañadas, G. & Contreras, M. (2011). Las Tablas y Gráficas como Objetos Culturales. *Números. Revista de Didáctica de las Matemáticas*, 76, 55-67

Arteaga, P., Batanero, C., Contreras, M. & Cañadas, G. (2015). Statistical Graphs complexity and reading levels: a study with prospective teachers. *Statistique et Enseignement*, 6 (1), 3-23

Arteaga, P., Díaz-Levicoy, D. & Batanero, C. (2018). Investigaciones sobre gráficos estadísticos en Educación Primaria: revisión de la literatura. *Matemática, Educación e Internet*, 18 (1)

DOI: 10.20396/zet.v28i0.8656854

Avila, A. (1998). *La ensenañza oficial de las matemáticas elementales en México; su psicopedagogía y transformación (1944-1986)*. México: Universidad Pedagógica Nacional.

Batanero, C. (2002) Los retos de la cultura estadística, *Jornadas Interamericanas de Enseñanza de la Estadística*, Buenos Aires. Conferencia inaugural.

Batanero, C., Arteaga, P. & Ruiz, B. (2009). 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 la Ciencia*, 28 (1), 141-154

Ben-Zvi, D. & Garfield, J. (2004) Statistical Literacy, Reasoning, and Thinking: Goals, Definitions, and Challenges. In Ben-Zvi, D., Garfield, J. (Eds.) *The Challenge of Developing Statistical Literacy, Reasoning and Thinking* (pp. 3-15). Netherlands: Kluwer Academic Publishers

Estrella, S. (2014). El formato tabular: una revisión de literatura. *Revista Actualidades Investigativas en Educación*, 14 (2)

Estrella, S., Olfos, R. & Mena-Lorca, A. (2015). El conocimiento pedagógico del contenido de estadística en profesores de primaria. *Educação e Pesquisa*, 41 (2), 477-493

Gal, I. & Murray, S.T. (2011). Responding to diversity in users' statistical literacy and information needs: Institutional and educational implications. *Statistical Journal of the IAOS*, 27, 185-195

Instituto de Educación de Aguascalientes (IEA) (2012). *Las cifras de la educación. Inicio de ciclo 2012-2013. Estadísticas de educación básica*. México: Dirección de Planeación del Instituto de Educación de Aguascalientes

Jacobbe, T. & Horton, R.M. (2010). Elementary School Teachers' Comprehension of Data Display. *Statistics Education Research Journal,* 9 (1), 27-45

Juárez, J.A. & Inzunsa, S. (2014). Comprensión y razonamiento de profesores de Matemáticas de bachillerato sobre conceptos estadísticos básicos. *Perfiles Educativos*, 36 (146), 14-29

Park, M.S., Park, M., Lee E.J. & Lee, K.H. (2016). Preservice Teachers' Difficulties with Statistical Writing. In Ben-Zvi, G. y Makar, K. (Eds). *The Teaching and Learning of Statistics. International Perspectives* (pp. 261-270). Switzerland: Springer.

Secretaría de Educación Pública (SEP) (2011). *Plan de estudios 2011. Educación Básica.* México: Dirección General de Desarrollo Curricular, Secretaría de Educación Pública.

Secretaría de Educación Pública (SEP) (2012 a). *Acuerdo número 649 por el que se establece el Plan de Estudios para la formación de Maestros de Educación Primaria*. México: Diario Oficial de la Federación, del 20 de agosto de 2012.

DOI: 10.20396/zet.v28i0.8656854

Secretaría de Educación Pública (SEP) (2012 b). Programa del curso: Procesamiento de Información Estadística. México: Secretaría de Educación Pública.

Secretaría de Educación Pública (SEP) (2017). *Aprendizajes Claves para la Educación Integral. Plan y programas de estudio para la educación básica*. México: Secretaría de Educación Pública.

Secretaría de Educación Pública (SEP) (2018). *Acuerdo número 14/07/18 por el que se establecen los planes y programas de estudio de las licenciaturas para la formación de maestros de educación básica*. México: Diario Oficial de la Federación, del 24 de julio de 2018.

Secretaría de Educación Pública (SEP) (2019). *Hacia una Nueva Escuela Mexicana. Taller de Capacitación. Educación Básica. Ciclo Escolar 2019-2010*. México: Secretaría de Educación Pública.