The analysis of statistical graphs constructed by primary school teachers

Maria Niedja Martins [a,b]
Carolina Carvalho [a]
Carlos Monteiro [b]

[a] Universidade de Lisboa, Instituto de Educação, Lisboa, Portugal
[b] Universidade Federal de Pernambuco, Departamento de Psicologia e Orientação Educacionais, Programa de Pós-Graduação em Educação Matemática e Tecnológica, Recife, Pernambuco, Brazil.

ABSTRACT

Background: Statistical graphs are widely used in society and are important in presenting data related to different topics. Statistics education aims improve the teaching and learning of statistics, which include the approaches to statistics graphs. **Objectives**: This article analyses the construction of statistical graphs by in-service teachers and their reasons for choosing these graphs. **Design**: It was conducted an empirical study from a qualitative perspective which investigated teachers' choices and constructions of statistical graphs. **Setting and Participants**: The participants were 22 primary school teachers from public schools located in Recife Metropolitan Area, Brazil. They expressed their intention to be interviewed when they responded to a questionnaire applied among a larger number of teachers. **Data collection and analysis**: The interviews were composed of two tasks comprised of statistical data in which participants had to choose a type of graph, justify the choice, and construct a graph of a chosen type. It was carried out a content analysis of speech protocols. **Results**: The reasons for participants' choices seemed to be related to their familiarity with the type of graph, and they constructed graphs with high levels of complexity. **Conclusions**: The results of this study emphasised the importance of primary school teachers developing specific knowledge about graphs through pre-service and in-service elementary school teacher education. The results from this research study offer new questions concerning the construction and the choice for statistical graphs, which include the influence of new technologies and the use of textbooks.

Keywords: statistics education; mathematics education; teacher education; primary school curriculum; statistical graphs.

_____

Corresponding author: Carlos Eduardo Ferreira Monteiro. Email:
carlos.fmonteiro@ufpe.br

Análises de gráficos estatísticos construídos por professores dos anos iniciais

RESUMO

Contexto: Os gráficos estatísticos são amplamente utilizados na sociedade e são importantes na apresentação de dados relacionados a diversos temas. A educação estatística visa melhorar o ensino e a aprendizagem da estatística, o que inclui as abordagens aos gráficos estatísticos. **Objetivos**: Este artigo analisa a construção de gráficos estatísticos por professores do anos iniciais do ensino fundamental e as razões para a escolha desses gráficos. **Design**: Com o objetivo de discutir o problema de pesquisa, foi realizado um estudo empírico a partir de uma perspectiva qualitativa que investigou as escolhas de tipos de gráficos por professores e a construção de gráficos estatísticos. **Ambiente e participantes**: Os participantes foram 22 professores do ensino fundamental de escolas públicas localizadas na Região Metropolitana de Recife, Brasil. Eles expressaram a intenção de serem entrevistados quando responderam a um questionário aplicado a um maior número de professores. **Coleta e análise de dados**: As entrevistas foram compostas por duas tarefas compostas por dados estatísticos em que os participantes tinham que escolher um tipo de gráfico, justificar a escolha e construir um gráfico do tipo escolhido. Foi realizada uma análise de conteúdo de protocolos das entrevistas. **Resultados**: Os resultados deste estudo enfatizaram a importância de professores do ensino fundamental desenvolverem conhecimentos específicos sobre gráficos por meio da formação inicial e contínua de professores do ensino fundamental. Os resultados desta pesquisa oferecem novas questões relativas à construção e escolha de gráficos estatísticos, que incluem a influência das novas tecnologias e o uso de livros didáticos.
Palavras-chave: educação estatística; educação matemática; formação de professores; currículo dos anos iniciais; gráficos estatísticos.

INTRODUCTION

This article aims to discuss aspects of a study that analysed the construction of statistical graphs by in-service teachers and their reasons for choosing these graphs. It is expected to contribute to understanding the difficulties experienced by teachers while they are teaching about graphing.

Statistical graphs are widely disseminated in society and play an important role in presenting data related to different topics. This type of statistical representation can be considered a cultural artefact given its level of insertion and diffusion in our society (Vygotsky, 1978). Batanero et al. (2009) explain that the ability to read and understand graphs and tables is essential for literate citizens. Several countries have included graphing as curriculum content in the early years of primary school due to its importance in enabling the interpretation of statistical data. González et al. (2011) carried out a

literature review on this topic which indicated that most studies approached pre-service teachers and concluded that they have low levels of statistical knowledge, including limited ability in graphing. Teacher training in statistics represented a challenge in order to support educators' teaching of statistical graphs.

In Brazil, statistics topics are part of the mathematics curriculum for early years education (Ministério da Educação, 2018). The thematic unit called probability and statistics consists of prescriptions on the use of graphs in statistical research conducted by teachers. Therefore, it is important to develop teachers' competences to involve their students in an entire statistical research process, which includes the choice and construction of graphs to communicate statistical data. Statistics teacher education for primary school teachers does not seem to provide the necessary knowledge to teach statistics in the initial years.

Angra and Gardner (2018) argue that "the multifaceted and complex nature of graphing makes it difficult for instructors to diagnose student difficulties and for students to master the skill of graphing. There have been scattered efforts to identify and address student difficulties with graphing" (p.2). Angra & Gardner (2016, 2017) emphasise that the competence to choose and construct statistical data is much more complex than the ability to interpret it.

In general, these artefacts have a highly semiotic complexity, and it is common for students and teachers to show difficulties in properly choosing, building and interpreting representations. Other difficulties experienced by teachers associated with the construction of graph scale have also been listed in previous studies (Arteaga et al. 2015; Souza & Monteiro, 2020).

Arteaga (2011) classified the complexity of graphs produced by Spanish pre-service teachers and found different degrees of difficulty presented by these participants in constructing graphs. The majority of studies involving Brazilian teachers are only focused on the interpretation of graph skills. There are only a few studies that have investigated in-service teachers' skills in choosing and constructing graphs.

Investigating the competence in building and choosing statistical graphs among teachers who work in the early years is relevant because these professionals are responsible for teaching statistics to children in the elementary years. Understanding graphs has been considered an important part of statistical training for all citizens, which implies the acquisition of knowledge about these representations throughout basic education. That is, for

students to be able to appropriate statistical reasoning to act in the world, a solid knowledge base is necessary, including those related to statistical graphs.

For this to happen at the beginning of basic education, however, it is crucial to strengthen pedagogical knowledge and content on statistical graphs among teachers. It is through knowledge focused on these artefacts that teachers will be able to choose tasks and representations that are more appropriate to the statistical research they conduct in the classroom with their students. This does not claim that just studying these themes is sufficient, but it undoubtedly allows the teacher to develop important skills for conducting their work to guarantee students' learning from an early age. It would not be appropriate for us to wait for the student to reach the end of basic education to begin essential content for the development of his worldview.

The difficulties in teacher training found in Brazilian contexts lead us to consider that teachers in initial training when acting as teachers cannot organically overcome the challenges found in the classroom, especially those related to the teaching of Statistics. According to Lopes at al. (2020) one of the problems that need urgent attention refers to the initial and continuous training of teachers who teach statistics and the difficulty of teachers in working with graphic representations already mentioned by numerous studies (Arteaga et al., 2015; Batanero et al., 2009; Souza & Monteiro, 2020)

Therefore, this article presents aspects taken from an empirical study which investigated in-service teachers' choices and constructions of statistical graphs. The research study was developed with 22 teachers who responded to tasks prompting them to choose a type of graph that would represent specific statistical data, justify their choices and construct a graph of a chosen type.

LITERATURE REVIEW

The statistical graph and its semiotic complexity

Statistical literacy has been recognised as people's ability to critically interpret statistical data in a variety of contexts and communicate this understanding in ways that can impact their decision making (Gal, 2002). Statistical literacy is also one of the main concerns of statistical education, a research field which investigates epistemological, methodological, and didactical aspects of the learning and teaching of statistics.

Statistical graphs are different from mathematical graphs because they model non-deterministic functions. Cazorla (2002) emphasises that this

distinction is important in the teaching of statistical graphs because these representations are more suitable to support the presentation of data organised on a frequent perspective and based on transdisciplinary themes. Ainley (2000) argues that the frequent use of graphs in our society is supported by the idea that graphs are transparent when communicating their meanings. Also, mathematics and statistics curricula around the world still maintain conceptions of statistical graphs based on their neutrality. However, this objectivity perspective about graphs can mislead people to think that these representations cannot be unambiguous. On the other hand, if statistical education contributes to developing a wider perspective of statistical literacy (Gal, 2002), citizens would be able to identify and criticise intentional manipulations of statistical data presented in different media (Monteiro & Ainley, 2007, Cazorla & Castro, 2008). Therefore, it is becoming increasingly important to enable people to critically interpret inconsistent statistical information displayed on graphs.

Ainley (2000) explains that the types of graphs chosen to compose basic school curricula are based mainly on evidence about students' difficulties to construct graphs based on the utilisation of pencil and paper. Therefore, bar graphs, line graphs and pie charts are still the most common types of graphs in school classrooms (Carvalho, 2009; Arteaga, 2011) and in textbooks (Guimarães et al., 2007). On the other hand, important types of graphs could be included, such as scatter graphs, if there was a consideration of computer-based activities (Ainley at al., 2001; Martins et al., 2017).

Font et al. (2007) argue that in mathematical practices related to solving problems, there are multiple semiotic functions of reading or representation due to the need to use and operate with mathematical objects which are immaterial. These authors consider a typology of mathematical objects (verbal or symbolic expressions, properties, procedures, problems, arguments, and concepts) which intervene in mathematical practices, each of which can play the role of antecedent or consequent of a semiotic function. From these ideas, they defined the following levels of complexity:

- o **Level 1:** Graphs based on less complex concepts and procedures. It is a representation with values related to the variable of a particular case. This type of graph only allows the reading of the data (Curcio, 1989).

- o **Level 2:** It is possible to understand the structure of the data presented. However, although the graph provides information to answer questions at the level of reading data, it is not possible to go beyond that, or to reach the level of identifying trends.

- o **Level 3:** Graph presents distributions for two variables separately in two graphs. Otherwise, each pair of variables is represented by two graphs. This strategy generally makes it difficult to compare variables, especially when the person constructs very different representations. It is considered a representation superior to level 2 because the person can identify trends.

- o **Level 4:** The person can produce a graph that aggregates two distributions of variables. This type of graph is more complex and therefore allows more sophisticated types of reading, considering the analysis of structure, the comparison of trends between two variables and their variability on the same graph.

These levels of complexity were based on empirical data and analysis of different types of interpretation of graphs. Analyses of these levels can be used as a tool to teacher education to promote to statistics educators their knowledge and attitudes associated with this type of statistical data represented in context. Arteaga (2011) articulated the types of reading pointed out by Curcio (1989) and the reading levels of graphs to elaborate his classification on the construction of graphs by pre-service teachers. Arteaga's approach is based on an ontosemiotic perspective connected with different theoretical perspectives about mathematical knowledge and the teaching and learning of mathematics (Godino, 2017; Godino et al., 2007). The perspective developed by Arteaga (2011) emphasises the need for teacher education to explore specific contents related to the learning of statistics. Therefore, teachers need to produce and expose their own ideas and skills regarding statistics. Borelli et al. (2016) state that teachers' experiential practice in these teacher education environments can lead them to a closer relationship with specific knowledge because this knowledge would be used in their practices, creating a cyclical process between practice and teacher education.

Competence to choose and construct graphs

Although there are an increasing number of investigations on the difficulties associated with statistical graphs, most studies focus mainly on graphing comprehension skills, including the processes of reading, interpretation and construction among students and pre-service teachers. According to Ashraf (2014), there are only a few studies which have investigated in-service teachers' understanding of statistical graphs.

Gonzáles et al. (2011, p. 190) argue that graphing competence is related to teachers' knowledge which enables them to develop activities, such as:

- o Identifying data from different types of graphs and interpreting them based on reading between, beyond and behind the displayed data in order to make hypotheses about the phenomena represented.

- o Selecting and constructing adequate graphs for specific situations, with or without technological support.

- o Evaluating graphs critically, assessing the potential and limitations of specific graphs, recognising that the construction of graphs involves interpretation of original data.

Therefore, among other activities, the processes of choosing and constructing graphs are necessary for graphing competence, which involves reasoning about representations.

Angra and Gardner (2016) developed a theoretical approach which conceptualises graphing competence as part of Metarepresentational Competence (MRC), which is associated with knowledge to construct and reasoning with external representations, such as statistical graphs. The conceptual framework on MRC was developed based on research studies conducted with students. MRC is based on four areas related to students' abilities and inabilities to choose, construct, and critically interpret graphs: invention, criticism, functioning and learning/reflection. The area called invention is associated with students' abilities which allow them to concept innovative graphical representations from data. The second area, criticism, is related to students' critical knowledge to assess different types of graphs and their strengths and weaknesses. The third area is functioning, which refers to students' reasoning to understand the purpose of different types of graphs and their use depending on the types of data present. The fourth area, called learning or reflection, reveals the students' awareness of their understanding of graphs. Angra & Gardner (2016) conducted interviews with university professors and students to identify the differences between the reasoning and quality of graphs constructed by them. As part of the interview, the participants were asked to analyse the graphs produced. The results suggested that professors focused their analysis on the function of graphs and the experimental task design. On the other hand, most undergraduate students interpreted based on their intuitions and data provided in the task. The results suggested that professors seemed to have difficulties in interpreting graphs, going no further than reading the data.

Arteaga et al. (2015) evaluated mistakes made by pre-service primary school teachers while constructing statistical graphs. The researchers analysed 207 graphs constructed from an open task, in which the participants had to collect the data through a random experiment, record it on a sheet, analyse it and then draw intuitive conclusions about it. After this stage, the participants were challenged to confront their initial analysis and present a written conclusion. Participants had the possibility of using statistical graphs. The researchers carried out a classification that sought to identify whether the produced graphs were correct or not. They recognised different semiotic conflicts to the interpretations of mathematical expressions that do not correspond to those intended to be taught. These conflicts are due to inadequate relations between semiotic function terms. Participants' conflicts related to the construction and selection of graphs were identified, such as choosing inappropriate graphs to represent data, representing inadequate variables, and constructing non-proportional scales.

Alacaci et al. (2011) investigated how future teachers choose appropriate graphs. The participants included 51 pre-service teachers who initially solved tasks related to three contexts of scientific problems, and afterwards interpreted four graphs that represented the same quantitative data associated with the tasks. Teachers were asked to evaluate the adequacy of each graph presented and indicate the reason for having chosen one of them. Based on the data analysis, the researchers elaborated four categories for choosing and not choosing graphs:

> **1. Conceptual explanation:** the explanation for choosing, or not choosing, a specific graph explicitly referred to the type of tasks implied in the context.

> **2. Indirect conceptual explanation:** the explanation for choosing, or not choosing, a certain graph made an indirect reference to the type of task.

> **3. Explanation of the structural components of graphs:** the explanation focused on the perception of the graphs' components (for example, bars, dots, lines, and subtitle).

> **4. Other explanations:** the explanation was based on personal opinions or preferences (for example, I just like it, this one is good, the graph is very clear).

Alacaci et al. (2011) concluded that the participants were able to recognise situations where the use of bar graphs, pie charts and line graphs was

appropriate, and to combine appropriate graphs to contexts. However, they had limited knowledge about scatterplots and did not recognise contexts in which this type of graph could be used. The authors identified that almost all participants commented on their favourite graph, the least preferred graph and some participants commented on other types of graphs as well. These results suggested that pre-service teachers' explanations were based on cognitive as well as affective aspects.

Carrión and Espinel (2006) carried out a comparative study between students of ages ranging between 10 and 12 from Spain and New Zealand. The authors gave the students a questionnaire hoping that they could assess the students' ability to pass information from one graph to another. Scatter plots, box plots, stem and leaf diagrams and histograms were included in the questionnaires. Students were asked to select the most appropriate representation for the statistical data given. In the results, it was observed that few students were able to understand the relationships between the different types of graphs and transport one representation to another. According to the authors, the most difficult transpositions involved stem and leaf diagrams and box plots.

Following the research carried out by Carrión and Espinel (2006), the same questionnaire was presented to a group of 40 future teachers before they had any classes involving statistics. The same questionnaire was then presented to a second group of 37 teachers after they had attended Statistics classes. One of more difficult tasks to be performed among teachers was to pass some information in text form to a graph. In this task, teachers found it less difficult to work with a bar graph and more difficult with a box plot. When comparing the responses between the two groups, an improvement was observed in the quality of the responses, with the second group presenting the best responses. From this research, the authors pointed out that, in general, there are many similarities between the difficulties faced by teachers and students.

The research studies of Alacaci et al. (2011), Arteaga et al. (2015) and Carrión and Espinel (2006) gave evidence that there is a need to improve teachers' knowledge about graphs.

METHOD

Research Design

This article aims to discuss an empirical study which investigated teachers' choices and constructions of statistical graphs from the results of interviews. The research study was developed from a qualitative perspective based on interpretive and qualitative paradigms, which consider the participants as part of the context (Bogdan & Biklen, 2007). The data was collected in a municipality located in the Metropolitan Area of Recife, Brazil[1].

Research Sample

Initially, this town had 202 teachers who were teaching 1st to 5th grade in 24 municipal public elementary schools. After visiting all of the schools, 132 teachers volunteered to give interviews which would investigate their choices and constructions of graphs. However, only 22 teachers were available to give interviews. These 22 teachers had an average age of 42 years with DP = 11.1, with the youngest teacher being 25 years old and the oldest being 66 years old. To preserve the participants' identities, a code was assigned to each teacher, formed by the letter P followed by a number from 1 to 22 that indicates the teaching time. Information on teaching experience time of each participant is presented in Table 1:

Table 1

The time of teaching experience in years of the participants of study 2

Teaching experience	Participants	Frequency
2-5 years	P1, P2	2
6-10 years	P3, P4, P5, P6, P7, P8, P9	7
11-15 years	P10, P11	2
> 15 years	P12, P13, P14, P15, P16, P17, P18, P19, P20, P21, P22	11
Total		**22**

---

[1] The study followed all ethical procedures related to research involving human beings, including the signing of free and informed consent form by the participants.

Most of the participants had a degree in Pedagogy. Only 3 teachers had other degrees, namely in Biological Sciences (P4 and P10) and Psychology (P8), while participant P18 had three different degrees (Mathematics, Pedagogy and Psychology) and participant P11 had only technical training in education. Among the participants, 18 teachers had a specialisation course (post-graduation course) and only one is an educational science master. None of the post-graduation certificates was related to statistical education or mathematical education.

Research Instruments and Procedures

The instruments used were semi-structured interviews about the choice of graphs used in the classroom and a task presenting statistical data for the construction of graphs by the participants. The interviews were recorded on audio and had an average time of 50 minutes. All interviews were transcribed as whole.

Figure 1

Task form to choose a graph: context 1.

Context 1

1) Observe the following information:

Day	Max. temperature (ºC)	Min temperature (ºC)
26-April	30	25
27-April	30	26
28-April	30	26
29-April	30	26
30-April	28	27
1-May	27	25
2-May	30	24
3-May	30	24
5-May	29	23
6-May	30	24

2) The data above show the record of maximum and minimum temperatures in the city of Recife over a 10-day period. If you had to represent this data on a graph, which one would you choose and why?

Figure 2

Task form to construct a graph: context 2.

<div align="center">

Context 2

</div>

1) The following data were grouped into classes and present scores from a test (which was scored from 0 to 100) taken by participants in a contest.

Scores	Absolute frequency
20 to 29	6
30 to 39	36
40 to 49	52
50 to 59	46
60 to 69	36
70 to 79	12
80 to 89	20
90 to 99	15
Total	223

2) If you had to represent this data on a graph, which one would you choose and why?

This article addresses the constructions of teachers based on the second instrument, in which teachers analysed two statistical data sets inserted in two different contexts: the first involving the measurement of the temperature of a city over 10 days, and the second related to the frequency of candidates' scores for a test (Figures 1 and 2). These situations were adapted from Martins and Ponte (2010).

Context 1 presents data related to continuous variables: the maximum and minimum temperature over 10 consecutive days. The most adequate graph to represent these data would be a line graph because it is appropriate to represent the way a variable evolves in relation to another variable (Martins & Ponte, 2010). The expected answer for context 2 would be a histogram because the data is represented in intervals.

Data Analysis

The audio-recorded interviews were transcribed and, with the speech protocols created, a content analysis was performed. The content analysis is, according to Bardin (1977), a set of data analysis techniques that aims to

identify the content of the messages through systematic procedures. The theme can be understood as the unit of meaning that is freed from a text according to the theory that serves as a guide to the researcher's interpretations. Thematic category analysis technique was used, which synthetically consists of the search for theoretical or empirical categories, responsible for the specification of a theme. Regarding the constructions of graphs, these were classified according to their level of semiotic complexity and the main reference was the study carried out by Arteaga (2011). The data analysis aimed to classify teachers' skills concerning their choice of graphs, as well as their construction, based on the framework of metarepresentational competences (MCR) by Angra and Gardner (2016).

Reliability of the Study

In this research, efforts were made to increase the reliability of the study. After transcribing the answers and before the graphs constructed by the teachers, the first investigator attributed an open code to the fragments of each answer. These codes, of a descriptive nature, sought to highlight the central meaning of the responses of each participant. Then, the first investigator, together with a second investigator, organised these codes into categories. Then, the two researchers discussed these categories to verify whether they were valid, given the complete responses of the study participants. From this, the categories were revised and refined, generating clearer definitions. After a consensus among the investigators, the first investigator analysed all the responses of the participants in the last stage of the codification process. The second investigator confirmed the analysis of the first coder, discussing all divergences until reaching a consensus. From this, the categories built around the teachers' reasons for choosing the graphs and also the levels of complexity of each representation were confirmed and validated.

RESULTS

Teachers' justifications for choosing statistical graphs

Regarding the reasons listed by teachers for choosing graphs, two categories were found. Only one teacher claimed not to know for sure what the reason for choosing the graph for context 2 was. All other teachers justified their choices of graphs related to contexts 1 and 2.

For context 1, involving the measurement of temperatures in a city, only one teacher chose to construct a pie chart, 9 teachers chose line graphs, and 11 teachers chose bar graphs. Only one teacher did not choose any type of graph. Regarding the results associated with context 2, the frequencies for each type of graph were: bar (13), pie (4), line (1) and dot (1). Three teachers were unable to construct the type of graph which they chose. Although the adequate graph would have been a histogram, nobody chose to construct this type of graph.

Despite some teachers being unable to represent the type of graph they had chosen, most teachers justified their choices. Teachers' arguments for indicating the types of graphs in the two contexts were grouped into two categories: Firstly, familiarity and facility to represent the type of graph, and secondly, the elements which comprise the type of graph. Those who did not explicitly recognise reasons for choosing the representation were considered in the category *do not know*.

The first category relates to the idea of choosing due to a previous experience with the representation. The following excerpts from the interview exemplify this category of teachers' arguments:

> I only work with the bar, I choose the bar. It is easier for me too (P17, context 1).

> It's easy for me to read. Maybe the circled one, I couldn't read. But the line for me is easy to read (P7, context 1).

The second category relates to how a graph could meet the representational needs of that statistical data set. The teachers emphasised that the structural elements of the graph allow better representation, such as bars and oscillations. The following excerpts illustrate those responses:

> The bar one goes well. Because, precisely, you will be able to name each bar as high as it went and put the score...You can put the frequency of people and in the horizontal part, (you put) the score. I understand that you can do it like this. That is why I think the bar one would fit well. (P11, context 2)

> Because the bars...at least horizontally, it will give a greater sense of temperature, especially if they are coloured. One blue and one orange for the maximum. It could make understanding easier. (P4, context 1)

Precisely because it will be able to show these oscillations, right, time goes... time goes up, time goes down, right? (P11, context 1)

Figure 3

The distributions of responses related to reasons for choosing graphs by type of graph context 1.
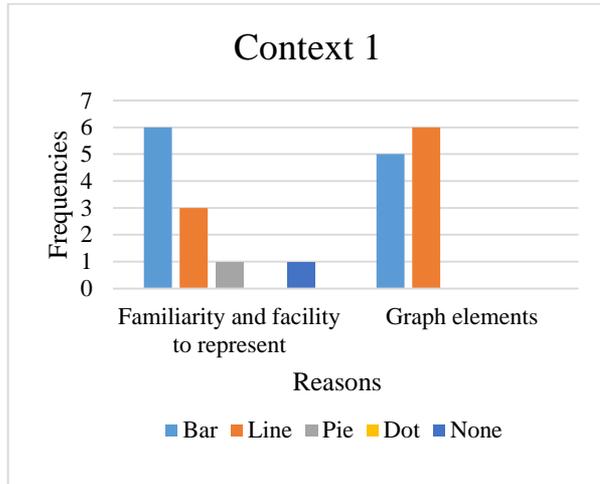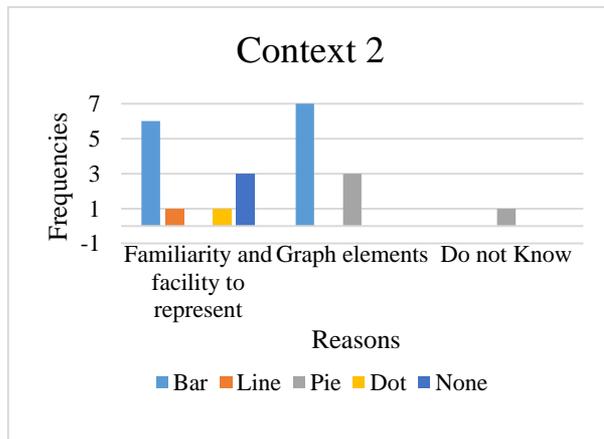

Figure 4

The distributions of responses related to reasons for choosing graphs by type of graph context 2.

Graphs on Figures 3 and 4 present results which combine information from interviews and the chosen graph to represent data from contexts 1 and 2:

Regarding context 1, which presented the temperatures of a city, there were a total of 22 justifications offered by teachers for choosing a graph to represent such data, 11 of which were associated with category 1 related to familiarity and facility, and 11 related to category 2 on graph elements.

Graphs constructed by participants

Only three teachers were unable to construct a graph: one participant did not construct any graph for the two contexts, while two participants were only able to represent one of the contexts. A total of 40 representations were produced by the participants: 21 graphs were related to context 1, and 19 graphs associated with context 2. The information related to the construction of graphs in the two contexts is summarized in Table 2:

Table 2

Types of Graphs	Contexts				Total
	Context 1	**(f)**	**Context 2**	**(f)**	
Bar graph	P1, P4, P8, P13, P17, P22	11	P4, P7, P15, P21, P22, P2, P5, P9, P10, P11, P14, P18, P19	13	24
Line graph	P3, P7, P15	9	P13	1	10
Pie chart	P21	1	P1, P8, P17, P6	4	5
Dot plot	-	0	P3	1	1
Do not Know/None	P16	1	P16, P12, P20	3	4
Total		**22**		**22**	**44**

Frequencies of types of graphs constructed by the participants in contexts 1 and 2.

The analysis of 40 representations constructed by teachers related to the two contexts provides classification at levels 2, 3, 4 (Figures 5, 6 and 7), and a graph was identified at an idiosyncratic level (Figure 8).

Figure 5

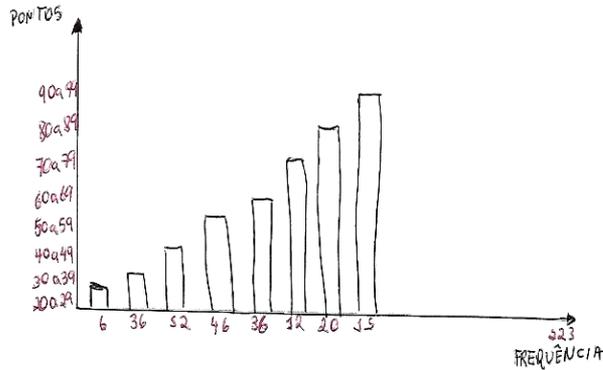*Example of graph classified on level 2 according to the semiotic complexity levels.*

Figure 6

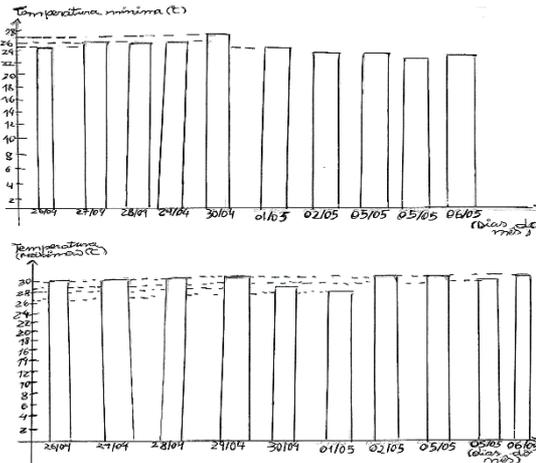*Example of graph classified on level 3 according to the semiotic complexity levels.*

Figure 7

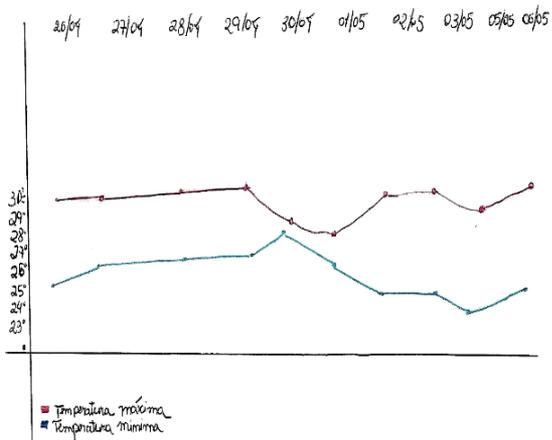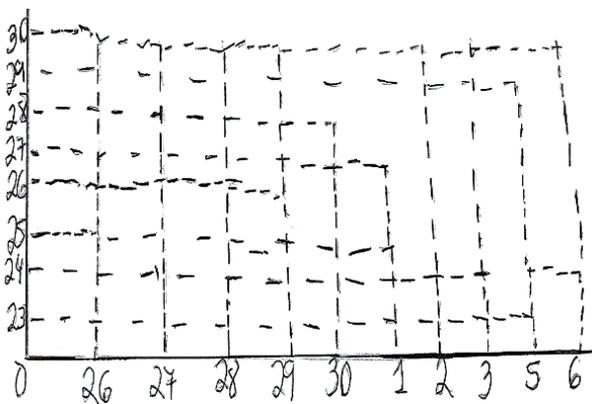*Example of graph classified on level 4 according to the semiotic complexity levels.*

Figure 8

Example of graph classified on idiosyncratic level according to the semiotic complexity levels.

An analysis of Figure 8 shows that the representation classified at the idiosyncratic level is logical in the form of presentation of the values for each axis, with a connection between the values of the X and Y axes.

Figure 9

The distributions of semiotic complexity levels for types of constructed graphs for context 1.
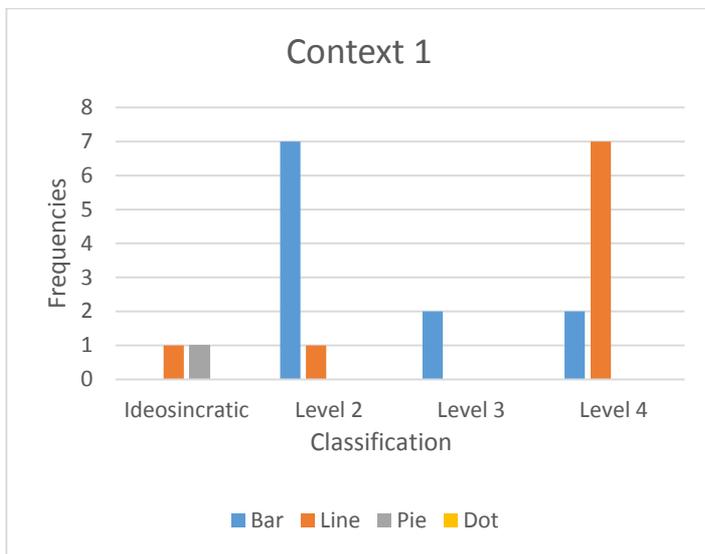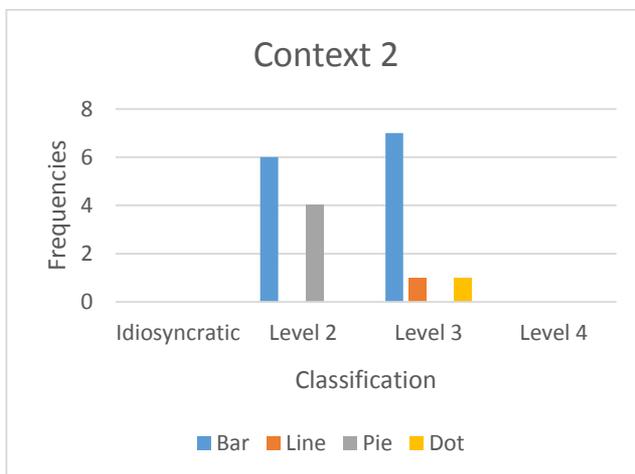

Figure 10

The distributions of semiotic complexity levels for types of constructed graphs for context 1.

In context 1, a greater number of teachers who constructed line graphs were able to construct more complex representations (Figure 9). On the other hand, the teachers who constructed bar graphs presented graphs with more elementary levels of complexity.

A significant number of participants (13) constructed bar graphs in context 2. Therefore, the criterion adopted to classify the same type of graph at level 2 or level 3 was related to the teachers' knowledge concerning the representation of data distribution. Some teachers, when inverting the order of axes, presented a trend that did not match the values presented in context 2 (Figure 10).

Figure 11 (A) is an example of a bar graph made by professor P18 which, despite adding two variables, does not allow the differentiation of maximum and minimum temperatures, since there are no distinctions between bar colours. Other elements that make it difficult to read the graph refer to the absence of labels, axis title, graph title, and proportionality of scale.

Figure 11

A) Bar graph produced in context 1 and classified at level 2; B) Line graph produced in context 1 and classified at level 2.

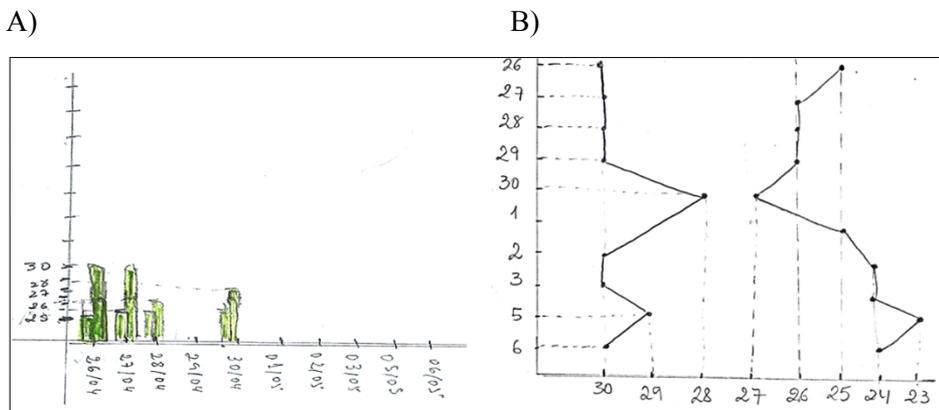A)   B)

Figure 11 (B) presents a line graph constructed by teacher P5, which represents two continuous variables in a suitable type of graph. The reading of

data trends was difficult due to scale with temperature having been represented in the class axis, as well as for not having differentiated lines or determining a title for the axes. Some teachers did not understand how the scale should be presented on the X or Y-axis. This difficulty was seen among teachers who chose bar graphs, dot plots and line graphs.

According to results related to *the reasons for teachers to choose statistical graphs*, presented in the previous section (Figures 3 and 4), it was observed that for context 1, teachers also justified that their familiarity with the representation was the reason for having chosen the bar graph as the most appropriate to represent the data. However, when viewing the distribution of reasons concerning the types of graphs chosen, it is clear that the arguments for choosing the bar graph in context 1 are more related to category 1, suggesting that familiarity or facility in constructing this type of graph was an element that influenced the choice. Also, this type of graph was the most chosen among teachers in both contexts, even when the data showed trends and variations that could be better viewed with other types of graphs.

Teachers' responses classified in category 1 in context 1, included not only bar graphs, but also line graphs and pie charts. Teachers' justifications included in category 2 related to only two types of graphs: bar and line. There were also more teachers who chose line graphs among those teachers' justifications classified in category 2, this graph being the most adequate to represent the data in context 1. It suggests that teachers' justifications associated with graph elements were more precise in their choice.

Most of the teachers' justifications to context 2 included bar graphs. This representation was the most chosen both among teachers who gave reasons related to familiarity and among those who justified their choices by graph elements. However, in contrast to context 1, the choice for the bar graph was more associated with category 2. This may be related to the types of data presented in context 2 (Figure 2). Therefore, in context 2, it seemed more evident that teachers' choices were associated with the relationship between the data presented and elements of the chosen graph.

Some teachers chose representations that they were unable to construct. These teachers, even considering that the selected type of graph was familiar or easier to represent, were unable to construct it. Even graphs which are considered as easier have a level of complexity to construct (Contreras et al., 2017). Therefore, being familiar with a representation does not enable people to know how to construct or critically evaluate it.

However, the representation does not have components such as titles, lines or bars, which prevents adequate reading and classification. The graph classified as level 2 indicates one or more variables, but it is not possible to identify the frequency distribution, as the data set suggested. Therefore, it is possible to have scale inversion (Y-axis) for categories (X-axis). The example of the graph at level 3 represents trends, but instead of representing all variables in one graph, there are two different representations, each one with a variable. Finally, the example of level 4 includes all variables from the same data set in a single representation and presents the frequency distribution correctly.

After analysing the participants' responses for context 1, 21 graphs were identified: 8 (double bar graphs), 3 (single bar graphs), 9 (line graphs) and 1 (pie chart). The classification by level of complexity in context 1 identified the following: 2 graphs had idiosyncratic complexity level, 7 graphs (level 2), 2 graphs (level 3) and 10 (level 4).

19 graphs were classified for context 2: 13 (bar), 4 (pie), 1 (line) and 1 (dot). In the classification by level of complexity: 10 graphs (level 2), and 8 (level 3). Since it was only necessary to use one variable to answer the task of context 2, level 3 would be the highest that could be reached. In Figure 4, the distribution of types of graphs based on semiotic complexity levels in each context can be observed.

The classification presented in Figure 4 is further supported by the fact that the bar graphs constructed in context 1 did not correctly allow an analysis of data trends, both because of continuous variables requiring another type of representation and because of some teachers' difficulties to differentiate context 1 variables when constructing their graphs.

Regarding responses to context 2, there was a higher frequency of pie charts. It is possible to infer two possible factors that caused these results: context 2 proposes a variable, and teachers have difficulty in representing interval data on a scale. Therefore, the option to represent data in a pie chart overcame the need to add scale to representation. Despite this, the choice for this type of graph did not easily allow the visualisation of trends in the context 2 data, as these constructions were classified in level 2 of semiotic complexity.

DISCUSSION AND RECOMMENDATIONS

This article aimed to analyse aspects of in-service teachers' knowledge about statistical graphs based on their reasons for constructing statistical graphs.

The research study intended to contribute to understanding the difficulties experienced by teachers while they are teaching about graphing.

The methodology was comprised of the following data collection instruments: semi-structured interviews about the choice of graphs to use in the classroom and a task presenting statistical data for the construction of graphs by teachers. This article addresses the constructions of teachers based on the second instrument.

The analysis of the semiotic complexity of graphs constructed by teachers and their reasons to choose such graphs indicated participants' difficulties associated with the construction of the chosen graphs since 20 of 40 representations had low levels of semiotic complexity (from idiosyncratic level to level 2). Also, some participants were unable to produce any type of graph and had difficulties to read trends from data. This result corroborates the idea pointed out in previous studies that the construction of graphs is not a simple task (Arteaga, 2011; Angra & Gardner, 2016, 2017, 2018).

Teachers chose bar graphs even when the representation was not adequate to display the data. Their justifications for such choices were based on their familiarity with the bar graph or its supposed ease of construction. Different mistakes were observed in the participants' chosen representations, such as the proportions of bars and scales, the absence of naming titles and graph axes. These difficulties, associated with the construction of graphs, can be complementary to the analyses of their semiotic complexity since such errors, when identified, can help explain difficulties in using a graph.

Recommendations

When comparing the results from this study with the results of Arteaga (2011), it is possible to corroborate his categories to classify the semiotic complexity of graphs produced by teachers and another category could be added related to an idiosyncratic level. The data analyses emphasise the idea that the difficulties faced by teachers in pre-service teacher education do not seem to be overcome when they are in-service. Based on these results, it is recommended that mathematics and statistics teacher education should develop the understanding of aspects related to the choice and construction of graphs associated with different types of statistical data. Teachers need to have continuing teacher education which promotes the analysis of actual teaching practices to allow teachers to understand and overcome their difficulties related to the tasks of choosing and producing graphs with their students.

Further research

The present research offers a series of new questions about the investigated problem. One of the aspects that could be further explored in further studies might be the influence of textbooks regarding activities with statistical graphs to identify which types of graphs and activity approaches are suggested to teachers. It is crucial to understand teachers' possibilities to choose such activities to understand how to propose and develop tasks which could be improved.

Further studies should apply new questionnaires which explore teachers' knowledge about statistical graphs, including elementary graphs such as bar graphs, line graphs, pie charts, and dot plots, as well as more complex graphs, such as histograms, box plots and scatter plots.

It is also interesting that new studies can develop analyses which connect teachers' technical skills with pedagogical skills based on data collected from actual classroom activities involving the choosing and constructing of graphs. This further research should be based on methodological procedures of observation which enable researchers to identify other kinds of teachers' justifications for choosing graphs. Therefore, it would be possible to compare whether having a more sophisticated competence for choosing and constructing graphs contributes or facilitates the process of conducting statistical research in the classroom.

The study was comprised of research tasks to be solved only with the support of pencil and paper, since this was the major resource among teachers to conduct their classes with graphs. However, other resources can differently support the construction of graphs. For example, computer-based research tasks can provide different possibilities to construct statistical graphs.

Limitations

The data analysis suggested three limitations for this study which need further investigation in future research studies. The first limitation is related to the context 2 task, for which no teacher chose the most appropriate graph to represent the data: namely the histogram. This result raises questions concerning teachers' knowledge levels about histograms. Similarly, this can be a questioning element for other types of graphs chosen by teachers. Therefore, further studies should contribute to explain whether teachers' knowledge levels about the types of graphs influences their choices.

The second limitation of this study is associated with the fact that it was requested that teachers respond to a protocol for choosing graphs with specific scenarios and predetermined by the researchers. We believe that this can be a limiting aspect to understand the complexity of teachers' choices. It is possible that other variables could influence the process of choosing statistical representations by teachers, such as issues that involve the curriculum and students' learning.

A third limitation is related to the research sample itself. The participants represent a small sample of Brazilian teachers when the vast diversity of national territory is considered. For example, the fact that state and municipal elementary education networks conduct in-service teacher education autonomously can be an important variable. It is possible that certain municipalities may prioritise more training related to statistical topics than others. If there were samples from different regions of the country, it can be an important methodological procedure to identify elements that can contribute to understanding other specificities about the activity of choosing and constructing graphs in classroom tasks.

CONCLUSIONS

The main objective of this article was to analyse the choices and the construction of statistical graphs among early years teachers. The analyses of participants' choices were not only based on the identification of chosen graphs but also considering teachers' arguments to validate those choices. In order to analyse the construction of the chosen graphs, the level of semiotic complexity of graphs was assessed. We observed that the participants had difficulties constructing and choosing graphs based on statistical data, corroborating previous investigations on the complexity of these tasks.

The participants chose statistical graphs considering their familiarity and the ease with which some types of graphs are constructed, as well as analytical aspects associated with the representation of data. It seems that the justification related to how difficult it was to construct certain types of graphs influenced choices of graphs that were not adequate to represent the data associated with the tasks. This suggested that the specific knowledge about graphs needs to be developed in mathematics and statistics teacher education to allow teachers to make more conscious decisions regarding graphing tasks.

This study motivated further research to discuss the most effective ways of developing teachers' graphing understanding which considers different

types of graphs and their selection for pedagogical purposes. The improvement of early years teachers' understanding of statistics is relevant to promote statistics teaching which relies on a deeper knowledge of this area of study. The results emphasised the idea that it is necessary to discuss the work with graphs through pre-service and in-service elementary school teacher education.

ACKNOWLEDGEMENTS

AUTHORS' CONTRIBUTIONS STATEMENTS

MNM conceived and conduct this research project supervised by CC. CM participated in the discussion of main results, implications and translation. All authors actively participated in the writing, review and editing of the final version of work.

DATA AVAILABILITY STATEMENT

Authors agree to make their data available upon reasonable request from a reader. It is up to the authors to determine whether a request is reasonable or not.

REFERENCES

Ainley, J. (2000) Transparency in graphs and graphing tasks: an iterative design process. *Journal of Mathematical Behavior, 19*(3), 365-384. https://doi.org/10.1016/S0732-3123(00)00052-3

Ainley, J., Pratt, D., & Nardi, E. (2001). Normalising: Children's activity to construct meanings for trend, *Educational Studies in Mathematics,*

45(1-3), 131-146.
https://link.springer.com/article/10.1023/A:1013822512833

Alacaci, C., Lewis, S., O'Brien, G. E., & Jiang, Z. (2011). Pre-service elementary teachers' understandings of graphs. *Eurasia Journal of Mathematics, Science y Technology Education, 7*(1)*,* 3-14. https://doi.org/10.12973/ejmste/75171

Angra, A., & Gardner, S. M. (2016). Development of a framework for graph choice and construction. *Advances in Physiology Education Published*, *40*(1), 123-128. https://doi.org/10.1152/advan.00152.2015

Angra, A., & Gardner, S. M. (2017). Reflecting on Graphs: Attributes of Graph Choice and Construction Practices in Biology. CBE - Life Sciences Education, *16*(3), 1-15. https://doi.org/10.1187/cbe.16-08-0245

Angra, A., & Gardner, S. M. (2018). The Graph Rubric: Development of a Teaching, Learning, and Research Tool. *CBE - Life Sciences Education*, *17*(4), 1-18. https://doi.org/10.1187/cbe.18-01-0007

Arteaga, P. (2011). *Evaluación de conocimientos sobre gráficos estadísticos y conocimientos didácticos de futuros profesores* [Doutoral Thesis, University of Granada]. https://www.ugr.es/~batanero/pages/ARTICULOS/arteaga.pdf

Arteaga, P., Batanero, C., Contreras, J. M., & Cañadas, G. R. (2015). Statistical graphs complexity and reading levels: a study with prospective teachers. *Statistique et Enseignement, 6*(1), 3-23. http://statistique-et-enseignement.fr/article/view/430/407

Ashraf, M. (2014). *Teacher and student categorization of graphics into graphical genres* [Doctoral Thesis, Université de Grenoble]. https://tel.archives-ouvertes.fr/tel-01200726/document

Bardin, L. (1977). *Análise de conteúdo.* Edições 70.

Batanero, C., Arteaga, P., & Ruiz, B. (2009). Statistical graphs produced by prospective teachers in comparing two distributions. In V. Durand-Guerrier., S. Soury-Lavergne., F. Arzarello (Eds.), *Proceedings of 6ᵗʰ Congress of the European Society for Research in Mathematics Education* (pp. 368-377). Instituit Français d'éladucation.

Bogdan, R., & Biklen, S. (2007). *Qualitative research for education: An introduction to theories and methods*. Pearson.

Borelli, S., Pacheco, D., & Pires, C. (2016). Professores do 1º ano que ensinam Matemática no Ensino Fundamental da rede estadual de São Paulo: uma possibilidade de mudança na prática a partir da formação. *REnCiMa, Edição Especial: Educação Matemática, 7*(4), 49-62. https://doi.org/10.26843/rencima.v7i4.1205

Carrión, J .C., & Espinel, M. C. (2006). Una investigación sobre la traducción e interpretación de gráficas y tablas estadística por estudiantes de educación primaria. In A. Rossman & B. Chance (Eds.), *Proceedings of the Seventh International Conference on Teaching Statistics.* International Statistical Institute and International Association for Statistical Education.

Carvalho, C. F. (2009). Reflexões em torno do ensino e da aprendizagem da Estatística. In J. A. Fernandes, F. Viseu, M. H. Martinho & P. F. Correia (Eds.) *Atas do Segundo Encontro de Probabilidades e Estatística na Escola* (pp. 22-36) Centro de Investigação em Educação da Universidade do Minho.

Cazorla, I. (2002). *A relação entre a habilidade viso-pictórica e o domínio de conceitos estatísticos na leitura de gráficos* [Doctoral thesis, Universidade Estadual de Campinas]. Repositório Unicamp. http://repositorio.unicamp.br/jspui/handle/REPOSIP/250925

Cazorla, I. M., & Castro, F. C. (2008). Papel da Estatística na leitura do mundo: o letramento estatístico. *Publicatio UEPG: Ciências Humanas, Ciências Sociais Aplicadas, Lingüística, Letras e Artes, 16*(1), 45-53. http://dx.doi.org/10.5212/PublicatioHum.v.16i1.045053

Contreras, J.M., Molina, E., Arteaga, P., Godino, J.D., & Rodrígues-Pérez,C. (2017). Funciones semióticas críticas en el uso de diagramas de barras por los médios de comunicación. In J. M. Contreras, P. Arteaga, G. R. Cañadas, M. M. Gea, B. Giacomone & M. M. López-Martín (Eds.), *Actas del Segundo Congreso International Virtual sobre el Enfoque Ontosemiótico del Conocimiento y la Instrucción Matemáticos* (pp. 1-11). Digibug, Universidad de Granada.

Curcio, F. R. (1989). *Developing graph comprehension: elementary and middle school activities.* NCTM.

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), 2-7. https://www.jstor.org/stable/40248564

Gal, I. (2002) Adult statistical literacy: Meanings, components, responsibilities. *International Statistical Review, 70*(1), 1-25. https://doi.org/10.2307/1403713

Godino, J. D. (2017). Contruyendo un sistema modular e inclusivo de herramientas teóricas para la educación matemática. In J. M. Contreras, P. Arteaga, G. R. Cañadas, M. M. Gea, B. Giacomone & M. M. López-Martín (Eds.), *Actas del Segundo Congreso International Virtual sobre el Enfoque Ontosemiótico del Conocimiento y la Instrucción Matemáticos* (pp. 1-20). Universidad de Granada.

Godino, J. D., Batanero, C., & Font, V. (2007). The onto-semiotic approach to research in mathematics education. ZDM. *The International Journal on Mathematics Education, 39*(1-2), 127-135. https://link.springer.com/article/10.1007/s11858-006-0004-1

González, T., Espinel, C., & Ainley, J. (2011). Teachers' graphical competence. In C. Batanero, G. Burrill & C. Reading (Eds.). *Teaching Statistics in School Mathematics. Challenges for Teaching and Teacher Education*: *A Joint ICMI and IASE Study*. Springer.

Guimarães, G.L., Gitirana, V., Cavalcanti, M., & Marques, M. (2007). Livros didáticos de matemática nos anos iniciais: Análise das atividades sobre gráficos e tabelas. In C. S. Marcelo (Coord.), *Anais do 9º Encontro Nacional de Educação Matemática* (pp. 1-17). SBEM.

Lopes, C., Powell, A., Ribeiro, R., & Pazuch, V. (2020). Continuing professional development of a Mathematics teacher. *Revista de Ensino de Ciências e Matemática*, 11(7), 64-80. https://doi.org/10.26843/10.26843/rencima.v11i7.2790

Martins, M. E. G., & Ponte, J. P. (2010). *Organização e tratamento de dados*. DGIDC.

Martins, M. N., Monteiro, C., Prodromou, T. (2017). Teachers Analyzing Sampling with TinkerPlots: Insights for Teacher Education. In T. Prodromou (Ed.), *Data Visualization and Statistical Literacy for Open and Big Data* (pp. 194-222). IGI Global. https://doi.org/10.4018/978-1-5225-2512-7.ch008

Monteiro, C., & Ainley, J. (2007). Investigating the interpretation of media graphs among student teachers. *International Electronic Journal of Mathematics Education, 2*(3), p. 188-207.

https://www.iejme.com/article/investigating-the-interpretation-of-media-graphs-among-student-teachers

Ministério da Educação. (2018). *Base Nacional Comum Curricular (BNCC). Educação é a Base*. MEC/CONSED/UNDIME.

Souza, J., & Monteiro, C. E. F. (2020). Understanding about graphs by field school teachers. Zetetike, 28, e020005-e020005. https://doi.org/10.20396/zet.v28i0.8657061

Vygotsky, L. (1978). *Mind in Society: The development of higher psychological processes*. Harvard University Press.