http://doi.org/10.14393/ER-v31e2024-04

# Registers of semiotic representation in statistics teaching: an analysis of textbook activities[1]

*Dienifer Ferner Fernandes*[2]
*Carlos Maximiliano Dutra*[3]
*Maria Arlita da Silveira Soares*[4]

**ABSTRACT**

Given the importance of interpreting statistical information and the textbook, this work aims to analyze, in activities proposed by two collections of textbooks in mathematics and its technologies (high school), approved by the PNLD, how statistical knowledge is approached, through the content analysis, based on the BNCC indications combined with assumptions of the theory of registers of semiotic representation. Ninety-eight activities were analyzed in Book A and 164 in Book B. In the analyses, we found that contexts about everyday life predominate and skills provided in the BNCC are little worked. Different types of semiotic representations were explored in the analyzed activities. The conversion activity stood out in both works, with an emphasis on graphic to numeric and graphic to natural language conversions. The identification of the distribution of semiotic representations and the proposed transformations allows the teacher to define better the statistical skills to be worked on.

**KEYWORDS:** Common National Curriculum Base; Statistical Education; Mathematics Education.

[1] English version by Maria Isabel de Castro Lima. *E-mail*: baulima@gmail.com.

[2] Doutoranda em Educação em Ciências. Universidade Federal do Pampa, Uruguaina, Rio Grande do Sul, Brasil. Orcid: https://orcid.org/0000-0002-4551-0763. *E-mail*: dieniferlferner@gmail.com.
[3] Pós-Doutor em Astronomia. Universidade Federal do Pampa, Uruguaina, Rio Grande do Sul, Brasil. Orcid: https://orcid.org/0000-0003-4743-874X. *E-mail*: carlosdutra@unipampa.edu.br.
[4] Doutora em Educação nas Ciências. Universidade Federal do Pampa, Caçapava do Sul, Rio Grande do Sul, Brasil. Orcid: https://orcid.org/0000-0001-5159-8653. *E-mail*: mariasoares@unipampa.edu.br.

*Registros de Representação Semiótica no Ensino de Estatística: uma análise de atividades propostas em livros didáticos*

**RESUMO**

Dada a importância da interpretação de informações estatísticas e do livro didático este trabalho objetiva analisar, em atividades propostas por duas coleções de livros didáticos da área da Matemática e suas Tecnologias (Ensino Médio), aprovadas pelo PNLD, como são abordados conhecimentos estatísticos, através da Análise de Conteúdo, tendo como base as indicações da BNCC aliada a pressupostos da teoria dos Registros de Representação Semiótica. Foram analisadas 98 atividades no Livro A e 164 no Livro B. Nas análises constatou-se que predominam os contextos sobre cotidiano e existem habilidades previstas na BNCC que são pouco trabalhadas. Foram explorados diferentes tipos de representações semióticas nas atividades analisadas. A atividade de conversão se sobressaiu em ambas às obras, com ênfase às conversões gráfica para numérica e gráfica para língua natural. A identificação da distribuição das representações semióticas e as transformações propostas permitem ao professor definir melhor as competências estatísticas a serem trabalhadas.

**PALAVRAS-CHAVE:** Base Nacional Comum Curricular; Educação Estatística; Educação Matemática.

*Registros de Representación Semiótica en la Enseñanza de la Estadística: un análisis de las actividades propuestas en los libros de texto*

**RESUMEN**

Dada la importancia de interpretar la información estadística y el libro de texto, este trabajo tiene como objetivo analizar, en actividades propuestas por dos colecciones de libros de texto del área de Matemáticas y sus Tecnologías (Escuela Secundaria), aprobadas por el PNLD, cómo se aborda el conocimiento estadístico. , a través del Análisis de Contenido, a partir de las indicaciones de la BNCC combinadas con supuestos de la teoría de los Registros de Representación Semiótica. Se analizaron 98 actividades en el Libro A y 164 en el Libro B. En los análisis se encontró que predominan los

contextos sobre la vida cotidiana y hay habilidades brindadas en la BNCC que son poco trabajadas. En las actividades analizadas se exploraron diferentes tipos de representaciones semióticas. La actividad de conversión se destacó en ambos trabajos, con énfasis en las conversiones de gráfico a numérico y de gráfico a lenguaje natural. La identificación de la distribución de las representaciones semióticas y las transformaciones propuestas permiten al docente definir mejor las habilidades estadísticas a trabajar.

**PALABRAS CLAVE:** Base Curricular Nacional Común; Educación estadística; Educación Matemática.

* * *

## Introduction

Statistical information is present in people's daily lives. During the COVID-19 pandemic, the population saw the use of statistics more clearly. Different statistical analyses were published daily in the news, highlighting the number of cases, regions with the highest contamination rates, mortality and recovery rates as the days went by, and the number of people vaccinated, among countless other highlights. Therefore, there is a need for people who know how to carry out this type of research: collect data and analyze it to make it meaningful for the population. It is also important that the population knows how to interpret published research, as it influences the organization and progress of society in different ways.

To this end, the National Common Curricular Base (Base Nacional Comum Curricular - BNCC) (Brasil, 2018) proposes the study of statistical content in the area of mathematics and its technologies in the thematic unit entitled "Estatística e Probabilidade" [Statistics and Probability] throughout basic education. In early childhood education, the first notions are, for example, constructing graphs to express measurements (weight, height, among others). In elementary education, planning and carrying out data collection and organization are indicated,

based on a survey of interest to students to mobilize interpretation and construction of tables and graphs, calculation of central tendency measures, and subsequently, preparation of research reports.

In secondary education, all the ideas covered during basic education are unified and expanded. At this level, the BNCC highlights not only the interpretation of statistical information released by the media but also the planning and execution of sample research, including the interpretation of central tendency measures, the communication of results through reports, and the mobilization of appropriate graphical representations.

The BNCC underscores specific skills for each area of knowledge and level of education, followed by skills described to develop each skill. In secondary education, two competencies[5] explicitly highlight the work with situations arising from different contexts. Therefore, the focus of secondary education is on "[...] constructing an integrated vision of mathematics, applied to reality, in various contexts" (Brasil, 2018, p. 528). In this sense, there is a need to approach statistical content through various contexts. In other words, explore situations that mobilize contexts of mathematics, everyday life, or other areas of knowledge.

When referring to statistics, we often think of tables and graphs. There are different types of such representations, i.e., different ways of organizing and presenting tables and graphs (sector, bars, line, box). However, statistical information can also be communicated through other representations, for example, natural language.

The fourth specific competence of mathematics for secondary education indicated by the BNCC (2018, p. 531) emphasizes that the student should be able to: "Understand and use, with flexibility and precision, different registers of mathematical representation (algebraic, geometric, statistical, computational etc.), in the search for solutions and communication of problem results". Duval (2011) also defends this

[5] Specific competence one and three in the area of mathematics and its technologies.

idea, understanding that there is only a complete understanding of a mathematical object when the subject can recognize and understand the different forms of representation that it can take.

Raymond Duval (2011), in his theory of the registers of semiotic representation (RSR), investigates the influence of several types of representations of mathematical objects in the teaching and learning process. According to Duval (2013), students' difficulties are not directly related to the concepts covered but to the multiple forms of semiotic representations that can be used for mathematical objects and the "confused" use they make when using them.

Several studies (Bayer *et al*., 2004; Coast; Nacarato, 2011; Pietropaolo; Silva; Amorim, 2019) report difficulties teaching statistical content. The authors discuss the obstacles teachers face when teaching classes in this area and the lack of priority in developing these contents, whether due to lack of affinity with the topic, insignificant experiences or even lack of time. These reasons may be combined with the difficulty of understanding that statistics is not limited to applying formulas; in other words, the data needs to be part of a context to become understandable and meaningful.

Besides the obstacles mentioned, the teacher needs to adapt their classes to develop the skills and abilities indicated in the BNCC. We must note that books approved by the National Book and Teaching Material Program (Programa Nacional do Livro e Material Didático - PNLD) can be a good resource for organizing classes, as they aim to follow the BNCC guidelines. However, it is necessary to analyze the works to identify their focuses and individualities. Due to basic education teachers' demands, we understand that conducting an in-depth study of the works may be unfeasible. For this reason, research that analyzes textbooks is so relevant, as by contributing to teachers, they indirectly collaborate in improving students' teaching and learning processes.

Given the above, this investigation aims to analyze how statistical knowledge is approached in activities proposed by two collections of textbooks in mathematics and its technologies (high school) approved by the PNLD. The analyses sought to verify the skills and abilities related to statistics indicated by the BNCC and the identification of the assumptions of the theory of registers of semiotic representation, such as the mobilization of different representations and the proposed cognitive transformations.

## Registers of semiotic representation in statistics teaching

Semiotic representations are "productions constituted by the use of signs belonging to a representation system, which have their own difficulties in meaning and functioning" (Duval, 1993, p. 39, our translation). And registers are "a field of variation in semiotic representation depending on cognitive factors that are specific to them" (Duval, 2011, p. 97).

Just like in mathematics, in statistics, we can represent the same information through different semiotic representations. For example, a frequency distribution can be represented through a bar chart, a table, or even described in natural language.

Each mobilized representation has particular properties that can favor the resolution of a problem to the detriment of other representations. This is because two representations of the same object, formed from different records, may not present the same content. Therefore, the student must be able to understand and identify the efficiency of each of them. However, most of these subjects do not have control over some types of representations because some are more privileged than others in mathematics teaching, particularly statistics content.

One of the central points of the RSR theory is that understanding conceptual content lies in the simultaneous coordination of at least two registers of semiotic representation (Duval, 2003). In this sense, it is

pertinent to prioritize teaching situations from different contexts that seek to mobilize diverse representations, as they enhance the understanding of objects, in this case, statistics.

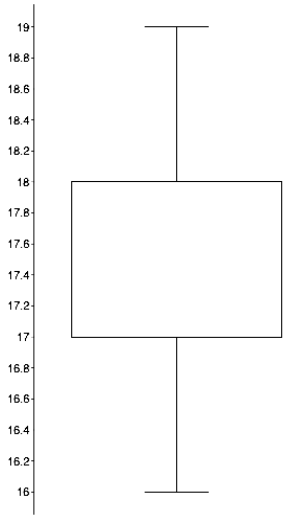
This point underlines the cognitive transformations that can occur between semiotic representations. They change the representation but maintain the reference to the same object as the initial representation. Cognitive transformations are categorized into two types: treatment and conversion. The treatment occurs when the transformation operation is internal to the set-off register; that is, the semiotic representation is modified, but the type of register remains. Conversion is performed when the transformation operation is external to the set-off register, i.e., the semiotic representation is changed, and the type of register is also modified. Cognitive transformations are exemplified in Chart 1.

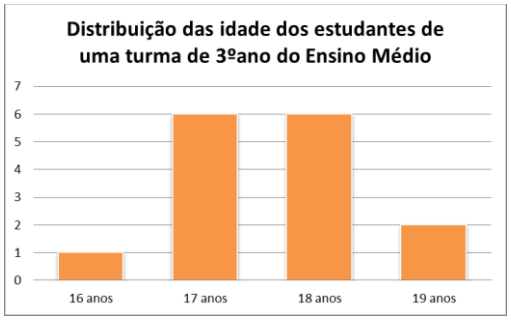
The dotted arrows (Chart 1) indicate the treatment cognitive transformations, while the solid arrows represent the conversions that can be carried out. The direction of the arrows indicates the departure representation and the arrival representation. We observe that some representations allow conversion in both directions and others do not; for example, it will not be possible to construct a bar graph from a box-plot graph. It highlights that the conversion activity is not as simple as it seems.

For Duval (2004, p. 47, our translation), "[...] conversion of semiotic representations constitutes the least spontaneous and most difficult cognitive activity to acquire for the vast majority of students". This fact may be associated with teaching and learning, which sometimes privileges specific types of representations and meanings of conversions. In the case of statistics, tabular and graphical representations and conversions starting from a table and having the graph as the final representation are privileged.

http://doi.org/10.14393/ER-v31e2024-04

**CHART 1**: Cognitive transformation in statistics

<table>
<tr><th>Tabular register</th><th>Graphic register</th></tr>
<tr><td>
Table A: Ages of students in a 3rd-grade high school class
<table>
<tr><th>Students</th><th>Ages</th></tr>
<tr><td>Fernanda</td><td>18</td></tr>
<tr><td>Ana Clara</td><td>18</td></tr>
<tr><td>Bruno</td><td>17</td></tr>
<tr><td>Flávia</td><td>18</td></tr>
<tr><td>Miguel</td><td>19</td></tr>
<tr><td>Alexandre</td><td>17</td></tr>
<tr><td>João</td><td>17</td></tr>
<tr><td>Gabriela</td><td>17</td></tr>
<tr><td>Julia</td><td>18</td></tr>
<tr><td>Eduardo</td><td>16</td></tr>
<tr><td>Ricardo</td><td>19</td></tr>
<tr><td>Luana</td><td>17</td></tr>
<tr><td>Amanda</td><td>17</td></tr>
<tr><td>Luiz</td><td>18</td></tr>
<tr><td>Tiago</td><td>18</td></tr>
</table>
Source: Fictitious data.
</td><td>
Box-plot based on the ages of students in a 3rd-grade high school class


Source: Fictitious data.
</td></tr>
<tr><td>
Table B: Age distribution of students in a 3rd-grade high school class
<table>
<tr><th>Ages</th><th>Qty. of students</th></tr>
<tr><td>16</td><td>1</td></tr>
<tr><td>17</td><td>6</td></tr>
<tr><td>18</td><td>6</td></tr>
<tr><td>19</td><td>2</td></tr>
<tr><td>Total</td><td>15</td></tr>
</table>
Source: Fictitious data.
</td><td>


Source: Fictitious data.
</td></tr>
</table>

**Source**: Organized by the authors.

However, Duval (2003) reinforces that the problem may lie in the erroneous understanding that carrying out activities in a conversion direction also automatically serves to "train" conversion in the opposite direction. In other words, the fact that students manage to mobilize a transformation external to the starting register does not mean that they can perform this exact conversion in reverse, as each type of conversion between representations has its own rules that, when changing the direction of the conversion, they, for the most part, also change.

Thus, the need to approach activities exploring the most varied arrangements of representations that involve conversions in both directions is emphasized. Given the above, we understand that through the mobilization of different semiotic representations, a more comprehensive field of study will be organized for students in statistics, providing an understanding of the object in its entirety through the semiotic functioning of each representation that this object can assume.

## Methodological procedures

For this investigation, we adopted the assumptions of qualitative research since, according to Borba (2004), this prioritizes descriptive procedures to deepen the understanding of something. However, this does not rule out the possibility of using quantitative procedures. As for the research procedures, this qualifies as documentary, as its source of data production is textbooks approved by the PNLD.

To assist in organizing and analyzing the data, we used some assumptions from content analysis proposed by Bardin (2022). This technique consists of three stages: pre-analysis, exploration of the material, and treatment of results and interpretations. The pre-analysis stage aims to organize the research, i.e., establishing the objectives, choosing the data source, and preparing the indicators.

For this investigation, in December 2022, we accessed the *Relatório de escolas participantes da escolha de livros* [Report of schools participating in the choice of books][6], which allowed us to check the works approved by the PNLD 2021 of Object 2[7] (Brasil, 2020), chosen by each school in each municipality in Brazil. In this, we searched for didactic works in mathematics and its technologies intended for high schools selected by public schools in Rio Grande do Sul[8]. To choose a

[6] http://simec.mec.gov.br/livros/publico/index_escolha.php
[7] It consists of didactic works by area of knowledge and specific works.
[8] The choice of works is restricted to the RS because this investigation is part of the first author's thesis project and is carried out in this state.

book, each school must indicate 1st and 2nd options from different publishers or suggest they do not want to receive material for a specific area. Each selection was organized in an electronic spreadsheet,[9] which allowed us to verify that, without making a distinction between 1st and 2nd choice, with 50.2% and 32.5% of schools (570 and 369 selections), respectively, the works *Prisma matemática* [Mathematics Prisma] and *Matemática em contextos* [Mathematics in contexts] were the ones most chosen in the schools we mapped.

Thus, considering that the works in mathematics and its technologies have six volumes, we selected the book that addresses statistical content in the versions of the teacher's handbook: *Prisma matemática: estatística, combinatória e probabilidade: área do conhecimento: matemática e suas tecnologias* [Mathematical prism: statistics, combinatorics and probability: area of knowledge: mathematics and its technologies] (Bonjorno; Giovanni Júnior; Sousa, 2020) and *Matemática em contexto: estatística e matemática financeira* [Mathematics in context: statistics and financial mathematics] (Dante; Viana, 2020), which we will refer here as Book A and Book B, respectively.

Once selected the documents, we carried out a floating reading to get to know the works, "[...] allowing oneself to be invaded by impressions and guidelines" (Bardin, 2022, p. 122). The referencing and creation of indicators are determined from excerpts of texts exposed in the documents. Thus, in this work, the indices are the statistical concepts presented in the activities in the books, and the indicators are the number of times these concepts appear in the activities (frequency). The preparation of the material involved organizing the activities mapped out in Excel spreadsheets[10].

The data exploration stage consists of "[...] systematic application of the decisions made" (Bardin, 2022, p. 127). At this stage, coding, classifying

[9] Disponíveis em: https://docs.google.com/spreadsheets/d/1Sz4iSs04cS7ElmZaFTG0MuXWzKBJ-xTM/edit?usp=sharing&ouid=110912744028284594015&rtpof=true&sd=true.
[10] Available at: https://docs.google.com/spreadsheets/d/1xIFKdO-r3uYH-YNSiQlNeT2A36PrchqO/edit?usp=sharing&ouid=110912744028284594015&rtpof=true&sd=true.

and categorizing are essential. Coding was carried out based on the selection of activities (solved and proposed) that involved statistical concepts, highlighting the chapter, page, and number of the activity in the work and the content involved. Afterwards, the activities were analyzed based on the analysis categories presented in Chart 2.

**CHART 2**: Categories of analysis

| Category | Description |
|---|---|
| Context | Identify which contexts are covered. |
| BNCC skills | Analyze which and how the skills related to the thematic unit "Probabilidade e Estatística" [Probability and Statistics], listed in the BNCC, are explored. |
| Registers of semiotic representation | Verify which registers of semiotic representation are mobilized, and the types of cognitive transformations most explored. |

**Source**: Organized by the authors.

Subsequently, in the next section, in the stage of processing results and interpretations, we described the works and analyzed their activities according to the categories listed.

## Data analysis

Book A is organized into four chapters, two aimed explicitly at statistical objects, "Noções de Estatística" [Notions of Statistics] and "Pesquisa Estatística" [Statistical Research] and occupy 67 pages of the work. Book B is organized into two chapters, with 77 pages dedicated to the "Estatística" [Statistics] chapter. In the specific chapters, both works open with texts with examples about the need and usefulness of statistics, for example, voters' ballot intentions and statistical surveys carried out by the Brazilian Institute of Geography and Statistics (Instituto Brasileiro de Geografia e Estatística - IBGE). According to the instructions for the teachers at the end of each book, those texts should be discussed with the class to map the students' previous knowledge of this field of knowledge.

Each chapter is organized into topics and subtopics, which contain content such as population, sample, variables, relative and absolute frequencies, graphical representations, errors in graphs (Book A), construction of graphs (Book B), measures of central tendency and dispersion, stem and leaf diagram and box-plot. Only Book A presents a more in-depth discussion of statistical research, covering the stages of research, making notes on sample research and exploring examples from the human development index (HDI). Only Book B has a section in the "Estatística" [Statistics] chapter called "Estatística e Probabilidade" [Statistics and Probability] that seeks to introduce the first notions of probability through interpretations of statistical tables and graphs.

The sections "Atividades Resolvidas" [Solved Activities] and "Atividades" [Activities] are proposed throughout the chapters dedicated to statistics and are where the analyses[11] of this investigation fall. The books present 47 and 87 activities, respectively, Books A and B (among these, ten –Book A– and six –Book B– are resolved in detail during the chapter). We analyzed each activity statement separately; for example, activities 1-a and 1-b received separate analyses and situations that presented two instructions, such as "Activity 23: Build a sector graph. After the construction, interpret this graph"[12]. Thus, at the end of the organization, 98 activities were classified in Book A and 164 activities in Book B.

To analyze the contexts presented in each activity, they were classified into mathematics subjects (MS), situations from the student's daily life (SDL) and themes from other areas of knowledge (OAK) (Chart 3). Also, it was necessary to list the "Não Identificado" [Unidentified] category, as four activities in Book A and five in Book B ask students to prepare questions based on presented data or organize statistical research without delimiting the theme; therefore, it is not possible to predict the context that will be addressed.

[11] These will be based on the resolutions presented in the teacher's manual.
[12] These will be referred to as activity 23-1 and 23-2 in the analyses.

**CHART 3**: Contexts identified in the activities analyzed

| Works | Context | | | | Total |
|---|---|---|---|---|---|
| | MS | SDL | OAK | Unidentified | |
| **Book A** | 0 | 43 | 50 | 5 | 98 |
| **Book B** | 13 | 89 | 58 | 4 | 164 |
| **Total** | 13 | 132 | 108 | 9 | 262 |

**Source**: Organized by the authors.

The mathematics subject was identified only in Book B, representing approximately 7.9% of the activities analyzed in the work. The activities have statements such as *"Considering the following numbers, calculate the measures of central tendency"*, in which it is necessary to mobilize the contents explored during the chapter directly.

Activities that explore another area of knowledge as context are more present in Book A, representing approximately 51% of statistics-related activities. Different themes were identified: IBGE surveys, electric power consumption, health, and object manufacturing process, among others. We understand that working with this type of activity allows students to realize the importance of statistics in solving problems in other areas of knowledge. It is also important to highlight that situations arising from different contexts, in particular, from other areas of knowledge, require the student to mobilize and coordinate different registers of semiotic representation, as the statements and/or data used to solve the activities are expressed by through tables, natural and figural language, among others.

In Book B, most activities, 53.9%, are related to the students' daily lives. These activities mostly explore issues related to school and/or the classroom, for example, organizing school work, the age of students, surveys carried out among students, school elections, and test scores, among others. Activities that address situations in students' daily lives are necessary when exploring statistical content, as they favor attributing meaning to objects in this area of knowledge.

For the thematic unit “Probabilidade e Estatística” [Probability and Statistics], the BNCC indicates ten skills; five are directly related to statistics. These aim to highlight how specific skills in mathematics and its technologies, also indicated by the BNCC, can be mobilized when exploring statistical content. Chart 4 presents the description of the skills and the classification of the activities analyzed. Considering that the skills are broad, for example, “Solve and create problems [...]” (BRASIL, 2018, p. 546), we understand that its development requires work with several and diverse activities. Therefore, activities were classified into one of the skills when they met at least one of the characteristics highlighted in the description. In other words, for the example mentioned above, it would be when the activity proposes resolving or creating a problem.

**CHART 4**: Skills related to statistics highlighted by the BNCC, mobilized in the activities analyzed

| **Statistics-related skill** | **Number of activities** | |
|---|---|---|
| | **Book A** | **Book B** |
| **EM13MAT102**: Analyze tables, graphs, and samples of statistical research presented in reports published by different media, identifying, when applicable, inadequacies that may lead to interpretation errors, such as inappropriate scales and samples. | 21 | 65 |
| **EM13MAT202**: Plan and execute sample research on relevant issues, using data collected directly or from different sources, and communicate the results through a report containing graphs and interpretation of measures of central tendency and measures of dispersion (amplitude and standard deviation), using or not technological resources. | 2 | 0 |
| **EM13MAT316**: Solve and elaborate problems in different contexts involving calculating and interpreting central tendency measures (mean, mode, median) and dispersion measures (amplitude, variance, and standard deviation). | 30 | 44 |
| **EM13MAT406**: Build and interpret frequency tables and graphs based on data obtained in statistical sample research, including or not the use of software that interrelates statistics, geometry, and algebra. | 12 | 23 |
| **EM13MAT407**: Interpret and compare statistical data sets through different diagrams and graphs (histogram, box-plot, stem and leaf, among others), recognizing the most efficient ones for your analysis. | 0 | 6 |

**Source**: Organized by the authors.

A total of 33 and 27 activities analyzed from Books A and B, respectively, were not classified according to skills, which does not mean they do not contribute to teaching statistics, they just did not fit into the skill descriptions. However, the authors of the works can understand them as a necessary path for developing skills.

The first skill mentioned, EM13MAT102, was highlighted in Book B, representing 39.4% of the activities related to statistics content in this work, and in Book A, 21.4% of the activities analyzed in the work. The activities classified, for the most part, present questions that must be solved based on the analysis of the data exposed in the statement, which, for the most part, mobilize the graphic and/or tabular semiotic register. For example, in an activity that displays a bar graph that indicates votes by sex in the election for class representative, the questions are of the type *"How many votes did candidate X receive? How many women voted for candidate X?"*. Few activities involved identifying errors in graphics; we located three (one in Book A and two in Book B).

The EM13MAT202 skill was identified only in two activities of Book A. In one of the activities, which must be developed in a group, students are asked to carry out statistical research, following the steps: choosing the topic, target audience, type of research, data collection, processing and analysis of these data and, finally, discussion and conclusion of the research. In this example, the importance of the student knowing how to mobilize different types of semiotic registers becomes clear since only in this activity, interdependent on the chosen context, natural language, tabular, graphic and numerical registers must be mobilized.

Only Book A brought discussions on how statistical research is organized. We realize that students must have contact with this type of activity since society is full of information, generally organized through statistics. Therefore, it is essential to understand how research works to avoid being fooled, as it can be used to emphasize only what one wants to show.

EM13MAT316 was proposed in 30.6% and 26.7% of the activities in Books A and B, respectively. Among the activities identified, only two in each work are related to elaborating problems based on statistical data (tabular and/or graphic representations). We must propose more activities in this format during classes, as these differ from most usual activities in which data is provided, and students only need to solve what is requested. In this way, challenging students to elaborate problems takes them out of their comfort zone as they need to understand, for example, the elaboration of data and what they provide regarding information that can be questioned through problem formulation.

Of the activities, 12.2% in Book A and 13.9% in Book B were classified in the EM13MAT406 skill, which focuses on constructing and interpreting tables. Book A has only one activity that requests students to build a frequency table and four, a graph. The other activities explore the interpretation of graphic and tabular representations. In Book B, the number of activities that require the construction of these statistical objects is higher: four and 11 activities explore the construction of tables and graphs, respectively. This result may be related to a topic in Book B that discusses the construction of graphs.

We identified EM13MAT407 skill only in activities from Book B, six in total. These explore the efficiencies of diagrams and graphs, in which one of the activities presents a set of numbers that represent the number of calls made by 40 employees of a telemarketing company and requests the construction of a stem and leaf diagram and a box diagram (Box plot). The following questions ask which representations make it easier to identify the median, the number of calls most repeated among employees (mode), and other measures.

Regarding the registers of semiotic representation, in the analysis of the cognitive transformations required in activities related to statistics in the works analyzed, we identified representations in natural language (LN),

graphic (G), tabular (T)[13], numerical (N), figural (F) and diagram (D)[14]. Five and four activities from Books A and B, respectively, were not classified, as they are personal, and the teacher's manual does not provide a possible answer for them.

Cognitive treatment transformation was promoted in 25.5% and 17.6% of the activities analyzed in Books A and B, respectively. Chart 5 indicates the representations mobilized in this type of transformation.

**CHART 5**: Cognitive treatment transformations identified in the activities analyzed

| Works | Treatment | | | | | Total |
|---|---|---|---|---|---|---|
| | LN | N | G | T | F | |
| **Book A** | 25 | 0 | 0 | 0 | 0 | 25 |
| **Book B** | 18 | 4 | 3 | 3 | 1 | 29 |
| **Total** | 43 | 4 | 3 | 3 | 1 | 54 |

**Source**: Organized by the authors.

Treatment in natural language was highlighted in both books, representing 25.5% of the activities in Book A and 10.9% in Book B. Most of these activities involve issues related to interpreting information (Chart 6), explaining in your own words or justifying your answer.

Chart 6 shows an example of treatment in natural language covered in Book A. In this activity, the student must know how to interpret information described in natural language and about data analysis, specifically, the arithmetic mean and how the data influences it. Furthermore, the activity contributes to developing the EM13MAT316 skill through a situation that students can consider everyday life.

[13] Both books use the terminology chart and table interchangeably in the statements of the activities to present their data; for this reason, we decided to categorize them together with the tables.
[14] Branch and leaf diagram.

**CHART 6**: Treatment activity in natural language in Book A

<table>
<tr><td colspan="3">13. Um levantamento feito pelos professores de um colégio concluiu que a altura média dos 405 estudantes do Ensino Médio é 1,68 m. Sabendo que eles não têm a mesma altura, analise se cada afirmação a seguir é verdadeira ou falsa e justifique sua resposta.<br>I) Há, no grupo, pelo menos um estudante com altura maior do que 1,68 m e pelo menos um que mede abaixo de 1,68 m.<br>II) Com certeza, há, nesse grupo, mais de um estudante com mais de 1,68 m de altura e mais de um estudante com altura abaixo de 1,68 m.<br>Answer:<br>13. I) Verdadeira, pois se não houvesse estudante com altura maior do que 1,68 m, a média seria inferior a esse valor. Raciocínio semelhante pode ser feito para o caso de haver um estudante com menos de 1,68 m de altura.<br>II) Falsa, pois os dados são insuficientes para afirmar com certeza que há mais de um estudante nas condições dadas.</td></tr>
<tr><td>Context:<br>Everyday Life</td><td>Skill:<br>EM13MAT316</td><td>Cognitive transformation:<br>Natural language treatment</td></tr>
</table>

**Source**: Bonjorno, Giovanni Júnior, and Sousa (2020, p. 13, 240).

Book A explored the treatment activity only in natural language representation, while Book B also presented activities with treatments in numerical (2.4%), graphic (1.8%), tabular (1.8%) and figural (0.6%) representations. Activities with numerical treatment in their statements list a set of numbers and only request the calculation of measures of central tendency, remaining only in numerical representation. The activities involving treatments in graphical representation presented graphs, such as histogram, frequency polygon, pictogram, and box-plot. Their statements involve situations in which, based on a graph, students need to construct or analyze another graph. We found tabular treatments in only three activities in Book B, in which their statements present the data in tables, and from there, the organization of a frequency table is requested. Figural treatment was proposed in just one activity in Book B to solve a question based on a given a sector graph; students must mobilize the interpretation of the statement in natural language representation –according to Moretti (2022, p. 95), "[...] text comprehension is an essential prerequisite in mathematical activity"– then, perform percentage calculations (numerical representation) and, finally, return to the figural representation.

A greater emphasis was placed on conversion activities, with 69.4% and 78.8% in Books A and B, respectively. Different arrival and departure representation combinations were identified, 24 in total (considering both works). Chart 7 shows the most explored conversions in the analyzed activities. The "Outros" [Others] category groups the types of conversions that had fewer than five activities categorized in each book.

**CHART 7**: Cognitive conversion transformations identified in the analyzed activities

| Work | Conversion | | | | | | Total |
|---|---|---|---|---|---|---|---|
| | LN-N | G-N | G-LN | T-N | T-LN | Others | |
| **Book A** | 8 | 5 | 17 | 14 | 9 | 15 | 68 |
| **Book B** | 17 | 28 | 26 | 24 | 6 | 29 | 130 |
| **Total** | 25 | 33 | 43 | 38 | 15 | 44 | 198 |

**Source**: Organized by the authors.

The GN conversion[15] was highlighted in Book B, covering 17% of activities related to statistics, while in Book A, they represent 5.1% of the analyzed activities. For the most part, the activities identified in this type of conversion propose questions resolved through calculations based on data expressed in the graphs. An example is the activity submitted by Book B, presented in Chart 8.

In this activity (Chart 8), students must understand the data available in the graph, calculate the estimated production reduction for each highlighted state, calculate the average reduction, and resolve the issue. It involves a context from another area of knowledge, as it deals with soybean production, and mobilizes the EM13MAT316 skill, as it requires the resolution of a problem based on the calculation of measures of central tendency, in this case, the mean.

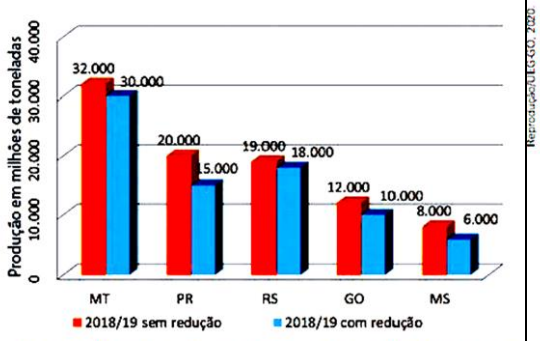
[15] It reads: Conversion from graphical to numerical representation.

**CHART 8**: Activity that involves converting graphical representation to numerical representation in Book B

**57. (UEG-GO)** A produção de soja em 2019, nos principais estados produtores do Brasil, tem uma previsão de redução em relação à produção de 2018. O gráfico a seguir mostra essa previsão de redução de maneira aproximada.

Caso a previsão de redução se confirme, a média de redução entre os estados de Mato Grosso, Rio Grande do Sul e Goiás, em milhões de toneladas, será aproximadamente igual a: Alternativa **e**.

**a)** 1304
**b)** 1027
**c)** 1532
**d)** 1120
**e)** 1667




Disponível em: https://coopadapsementes.com.br/site/impactos-da-estiagem-na-producao-da-soja/. Acesso em: 16 set. 2019. (Adaptado).

**Answer:**

57. Redução em MT: 32000 − 30000 = 2000
Redução em RS: 19000 − 18000 = 1000
Redução em GO: 12000 − 10000 = 2000
Assim, a média de redução em MT, RS e GO é:

$$\frac{2000 + 1000 + 2000}{3} \simeq 1667$$

Portanto, a média de redução entre esses estados é de 1667 milhões de toneladas.
Alternativa **e**.

| **Context:** | **Skill:** | **Cognitive transformation:** |
|---|---|---|
| Everyday Life | EM13MAT316 | Natural language treatment |

**Source**: Dante and Viana (2020, p. 56, 212).

We identified 43 activities exploring the G-LN conversion in the two books. In Book A, these correspond to 17.3% of the activities that involve statistical objects, with conversion receiving greater emphasis in this work and 15.8% in Book B. The activities that promote this type of conversion, for the most part, are those that explore the analysis and/or interpretation of graphs, that is, situations in which only the data displayed in the graphs solve the question (Chart 9) and/or that require students to express what they understand based on the information provided in this representation.

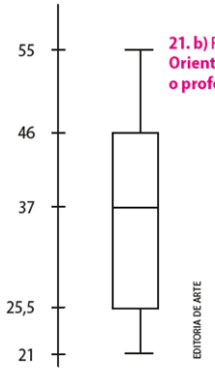
**CHART 9**: Activity that involves converting the graphical representation to natural language in Book A

**21.** Observe o *box-plot* a seguir, que indica dados referentes a uma turma de Educação de Jovens e Adultos (EJA), e responda às questões.

**b)** Podemos dizer que 75% dos estudantes dessa turma têm menos do que 46 anos? Explique sua resposta.

**c)** Há mais estudantes com idade entre 21 e 25,5 anos ou entre 46 e 55 anos? Justifique sua resposta.

**d)** O que 37 indica nesse gráfico?




**21. b)** Resposta nas **Orientações para o professor.**

**Answer:**

b] Sim. Resposta esperada: O terceiro quartil indica que 25% dos dados estão acima dele, portanto, 75% estão abaixo.

c] Há mais estudantes entre 46 e 55 anos. Resposta esperada: É possível concluir isso porque a linha do terceiro quartil até o valor máximo é maior do que a linha do primeiro quartil até o valor mínimo.

d] O número 37 indica que metade dos estudantes tem idade abaixo de 37 anos; a outra metade, acima de 37 anos.

| **Context:** Everyday Life | **Skill:** EM13MAT102 | **Cognitive transformation:** G-LN Conversion |
|---|---|---|

**Source**: Bonjorno, Giovanni Júnior, and Sousa (2020, p. 44, 242).

The activity in Book A (Chart 9) requires a G-LN conversion. To solve it, it is necessary to analyze a box plot and answer questions based precisely on the elements highlighted in this representation: quartiles, maximum value, minimum value, and median. It is worth noting that this activity allows the development of the EM13MAT102 skill, as it requires analyzing a graph, and regarding the context, it can be classified as students' everyday lives because it concerns the ages of a class.

When proposing their activities, the books analyzed differed in terms of the most used set-off representation. Book A used tabular representation in 26.5% of its activity statements. Book B privileged activities that present graphic representation in their statements, with 32.7% of the activities analyzed in the work. These results align with Pallauta and collaborators' (2021) findings,

who, when investigating Spanish textbooks, verified the predominance of activities with these set-off, graphic and tabular representations.

Duval (2011) points out in his theory the importance of exploring conversions in various ways, for example, activities that promote G-N conversion and activities that mobilize N-G conversion because, according to the author, carrying out conversion activities in just one sense is not associated with understanding this conversion the opposite way. We found inverse conversions between the representations: in Book A, "G and LN", "LN and D", "LN and N" and "LN and T", and in Book B, "G and LN", "G and N" and "LN and T". However, there are few activities that mobilize reverse conversion. For example, in Book A, 17 activities promote G-LN conversions, whereas only two LN-G conversions exist. None of the combinations of representations in which inverse conversions occur exceeds the amount of two activities in one of the directions. In other words, the reverse conversion occurs in the works analyzed, but even so, one of the meanings ends up being privileged.

## Final discussion

The contexts of everyday life and other areas of knowledge predominated among the activities. However, they were used in most activities only to illustrate the data, becoming just a "background context", as when solving the activity, it was directly related to the statistical content. For example, "What is the arithmetic mean? And what about moda?" There is a need for more activities that go beyond specific questions, exploring students' understandings and interpretations of data, graphs, and statistical information, making them attribute meaning to statistical concepts.

As already pointed out, the BNCC advises that high school students should be able to carry out sample research, with one of the forms of communicating results being the creation of graphs based on the data collected. The analyses indicate that books lack information on these subjects. Only Book A presents discussions on how to prepare a sample

survey, and only in Book B were notes on graph construction identified. It is important to highlight that, even though BNCC emphasizes it, we found very few activities that require the construction of graphs (three activities in Book A and nine in Book B).

We identified that the distribution of activities in terms of skills related to statistics indicated by the BNCC is very disproportional. The EM13MAT202 and EM13MAT407 skills are little explored, with two activities in Book A and six in Book B, respectively.

Different types of semiotic representations were identified in the activities, emphasizing natural, tabular, and graphic languages. The conversion activity stood out in both works, emphasizing graphical to numerical conversions in Book A, and graphical to natural language in Book B. Although there are few activities on graphic construction, this type of register is predominant as a set-off representation in activities, together with tables. Most activities involve observing the data expressed in tables and graphs and statistical calculations. The analyses are close to understanding Duval's theory (2011), except for one point: the inverse conversions. The author emphasizes the importance of exploring this type of conversion. However, in both books, these situations are rare.

Based on the results, the importance of analyzing textbooks is once again emphasized. It is essential to be careful when choosing a textbook and searching for other materials to plan and teach classes. Because, from the analysis of two works approved by the PNLD, we see that, concerning statistics based on the BNCC and the theory used, both works have positive points, but specific issues can be better explored in the activities proposals.

Finally, we point out that this research is the starting point of an ongoing investigation that seeks to assist mathematics teachers in organizing their statistics classes to provide the development of statistical knowledge proposed under the guidelines of the BNCC and the organization of the New High School.

## References

BARDIN, L. *Análise de Conteúdo*. 4. ed., Lisboa: Edições 70, 2022.

BAYER, A. *et al*. Formandos em Matemática x Estatística na Escola: Estamos Preparados?. *In:* XII Simpósio Sul brasileiro de Ensino de Ciências, 2004, Canoas. *Anais* [...], Canoas: ULBRA, v.1, p. 1-12, 2004.

BONJORNO, J. R.; GIOVANNI JÚNIOR, J. R.; SOUSA, P. R. C. de. *Prisma matemática*: estatística, combinatória e probabilidades: érea do conhecimento: matemática e suas tecnologias. 1. ed. São Paulo: FTD, 2020.

BORBA, C. M. A Pesquisa Qualitativa em Educação Matemática. *In*: 27ª reunião anual da ANPEd, 2004, Caxambu. *Anais da 27ª reunião anual da Anped*, Caxambu, 2004.

BRASIL. *Base Nacional Comum Curricular*. Brasília: Ministério da Educação, Secretaria de Educação Básica, 2018.

BRASIL. *Guia Digital PNLD 2021*: Matemática e suas Tecnologias/Brasília: Ministério da Educação, Secretaria de Educação Básica, 2020.

COSTA, A.; NACARATO, A. M. A Estocástica na Formação do Professor de Matemática: percepções de professores e de formadores. *Bolema*, Rio Claro, v. 24, n. 39, p. 367-386, nov. 2011.

DANTE, L. R.; VIANA, F. *Matemática em contexto: estatística e matemática financeira.* 1. ed. São Paulo: Ática, 2020.

DUVAL, R. Registres de représentation sémiotique et fonctionnement cognitif de la pensée. *Annales de Didactiques et de Sciences Cognitives*, Strasbourg, v. 5, p. 35-65, 1993.

DUVAL, R. Registros de representação semiótica e funcionamento cognitivo da compreensão em matemática. *In*: MACHADO, S. D. A. *Aprendizagem em matemática:* Registros de representação semiótica. Campinas: Papirus, 2003. p. 11-33.

DUVAL, R. *Semiosis y Pensamiento Humano*. Registres sémiotiques et apprentissages intellectuels: Santiago de Calai, Colômbia: 2004.

DUVAL, R. *Ver e ensinar a matemática de outra forma: entrar no modo matemático de pensar:* os registros de representações semióticas. Org. Tânia M. M. Campos. 1. ed. São Paulo: PROEM, 2011.

DUVAL, R. Entrevista: Raymond Duval e a Teoria dos Registros de Representação Semiótica. Concedida a FREITAS, de. J. L. M.; REZENDE, V. *Revista Paranaense de Educação Matemática*, Campo Mourão, v. 2, n. 3, 2013. DOI: https://doi.org/10.33871/22385800.2013.2.3.10-34.

IBGE. *Centro de Documentação e Disseminação de Informações*. Normas de apresentação tabular I Fundação Instituto Brasileiro de Geografia e Estatística, Centro de Documentação e Disseminação de Informações. 3. ed., Rio de Janeiro: IBGE, 1993.

MORETTI, M. T. A COMPREENSÃO DE TEXTO SEGUNDO RAYMOND DUVAL. *Educação Matemática Sem Fronteiras*: Pesquisas em Educação Matemática, Chapecó, v. 4, n. 2, p. 92 - 112, 2022. DOI: https://doi.org/10.36661/2596-318X.2022v4n2.13229.

PALLAUTA, J. D. et al. Significado de la tabla estadística en libros de texto españoles de educación secundaria. *Bolema*, Rio Claro, v. 35, n. 71, p. 1803-1824, 2021. DOI: https://doi.org/10.1590/1980-4415v35n71a26.

PIETROPAOLO, R. C.; SILVA, A. da F. G.; AMORIM, M. É.. Conhecimentos de Professores de Matemática para o ensino de noções relativas à Estatística na Educação Básica. *Revista Eletrônica de Educação Matemática – REVEMAT*, Florianópolis, v.14, Edição Especial Educação Estatística, p.1-20. 2019. DOI: http://doi.org/105007/1981-1322.2019.e62757.

Received in June 2023.

Approved in November 2023.