

Characterisation of the reading levels of information presented by Chilean secondary students in tasks with statistical tables

Jocelyn D. Pallauta ^a 
 Carmen Batanero ^b 
 Pedro Arteaga ^b 

^a Universidad de Los Lagos, Departamento de Ciencias Exactas, Osorno, Chile
^b Universidad de Granada, Departamento de Didáctica de la Matemática, Granada, España

ABSTRACT

Background: Developing the ability to interpret statistical information at school critically will allow students to make informed decisions in uncertain contexts and to participate reflectively in society. **Objectives:** In this work, the ability of Chilean students to interpret information presented in statistical tables is analysed in the transition from elementary to high school. **Design:** An interpretive paradigm with a qualitative approach is followed, employing content analysis to identify, classify, and evaluate qualitative categories, thereby enabling an in-depth understanding of the data. **Setting and participants:** 191 Chilean students from the 7th grade of elementary school (14 to 15 year-olds) and the 1st grade of secondary education (15 to 16 year-olds) participated. **Data collection and analysis:** Students individually solved two tasks on statistical tables. Data analysis was carried out through the written productions, considering the interpretation and the reading level reached. **Results:** Most answers were correct or partially correct, although the results were lower than those of previous studies, and the most frequent reading level was intermediate, focused on value comparisons. Notably, a significant proportion of students do not reach a minimum reading level. **Conclusions:** It is important to propose to secondary school students tasks that develop their capacity for critical reading of the information presented in statistical tables, promoting reflective interpretations linked to their daily experiences

Keywords: reading levels; statistical tables; secondary education; critical interpretation.

Corresponding author: Jocelyn D. Pallauta. Email: jocelyn.diaz@ulagos.cl

Caracterización de los niveles de lectura de la información presentada en tablas estadísticas por estudiantes chilenos de educación secundaria

RESUMEN

Contexto: Desarrollar en la escuela la capacidad de interpretar críticamente la información estadística permitirá a los estudiantes tomar decisiones informadas en contextos de incertidumbre y participar de manera reflexiva en la sociedad. **Objetivos:** En este trabajo se analiza la capacidad de los estudiantes chilenos para interpretar información presentada en tablas estadísticas al pasar de la educación básica a la secundaria. **Diseño:** Se sigue un paradigma interpretativo con un enfoque cualitativo, empleando el análisis de contenido para identificar, clasificar y evaluar categorías cualitativas, lo que permite una comprensión en profundidad de los datos. **Ambiente y participantes:** Participaron 191 estudiantes chilenos de 7° curso de educación básica (14 a 15 años) y de 1° curso de educación secundaria (15 a 16 años). **Recolección y análisis de los datos:** Los estudiantes resolvieron individualmente dos tareas sobre tablas estadísticas. El análisis de los datos se realizó a través de las producciones escritas, considerando la interpretación y el nivel de lectura alcanzado. **Resultados:** La mayor parte de las respuestas fue correcta o parcialmente correcta, aunque los resultados fueron inferiores a los de estudios previos y el nivel de lectura más frecuente es intermedio, centrado en la comparación de valores. Cabe destacar que una proporción importante de estudiantes no alcanza un nivel mínimo de lectura. **Conclusiones:** Es importante proponer a los estudiantes de secundaria tareas que fomenten el desarrollo de la capacidad de lectura crítica de la información presentada en tablas estadísticas, promoviendo interpretaciones reflexivas vinculadas con sus experiencias cotidianas.

Palabras clave: niveles de lectura; tablas estadísticas; educación secundaria; interpretación crítica.

Caracterização dos níveis de leitura das informações apresentadas por alunos chilenos em tarefas com tabelas estatísticas.

RESUMO

Contexto: O desenvolvimento da capacidade de interpretar criticamente a informação na escola permitirá aos alunos tomar decisões informadas em contextos de incerteza e participar de forma reflexiva na sociedade. **Objetivos:** Este artigo analisa a capacidade dos alunos do ensino secundário chilenos de interpretar informações apresentadas em tabelas estatísticas. **Desenho:** Este estudo adota um paradigma interpretativo, com abordagem qualitativa, empregando a análise de conteúdo para identificar, classificar e avaliar categorias qualitativas, o que permite uma compreensão em profundidade dos dados. **Cenário e participantes:** Participaram 191 estudantes chilenos do 7° ano do ensino básico (14 a 15 anos) e do 1° ano do ensino secundário (15 a 16 anos). **Recolha e análise de dados:** Os alunos resolveram individualmente duas tarefas com tabelas estatísticas. A análise dos dados foi realizada por meio de

produções escritas, considerando a interpretação e o nível de leitura atingido. **Resultados:** A maioria das respostas foi correta ou parcialmente correta, embora os resultados tenham sido inferiores aos de estudos anteriores e o nível de leitura mais frequente tenha sido o intermédio, centrado na comparação de valores. Destaca-se que uma percentagem significativa de alunos não atingiu um nível mínimo de leitura. **Conclusões:** É importante propor aos alunos do ensino secundário tarefas que desenvolvam a capacidade de leitura crítica da informação apresentada em tabelas estatísticas, promovendo interpretações reflexivas ligadas às suas vivências cotidianas. **Palavras-chave:** níveis de leitura; tabelas estatísticas; ensino médio; interpretação crítica.

INTRODUCTION

The abundance of information on social, political, and economic issues disseminated through the press, television, and the internet demands that citizens interpret and critically analyse the data (Gal, 2022). Thus, the ability to read and interpret tables and charts is part of statistical literacy (Weiland, 2017). This competence is necessary for developing students' critical thinking and their ability to make judgments when faced with data-based statements (Friedrich et al., 2024).

Statistical tables and charts are used to summarise different types of data and variables (Bargagliotti et al., 2020), to identify trends and correlations, and to help recognise when data are misused (OECD, 2021). In particular, statistical tables are widely used to organise and record information clearly and accessibly, both in everyday life and in professional and scientific fields (Bartram et al., 2022). In addition, statistical tables play an important role in data registration and organisation, as well as serving as a preliminary step in the construction of various types of graphs (Vásquez et al., 2022).

Given its relevance, the study of statistical tables is included in the curricula of different countries, starting in primary education (Gea et al., 2022), as an object of study in its own right and as a means of addressing various topics, especially statistics and probability. Specifically, in Chile, the curriculum guidelines (MINEDUC, 2015) propose for the 7th and 8th grades of basic education (12 to 13 years) the study of the distribution tables of a variable with absolute and relative frequencies. For the 1st grade of secondary education (13 to 14 years), the double-entry table is used to analyse two-dimensional distributions, whereas in the 2nd grade of secondary education (14 to 15 years), the statistical table is used to represent the distribution of a discrete random variable.

Despite the widespread presence of tables in curriculum guidelines and textbooks across educational levels (Gea et al., 2022; Pallauta et al., 2021b; 2023b), few studies examine their interpretation, especially among high school students. Most research studies on this topic have been conducted with elementary students (e.g., Díaz-Levicoy et al., 2020; Guimarães et al., 2021), focus on double-entry tables (Castellaro & Roselli, 2020; Gabucio et al., 2010), or provide information from small samples of students (Pallauta et al., 2023a).

To complete this information, the objective of this work is to analyse the ability of Chilean secondary school students to interpret information presented in statistical tables. For this, we plan the following research questions:

1) How do Chilean students interpret the information presented in statistical tables?

2) What reading levels do Chilean high school students achieve in tasks that involve statistical tables?

The following sections present the rationale behind the work, followed by the method and the results. It ends with the conclusions and didactic suggestions.

FUNDAMENTALS

Structure and types of statistical tables

In this work, we will understand the table as a representation that organises quantitative information in rows and columns, enabling us to structure data or elements related to each other (Campbell-Kelly et al., 2003). The numerical values located in the inner cells are interpreted by combining the concepts arranged in the headings of the corresponding rows and columns (Martí, 2009).

It is possible to distinguish the following variety of statistical tables, each with specific characteristics and cognitive complexity, whose understanding involves processes of interpretation, generalisation, representation, and symbolisation (Pallauta et al., 2023b):

- *Data table.* It is used to record the values of variables for each individual or element of the study. Its use is simple and allows for identifying and associating, in sequence, the values of each variable with each individual sequentially. This type of table does not present the concept of frequencies or frequency distribution.

- *Table of the frequency distribution of a variable.* It allows us to organise the information by the frequency of each value of the represented variable. This type of table is divided into three types: a) *Table with ordinary frequencies.* Represents the distribution of absolute, relative, or percentage frequencies of a qualitative or discrete variable; b) *Table with cumulative frequencies.* Represents the distribution of cumulative frequencies of a variable. usually discrete; c) *Table with data grouped into intervals.* In this case, the values of the continuous or discrete variable are presented through data grouped in intervals. In addition to the frequencies and distribution, we work with the concepts of interval, end, and class mark.
- *Double-entry table.* It represents a two-dimensional variable, so its handling is more complex, since it requires identifying the values of each variable and its crossing (in rows and columns). In this way, three types of distributions are represented: joint, marginal, and conditional. Two types are differentiated: a) *Double-entry table with categorical data.* When qualitative or discrete variables are used; and b) *Double-entry table with data grouped into intervals.* When the values of one or both variables are presented through data grouped in class intervals.

Reading levels of charts and statistical tables

The interpretation of statistical tables and graphs involves addressing questions of varying difficulty, whose complexity has been characterised by various authors. Curcio (1989) developed a three-level hierarchy to analyse students' graphic understanding, which Shaughnessy (2007) expanded to four levels. The four levels are collected in Friel et al. (2001), which is the model we use in this work, and is described below:

- *N1. Reading the data.* It focuses on interpreting direct questions about the structure and identification of elements in statistical tables. For example, indicating the frequency of a certain modality or the value of the variable represented in a table.
- *N2. Reading between the data.* It involves establishing relationships or performing operations with the information presented in the table. An example is determining the mode or median.
- *N3. Reading beyond the data.* It requires making inferences from the information in the table, such as interpolating or extrapolating a value that is not explicitly given within a series of data over time.
- *Level 4. Reading behind the data.* It involves a critical evaluation of the information presented in the table, establishing connections with the

context and reliability of the data. For example, questioning possible biases in the claims derived from the table, or reflecting on the implications of the data in their context.

Background

The first group of investigations that support our work has analysed high school students' interpretation of statistical tables. Since the level of reading achieved by students in the statistical tables will depend on their familiarity with the task, research on the level of reading of statistical tables promoted in the activities proposed in the textbooks has also been considered.

Regarding the interpretation of the statistical tables by high school students, Gabucio et al. (2010) analysed the reading level achieved by 200 Spanish students from 5th grade to 2nd grade of high school (10 to 14 years) through a questionnaire composed of 12 multiple-choice items, related to a double-entry table. The authors, starting from Friel et al. (2001) 's framework, proposed four levels of table reading: (a) understanding the tabular structure, or knowledge of the different elements of the table; (b) direct reading of the data (analogous to level N1 by Friel et al., 2001); (c) data inference (allows finding information not directly represented in the table and combines levels N2 and N3 by Friel et al., 2001; and (d) global interpretation (equivalent to level N4 by Friel et al., 2001). The items that were most difficult were those that involved inference of the data (levels N2 and N3), with a percentage of achievement between 29% and 47%, and the simplest corresponded to the direct reading of the data (level N1), with a percentage of success between 80.5% and 89.5%. On the other hand, there was hardly any increase in the correct answers as the course progressed. Similar results were obtained by Castellaro and Roselli (2020) in a replication with 90 Argentine students aged 12 and 13.

Prodromou (2015) characterised the interpretation of data tables on different social issues (e.g., demographic change, migration) by eight groups of four secondary school students (15 years old) from Cyprus, using the reading levels by Friel et al. (2001). The results showed that students' interpretations of the information in the tables reached four reading levels, with some groups focusing on basic levels (N1 and N2) and others on more advanced levels (N3 and N4). They do not report the number of students who reason at each level.

González et al. (2021) analysed how ten Chilean 4th-grade high school students (15 to 16 years) interpreted frequency tables with real-world data grouped into intervals in the context of the COVID-19 pandemic. The investigators posed three questions linked to the reading levels proposed by

Curcio (1989): reading the data (N1), reading between the data (N2), and reading beyond the data (N3). The researchers obtained good results at the N1 level. In contrast, the N3 level was more complex, as students relied on personal opinions rather than information, underscoring the need to reinforce the development of more sophisticated reading levels.

Pallauta et al. (2023a) developed a questionnaire, from which the items considered in this work are drawn, that addresses the main variables and skills identified in the research presented in the statistical tables included in the textbooks, as well as the curriculum guidelines. In addition to describing these variables and the construction and validation of the questionnaire, data are provided on the difficulty indices from a sample of 128 Spanish 3rd-grade high school students (14 to 16 years). The authors describe the main difficulties students had in interpreting statistical tables, as well as the reading levels achieved in some items of the questionnaire. In the first task proposed in our work (Figure 1), the percentage of correct answers was 52%, while 55.5% of the sample reached the N2 level (reading between the data) according to Friel et al. (2001), and 40.6% did not reach the N1 level. In the second task (Figure 2), 67.2% and 68% did not reach the N1 level in sections a and b, with correct answer rates of only 16% and 10% in sections a and b, respectively. Given that the sample was limited to students from one course and country, this work analyses in more detail the level of reading and the answers to some of these items from a larger sample of 1st- and 3rd-grade high school students (14 to 16 years) from Chile.

Other research addresses errors and difficulties in the construction of tables (Álvarez et al., 2020), the translation of charts into tables and their subsequent interpretation (Pallauta et al., 2021a), providing information that contributes to establishing the categories of analysis of our study.

In relation to the studies on the tasks on statistical tables in the textbooks, Pallauta et al. (2021b) analysed the reading levels (Friel et al., 2001) promoted in the tasks on statistical tables of six Chilean and six Spanish textbooks in the first and second grades of high school (12 to 13 years). In both countries, a predominance (approximately 90%) of tasks that required levels N2 (reading between the data) and N1 (reading the data) were observed. The most advanced levels, N3 (reading beyond the data) and N4 (reading behind the data), had a low presence; the highest percentage of tasks with the N4 level (15%) appeared in the 2nd grade of secondary education in Chile. Similar results are reported in Gea et al. (2024), who analysed ten textbooks from 1st to 4th grades of secondary education published in accordance with the current

curricular regulations in Spain (MEFP, 2022). Most tasks focus on the N2 level (between 67.1% to 93.5%, depending on the course and editorial). These investigations warn that most of the tasks in statistical tables in textbooks only promote basic reading levels, neglecting the development of critical thinking and students' statistical culture (Weiland, 2017).

METHODOLOGY

Participants

The work was carried out with an intentional sample of 191 Chilean students from five schools in different provinces of the country: one public, three concerted and one private. The research was carried out in accordance with ethical principles, with the informed consent of the teachers and directors of the participating schools. The group consisted of 103 students in the 7th grade of basic education (BE) (14 to 15 years) and 88 students in the 1st grade of secondary education (SE) (15 to 16 years). The distribution of the sample by gender and course is presented in Table 1. The distribution by gender was similar, with slightly higher participation by women (52.9%) than by men (47.1%).

Table 1

Frequency (F) and percentage (%) distribution of the sample by gender and course.

Course	Male N = 90		Female N =101	
	F	%	F	%
7th BE	49	25.7	54	28.3
1st SE	41	21.5	47	24.6

Data collection instrument

Students were asked two tasks from the questionnaire developed by Pallauta et al. (2023a). The tasks were designed based on an analysis of the curriculum guidelines, previous research, and Chilean and Spanish textbooks. The selection of the tasks was based on expert judgment to ensure the validity of the content. For this study, a task based on a data table (Task 1) and another based on a distribution table of a variable with data grouped into intervals (Task 2) are considered, both with a reading level N4 (Friel et al., 2001).

Figure 1

Task 1: Reason from a data table. (Pallauta et al., 2023a, p.111)

Al medir la altura en cm. que pueden saltar un grupo de niñas, antes y después de haber efectuado un cierto entrenamiento deportivo, se obtuvieron los siguientes valores:

Altura saltada en cm.										
Estudiante	Ana	Bea	Carol	Diana	Elena	Fanny	Gía	Hilda	Inés	Juana
Antes del entrenamiento	115	112	107	119	115	138	126	105	104	115
Después del entrenamiento	128	115	106	128	122	145	132	109	102	117

Responde
¿Piensas que el entrenamiento es efectivo? Justifica tu respuesta.

Task 1 (Figure 1), which was also used by Batanero (2000), was presented to students in both groups. They received a table showing the heights (in cm) reached by a group of girls before and after training. The exercise asks students to reason about whether the sports training was effective and to justify their answers, which requires an N4-level reading (Friel et al., 2001).

Task 2 (Figure 2), proposed only in the 1st grade of SE, requires a critical analysis of the information presented in a distribution table for a single variable with data grouped into intervals. It consists of two sections: the first (2.a) involves reading an accumulated frequency, while the second (2.b) requires analysing the data trend. Both issues demand a critical reading of the information, corresponding to an N4-level reading (Friel et al., 2001).

Figure 2

Task 2: Critically analyse information in a table with data grouped in intervals. (Pallauta et al., 2023a, p.111)

Se realiza un estudio sobre el tiempo de eficacia al aplicar un medicamento a un grupo de hombres. La siguiente tabla muestra los tiempos que demoran en hacer efecto.

Responde

a) Existen en total 20 hombres que reaccionan al medicamento en menos de 60 minutos
¿Estás de acuerdo con esta afirmación? ¿Por qué?

b) ¿Te parece que el medicamento tiene un efecto rápido o lento? ¿Por qué?

Tiempo de demora en un grupo de hombres		
Tiempo (min.)	Frecuencia absoluta	Frecuencia acumulada
[0, 15[3	3
[15, 30[12	15
[30, 45[18	33
[45, 60[20	53

Analysis procedure

This study is developed within the interpretative paradigm and adopts a qualitative approach (Cohen et al., 2018). For data analysis, the content analysis technique is used (Schreier, 2014), enabling the identification, organisation, and evaluation of qualitative categories. This approach facilitates an in-depth interpretation of the collected information, ensuring informed, contextualised conclusions.

Each answer in the different sections of the tasks constitutes the unit of analysis. To illustrate the identified categories, some examples of these responses are included, each of which was assigned a code of the form Ex, where x corresponds to 1–191.

For the analysis of the answers, an inductive and cyclical procedure included the following stages: 1) the answers to each of the two tasks proposed were coded according to the categories: correct, partially correct, and incorrect, and additionally the level of reading achieved was identified (Friel et al., 2001); 2) To ensure the reliability of the coding, 15% of the answers were analysed independently by another researcher and, in case of disagreement, they were discussed with the other authors until an agreement was reached; 3) Summarised tables of the results were prepared to facilitate extracting conclusions.

RESULTS

The results for the students' interpretation and reading levels on the tasks set are presented below.

Interpretation of the tables

To answer the first investigative question (How do Chilean students interpret the information presented in statistical tables?), the following sections present the results of the two tasks posed to Spanish and Chilean students.

In Task 1, to analyse the information presented in a data table and reason in an argued way about whether there is an effect of the training, Batanero (2000) suggests comparing the averages of the two groups or considering another statistic that takes into account both complete datasets. Student answer categories are described below.

Correct answer. The information presented in the table is analysed, showing an improvement in the girls' performance after the training. For

example, student E74 reports the percentage of girls who improved their previous mark.

E74: I believe that training works, since 80% of the girls have increased the height of their jump, but 20% have decreased in height.

Partially correct answer. An answer that disregards some of the data set; it considers only a part to decide whether the training is effective. E90, for example, shows evidence of a local conception of the association (Estepa et al., 1999), which involves judging the relationship between two variables based only on a subset of the available data.

E90: According to the person, because Carla and Inés lowered the jump by centimetres.

Incorrect answer. Answers that indicate an inadequate interpretation of the information and the context of the data, for example, E15, which refers to the number of jumps rather than their height.

E15: Because in most cases, after training, the students could jump more times.

Table 2 presents the results per course; we note that the SE 1st-grade students achieve better results, both in the percentage of hits and in partially correct answers. In the 7th grade of BE, a significant percentage of students answer incorrectly. In both courses, only a few students fail to complete the task. When comparing the proportion of correct answers from the 1st grade of the SE with Pallauta et al.'s (2023a) results, we observe lower performance (51.6%). However, there are better results in the percentage of incorrect answers (21.8%) and unanswered (21.1%).

Table 2

Frequency (F) and percentage (%) of answers of Task 1, according to the course.

Assessment	7th BE N = 103		1st SE N = 88	
	F	%	F	%
Correct	22	21.4	27	30.7
Partially correct	39	37.9	46	52.3
Incorrect	35	34.0	14	15.9
No answer	7	6.8	1	1.1

Task 2 was proposed only to students in the 1st grade of SE, a datum that proved problematic in Pallauta et al.'s (2023a) study. It consists of two sections in which students must justify the truthfulness of statements based on information presented in a distribution table of a single variable, with data grouped into intervals. To do this, students must apply the concepts of absolute and accumulated frequencies and inequality directly associated with working with class intervals (Pallauta et al., 2023b). The categories for evaluating the answers in both sections are detailed below.

Correct answer. In the first section (2.a), the statement must be contrasted with the information presented in the table, establishing the frequency associated with a set of values less than a specific value. The student is expected to decide whether the statement is false, because 53 men react to the drug in less than 60 minutes, as indicated by E12, who identifies the addends that allow obtaining the total of the sample.

In the second section (2.b), the reaction to the drug is considered slow, according to the information in the table. For example, E35 notes that the number of people who react in a short time is minimal compared to those who react in the maximum time.

E12: No, because $3+12+18+20$ does not equal 20 (Section 2.a).

E35: Slow, because for three people the drug takes less than 15 minutes to be effective, and for 20 people, it takes 45 minutes or more (Section 2.b).

Partially correct answer. This answer in 2.a evidences a limited reading of the information in the table, because it only considers a part of the required data. For example, E8 only considers a part of the intervals that meet the condition.

In section 2.b, the decision is based on the observation of a too wide span of values, as in E4's answer.

E8: Thirty-three men react to the drug in less than 60 minutes (Section 2.a).

E4: Slow, since most react in more than 15 minutes (Section 2.b).

Incorrect answer. These answers indicate an inadequate interpretation of the table or the question posed. In 2.a, they agree with the statement, as we note in E48's answer. In 2.b, in turn, E41 applies a personal criterion to justify their response.

E48: Yes, because the absolute frequency is 20 (Section 2.a).

E41: Slow, because the medication, as it is swallowed and takes effect, takes a little while; as long as it works, it is fine (Section 2.b).

The results of Task 2 are presented in Table 3. We can observe that in the first section of the task (2.a), the percentage of correct answers is slight, with the most frequent being the partially correct answers. Incorrect and unanswered answers accounted for a significant percentage.

In the second section (2.b), the percentage of correct answers is increased, decreasing the partially correct ones. However, the number of incorrect answers increases significantly. The percentage of correct answers in Pallauta et al. (2023a) was lower in both sections of the task (2.a: 16.4%; 2.b: 11.7%); in addition, a large percentage of students did not respond to the task (2.a: 43.8%; 2.b: 43.0%).

Table 3

Frequency (F) and percentage (%) of answers of Task 2, according to the section.

Assessment	Task 2 1st SE N = 88			
	Section 2.a		Section 2.b	
	F	%	F	%
Correct	21	23.9	32	36.4
Partially correct	32	36.4	13	14.8
Incorrect	26	29.5	36	40.9
No answer	9	10.2	7	8.0

Table 4 presents the difficulty index values for both tasks, calculated by assigning a score of 1 to correct or partially correct answers and 0 to incorrect or partially incorrect answers, along with their standard deviations.

Table 4*Difficulty index of Task 1.*

	Task 1		Task 2	
	7th BE N = 103	1st SE N = 88	1st SE N = 88	
			2.a	2.b
Mean	0.59	0.83	0.60	0.51
Standard Deviation	0.494	0.378	0.492	0.503

Task 1 was more complex for the 7th-grade students of BE, as the analysis of differences in the means of successes by course showed a statistically significant difference ($t = -3.76$; $p = 0.000$). In Task 2, proposed only for the SE 1st graders, both sections have a similar difficulty, with section 2.b being slightly more complex, possibly because all the information presented in the table had to be considered to argue the answer.

Reading levels of tables

To answer the second research question (What reading levels do Chilean high school students achieve in tasks that involve statistical tables?), the reading levels proposed by Friel et al. (2001) were used to analyse how Spanish and Chilean students justified their answers to the two tasks.

Task 1 reaches the N4 reading level (Friel et al., 2001) because it requires critical reading of the information in the table. It is worth noting that the N3 level (reading beyond the data) is not considered, as data interpolation or extrapolation is not performed in this task. Here are the answers that characterise the different levels achieved by the students.

N0. They do not read data. In this category, we considered the justifications that reveal the roots of the personal beliefs described by Gal y Trostianitser (2016), rather than the information presented in the table, when making judgments, as E8 does.

E8: Yes, because when training, the muscles develop, improving and adapting to the type of sport or exercise.

N1. Reading the data. When the justification evidences a reading of particular values, for example, E129's answer. Responses that fall in this level reflect the local conception of the association (Estepa et al., 1999) and the

difficulty of comparing groups of data, treating each group as a unit or distribution.

E129: Fanny's jumping from 138 to 145.

N2. Reading between the data. The answers are based on the comparison of the values corresponding to the heights reached by the girls' group in the jump before and after training. For example, student E153 notes that most of the group has improved their initial mark.

E153: I think that training is effective, since most improved the height of their jump, except for Carol and Inés.

N4. Reading behind the data. Responses that reach this reading level; for example, E125 compares the sum of the heights before and after training to conclude its effectiveness. We should note that in none of the answers did the average support the effect of training.

E125: Yes, because if you add everything before and everything after, it gives you higher numbers afterwards, which means that they have improved.

Table 5 shows the reading levels students achieved in Task 1. We note that few students reach the highest level, N4, associated with critical reading, and their proportion is very similar in both courses.

The most frequent level was N2, in which the values of both data series are compared, reached by the majority of students in both courses and by a high percentage in the 1st grade of SE. Level N1, framed in a literal reading, appears similarly in both courses and slightly higher in the 7th grade of BE. The percentage of students who do not respond is low, while in the 7th grade of BE, a proportion of answers do not reach a minimum reading level (N1). The association between reading level and course was statistically significant in the Chi-square test ($\chi^2 = 15.084$, $g.l = 3$, $p = 0.002$), indicating that students in the 1st grade of SE achieved a higher level than younger students.

The results are similar to those obtained in Pallauta et al. (2023a), although in that study, a higher percentage of students did not answer the task (around 20%).

Table 5

Frequency (F) and percentage (%) of reading levels of Task 1, according to the course.

Reading level	7th BE N = 103		1st SE N = 88	
	F	%	F	%
N0. They do not read the data	31	30.1	13	14.8
N1. Reading the data	25	24.3	17	19.3
N2. Reading into data	32	31.1	49	55.7
N4. Reading behind the data	8	7.8	8	9.1
No answer	7	6.8	1	1.1

Task 2, composed of two sections, involves justifying the veracity of statements made from the information presented in a frequency distribution table with data grouped into intervals. Like the previous task, it also reaches the N4 level, while the N3 level is not considered. The different reading levels achieved by the students are presented below.

N0. Do not read the data. Answers that show students do not read the information in the table. For example, in section 2.a, E36 indicates his/her agreement because the information appears in the chart, confusing the graphical and tabular representations. In section 2.b, participant E1 does not justify his/her decision.

E36: Yes, because it is represented in the chart (Section 2.a).

E1: Quick (Section 2.b).

N1. Reading the data. The answer shows a literal reading of the information provided by the table. For example, E40, in 2.a, only considers the absolute frequency of the interval in which the value indicated in the statement appears, evidencing a lack of understanding of both the question and the concept of interval. In section 2.b, E126 makes an inadequate interpretation of an interval, confusing one of the extremes with the time expressed in seconds.

E40: Yes, because men have taken between 45 and 60 minutes to react (Section 2.a).

E126: Medium, since in thirty minutes and 45 seconds, 33 men are already affected (Section 2.b).

N2. Reading between the data. Although the answer suggests a correct reading of the table and the performance of operations or comparisons of the values, the justification denotes a limited understanding of the statement. For example, in 2.a, E129 considers only a subset of values below the indicated extreme value. In 2.b, student E4 only reads a limited part of the intervals in the table.

E129: No, it is 33 because 20 men are the ones who react between 45 and 60 minutes (Section 2.a).

E4: Slow, since most react in more than 15 minutes (Section 2.b).

N4. Reading behind the data. These answers demonstrate knowledge of the structure of the table and the context of the data, enabling adequate responses to each section. For example, in 2.a, E14 identifies that the frequency indicated in the statement corresponds to another set of values. In 2.b, E12 compares the frequencies of the intervals presented in the table to justify their response.

E14: No, because there are 53 men, not 20, who take between 45 and 60 minutes (Section 2.a).

E12: Slow, because most men have reacted to the drug between 45 and 60 minutes (Section 2.b).

The reading levels achieved in Task 2 are presented in Table 6. It should be noted that the percentage of answers that reached the N4 level is slightly higher than in Task 1. A higher percentage of Level N4 is reached in section 2.b. On the other hand, level N2 is presented similarly in both sections, being more frequent in section 2.a. Level N0 (they do not read the data) appears with great force, especially in 2.b., evidencing that students do not reach even a minimum reading level (N1: literal reading) in both sections. The results are better than those reported by Pallauta et al. (2023a) for the percentage of responses reaching the N4 level (Section 2.a: 15.6%; Section 2.b: 11.6%).

Table 6

Frequency (F) and percentage (%) of reading levels of Task 1, according to the section.

Reading level	Task 2 1st SE N = 88			
	Section 2.a		Section 2.b	
	F	%	F	%
N0. Do not read the data	19	21.6	29	33.0
N1. Reading the data	13	14.8	9	10.2
N2. Reading between the data	26	29.5	18	20.5
N4. Reading behind the data	21	23.9	25	28.4
No answer	9	10.2	7	8.0

CONCLUSIONS

This work aimed to analyse the ability of Chilean secondary school students to interpret information presented in statistical tables, thereby completing the research on this topic, which has been practically focused on double-entry tables or on children (e.g., Gabucio et al., 2010; Guimarães et al., 2021). The research questions raised are discussed below.

Regarding the first question (How do Chilean students interpret the information presented in statistical tables?), we concluded that students' interpretations of data tables were reasonable, with the task centred on a frequency table with data grouped into intervals and on accumulated frequencies being more difficult. Thus, in Task 1, based on a data table, most students provided correct or partially correct answers, although a significant percentage of 7th-grade BE students answered incorrectly.

When interpreting the accumulated frequencies of data grouped in intervals (Task 2), in the first section (2.a), most students respond correctly or partially correctly. Although in the second section (2.b), associated with the analysis of the trend of the data, the percentage of correct answers is increased, the percentage of incorrect answers is considerable.

Regarding the second research question (What reading levels do Chilean high school students achieve in tasks that involve statistical tables?), few students achieve the highest reading level N4, with level N2 being the most frequent in both tasks. On the other hand, he warns that none of the participants

in this study used the arithmetic mean as a strategy to answer the first task, confirming the difficulty associated with this concept (Batanero, 2000; Rondero & Font, 2015).

These results align with Gabucio et al. (2010) in interpreting double-entry tables with data grouped into intervals. In this line, there are previous works with high school students, such as Pallauta et al.'s (2023a) with a sample of Spanish students and the same tasks, and Prodromou (2015) in the analysis of data tables. These studies show that students tend to reach mainly the most basic levels of reading (N1 or N2), with difficulties in reaching more advanced levels. This trend could be influenced by the characteristics of the tasks performed in the classroom and the approach of the textbooks (Gea et al., 2024; Pallauta et al., 2021b). In this sense, González et al. (2021) indicate that tasks that require decision-making based on information presented in statistical tables can be complex for students, as they prioritise personal experiences over data.

It is striking that even though statistical literacy has been widely addressed in the literature and considered in various curriculum guidelines, different studies warn that students actually improve their statistical skills throughout their academic trajectory; however, few reach high levels of interpretation of graphic representations in secondary school (Aoyama & Stephens, 2003). One possible explanation is that the curriculum guidelines do not prioritise the development of the ability to analyse information across different contexts critically, or that this is only addressed superficially (Callingham & Watson, 2017).

The increasing use of data for decision-making across various areas makes it necessary for students to develop, in school, the ability to understand, analyse, and interpret various representations of data, for example, statistical tables, as part of the statistical literacy that every citizen should have (Gal, 2002).

We agree with Sharma (2013) that it is important to pose tasks to students that allow them to reinforce their critical reading and interpretations of the information presented in statistical tables, connecting them to their daily experience. There is also a need to incorporate into teaching situations that promote justifications based on the information presented, especially in contexts of uncertainty (Batanero & Borovcnik, 2016; Daniel & Braasch, 2013). In this sense, statistics teaching constitutes a challenge for teachers, given that the understanding of concepts should be promoted in the classroom, along with critical reflection on how data are collected, the variables, and the contexts that can influence research (Watson & Smith, 2022).

We conclude by recognising the limitations of the study, due to the moderate sample of students and the small number of tasks. Therefore, we recommend following up on the research by proposing additional similar tasks and designing teaching activities that promote the critical reading of statistical tables and other types of representations.

ACKNOWLEDGEMENTS

Project PID2022-139748NB-I00 funded by the Ministry of Science, Innovation and Universities of the Government of Spain (State Research Agency) (MICIU/AEI/10.13039/501100011033) and by the UE ERDF.

AUTHORSHIP CONTRIBUTION STATEMENTS

JDP and CB conceived the idea presented. JDP worked on theory development, and on the original draft, editing, and visualisation. CB was responsible for the review, formal analysis, and methodology. PA was responsible for the results and the discussion.

DATA AVAILABILITY DECLARATION

The data supporting the results of this study will be made available by the corresponding author, [JDP], upon reasonable request.

REFERENCES

- Álvarez, I., Guerrero, Y. y Torres, Y. D. T. (2020). Taxonomía de errores y dificultades en la construcción e interpretación de tablas de frecuencia. *Zetetike*, 28, e020012. <https://doi.org/10.20396/zet.v28i0.8656553>
- Aoyama, K. y Stephens, M. (2003). Graph interpretation aspects of statistical literacy: A Japanese perspective. *Mathematics Education Research Journal*, 15, 207–225. <https://doi.org/10.1007/BF03217380>
- Bargagliotti, A., Franklin, C., Arnold, P., Gould, R., Johnson, S., Perez, L. y Spangler, D. (2020). *Pre-K-12 Guidelines for assessment and instruction in statistics education II (GAISE II) A framework for statistics and data science education*. American Statistical Association and National Council of Teachers of Mathematics.
- Bartram, L., Correll, M. y Tory, M. (2022). Untidy data: The unreasonable effectiveness of tables. *IEEE Transactions on Visualization and*

Computer Graphics, 28(1), 686-696.
<https://doi.org/10.1109/TVCG.2021.3114830>

- Batanero, C. (2000). Significado y comprensión de las medidas de posición central. *Uno. Revista de Didáctica de las Matemáticas*, 25, 41-58.
- Batanero, C. y Borovcnik, M. (2016). *Statistics and probability in high school*. Sense Publishers.
- Callingham, R. y Watson, J. (2017). The development of statistical literacy at school Statistics. *Education Research Journal*, 16(1), 181-201.
<https://doi.org/10.52041/serj.v16i1.223>
- Campbell-Kelly, M., Croarken, M., Flood, R. y Robson, E. (2003). *The history of mathematical tables. From summer to spreadsheets*. Oxford University Press.
- Castellaro, M. y Roselli, N. (2020). Comprensión individual y diádica de tablas de frecuencias en alumnos de escolaridad primaria. *Pensamiento Psicológico*, 18(1), 57-70.
<https://doi.org/10.11144/Javerianacali.PPSI18-1.cidt>
- Cohen, L., Manion, L. y Morrison, K. (2018). *Research methods in education* (8th Edition). Routledge.
- Curcio, F. R. (1989). *Developing graph comprehension. elementary and middle school activities*. National Council of Teachers of Mathematics.
- Daniel, F. y Braasch, J. L. G. (2013). Application Exercises Improve Transfer of Statistical Knowledge in Real-World Situations. *Teaching of Psychology*, 40(3), 200-207.
<https://doi.org/10.1177/0098628313487462>
- Díaz-Levicoy, D., Morales, R., Arteaga, P. y López-Martín, M. M. (2020). Conocimiento sobre tablas estadísticas por estudiantes chilenos de tercer año de educación primaria. *Educación Matemática*, 32(2), 247-277. <https://doi.org/10.24844/EM3202.10>
- Estepa, A., Batanero, C. y Sánchez, F. T. (1999). Students' intuitive strategies in judging association when comparing two samples. *Hiroshima Journal of Mathematics Education*, 7, 17-30.
- Friel, S. N., Curcio, F. R. y Bright, G. W. (2001). Making sense of graphs: critical factors influencing comprehension and instructional

- implications. *Journal for Research in Mathematics Education*, 32(2), 124-158. <https://doi.org/10.2307/749671>
- Friedrich, A., Schreiter, S., Vogel, M., Becker-Genschow, S., Brünken, R., Kuhn, J., Lehmann, J. y Malone, S. (2024). What shapes statistical and data literacy research in K-12 STEM education? A systematic review of metrics and instructional strategies. *International Journal of STEM Education*, 11(58), 1-24. <https://doi.org/10.1186/s40594-024-00517-z>
- Gabucio, F., Martí, E., Enfedaque, J., Gilabert, S. y Konstantinidou, A. (2010). Niveles de comprensión de las tablas en alumnos de primaria y secundaria. *Cultura y Educación*, 22(2), 183-197. <https://doi.org/10.1174/113564010791304528>
- Gal, I. (2002). Adults' statistical literacy: Meanings, components, responsibilities. *International statistical review*, 70(1), 1-25. <https://doi.org/10.1111/j.1751-5823.2002.tb00336.x>
- Gal, I. (2022). Critical understanding of civic statistics: engaging with important contexts, texts, and opinion questions. En J. Ridgway (Ed.), *Statistics for empowerment and social engagement* (pp. 323-343). Springer. https://doi.org/10.1007/978-3-031-20748-8_13
- Gal, I., y Trostianitser, A. (2016). Understanding basic demographic trends: Connecting table reading, task design, and context. En J. Engel (Ed.), *Promoting understanding of statistics about society*. IASE.
- Gea, M. M., Pallauta, J. D., Batanero, C. y Valenzuela-Ruiz, S. M. (2022). Statistical tables in Spanish primary school textbooks. *Mathematics*, 10(15), <https://doi.org/10.3390/math10152809>
- Gea, M. M., Pallauta, J. D. y Arteaga, P. (2024). Análisis del nivel de lectura y contexto de tablas estadísticas en textos de secundaria españoles para desarrollar el sentido estocástico. En D. Díaz-Levicoy y A. Salcedo (Eds.), *Avances de investigaciones sobre libros de texto para el desarrollo de la cultura estadística y probabilística*. Centro de Investigación en Educación Matemática y Estadística, Universidad Católica del Maule.
- González, C. P., Muñoz, R. y Muñoz, J. (2021). Características argumentativas de la interpretación de tablas de frecuencia en estudiantes chilenos de segundo año medio. *Revista Chilena de Educación Matemática*, 13(1), 17-29. <https://doi.org/10.46219/rechiem.v13i1.39>

- Guimarães, G., Evangelista, B., y Oliveira, I. (2021). What students in the first grades of elementary school know about tables. *Statistics Education Research Journal*, 20(2), 9-9. <https://doi.org/10.52041/serj.v20i2.358>
- Martí, E. (2009). Tables as cognitive tools in primary education. En C. Andersen, N. Scheuer, M. P. Pérez Echeverría, y E. Teubal (Eds.), *Representational systems and practices as learning tools in different fields of learning* (pp. 133–148). Sense. https://doi.org/10.1163/9789087905286_009
- Ministerio de Educación, MINEDUC. (2015). *Bases curriculares 7° básico a 2° medio*. Unidad de Currículum y Evaluación.
- Ministerio de Educación y Formación Profesional (MEFP). (2022). *Real Decreto 217/2022, de 29 de marzo, por el que se establece la ordenación y las enseñanzas mínimas de la Educación Secundaria Obligatoria*.
- OECD (2021). *21st-century readers: Developing literacy skills in a digital world*. OECD Publishing. <https://doi.org/10.1787/a83d84cb>
- Pallauta, J. D., Arteaga, P. y Garzón-Guerrero, J. A. (2021a). Secondary school students' construction and interpretation of statistical tables. *Mathematics*, 9(24), 3197. <https://doi.org/10.3390/math9243197>
- Pallauta, J.D, Batanero, C. y Gea, M. M. (2023a). Un instrumento para evaluar la comprensión de tablas estadísticas en educación secundaria. *Enseñanza de las Ciencias*. 41(3), 89-112. <https://doi.org/10.5565/rev/ensciencias.5926>
- Pallauta, J. D., Batanero, C., Gea, M. M. y Arteaga, P. (2021b). Niveles de lectura y contextos en las actividades sobre tablas estadísticas en libros de texto chilenos y españoles. *Revista Chilena de Educación Matemática*, 13(4), 119-133. <https://doi.org/10.46219/rechiem.v13i4.74>
- Pallauta, J., Gea, M., Batanero, C. y Arteaga, P. (2023b). Algebraization levels of activities linked to statistical tables in Spanish secondary textbooks. En G. Kaiser y B. Sriraman (Eds.), *Data and statistical thinking: An international perspective*. Springer. https://doi.org/10.1007/978-3-031-29459-4_23

- Prodromou, T. (2015). Students' emerging reasoning about data tables of large-scale data. *International Journal of Statistics and Probability*, 4(3), p181. <https://doi.org/10.5539/ijsp.v4n3p181>
- Rondero, C. y Font, V. (2015). Articulación de la complejidad matemática de la media aritmética. *Enseñanza de las Ciencias*, 33(2), 29. <https://doi.org/10.5565/rev/ensciencias.1386>
- Schreier, M. (2014). Qualitative content analysis. En U. Flick (Ed.), *The SAGE handbook of qualitative data analysis* (pp. 170-183). SAGE. <https://doi.org/10.4135/9781446282243>
- Sharma, S. (2013). Assessing students' understanding of tables and graphs: Implications for teaching and research. *International Journal of Educational Research and Technology*, 4(4), 61-69.
- Shaughnessy, J. M. (2007). Research on statistics learning and reasoning. En F. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 957-1010). Information Age y NCTM.
- Vásquez, C., Arredondo, E. H. y García-García, J. I. (2022). Representaciones estadísticas a temprana edad: una aproximación desde los libros de texto de Chile y México. *Bolema*, 36(72), 116-145. <https://doi.org/10.1590/1980-4415v36n72a06>
- Watson, J. y Smith, C. (2022). Statistics education at a time of global disruption and crises: A growing challenge for the curriculum, classroom and beyond. *Curriculum Perspectives*, 42(2), 171-179. <https://doi.org/10.1007/s41297-022-00167-7>
- Weiland, T. (2017). Problematizing statistical literacy: An intersection of critical and statistical literacies. *Educational Studies in Mathematics*, 96(1), 33-47. <https://doi.org/10.1007/s10649-017-9764-5>