EMP Educação Matemática Pesquisa
ISSN 1983-3156
Qualis A1

http://dx.doi.org/10.23925/1983-3156.2026.v28.e77412

Teaching and Learning of the Boxplot: A Systematic Review

Ensino e aprendizagem do gráfico de caixa: uma revisão sistemática

Enseignement et apprentissage du diagramme en boîte : une revue systématique

Enseñanza y aprendizaje del Boxplot: Una revisión sistemática

Rodrigo Gutiérrez-Martínez[1]
Universidad Católica del Maule
Doctorando en Didáctica de la Matemática
https://orcid.org/0000-0003-3890-3347

Audy Salcedo[2]
Universidad Autónoma de Chile
Doctor en Educación
https://orcid.org/0000-0002-9783-8509

Danilo Díaz-Levicoy[3]
Universidad Católica del Maule
Doctor en Ciencias de la Educación
https://orcid.org/0000-0001-8371-7899

Abstract

This article presents a systematic review of empirical research published between 2010 and 2025 on the teaching and learning of boxplots. Following the PRISMA 2020 guidelines, a search was conducted in Web of Science, Scopus, SciELO, and Dialnet on May 31, 2025, using search terms in Spanish, English, and Portuguese. After removing duplicates and applying inclusion/exclusion criteria, 19 studies were analyzed. Data were extracted and synthesized using R-4.2.2. Bibliometric results show a predominance of articles in English (74%) and a geographic focus on Belgium, Germany, and Australia. Theoretically, the area heuristic underpins 47% of the studies, while conceptual change frameworks based on refutational text and multiple representations (MRs) proved to be the most effective intervention. However, reaction times indicate that visual bias persists. Only three studies involved practicing teachers,

[1] rgutierrezm@ucm.cl
[2] asalcedo@ucm.cl
[3] ddiazl@ucm.cl

Submetido em:12/07/2026 *Aceite em: 12/07/2026* *Publicado em: 30/07/2026*

and none evaluated effects beyond one week. It is concluded that understanding boxplots requires explicit and sustained instruction. Gaps are identified in longitudinal studies, methodological triangulation, and teacher training, which will guide future research and professional development agendas.

Keywords: Box plot, Systematic review, PRISMA, Teaching, Learning.

Resumo

Este artigo apresenta uma revisão sistemática de pesquisas empíricas publicadas entre 2010 e 2025 sobre o ensino e a aprendizagem de diagramas de caixa. Seguindo as diretrizes PRISMA 2020, os documentos foram pesquisados nas bases de dados Web of Science, Scopus, SciELO e Dialnet (31 de maio de 2025) utilizando termos de busca em espanhol, inglês e português. Após a remoção de duplicatas e a aplicação de critérios de inclusão/exclusão, 19 estudos foram analisados. Os dados foram extraídos e sintetizados utilizando o software R-4.2.2. Os resultados bibliométricos mostram uma predominância de artigos em inglês (74%) e um foco geográfico na Bélgica, Alemanha e Austrália. Teoricamente, a heurística da área fundamenta 47% dos estudos; enquanto as estruturas de mudança conceitual baseadas em texto refutacional e múltiplas representações (MRs) mostraram-se a intervenção mais eficaz, embora os tempos de reação indiquem que o viés visual persiste. Apenas três estudos envolveram professores em exercício, e nenhum avaliou os efeitos além de uma semana. Conclui-se que a compreensão de diagramas de caixa requer instrução explícita e contínua. Foram identificadas lacunas em estudos longitudinais, triangulação metodológica e formação de professores, que orientarão futuras pesquisas e agendas de desenvolvimento profissional.

Palavras-chave: Diagrama de caixa, Revisão sistemática, PRISMA, Ensino, Aprendizagem.

Résumé

Cet article présente une revue systématique des recherches empiriques publiées entre 2010 et 2025 sur l'enseignement et l'apprentissage des diagrammes en boîte. Conformément aux recommandations PRISMA 2020, les documents ont été recherchés dans Web of Science, Scopus, SciELO et Dialnet (31 mai 2025) à l'aide de termes de recherche en espagnol, en anglais et en portugais. Après suppression des doublons et application des critères d'inclusion/exclusion, 19 études ont été analysées. Les données ont été extraites et synthétisées à l'aide du logiciel R-4.2.2. Les résultats bibliométriques montrent une prédominance d'articles en anglais (74 %) et une concentration géographique sur la Belgique, l'Allemagne et l'Australie. Sur le plan théorique, l'heuristique de la surface sous-tend 47 % des études; tandis que les cadres de changement conceptuel basés sur un texte réfutationnel et des représentations multiples (RM) se sont avérés être l'intervention la plus efficace, bien que les temps de réaction indiquent la persistance d'un biais visuel. Seules trois études impliquaient des enseignants en exercice, et aucune n'a évalué les effets au-delà d'une semaine. En conclusion, la compréhension des diagrammes en boîte nécessite un enseignement explicite et soutenu. Des lacunes sont identifiées dans les études longitudinales, la triangulation méthodologique et la formation des enseignants, ce qui orientera les futurs programmes de recherche et de développement professionnel.

Mots-clés : Diagramme en boîte, revue systématique, PRISMA, enseignement, apprentissage.

Resumen

Se presenta una revisión sistemática de la investigación empírica publicada entre 2010 y 2025 sobre la enseñanza y aprendizaje del gráfico de caja (boxplot). Siguiendo las directrices PRISMA 2020, se buscaron documentos en Web of Science, Scopus, SciELO y Dialnet (31-mayo-2025) mediante ecuaciones en español, inglés y portugués. Tras eliminar duplicados y aplicar criterios de inclusión/exclusión, se analizaron 19 estudios. Los datos se extrajeron y sintetizaron con R-4.2.2. Los resultados bibliométricos muestran predominio de artículos en inglés (74 %) y foco geográfico en Bélgica, Alemania y Australia. Teóricamente, la heurística de área sustenta el 47 % de los trabajos; mientras que los marcos de cambio conceptual basados en texto refutacional + representaciones múltiples (MERs) mostraron ser la intervención más eficaz, aunque los tiempos de reacción indican que el sesgo visual persiste latente. Sólo tres estudios involucran a docentes en ejercicio y ninguno evalúa efectos más allá de una semana. Se concluye que la comprensión del boxplot exige enseñanza explícita y sostenida. Se identifican vacíos en estudios longitudinales, triangulación metodológica y formación docente, lo que orienta futuras agendas de investigación y desarrollo profesional.

Palabras-clave: Gráfico de caja, Revisión sistemática, PRISMA, Enseñanza, Aprendizaje.

Teaching and Learning of the Boxplot: A Systematic Review

Introduction

In contemporary society, statistics are present in virtually all areas of society, being a key element in the natural sciences, social sciences, economics, public policies, and playing a crucial role in strengthening a healthy democracy (Schoen et al., 2025). In this context, the training of citizens in statistical issues has been a problem of growing interest in recent years, with the concept of statistical literacy, understood as the ability to interpret, understand, and critically evaluate data (Castillo, 2024; Gal, 2002).

Gal (2002), in his definition of statistical literacy, highlighted the interpretation of graphic representations as an essential skill within it. Along the same lines, Kim et al. (2020) mention that the visualization of data in graphs is not only present in scientific practices, but also in everyday contexts. Therefore, developing skills in the construction and interpretation of graphs must occupy a central place in statistical education, so that students can make decisions when facing different problems.

In this scenario, the box plot—also known as a *boxplot*—is a powerful tool for visualizing the distribution of data, allowing the visualization of information about its central trend, dispersion, and symmetry. However, for its correct interpretation, it is necessary to understand statistical concepts, such as the median, quartiles and interquartile range (Bakker et al., 2004).

In addition, this graph deserves particular attention because, despite its relevance to describe the variability and distribution of data, it has been little researched in the educational field, especially in the case of teachers and future teachers. Most of the research carried out has focused on school students, evidencing serious difficulties in their interpretation (Abt et al., 2024; Lem et al., 2013; Madrid et al., 2022), even more marked than in other graphs (Lem et al., 2015). Added to this is the inherent difficulty of statistical graphs, as their understanding involves processes of translation, interpretation, and extrapolation/interpolation (Sánchez-Acevedo et al., 2021).

Despite its relevance, the literature on the understanding of the box graph is dispersed in different fields, such as the Didactics of Statistics, psychology, health, finance, and public policies (Videla-Jiménez et al., 2025). Therefore, it is necessary to systematize the existing empirical and theoretical findings, in order to identify common patterns in comprehension difficulties, the factors that explain them and the pedagogical strategies that have shown effectiveness in overcoming them.

Therefore, this review seeks to gather and analyze the available empirical evidence on the understanding of the box plot. To this end, it is intended to contribute to answering the following questions:

a. What are the bibliometric characteristics of the studies?
b. What theoretical frameworks do the studies use?
c. What are the objectives of the studies, and at what educational levels have they focused on the teaching and learning of box plots?
d. What methodologies have been used to investigate the teaching and learning of box plots?
e. What results and conclusions do the studies show?

Methodology

To address the objective, a systematic literature review was conducted, following the guidelines provided by the PRISMA (*Preferred Reporting Items for Systematic Reviews and Meta-Analyses*) statement, in order to rigorously document each stage of the process (Page et al., 2021).

The search for articles was carried out on May 31, 2025, in four databases: Web of Science (WoS), Scopus, SciELO and the bibliographic portal Dialnet. The selection of these databases was based on the need to ensure a broad, relevant and multilingual coverage of the scientific literature related to the box graph in educational contexts. The incorporation of WoS and Scopus is due to their recognized high editorial quality and international coverage, which ensures access to relevant research with high standards (Chadegani et al., 2013). For their part, SciELO and Dialnet were incorporated due to their reach within the scientific communities of Ibero-America, particularly to identify studies published in languages other than English, such as Spanish and Portuguese, which are not always adequately represented in English-language databases (Mongeon & Paul-Hus, 2016).

To identify the articles, specific search equations were constructed for each database, using terms in English, Spanish and Portuguese related to the box graph (*boxplot, box plot, box-plot*, box plot, box plot, diagrama de caixas) and its teaching and learning (*teaching* and *learning*). To broaden the search and avoid excluding relevant studies, terms commonly used in articles that address statistical graphs (*reading*, *construction*, *interpretation*, *knowledge*, *understanding*) were incorporated. The equations are detailed in Table 1.

Table 1

Search equations

Database	Language	Equation
WoS	English	TS=((boxplot* OR "box plot*" OR box-plot*) AND (teach* OR learn* OR reading OR construction OR interpreting OR knowledge OR understanding))
	Spanish	TS=(("box chart*" OR "box plot*") AND (teaching OR learning OR reading OR construction OR interpretation OR knowledge OR comprehension))
	Portuguese	TS=(("box plot*" OR "box plot*") AND (teaching OR learning OR reading OR construction OR interpretation OR knowledge OR understanding))
Scopus	English	TITLE-ABS-KEY ((boxplot* OR "box plot*" OR box-plot*) AND (teach* OR learn* OR reading OR construction OR interpretation OR knowledge OR understanding))
	Spanish	TITLE-ABS-KEY(("box chart*" OR "box plot*") AND (teaching OR learning OR reading OR construction OR interpretation OR knowledge OR comprehension))
	Portuguese	TITLE-ABS-KEY(("box plot*" OR "box plot*") AND (teaching OR learning OR reading OR construction OR interpretation OR knowledge OR understanding))
SciELO	English	(boxplot OR "box plot" OR box-plot) AND (teach OR teaching OR learn OR learning OR reading OR construction OR interpretation OR knowledge OR understanding)
	Spanish	("box chart" OR "box plot") AND (teaching OR learning OR reading OR construction OR interpretation OR knowledge OR comprehension)
	Portuguese	("box chart" OR "box plot") AND (teaching OR learning OR reading OR construction OR interpretation OR knowledge OR understanding)
Dialnet	English	("box plot" OR boxplot OR box-plot) AND (teach OR teaching OR learn OR learning OR reading OR construction OR interpretation OR knowledge OR understanding)
	Spanish	("box chart" OR "box plot") AND (teaching OR learning OR reading OR construction OR interpretation OR knowledge OR comprehension)
	Portuguese	("caja chart" OR "box plot") AND

Database	Language	Equation
		(teaching OR learning OR reading OR construction OR interpretation OR knowledge OR understanding)

According to the inclusion and exclusion criteria presented in Table 2, the search included all studies related to the teaching or learning of box plots. No restrictions were established regarding publication year or educational level, and no filters by subject area or journal were applied.

Table 2

Review inclusion and exclusion criteria

Criteria	Inclusion	Exclusion
Object of study	IC1: Focused on Teaching or Learning Box Chart	EC1: Not focused on the teaching or learning of the box graph
Types of documents	IC2: Articles, books, book chapters or conference proceedings	EC2: Editorials, prologues, theses or others not included in CI2
Language	IC3: Spanish, English and Portuguese	EC3: A language other than Spanish, English or Portuguese
Type of study	IC4: Empirical studies	EC4: Systematic reviews, theoretical and methodological studies or didactic proposals
Accessibility	IC5: Full-text articles available	EC5: Inaccessible articles

After the application of the search equations, 5281 articles were found. However, of the 3202 found on Dialnet, only 1000 could be accessed, due to server failures, it was not possible to view the remaining ones.

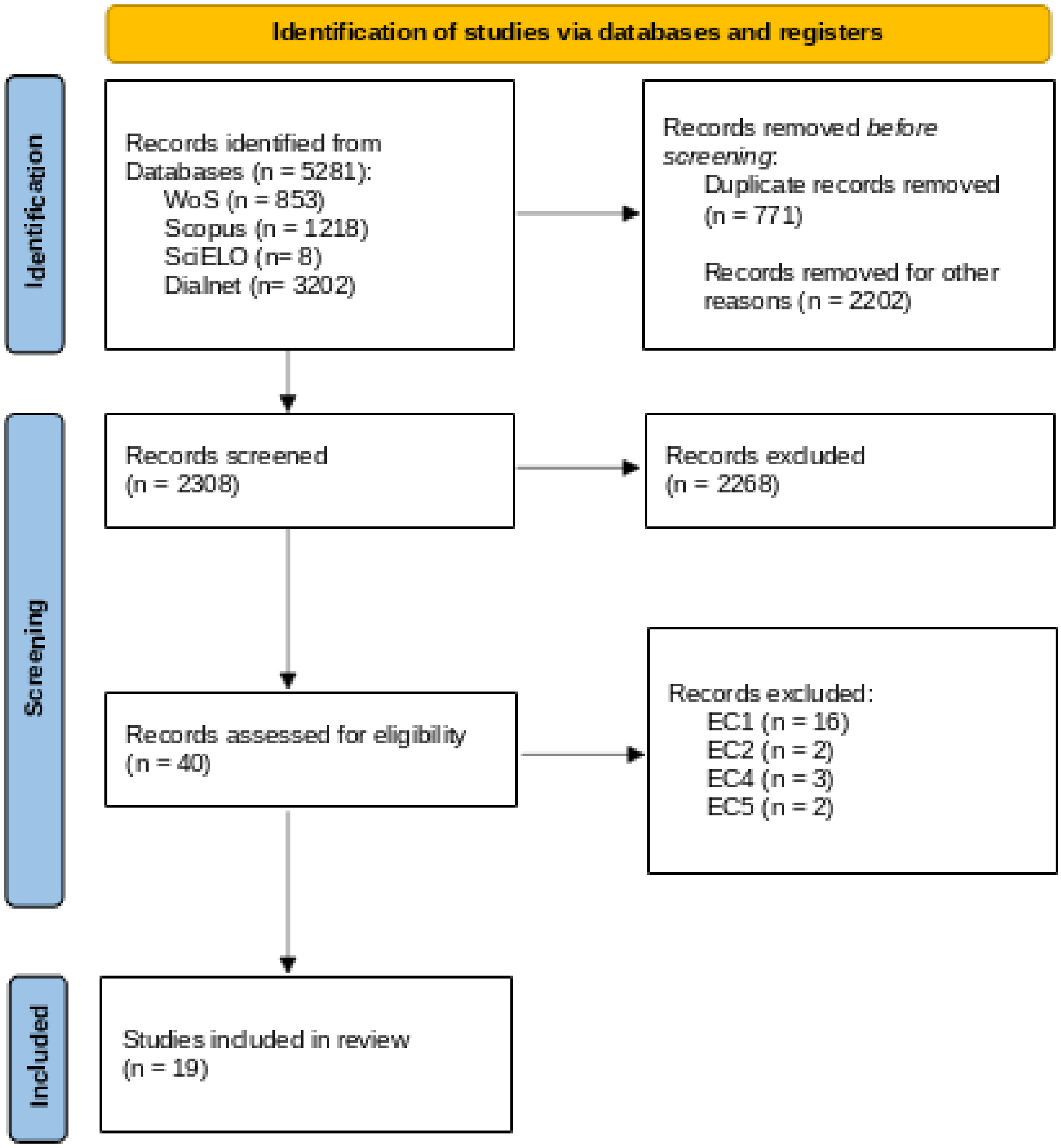
All records were imported and identified in a spreadsheet using Microsoft Excel. After removing duplicates (and four triplicate records) detected, 2308 articles remained, the titles, keywords, and/or abstracts were screened. After this reading, 2268 articles were discarded for not focusing on the teaching or learning of the box chart, leaving a sample of 40 articles.

Subsequently, the Inclusion and Exclusion Criteria previously described were applied to these 40 articles, which resulted in 21 excluded articles, leaving a total of 19 studies to be included in the review. This entire process was carried out by the principal investigator, who, in case of doubts, consulted the second researcher, accessing the full text when necessary to support the decision. The diagram in Figure 1 summarizes this process of exclusion and selection of articles.

Finally, the selected articles were categorized and analyzed according to their bibliometric characteristics, objectives, theoretical justification, methodologies, results and conclusions, in relation to the research questions. These analyses were carried out with the *software* R version 4.2.2 (R Core Team, 2022), using the RStudio development environment (RStudio Team, 2024).

Figure 1

Study selection process according to PRISMA 2020



Results

This section presents the results obtained in relation to the research questions posed.

As previously described, a total of 19 articles addressing the teaching and/or learning of box plots were analyzed. Table 3 presents the authors, titles, and reference codes assigned to each study throughout the paper. (Table 3).

Table 3

Articles selected for review

Code	Authors	Year	Title
A01	Abt, Leuders, Loibl, Reinhold	2024	Developing initial notions of variability when learning about box plots
A02	Abt, Leuders, Loibl, Strohmaier, Van Dooren, Reinhold	2024	How can eye-tracking data be used to understand cognitive processes when comparing data sets with box plots?
A03	Abt, Loibl, Leuders, Van Dooren, Reinhold	2025	Understanding student errors in comparing data sets with boxplots
A04	Carvalho, Fernandes, Freitas	2019	Building and interpretaing boxplots by 9th grade Portuguese students
A05	Edwards, Özgün-Koca	2017	Interpretations of boxplots: Helping middle school students to think outside the box
A06	Kortenkamp, Rolka, DurandGuerrier, SouryLavergne, Arzarello	2010	Using technology in the teaching and learning of box plots
A07	Lem, Baert, Ceulemans, Onghena, Verschaffel, Van Dooren	2017	Refutational text and multiple external representations as a method to remediate the misinterpretation of box plots
A08	Lem, Kempen, Ceulemans, Onghena, Verschaffel, Van Dooren	2015	Combining Multiple External Representations and Refutational Text: An Intervention on Learning to Interpret Box Plots
A09	Lem, Onghena, Verschaffel, Van Dooren	2013	On the misinterpretation of histograms and box plots
A10	Lem, Onghena, Verschaffel, Van Dooren	2013	The heuristic interpretation of box plots
A11	Lem, Onghena, Verschaffel, Van Dooren	2017	The power of refutational text: changing intuitions about the interpretation of box plots
A12	Lem, Onghena, Verschaffel, Van Dooren	2014	Experts' Misinterpretation of Box Plots - a Dual Processing Approach
A13	Lem, Vandebeek, Onghena, Verschaffel, Van Dooren	2014	Heuristic reasoning in interpreting box plots: The influence of orientation
A14	Madrid, Valenzuela-Ruiz, Batanero, Garzón-Guerrero	2022	Interpretation of boxplot by university students of physical activity and sport sciences

Code	Authors	Year	Title
A15	Pierce, Chick	2013	Workplace statistical literacy for teachers: Interpreting box plots
A16	Pierce, Chick, Wander, Phillips, MacGillivray, Martin	2014	Improving teachers' professional statistical literacy
A17	Verdugo-Hernandez, Spinoza-Vasquez, Cumsille	2024	Mathematical work of a future teacher in teaching the box and whisker plot
A18	Ferreira, Carvalho, Eugênio	2024	Statistical Literacy of High School Students: An Experience with the Boxplot Graph
A19	Gea, Artega, Cañadas	2017	Interpretation of statistical graphs by future Secondary Education teachers

Bibliometric characteristics of the studies

This section presents the general characteristics of the analyzed studies, such as years of publication, type of document, language, country of affiliation of the authors and the journals in which they were published.

Document types, years of publication and language

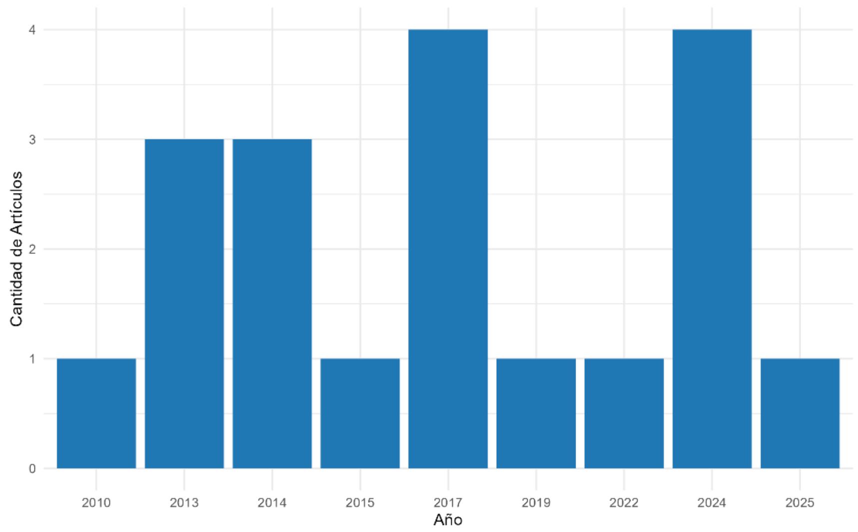
Of the 19 articles analyzed in this review, 17 were journal articles published in scientific journals and 2 to papers presented at academic conferences (A06 and A16). Nos books or book chapters addressing the topic under investigation were identified in the databases reviewed.

Regarding the years of publication (Figure 2), the studies were published between 2010 and 2025. However, it should be noted that the search was conducted during the first half of 2025; therefore, publications from that year were not fully represented.

In terms of publication language, the search considered studies published in English, Spanish, and Portuguese. Of the 19 articles, 14 were written exclusively in English, 2 were published in both English and Spanish (A14 and A17), 1 exclusively in Spanish (A19), and 2 exclusively in Portuguese (A04 and A18).

Figure 2

Distribution of studies by year of publication



Geographical distribution of authors and journals

Figure 3 shows the geographical distribution of the authors. The studies included authors affiliated with institutions across three continents. Authors were affiliated with institutions in the United States (3), Brazil (3), Chile (3), Spain (3), Australia (3), Belgium (8), Portugal (3), and Germany (7). Few international collaborations were identified, with collaborations observed only between Belgium and Germany.

Regarding the journals in which the studies were published, Educational Psychology (ISSN 1469-5820) stands out with two articles (A07 and A09), both authored by Lem et al. The remaining journals each published one article, as shown in Table 4.

Figure 3

Geographical distribution of authors

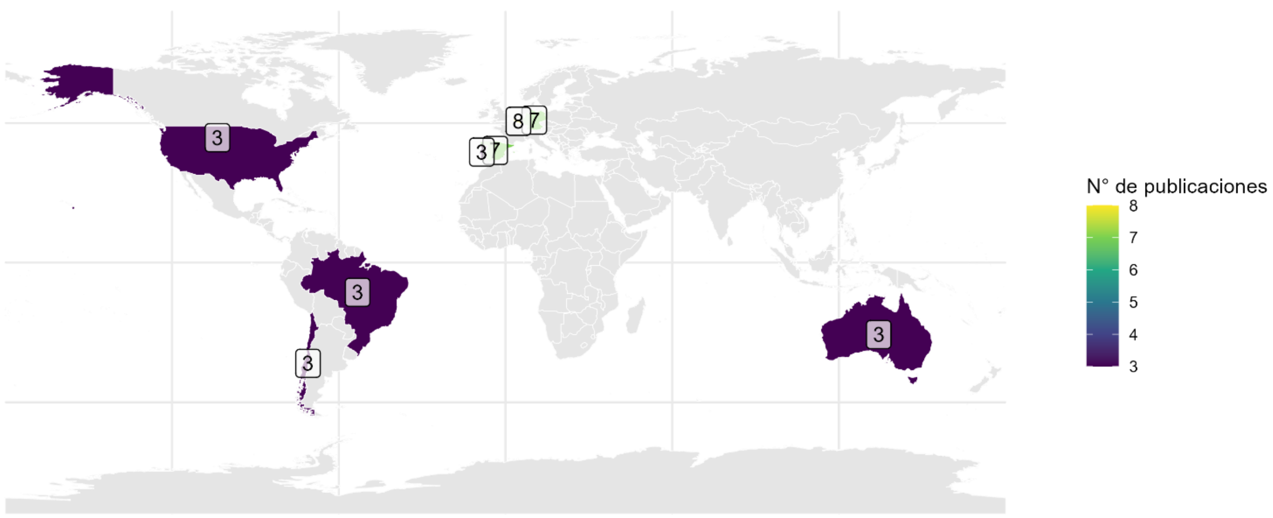


Table 4

Journals of publication

Name Magazine	Number of Articles
Journal of Statistics Education	1
Educational Psychology	2
Areté: Digital Journal of the Doctorate in Education	1
Research advances in mathematics education: AIEM	1
Bolema - Mathematics Education Bulletin	1
Mathematics Education	1
Educational Studies in Mathematics	1
European Journal of Psychology of Education	1
Frontiers in Education	1
International Journal of Science and Mathematics Education	1
Learning and Instruction	1
Mathematical Thinking and Learning	1
Mathematics Education Research Journal	1
Psychology Belgium	1
Journal of Social and Environmental Management	1
Studia Psychologica	1
ZDM - Mathematics Education	1

Regarding the conference papers, these were presented at the 6th Congress of the European Society for Research in Mathematics Education (A06) and the Eighth Australian Conference on Teaching Statistics (A16). The manuscript of the latter was later published as a book chapter in Topics from Australian Conferences on Teaching Statistics (2014).

Theoretical justification of the studies

In relation to the second research question, this section presents a synthesis of the theoretical frameworks used in the analyzed studies.

Table 5

Theoretical frameworks used in the analyzed studies

Conceptual framework / theoretical element	Frequency	Article
Area Heuristics or ARFS	8	A02, A03, A09, A10, A11, A12, A13, A14
Rebuttal text	4	A07, A08, A10, A11
Multiple External Representations (MERs)	2	A07, A08
Variability-Focused Cognitive Activities (VFCA)	1	A01
Ontosemiotic approach	2	A04, A14
Levels of graphic reading (Curcio 1987; Friel et al. 2001)	4	A04, A05, A06, A19
Math workspaces	1	A17
Statistical Literacy	3	A15, A16, A18

As can be seen in Table 5, there is heterogeneity in the theoretical frameworks addressed by the studies, with some frameworks being considerably more prevalent than others.Articles addressing area heuristics or *Area-Represents-Frequency Schemas* (ARFS) account for about 47% of the included studies. Although there are 8 articles, these are by two main authors: Abt (A02 and A03) and Lem (A09, A10, A11, A12, A13, A14). These studies have focused mainly on the university population.

Conceptual change frameworks (refutational text and MERs) concentrate on 4 articles. Articles that address MERs use the overlay of the box graph with other representations.

On the other hand, 4 articles (A04, A05, A06, A19) use the Curcio/Friel reading levels. This framework allows students to be classified according to levels of graph comprehension complexity. This research has been applied to secondary school students and future teachers.

Variability-Focused Cognitive Activities (A01), Ontosemiotic Approach (A04, A14) and Mathematical Workspaces (A17) appear only in one or two articles each, while statistical literacy has been explicitly addressed by 3 manuscripts (A15, A16, A18), all focusing on in-service or pre-service teachers.

Objectives of the studies

From the analysis of the research objectives, it is observed that there is a predominance of studies aiming to "evaluate" (A11, A12, A14, A19) or "verify" (A07, A10), totaling six studies. These articles are associated with research on the heuristics of the area or the use of the Curcio/Friel Reading Levels.

On the other hand, 5 studies identify "analyze" or "examine" as objectives. Most of these focus on statistical literacy in in-service teachers (A15 and A16) or in secondary school students (A18).

"Identify" (A03) and "explore" (A09) appear in only one study each. In these articles, these studies seek to examine factors of difficulty in the interpretation of the box graph, but without proposing interventions.

In 4 articles (A01, A02, A05, A13) no explicit objective statement was identified through a declarative phrase.

Finally, only 1 article states that it seeks to "characterize". This work is based on a case study of a pre-service teacher.

Methodological aspects

For the analysis of the methodological aspects of the studies, 4 dimensions were reviewed: Type of methodology declared, educational level, data collection instruments, analysis techniques and participant type (Table 6).

Table 6

Methodological aspects of the studies

Dimension	Category	Frequency	Article
Stated methodology	Quantitative	1	A04
	Qualitative	2	A17, A18
	Mixed	1	A15
Level of study	Experimental / quasi-experimental	7	A01, A07, A08, A10, A11, A12, A13
	Exploratory	2	A02, A06
	Case Study	1	A17
	Description	5	A04, A05, A14, A16 A19
Data collection instruments	Closed-response test	11	A01, A03, A05, A07, A08, A09, A10, A11, A12, A13, A15
	Open-Response Test	7	A04, A06, A07, A08, A11, A18, A19
	Eye-tracking	1	A02
	Reaction times	3	A10, A11, A13
	Classroom Observation Guideline / Video	1	A17
Analysis techniques	Frequencies and percentages	14	A01, A03, A04, A05, A06, A07, A08, A09, A11, A12, A14, A15, A16, A19
	Parametric tests (t, ANOVA)	4	A04, A14, A15, A19
	Mixed / GLMM models	6	A03, A07, A08, A10, A11, A12
	Content Analysis	5	A04, A06, A17, A18, A19
	Oculometric measurements	1	A02
Type of participant	High School Students	5	A01, A04, A05, A06, A18
	University (non-teaching degree)	8	A03, A07, A08, A09, A10, A11, A13, A14
	Teachers in training	3	A02, A17, A18
	In-service teachers	2	A15, A16
	Expert professionals	1	A12
	Adults not specified / general public	1	A03

Only four studies explicitly stated their methodological approach, including quantitative (A04), qualitative (A17 and A18), and mixed-methods studies (A15). Regarding the second dimension analyzed, namely study design, most studies were experimental or quasi-experimental (A01, A07, A08, A10, A11, A12, A13), followed by descriptive studies (A04, A05, A14, A16, A19).

The most common data collection instruments were open- and closed-response tests administered in paper-based or computer-based formats, appearing in a total of 18 studies. Less frequently used instruments included eye-tracking (A02) and classroom observation guidelines based on video recordings (A17), each appearing in one study.

Regarding analysis techniques, excluding the widespread use of frequencies and percentages, mixed models were the most frequently used. However, these studies were primarily conducted by two researchers: Abt (A03) and Lem (A07, A08, A10, A11, A12). Content analysis was also commonly employed, appearing in five studies (A04, A06, A17, A18, A19). Only one study used eye-tracking techniques (A02).

Regarding participant type, nearly half of the studies focused on university students, although most of these were conducted by Lem et al. (A07, A08, A09, A10, A11, A13). The second most frequently studied group was high school students (26%). Studies involving teachers (both pre-service and in-service) accounted for 32% of the sample, with more studies focusing on pre-service teachers (16%) than on in-service teachers (11%). In the case of in-service teachers, both studies were conducted by Pierce et al. Finally, only one study (A12) focused on expert users of box plots, and one study (A03) focused on the general population, although it included a phase involving university students.

Study results and conclusions

Table 7 summarizes the main results and conclusions of the articles.

Table 7

Main results and conclusions of the studies

Code	Results	Conclusions
A01	The group that worked with the "variability-first" digital environment obtained higher scores in the post-test and generated more cognitive activities focused on dispersion.	Focusing on variability before the central trend favors the initial understanding of the box graph
A02	The *eye-tracking* showed differentiated fixation patterns: the ″ area" schema is concentrated in the box; the "median" schema is concentrated in the center line.	Ocular analysis allows the diagnosis of cognitive processes and confirms the validity of area heuristics; It is advisable to design tasks that redirect attention to the median and quartiles.
A03	The ARFS heuristic generated most of the incongruous errors; item difficulty factors (salience,	It is necessary to explicitly address the idea of frequency-area and to graduate the difficulty of

Code	Results	Conclusions
	congruence) predicted failure; The student profiles show an incomplete conceptual change.	items to accompany the conceptual change.
A04	Students build the box plot better when the data is in bar graph; The overall interpretation (dispersion, symmetry) was very low.	It is recommended to link construction and interpretation and to use intermediate representations (tables, dot-plots) to consolidate meanings.
A05	70% confused the length of the box with "amount of data"; few related percentages to quartiles.	The activities with *dot-plot* + *box plot* in TI-Nspire help, but they require teacher intervention for students to connect parts of the graph with real data.
A06	After manipulating the app, the students identified the median better, but they continued to omit the whiskers in their explanations.	Dynamic manipulation favors comprehension, as long as the teacher guides the reading of all the elements of the diagram.
A07	The combination of MERs + refutal text reduced the area heuristic by 40% compared to the control; Each technique separately produced partial improvements.	Joining aligned representations with an explicit refutation is the most effective intervention against heuristics.
A08	At one week, the "MERs+Refutation" group maintained the highest percentage of correct answers; MERs only showed an intermediate effect.	Rebuttal is an essential component. MERs enhance its effect, but do not replace it.
A09	The error was due more to the graphic salience (height/area) than to the lack of statistical definitions; Tversky's principles predicted failures.	Teaching how to "read the direction" of the graph and using designs that minimize misleading salience can reduce errors.
A10	With a limit of 5 seconds, incongruous errors increased; The brief intervention reduced failures, but reaction times still indicated heuristic processing.	Area heuristics can be inhibited, but not disappear; Prolonged experiences are needed to modify rapid processing.
A11	Rebuttal text improved accuracy, but reaction times showed internal conflict (ambivalent processing).	Conceptual change is fragile; It is advisable to combine refutation with supervised practice to consolidate the new intuition.
A12	Even experts (master's degree, teaching staff) frequently confused area with incongruent items (error ≈ 35%).	Advanced training does not guarantee to avoid heuristics; Specific professional development strategies are required.
A13	Portrait orientation reduced heuristics (~15% fewer errors) but did not eliminate it; Higher RTs	Modifying the visual design mitigates bias, but it should be supplemented with conceptual

Code	Results	Conclusions
	reveal conflict.	instruction.
A14	Sports students showed reading level 1 majority; Semiotic conflicts prevailed mid-center	Integrating tasks that articulate measurements and representation and using contexts of interest to students raises reading levels.
A15	38% of teachers misinterpreted quartiles; TPB (perceived control) attitudes predicted performance.	Professional statistical literacy requires workshops that combine content and attitudes; Perceived control is key.
A16	The half-day workshop improved correct interpretation by 20% at 6 weeks; greater effect on those who practiced with their own data.	Short workshops can be effective if they include contextualized application and follow-up.
A17	The future teacher focused the class on procedure and omitted the discursive genesis; variability was not problematized.	Initial training should promote discursive space and the data-graph connection to teach *boxplots*.
A18	After the unit contextualized in COVID-19, students improved in building the diagram, but only 35% correctly justified the dispersion.	Contextual statistical literacy motivates, but it is necessary to reinforce the interpretation of quartiles and interquartile range.
A19	Future teachers interpreted histograms better than boxplots; they confused whiskers with absolute minimum/maximum.	Initial training needs to compare multiple graphs and emphasize the meaning of "whiskers" and IQR.

Several articles agree on the persistence of errors in interpreting the box graph. For example, the "area heuristic", which refers to the belief that a wider box contains a more data points, is the most frequently studied error (A03, A04, A05, A09, A10, A12, A14), this misconception appears among secondary school students, persists among university students, and is also observed in teachers and expert users with postgraduate training (A12). Other reported errors include identifying the median as the "geometric center" of the box and the confusion between the whiskers and the absolute minimum and maximum values (A14, A19) are also described. These results suggest that current teaching emphasizes the procedural construction of the graph but fails to consolidate its statistical meaning.

Eight studies adopt conceptual change strategies, with the combination "refutational text + multiple external representations (MERs)" being the most effective (A07, A08). When the box chart is superimposed on a histogram or *dot-plot,* and the misconception is explicitly refuted, the correct responses increase by twenty to forty percentage points (A07, A08). However, evidence indicates that visual bias does not disappear completely: reaction times continue to reveal internal conflict (A10, A11), so refutation requires supervised practice and varied tasks to be fully consolidated.

Three studies (A02, A10, A13), all by Lem et al., provide instrumental data thatprovide insight into "reasoning in action". With a limit of five seconds per item, the proportion of inconsistent errors exceeds 50% (A10). When the time limit is removed, errors decrease, but they don't go away. *Eye-tracking* (A02) confirms that, when the heuristic is activated, attention is focused on the surface of the case and while avoiding the median line. These indications suggest that bias can be inhibited by instruction, but it is still triggered automatically.

Only three studies focus on teachers in practice or in continuing education (A15, A16, A18). About 40% of the teachers evaluated misinterpret the quartiles (A15), and their perception of control over the statistical content is a significant predictor of performance. Short short professional development workshops (A16) achieve immediate improvements – approximately twenty percentage points – but the authors stress the need for follow-up and contextualized implementation. A research gap regarding classroom practice remains, as there are no longitudinal designs that evaluate the actual transfer of these improvements to student learning.

Discussion and conclusions

The present study systematized nineteen empirical studies published between 2010 and 2025 on the teaching and learning of the box plot. Based on the bibliometric, methodological, theoretical, and results analyses, the following conclusions can be drawn:

Most studies focused on first-year university students (47%), while research with practicing teachers or with students in primary and lower secondary education remains limited. The dominant error of the "area heuristic" is maintained from secondary school students to experts with advanced statistical training, which shows that the mere accumulation of knowledge is not enough to eradicate it. This suggests that interpreting box plots requires explicit, not implicit, teaching, even in advanced programs.

Interventions that combine refutal text with multiple representations (histogram + *boxplot*) achieve the most significant and lasting improvements: increasing accuracy and leading to a significant reduction in heuristic explanations. However, reaction times and *eye-tracking* indicate that bias does not disappear, but is instead reduced, so repeated and guided practice is required.

Technological resources (*applets*, TI-Nspire) improve the identification of the median and interquartile range, but students continue to ignore the whiskers if teachers do not direct the comprehensive interpretation of the graph. Classroom evidence suggests that the teaching of statistical concepts related to this graph, such as variability, continues to focus on procedural aspects to the detriment of statistical thinking.

Studies show that perceived control and attitudes are good predictors of teachers' performance in statistical concepts. Short workshops appear to be an effective strategy for improving the interpretation of box plots.

No studies follow participants beyond one week or systematically triangulate classroom data, qualitative analyses, and process-based measures. This limitation restricts understanding of the long-term impact of interventions and their transfer to actual teaching practice.

Finally, it is necessary to highlight that many articles do not explicitly state the methodology used, which may create difficulties for reproducibility and methodological classification.

The search was limited to four databases and three languages; therefore, relevant studies published in other sources or languages may have been omitted. In addition, heterogeneity in the methodological reporting across studies made the final classification of approaches difficult. Nevertheless, the patterns identified provide a solid basis for guiding didactic interventions and professional development programs, as well as for generating new research questions about box plots and statistical literacy.

Acknowledgments

ANID National Scholarships/Doctorate 21252622

References

Abt, M., Leuders, T., Loibl, K., & Reinhold, F. (2024). Developing initial notions of variability when learning about box plots. *Mathematical Thinking and Learning,* ahead of print. https://doi.org/10.1080/10986065.2024.2421412

Bakker, A., Biehler, R., & Konold, C. (2004). Should young students learn about box plots? En G. Burrill & M. Camden (Eds.), *Curricular Development in Statistics Education: International Association for Statistical Education* (IASE) (pp. 163-173). International Statistical Institute.

Castillo, I. (2024). Improving Statistical Literacy through Evidence-Based Strategies Among First-Year Education Students in a State University. *Journal of Contemporary Educational Research, 8*(1), 246-259.

Chadegani, A. A., Salehi, H., Yunus, M. M., Farhadi, H., Fooladi, M., Farhadi, M., &

Ebrahim, N. A. (2013). A Comparison between Two Main Academic Literature Collections: Web of Science and Scopus Databases. *Asian Social Science, 9*(5), 18-26. https://doi.org/10.5539/ass.v9n5p18

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

Kim, E. M., Nabors Oláh, L., & Peters, S. (2020). A Learning Progression for Constructing and Interpreting Data Display. *ETS Research Report Series, 2020*(1), 1-27. https://doi.org/10.1002/ets2.12285

Lem, S., Kempen, G., Ceulemans, E., Onghena, P., Verschaffel, L., & Van Dooren, W. (2015). Combining Multiple External Representations and Refutational Text: An Intervention on Learning to Interpret Box Plots. *International Journal of Science and Mathematics Education, 13*(4), 909-926. https://doi.org/10.1007/s10763-014-9604-3

Lem, S., Onghena, P., Verschaffel, L., & Van Dooren, W. (2013). On the misinterpretation of histograms and box plots. *Educational Psychology, 33*(2), 155-174. https://doi.org/10.1080/01443410.2012.674006

Madrid, A. E., Valenzuela-Ruiz, S. M., & Batanero, C. (2022). Interpretación del diagrama de caja por estudiantes universitarios de ciencias de la actividad física y deporte. *Educación Matemática, 34*(3), 275-300. https://doi.org/10.24844/EM3403.10

Mongeon, P., & Paul-Hus, A. (2016). The journal coverage of Web of Science and Scopus: A comparative analysis. *Scientometrics, 106*(1), 213-228. https://doi.org/10.1007/s11192-015-1765-5

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ, 372*(71), 1-9. https://doi.org/10.1136/bmj.n71

R Core Team. (2022). *R: A language and environment for statistical computing* (Version 4.2.2). R Foundation for Statistical Computing.

RStudio Team. (2024). *RStudio: Integrated Development Environment for R*. Posit Software, PBC.

Sánchez-Acevedo, N., Toro-Barbieri, E., & Araya-Bastias, D. (2021). Interpretación y comprensión de gráficos estadísticos por profesores de Matemáticas en formación. *Revista Chilena de Educación Matemática, 13*(4), 230-243. https://doi.org/10.46219/rechiem.v13i4.86

Videla-Jiménez, P., Ruz, F., & López-Martín, M.M. (2025). Research on box plot knowledge in statistical education: a systematic review of the literature. *European Journal of Psychology of Education, 40*, 141. https://doi.org/10.1007/s10212-025-01048-z