revista de investigación educativa de la Rediech

• Volumen 16 • 2025 • e2300 • ISSN: 2448-8550 •

Evaluation of the effectiveness of a data visualization workshop in the development of graphical representation skills: A pre- and post-implementation comparative study

Valoración de la efectividad de un taller de visualización de datos en el desarrollo de habilidades de representaciones gráficas: estudio comparativo pre y post implementación

Dylan Soto Umaña • Eduardo Aguilar Fernández • José Andrey Zamora Araya

Abstract

The objective of this investigation was to compare the skills of prospective teachers at the Universidad Nacional de Costa Rica related to graphical representations of information before and after implementation of a data visualization workshop. The workshop was implemented and two questionnaires were developed to collect information about the students' perceptions of their skills in data visualization. Student perceptions were then compared using Student's t-test for paired measurements and the Wilcoxon-Pratt rank test. The analysis showed that student ratings of their graphical composition skills ($p < 0.0001$), graphical interpretation ($p < 0.0001$), optimal representation ($p < 0.0001$), relationship with other representations ($p = 0.0004$) and contextual relationship ($p = 0.0003$) increased significantly after implementation of the workshop, indicating that participating students had acquired fundamental tools for understanding, interpreting and properly using graphical representations of statistical data in different contexts involving large amounts of information.

Keywords: teacher skills, mathematics teachers, teacher education, graphics, visualization.

Resumen

El objetivo de la presente investigación fue comparar las habilidades relacionadas con representaciones gráficas de la información en docentes en formación de la Universidad Nacional de Costa Rica, antes y después de la implementación de un taller de visualización de datos. Se implementó un taller sobre visualización de datos y se elaboraron dos cuestionarios para recolectar la información de la percepción sobre las habilidades del estudiantado en la visualización de datos. La percepción estudiantil fue comparada mediante la prueba t-student para medidas pareadas y mediante la prueba de rangos de Wilcoxon-Pratt. Los análisis mostraron que las valoraciones estudiantiles sobre sus habilidades de composición gráfica ($p < 0,0001$), interpretación gráfica ($p < 0,0001$), representación óptima ($p < 0,0001$), relación con otras representaciones ($p = 0,0004$) y relación contextual ($p = 0,0003$) presentaron aumentos significativos luego de la implementación del taller. Se logró un progreso significativo en todas las áreas evaluadas luego de la implementación del taller, lo cual indica que el estudiantado participante adquirió herramientas fundamentales para comprender, interpretar y utilizar adecuadamente las representaciones gráficas estadísticas en diferentes contextos que involucren gran cantidad de información.

Palabras clave: competencias docentes, docentes de matemáticas, formación docente, gráficas, visualización.

Received: September 13, 2024 | *Approved: May 12, 2025* | *Published: June 10, 2025*
https://doi.org/10.33010/ie_rie_rediech.v16i0.2300

ie revista de investigación educativa de la Rediech
vol. 16 • 2025 • e2300 • ISSN: 2448-8550

Introduction

Vast amounts of information are now being made available in different media, which allows us to analyze social, economic, political, and other phenomena. As a result, having presentation tools that allow us to communicate information quickly and effectively is becoming increasingly important. Gal (2002) mentions that it is desirable for people to be able to interpret, critically evaluate, and communicate statistical information, and to develop skills that allow them to identify, interpret, and use information presented in lists, tables, and graphical presentations, among other media.

Statistical graphs, for example, can be used to efficiently summarize, analyze, explore, and discover relationships in data, to make diagnoses, and establish arguments, or even as a substitute for tables (Vanderplas et al., 2020). In other words, statistical graphs allow us to visually display the behavior of a set of data when describing, explaining, or analyzing a phenomenon in a particular context.

Traditional statistical graphics such as histograms or box diagrams often provide us with a useful perspective on the phenomenon under study. However, the large amount of information currently being generated poses a series of challenges to producing graphical representations that allow this information to be adequately interpreted, since as the amount of data increases, the possibility of capturing this information and explaining it using a traditional graphics decreases.

The tool known as data visualization provides an alternative approach to address the challenges to adequate representation of information that have arisen. Data visualization allows complex information to be displayed in a single graphic, which can be rapidly created, and clearly and efficiently conveys a message. Information is presented in a compact, accessible, and engaging way, stimulating people's interest in interpreting the information presented in the graphic (Grant, 2019). Furthermore, its

Dylan Soto Umaña is a Professor and Researcher at the Ministry of Public Education, Costa Rica. He holds a degree in Mathematics Teaching from the Universidad Nacional and works in statistics education research. Email: dylan1998u@hotmail.com. ID: https://orcid.org/0009-0004-8131-2627.

Eduardo Aguilar Fernández (corresponding author) is a Professor and Researcher at the School of Mathematics of the Universidad Nacional de Costa Rica. He holds a Master's degree in Statistics from the Universidad de Costa Rica. His recent publications include "Análisis de correspondencia simple para estudiar la relación entre factores del abandono escolar y el estrato del colegio de procedencia en la Universidad Nacional de Costa Rica" (2024) and "Las tarjetas como recurso didáctico para la enseñanza de la estadística: la experiencia en un curso de didáctica específica" (2024). Email: eduardo.aguilar.fernandez@una.ac.cr. ID: https://orcid.org/0000-0002-7864-2391.

José Andrey Zamora Araya is a Professor and Researcher at the School of Mathematics of the Universidad Nacional de Costa Rica, and a Professor in the School of Statistics of the Universidad de Costa Rica. He holds a PhD in Education from the Universidad Estatal a Distancia, Costa Rica, a Master's degree in Statistics, and a Professional Master's degree in Development Economics. His recent publications include the article "Las tarjetas como recurso didáctico para la enseñanza de la estadística: La experiencia en un curso de didáctica específica" (2024). Email: jzamo@una.ac.cr. ID: https://orcid.org/0000-0001-6050-5850.

Translation to English: **Xinia María Rodríguez Castillo**.

use permits the generation and communication of knowledge based on data from a fundamentally intuitive perspective (Asamoah, 2022). In other words, it is a tool that enhances skills in interpreting, critically evaluating, and communicating information through different types of representations. Data visualization helps people process large amounts of information (Llaha and Aliu, 2022), discover patterns, make inferences, and draw conclusions (Zalaan and Najimb, 2022), providing types of representations that help an audience to observe aspects of data sets that would otherwise not have been perceived when analyzing this data (Araujo, 2023). It can also be used in any area of knowledge to efficiently communicate research results (Grant, 2019). It is therefore clear that developing the skills that allow for the implementation of this tool in different fields of scientific knowledge is highly important.

The American Statistical Association (ASA) proposes that university programs should foster the development of skills in data presentation. It is important for students to acquire knowledge on how to visualize information and how to use this visualization for analysis (GAISE College Report ASA Revision Committee, 2016). Likewise, academia should provide individuals with the necessary training so that they can leverage visualization skills to explore and generate knowledge from data (Asamoah, 2022). Academic studies generate a large amount of information that must be attractive, accessible, and understandable to the general public, and data visualization can be a powerful tool for communicating research results in an understandable manner.

In Costa Rica, the Bachelor's and Master's Degree in Mathematics Teaching program at the Universidad Nacional contains three courses in statistics, including Statistics and Probabilities, which covers a range of knowledge related to the graphical representation of information through the use of bar, line, and pie charts, as well as the use of histograms, frequency polygons, and box plots in the field of descriptive statistics. Through the use of these charts, students are expected to develop the skills necessary to select the most appropriate of these tools for the representation of a given data set, as well as to interpret and communicate relevant aspects related to the distribution of data according to context (Universidad Nacional, 2017).

However, the course requirements do not consider other forms of information presentation, such as cartograms, word clouds, violin diagrams, which allow for the representation and analysis of larger and more complex databases. Complementary training in data visualization could be a tool that allows prospective teachers to analyze large amounts of data and strengthen the development of skills related to graphical representation, as described in the course syllabus.

Thus, the objective of this study was to compare skills related to graphical representations of information among prospective teachers in the Bachelor's and Master's Degree in Mathematics Teaching programs at the Universidad Nacional de Costa

ie revista de investigación educativa de la Rediech
vol. 16 • 2025 • e2300 • ISSN: 2448-8550

Rica before and after implementation of a data visualization workshop, to evaluate the workshop's effectiveness in developing these skills.

Graphical representations of information

According to Friel et al. (2001), a statistical graphic contains four structural components: (1) graphic frame, which consists of the axes, scales, reference marks on each axis, among other elements, allowing for display of information about the measures used and the measured data; (2) specifiers, which are the elements used to represent data and show the relationships between variables, typically lines, bars, or other marks; (3) labels that provide contextual information about the graphic and the variables represented, such as the type of measure used, the data to which that measure is applied, or the graph title; and (4) background, which includes the colors, grid, and images on which the graphic can be placed.

This investigation considers a statistical graphic to be any form of information presentation that complies with the structure described by Friel et al. (2001), which is intended to facilitate the understanding of a phenomenon in a synthesized manner that draws attention and allows for identification of patterns and trends without misleading others – that is, an illustration that attempts to present data in the best possible fashion.

Data visualization

Visualization constitutes "the study of how to represent data using a visual or artistic approach rather than a traditional reporting method" (Yuk and Diamond, 2014, p. 7). Unwin (2020) states that "data visualization involves drawing figures to display data" (p. 2), while the National Forum on Education Statistics (2016) states that data visualization involves transforming data so that the information it provides can be appreciated through a visual presentation. It can be characterized as the presentation of an image or figure that is the end result of a process involving a variety of communication methods, presentation technologies, and media formats to visually reveal the meaning of the data to an audience. That is, data visualization is not simply the presentation of a graphic figure, but rather the final product of a process with several stages.

In descriptive statistics, data visualization primarily fulfills two general functions: the exploratory function of the data, and the explanatory function of the data (Iliinsky and Steele, 2011). From an exploratory perspective, data visualization helps to observe elements that are not initially perceived when analyzing a data set (Araujo, 2023), helping to quickly identify characteristics or relationships embedded in the data by a consideration of the curves, patterns, trends, or outliers that are visible at first glance in graphical representations.

From the perspective of data explanation, visualization seeks to convey the final results of research to people unfamiliar with the data's context (Iliinsky & Steele, 2011). In other words, the explanatory function allows people to understand the story behind the data. Visualization thus not only aims to explain the data with a graphic but goes further, seeking to justify which graphic best communicates the results of the research. Therefore, the type of representation used should not be chosen at random, but should rather be selected using a series of principles that promote more effective communication, based on different models of cognitive information processing (Pascual, 2016).

Gestalt design principles

The Gestalt design principles developed by psychologists Max Westheimer, Kurt Koffka, and Wolfgang Köhler which can be used to validate data visualization designs are known as proximity, similarity, connectivity, continuity, closure, and common direction (Ware, 2013). These principles describe how patterns in visual representations are perceived, and define a series of characteristics that a graphical representation must meet to achieve its goal of conveying a message visually.

These design principles, while clarifying the characteristics that an adequate graphical representation must possess, must be complemented by other theoretical constructs that contribute to the information encoding process, such as visual properties and representational skills.

Visual properties of encoding

Visual properties of encoding are related to colors, shapes, sizes, lines, and text that act as common specifiers for data visualization graphics.

The appropriate selection of color must be carefully considered, since colors are sometimes selected irrationally, based on what is most striking to graphic designers, which is considered to be inappropriate practice (Iliinsky & Steele, 2011). One of the main conditions that must be taken into account when selecting graphic colors is color blindness, as the use of color shades that may not be distinguished by color-blind people should be avoided (Ware, 2013).

Size is used to represent the relative importance of each of the entities present in the graphic, since when one element is larger than another, it draws more attention, perceptually indicating that it is a more important element in the graphical representation (Iliinsky & Steele, 2011).

Labels also constitute an important part of the structure of any statistical graphical representation, so identifying the clarity of the letters involved in the graphic, and their size, position and typography is highly important in the data visualization process, since as Iliinsky and Steele (2011) state, words that are added to graphics must be clearly necessary and legible.

ie revista de investigación educativa de la Rediech
vol. 16 • 2025 • e2300 • ISSN: 2448-8550

Skills related to the graphical representation of information

According to Friel et al. (2001), there are five skills related to the construction, interpretation, and evaluation of a graphical representation: (1) recognizing the components of graphics, the interrelationships between them, and their effect on the graphical presentation of information; (2) understanding the relationships between a table, a graphic, and the data being analyzed; (3) interpreting the information displayed in a graphic by answering different levels of questions associated with graphical comprehension; (4) recognizing the usefulness of one graphic and its preference over another, considering the judgment criteria involved and the types of data being represented; and (5) establishing the relationship of the context of the graphic to the goal of interpreting information to make sense of what the data show in the graphic, while avoiding personalization of the data.

The skills of recognizing components, recognizing component usefulness, and understanding the relationships between different presentation formats are related to the process of constructing graphical representations. Based on an understanding of the data and structural components, necessary elements are determined so that the information can be transformed and displayed through figures that make it possible to appreciate the different characteristics presented by the data, incorporating design principles and data visualization theory.

Furthermore, the skills of interpreting information in a graphic and establishing its context are linked to the interpretation of graphics. They are associated with answering questions related to the information presented based on its relationship to the context in which the data is embedded, validating the principle that a graphic is not just an image.

Methodology

Type of research

A pre-experimental study was conducted with a single-group pre-test and post-test design. According to Salinas and Cárdenas (2008), this design detects changes that a treatment has brought about compared to an initial observation. This methodology is often applied to quantify the effectiveness of teaching techniques in the learning of concepts and the development of specific competencies.

Participants

The study population consisted of 36 students from the Quantitative Research and Teaching Development and Practice courses in the Bachelor's and Master's Degree in Mathematics Teaching program during the second semester of 2021. They participated in a workshop on data visualization that addressed theoretical foundations and the construction and interpretation of statistical graphical representations provided by this tool.

The workshop consisted of four two-hour sessions, held between September 1 and October 8, 2021, delivered virtually by the researchers using Google Meet and Zoom. A brochure was prepared to present the theoretical foundation of data visualization, and shared with participants. Microsoft Power BI software, freely accessible to the entire Universidad Nacional student body was used to create graphical representations. During the first 30 minutes, theoretical information related to statistical graphics was presented. Students were then provided with a database that they were asked to use to construct different graphics based on the specifications provided in the instructions. Finally, the results were presented, and discussions were held in which each participant's results were evaluated.

In addition, the students who participated in the workshop completed two questionnaires: the first before the workshop activities began and the second at the end of training. Of the total population, six students were unable to attend either the first or last session of the workshop. Therefore, information from both questionnaires for this group was not available. Due to incomplete information, they were not included in the study. Thus, a sample of 30 students was formed. They participated voluntarily and completed the two instruments involved in the research.

It is important to mention that, based on the population size (36), the sample of 30 students allowed us to assess the opinions of 83% of the workshop attendees, providing a significant sample for evaluating the workshop.

Instruments

The questionnaires used to collect information on students' perceptions of data visualization skills before attending the workshop—that is, with only prior statistical knowledge and skills—and after attending the workshop, had several sections and were constructed based on a literature review on the concept of data visualization.

The first part of the initial questionnaire was intended to collect general information from the participants, such as their ID number, age, sex, the academic year in which they passed the Statistics and Probability course, and whether or not they had repeated the course. The second part of the questionnaire was a series of Likert scale questions related to students' perceptions of their knowledge and skills related to the graphical representation of statistical information. This scale included a set of 29 statements or judgments, with five response options in which each person could express their level of agreement or disagreement with the statement regarding their knowledge and skills in the field of graphical representations. These items were divided into five dimensions or subscales corresponding to the five data visualization skills proposed by Friel et al. (2001), which were operationalized as graphic composition, relationships between forms of representation, graphic interpretation, optimal graphic representation, and contextual relationship. In the second questionnaire, the first questions, related to general information, only asked for the identification of each student, to establish correspondence with their first measurement. The second part consisted of the same Likert scale questions used in the first questionnaire, to obtain students' perceptions of their skills in the subject of graphic representation once the workshop was completed.

To validate the questionnaires, an instrument was constructed, and teachers from the School of Mathematics at the Universidad Nacional who specialized in statistics were asked to provide their opinions on the scale items. This was done to determine whether they were appropriately worded, relevant to the study objective, and adapted to each of the subscales of the graphical information representation skills mentioned by Friel et al. (2001). Additionally, a small-scale pilot study was conducted with five students graduating from the Mathematics Teaching program at the Universidad Nacional, who completed the questionnaire to determine whether the items were correctly worded or presented difficulties in understanding. Finally, the level of internal consistency for each of the subscales was evaluated using Cronbach's alpha coefficient. Observations of the individuals who validated the instruments allowed for improvements in the wording and structure of questionnaires.

Data analysis

The responses to the Likert scale questions included in the two questionnaires were analyzed, comparing the perceptions of participating students attending the workshop regarding their skills in graphically representing statistical information at the beginning and end of the workshop to determine whether the workshop's contribution was significant. The total values of responses to the items included in each subscale ranged from 4 to 40, representing each student's perception of each subscale.

To validate the workshop's contributions, a descriptive analysis of the changes obtained was performed to determine whether there was a significant increase or decrease in this total for each subscale. The differences in means between the responses to the second questionnaire and the first were determined, and the Anderson-Darling test was applied to identify whether these differences were normally distributed.

Student perceptions were then compared using Student's t-test for paired measures in cases where the differences showed a normally distributed pattern. In those cases in which it was determined that these differences were not normally distributed, the Wilcoxon-Pratt rank test was applied. Finally, to evaluate the reliability of each subscale, an internal consistency analysis was performed using Cronbach's alpha coefficient. For these statistical calculations, R software version 4.3.2 (R Core Team, 2023) was used, and p values less than 0.05 were considered significant.

Results

Cronbach's alpha coefficients for each of the questionnaires administered showed values greater than 0.65 (Table 1), which were considered satisfactory to ensure the reliability of the subscales for the purpose of the study.

Table 1
Reliability analysis for each of the subscales of skills for graphically representing information

Subscale	Alpha 1	Alpha 2
Graphical composition	0,82	0,80
Graphical interpretation	0,85	0,71
Optimal representation	0,85	0,85
Relation with other representations	0,76	0,70
Contextual relation	0,68	0,75

Source: Prepared by the authors.

Next, a comparison was made between the subscale responses for each participant in the initial questionnaire and the final questionnaire subscale to identify whether there was any significant increase or decrease in these scores after the workshop.

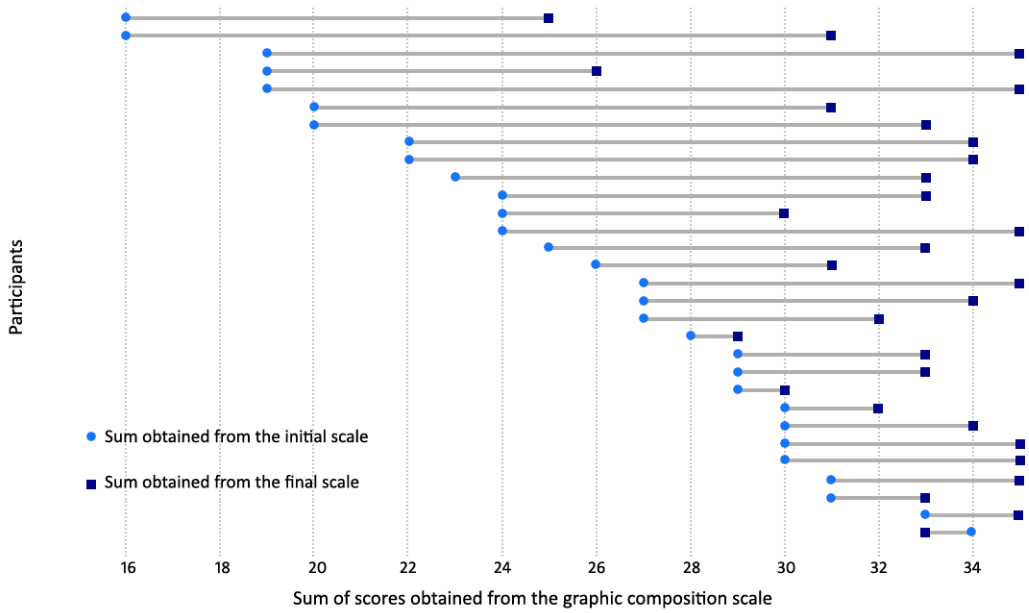
Comparison of the responses revealed that most participants showed increases in all subscales, in the sum of the scores obtained in the final questionnaire compared to the sum of scores in the initial questionnaire. In the case of the graphical composition subscale, for example, it was found that these increases are greater for some participants that for others. Decreases can also be observed in some participants; however, in the case of these variations, the decreases occur in one or two points of the subscale, mostly among students who initially perceived high scores in the area of graphical representation of statistical data (Figure 1).

The general perception of the behavior of values for each subscale tends to indicate that there was an increase in participants' perceptions of their abilities by the end of the workshop compared to their initial perceptions. To determine whether these increases were significant, the average scores obtained for each subscale were analyzed and the results compared.

The Anderson-Darling normality test was the most appropriate statistical test for the analysis of these data distributions, and the results of using this test indicated that the subscales of graphical composition ($p = 0.373$), graphical interpretation ($p = 0.119$), and optimal graphical representations ($p = 0.33$) were normally distributed, while the null hypothesis that the differences exhibit normal behavior was rejected for the subscales of relationship with other representations ($p = 0.003$) and contextual relationship ($p = 0.001$).

Figure 1

Universidad Nacional. Student differences in graphical composition subscale scores before and after the data visualization workshop



Source: Prepared by the authors.

For the graphical composition, graphical interpretation, and optimal graphical representation subscales, results of the paired-measures Student's t-test provided statistical evidence that the true average of the scores obtained for each subscale of the initial questionnaire was significantly lower than the true average of the scores obtained for the final questionnaire subscale. For the relationship with other representations and contextual relationship subscales, which are not normally distributed, results of the Wilcoxon-Pratt rank test indicated that the ranges of scores obtained for these two subscales in the initial questionnaire were significantly lower than the ranges of scores obtained for these subscales in the final questionnaire (Table 2).

Table 2

Universidad Nacional. Scores obtained by students on each of the subscales of skills for graphical representation of information before and after the data visualization workshop

Subscale	Mean 1	Deviation 1	Mean 2	Deviation 2	p-value
Graphical composition	25,47	5,00	32,57	2,56	< 0,0001
Graphical interpretation	23,20	4,38	27,60	2,25	< 0,0001
Optimal representation	28,63	5,79	37,30	2,99	< 0,0001
Relation with other representations	17,70	2,58	19,53	0,93	0,0004
Contextual relation	12,83	2,15	14,53	0,89	0,0003

Source: Prepared by the authors.

Students' responses to the questionnaires showed significant increases in the scores for their knowledge and skills related to the graphical representation of statistical information. These results indicate that the workshop contributed to the development of the five skills that Friel et al. (2001) consider necessary for adequate mastery of the processes of constructing, interpreting, and evaluating a graphical representation.

The results showed an improvement in students' perception of their graphical composition skills. The ability to recognize the components of a graph and how they are interrelated is perhaps a first step toward developing skills related to the use of tools such as data visualization for the study of graphical representations. Some research suggests that understanding a graphical representation of statistical data requires identifying its structure, the variables involved and their relationships, the type of axes, and the graphical elements used to represent the data (Batanero et al., 2021; Contreras et al., 2021). Use of statistical graphics for data analysis is an important component of statistics courses at different educational levels. Therefore, it is important for teachers who deliver courses in this discipline to understand the different components required for an adequate understanding and analysis of graphical representations. For this reason, the workshop was instrumental in strengthening this skill for prospective teachers.

The workshop contributed to the development of graphical interpretation skills, which is consistent with one of the main objectives of mathematics education: enabling students to interpret and understand graphical presentations of information (Romero-Ariza et al., 2024), and is also consistent with the content of the curriculum for the Bachelor's and Master's Degree in Mathematics Teaching program, which states that students are expected to be able to interpret information presented in a statistical table or graphic (Universidad Nacional, 2017). This information processing skill is essential for employment in the 21st century (Patahuddin and Lowrie, 2019), when workers seek to answer different levels of questions related to the information contained in data.

It is important to note that several factors can interfere with the proper interpretation and reading of statistical graphics, such as characteristics of the readers, the type of graphic (Guo et al., 2020), or lack of statistical training (Garzón-Guerrero and Jiménez, 2021). Therefore, educational processes related to the study of statistical graphics must respond to the need to train individuals capable of interpreting and communicating statistical information presented in different formats.

The study data showed that the aspects covered in the workshop improved the ability of the participants to identify the usefulness of one type of graphic over another. While there are documented recommendations on how to present data, the best version of the figure chosen is obtained as part of the work of the person who creates it, demonstrating their ability to distinguish the advantages and disadvantages

ie revista de investigación educativa de la Rediech
vol. 16 • 2025 • e2300 • ISSN: 2448-8550

of different options and to determine which one best fits the data they will be presenting (Midway, 2020), and making decisions about the coding and design that they consider will most accurately and convincingly communicate their interpretation of the data (Burns et al., 2020).

To construct or design a graph, it is necessary to consider the characteristics of the target audience as well as the type of information to be conveyed (Guo et al., 2020). Selecting the method used to represent a data set is not only based on documented technical aspects, but also requires appropriate judgment to select the graphic format that allows the information to be conveyed clearly and simply. Therefore, it is essential that academic research be carried out, providing both academic and non-academic staff with the ability to create well-designed graphs that are useful for conveying information (Vanderplas et al., 2020). This aspect is of utmost importance in the training of teaching staff who will work in secondary education, as they must be able to select appropriate graphic formats to represent and convey information from a data set, as well as possess the ability to construct, criticize, and analyze them adequately to guide their students in developing these skills.

The students' questionnaire responses also indicated that the workshop contributed to the development of their ability to establish relationships between different forms of data representation. Data visualization can be understood as the final product of a process that involves several stages (National Forum on Education Statistics, 2016), ranging from data collection to data representation. In this process, statistical tables play an important role, as they contribute to the classification and organization of data for subsequent representation through the use of graphics (Pallauta et al., 2021). On the other hand, statistical tables and graphics are objects that promote transnumeration (Wild and Pfannkuch, 1999), thereby allowing a greater understanding of the phenomenon under study, as students acquire skills including the ability to read, compare, infer, or make decisions (Arredondo et al., 2021). Statistical graphics and tables are essential tools for presenting data from the environment. It is important for prospective teachers to understand the different forms of representation and their relationships. These elements are part of secondary education curricula, which aim to encourage students to develop skills that allow them to display, summarize, and communicate information in various types of representations (graphical or tabular) (Ministry of Public Education [MEP], 2012) and use them as data analysis tools. Some representations can be used to analyze the mode, maximum, minimum, or pattern of variability of a data set (MEP, 2012).

Statistical graphics, as forms of information representation, require appropriate treatment in teacher training courses to achieve an adequate data transformation process, so that teachers responsible for teaching statistics can offer students the opportunity to understand and transform the data that they interact with.

The workshop also contributed to the development of students' ability to establish relationships between data in graphical representations and context, a fundamental issue in the particular case of statistics, in which context is always present. Graphics are not isolated elements but rather part of a whole, hence the importance of understanding the context for data visualization (Unwin, 2020). Achieving the goal of interpreting and understanding statistical information requires not only statistical knowledge per se but also other types of knowledge, among which contextual knowledge stands out (Gal, 2002).

Establishing relationships between data and context is closely related to teaching, and is one of the skills that the MEP educational programs aim to develop. The impact of the workshop on its development in prospective teachers is thus especially important, since the context cannot be separated from the content to be taught in statistics courses. Based on student questionnaire responses, the workshop activities contributed to the development of skills that allow students to understand the usefulness of data visualization in the analysis of social or economic phenomena. These activities were designed to help students appreciate the importance of making decisions contextualized with the information described in the activity, and the workshop methodology was aimed at providing tools that allowed them to examine patterns or analyze information from different contexts.

Conclusions

The research instruments developed in this investigation allowed students to evaluate their progress in developing graphical representation skills through responses to Likert scale questions at the beginning and end of the workshop. Analysis of their responses showed that the workshop achieved significant progress in all areas assessed, and that participating students acquired fundamental tools to adequately understand, interpret, and use statistical graphical representations in different contexts involving large amounts of information.

Currently, people receive a large amount of information through different media, so it is important to consider the possibility of establishing educational processes that promote the development of skills for working with large data sets, supporting the training of individuals so that they acquire the ability to present and analyze information efficiently.

The results obtained support the importance of these types of workshops as effective tools for learning statistics, as they allow for in-depth exploration of different aspects of the field and contribute to the development of skills in the field of data visualization, particularly those covered in a Statistics and Probabilities course, given its role as a tool for transforming information.

Initiatives such as the one presented in this article promote both the acquisition of skills in interpreting and using graphics and the development of critical thinking in individuals, by encouraging questions not only about the figures themselves but also about the sources from which they are derived. Since graphical representations are one of the most common ways of transmitting information, both in traditional and virtual media, adequate training in their proper interpretation and evaluation is essential for everyone.

It is also important to consider establishing a comprehensive approach, including all stakeholders in the educational process, to contribute to the development of skills that allow people to appropriately use the information generated in different media, so that they can identify, interpret, communicate, and use data effectively in activities that can benefit society.

Despite the workshop's benefits, the study has some limitations, such as the lack of a comparison group, which prevents us from identifying the factors that explain the variations between the first and second observations. Furthermore, the virtual format in which the workshop was conducted prevents us from generalizing its implications to other types of modalities, such as in-person learning.

Given these limitations, it is recommended that further research delve deeper into the topic by involving control groups that can help identify factors that explain changes in students' perceptions of graphic representation skills, thereby highlighting the impact that this type of workshop generates on the statistical training of those responsible for their future teaching. Further research is also encouraged to include this type of workshop in an in-person format, to replicate them in other educational contexts and expand the scope of this study.

The implementation of this type of workshop is also recommended to strengthen the training of teachers and those interested in the field of information presentation, allowing them to better stimulate their students' critical thinking skills to identify, interpret, and use information from graphical representations in decision-making.

In a world constantly generating data, it is not only appropriate but also necessary to distinguish between graphics that provide useful and relevant information and those that seek to confuse, deceive, or show only part of the picture.

References

Araujo, C. (2023). Data visualization in the Information Society. *Seminars in Medical Writing and Education*, *2*, 25. https://doi.org/10.56294/mw202325

Arredondo, E., Vásquez, C., & García-García, J. (2021). Análisis de las tablas y los gráficos estadísticos en libros de texto de Chile y España para la educación infantil. *Ridema, Revista de Investigação e Divulgação em Educação Matemática*, *5*(1), 1-26. https://periodicos.ufjf.br/index.php/ridema/article/view/35566/23526

Asamoah, D. (2022). Improving data visualization skills: A curriculum design. *International Journal of Education and Development using Information and Communication Technology (IJEDICT)*, *18*(1), 213-235. https://eric.ed.gov/?id=EJ1345404

Batanero, C., Garzón-Guerrero, J., & Valenzuela-Ruiz, S. (2021). Sentido gráfico y su importancia en la comprensión de la información sobre la COVID. *Paradigma*, *42*(1), 206-224. https://doi.org/10.37618/paradigma.1011-2251.2021.p206-224.id996

Burns, A., Xiong, C., Franconeri, S., Cairo, A., & Mahyar, N. (2020). How to evaluate data visualizations across different levels of understanding. *2020 IEEE Workshop on Evaluation and Beyond-Methodological Approaches to Visualization (BELIV)*, 19-28. https://arxiv.org/pdf/2009.01747.pdf

Contreras, J. M., Molina-Portillo, E., & Contreras, J. (2021). Nivel de lectura gráfica de futuros profesores de educación primaria como componente de la cultura estadística. *PNA: Revista de Investigación en Didáctica de la Matemática*, *15*(3), 137-158. https://revistaseug.ugr.es/index.php/pna/article/view/15271

Friel, S. N., Curcio, F. R., & 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. http://snoid.sv.vt.edu/~npolys/projects/safas/749671.pdf

GAISE College Report ASA Revision Committee (2016). *Guidelines for Assessment and Instruction in Statistics Education (GAISE) College Report 2016*. https://www.amstat.org/asa/files/pdfs/GAISE/GaiseCollege_Full.pdf

Gal, I. (2002). Adults' statistical literacy: Meanings, components, responsibilities. *International Statistical Review*, *70*(1), 1-51. https://www.stat.auckland.ac.nz/~iase/publications/isr/02.Gal.pdf

Garzón-Guerrero, J. A., & Jiménez, M. (2021). Un estudio exploratorio de la competencia gráfica de futuros profesores de Portugal e Italia a través de la interpretación de diagramas estadísticos de barras y sectores extraídos de la prensa escrita. *Números*, *106*, 32-42. https://hdl.handle.net/10481/88973

Grant, R. (2019). *Data visualization: Charts, maps and interactive graphics*. CRC Press.

Guo, D., Zhang, S., Wright, K. L., & McTigue, E. M. (2020). Do you get the picture? A meta-analysis of the effect of graphics on reading comprehension. *AERA Open*, *6*(1), 1-20. https://doi.org/10.1177/2332858420901696

Iliinsky, N., & Steele, J. (2011). *Designing data visualizations*. O'Reilly Media.

Llaha, O., & Aliu, A. (2022). Students' performance evaluation in higher education using data visualization techniques. *Annals of Philosophy, Social & Human Disciplines*, *2*, 43-55. http://www.apshus.usv.ro/arhiva/2022II/APSHUSDec2022_43_55.pdf

Midway, S. R. (2020). Principles of effective data visualization. *Patterns*, *1*(9), 100141. https://doi.org/10.1016/j.patter.2020.100141

MEP [Ministerio de Educación Pública] (2012). *Programas de estudios de Matemáticas*. https://www.mep.go.cr/sites/default/files/media/matematica.pdf

National Forum on Education Statistics (2016). *Forum guide to data visualization: A resource for education agencies (NFES 2017-016)*. National Center for Education Statistics. https://nces.ed.gov/pubs2017/NFES2017016.pdf

Pallauta, J. D., Arteaga, P., Begué, N., & Gea, M. M. (2021). Análisis de la complejidad semiótica y el contexto de las tablas estadísticas en los libros de texto españoles de secundaria. *Educação Matemática Pesquisa*, *23*(4), 193-220. https://doi.org/10.23925/983-3156.2021v23i4p193-220

Pascual, V. (2016). *Buenas prácticas en visualización de datos*. Universitat Oberta de Catalunya. https://hdl.handle.net/10609/59025

Patahuddin, S. M., & Lowrie, T. (2019). Examining teachers' knowledge of line graph task: A case of travel task. *International Journal of Science and Mathematics Education*, *17*, 781-800. https://doi.org/10.1007/s10763-018-9893-z

R Core Team (2023). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Romero, M., Quesada, A., & Estepa, A. (2024). Promoting critical thinking through mathematics and science teacher education: The case of argumentation and graphs interpretation about climate change. *European Journal of Teacher Education*, *47*(1), 41-59. https://doi.org/10.1080/02619768.2021.1961736

Salinas, P., & Cárdenas, M. (2008). *Métodos de investigación social* (2nd ed.). Ediciones Universidad Católica del Norte. https://biblio.flacsoandes.edu.ec/libros/digital/55372.pdf

Universidad Nacional (2017). *Plan de estudios de la carrera bachillerato y licenciatura en la enseñanza de la matemática*.

http://www.matematica.una.ac.cr/index.php/documentacion-digital/category/7-planes-de-estudio

Unwin, A. (2020). Why is data visualization important? What is important in data visualization? *Harvard Data Science Review*, *2*(1). https://doi.org/10.1162/99608f92.8ae4d525

Vanderplas, S., Cook, D., & Hofmann, H. (2020). Testing statistical charts: What makes a good graph? *Annual Review of Statistics and Its Application*, *7*, 61-88. https://doi.org/10.1146/annurev-statistics-031219- 041252

Ware, C. (2013). *Information visualization: Perception for design*. Morgan Kaufmann.

Wild, C. J., & Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. *International Statistical Review*, *67*(3), 223-265. http://www.iase-web.org/documents/intstatreview/99.Wild.Pfannkuch.pdf

Yuk, M., & Diamond, S. (2014). *Data visualization for dummies*. John Wiley & Sons.

Zalaan, Z. H., & Najimb, S. A. (2022). A comparative study for skeleton representation methods using data visualization. *Journal of Al-Qadisiyah for Computer Science and Mathematics*, *14*(2), 22-32. https://doi.org/10.29304/jqcm.2022.14.2.933

Cómo citar este artículo:

Soto Umaña, D., Aguilar Fernández, E., y Zamora Araya, J. A. (2025). Evaluation of the effectiveness of a data visualization workshop in the development of graphical representation skills: A pre- and post-implementation comparative study. *IE Revista de Investigación Educativa de la REDIECH*, *16*, e2300. https://doi.org/10.33010/ie_rie_rediech.v16i0.2300