



# Public engagement of underserved students with open civic data

Lucía Zapata-Cardona<sup>1</sup>

Accepted: 5 October 2024

© The Author(s) 2025

## Abstract

Today, various non-governmental organizations proclaim the need for public engagement of citizens, a concept that is integrated into the notion of global citizenship. However, much of the mathematics that students find in the primary school curriculum falls short to support that need. This lack of curriculum alignment with global needs has implications for the development of communities, but they are much more severe for students from underserved populations who systematically suffer different types of exclusions that increase their social disadvantages. This study explores the traits of global citizenship that students from underserved populations exhibit as they progress through a learning experience with open civic data on a global issue (climate change). The learning experience was used as the basis for an individual interview with five participants (9–10 years old) from public schools in low-income neighborhoods of a metropolitan area in northern Colombia. Participants were offered a dataset (data) ready to use and they were asked to undertake data handling actions (data-ing) using freely available web-based software. Participants were asked “What does this dataset say?” and their verbal statements were analyzed. The main findings suggest that in data and data-ing on open civic data reflecting global issues, participants exhibited understanding, belonging and action which are traits of global citizenship. In terms of understanding, participants described the behavior of the variables, suggested informal inferences, and made associations between variables. In terms of belonging and action, participants’ statements suggested including themselves as part of the community, taking responsibility, and proposing transformative actions.

**Keywords** Public engagement · Open civic data · Data and data-ing · Underserved populations · Climate change

## 1 Introduction

Data, understood as objects composed of information, permeate almost all aspects of globalized society, and have become an integral part of the daily lives of citizens of all ages (children, adolescents, and adults) in academic, cultural, work, and civic related environments. Today, a multiplicity of data is being generated at previously unimaginable speeds and volumes that citizens need to manage and interpret for their public engagement, that is, to understand the impact of daily human actions on a global world and to be able to collectively participate in society (Johnson & Morris, 2010). Within this multiplicity of data is open civic data, which typically takes the form of public records (Engel, 2019; Ridgway, 2022). These data reveal transcendental

aspects of society and are important sources of information for decision making and educating global citizens.

A globalized world is an asymmetrical network (the participation of communities is unequal) that connects societies and the lives of individuals, and consequently, environmental problems are also a source of connection (Gimeno, 2005). The global citizen referred to by UNESCO (2015, 2018)—a citizen immersed in a world characterized by global issues that increase social asymmetries (impoverishment, discrimination, migration, environmental degradation)—requires skills not only to manage this avalanche of political, economic, social, and cultural data, but also to develop a sense of understanding, belonging, and commitment to public engagement. It is not enough for citizens to recognize that they are immersed in a world driven by data (data); they also need to have data literacy (Gal, 2022) that enables them to prepare and handle open civic data (data-ing) for lifelong learning and informed participation in society (Makar et al., 2023).

---

✉ Lucía Zapata-Cardona  
lucia.zapata1@udea.edu.co

<sup>1</sup> Universidad de Antioquia, Medellín, Colombia

Everyday citizens need data literacy to successfully navigate a data-driven society. Today, citizens are in constant contact with search engines, streaming services, social media, wearables, fitness trackers, and news portals, where they leave their digital footprints. In addition, citizens provide civic information by living within communities. The use of algorithms, in the current demand for datafication of social practices, stratifies individuals and communities, exacerbating discriminatory practices in the most vulnerable populations. In this regard, O'Neil (2016) points out that students who grow up in low-income households are less likely to receive a high-quality college education later in life. The models used for college admissions and the college ranking system exacerbate social inequalities by disproportionately affecting those with the least economic opportunities. Facial recognition software is less accurate in African American populations, a condition that creates serious racial bias (Khalil et al., 2020). As if that were not enough, the consequences of global problems, especially those related to environmental degradation, are felt more acutely by vulnerable populations (Lanzetta, 2021). For example, heat waves are much more severe in low-income neighborhoods (Chen et al., 2023). In this sense, it is necessary to cultivate data literacy among vulnerable populations as a tool that allows them to understand, empathize, and act on global issues as global citizens.

Young children's ability to read and interpret complex data has often been underestimated (Zapata-Cardona, 2023), and school curricula oversimplify the complexity of data management by overstimulating the study of simple univariate datasets. Research has shown that the data management resources needed by today's citizens are much more complex than those provided in school curricula (Gould, 2021), and that the social experiences of young children are highly influential in providing future opportunities for critical reflection and transformative action (Santamaría-Cárdaba & Lourenço, 2021). There is a need to create opportunities for young children to engage with open civic data to develop skills that will enable them to participate in society (Gal, 2022) and support proposals for social change (Gutstein, 2006). Involving young children in the study of open civic data could contribute to both their personal development and the development of their communities. The study of open civic data, as real, situated and complex data, could help young children to (1) understand how their governments work and how decisions that affect their lives are made (Gal, 2022); (2) acquire skills to make sense of data (Gal et al., 2022); (3) develop critical and analytical thinking skills, which are essential for solving complex problems; (4) understand socio-scientific phenomena through objective information that empowers civic participation; (5) identify social, economic, and political inequalities to develop a sense of social justice to propose equitable changes (Gutstein, 2006);

and (6) gain an interdisciplinary and situated view of socio-scientific phenomena.

This study illustrates experiences with data and data-ing based on open civic data with elementary school children from underserved populations while establishing authentic connections to global citizenship. This is an opportunity to promote data literacy and engagement with complex data for global citizenship among underserved populations to highlight and advocate against social asymmetries. The goal is to investigate how, through the study of open civic data (data and data-ing), students from underserved populations exhibit traits of global citizenship (understanding, belonging, and action). The research question being explored is: what traits of global citizenship do underserved students exhibit in studying global issues using open civic data?

## 2 Literature review

A data-driven society requires its citizens to be data literate for decision-making and informed participation in society (Ridgway, 2022). Data handling skills (data-ing) are essential for educating global citizens who understand and act with empathy and responsibility on global issues. These skills should be cultivated from an early age, in contexts of global interest and using open civic data. However, the literature has shown that these attributes are rarely integrated into data literacy preparation, and that many of the skills that students need for civic participation are neglected in school curricula (Makar et al., 2023).

Open civic data is digital data of societal interest that is made available to citizens, without restrictions, that anyone can freely use, reuse, and redistribute for the purpose of understanding the world. Open civic data is observational in nature where it may be difficult to establish causal relationships (Engel, 2019), but it can suggest interesting associations, provide insights into the world, help develop dispositions, stimulate social awareness, and shape critical citizens (Engel, 2019; Wilkerson & Laina, 2017; Zapata-Cardona et al., 2022). In the field of statistics education, the use of civic open data has been defended as a call to action (Pro-CivicStat, 2018) to promote civic participation in society (Engel, 2019; Wilkerson & Laina, 2017), develop social justice (Lesser, 2007), encourage the development of awareness and connect to practices for life outside of school and beyond formal education (Rubin, 2022; Weiland, 2017). The literature in this area suggests that incorporating the study of open civic data for educational purposes is a beneficial factor in addressing inequities (Rowan, 2018), strengthening data literacy, promoting participation in society (Engel, 2019), and providing learning opportunities to underserved populations (Heinzman, 2022). In this study, open civic data

is used to design learning experiences that support global citizenship education.

Studies such as Hancock et al. (1992) have focused on data modeling and have suggested interesting tasks for students in grades five through eight. The tasks designed from an inquiry-based approach succeed in connecting students to key concepts of reasoning and data analysis, but the contexts behind those tasks fail to achieve an authentic connection to the global tensions facing today's citizens. In contrast to this approach to data modeling, but giving a leading role to social scenarios, Makar et al. (2023) worked with fourth grade students in a multidisciplinary data science project related to activities in cyberspace. The authors found that the participants developed citizenship skills and made sense of the data because of the social meaning that their data represented. In the same vein, Erwin (2015) showed that working with real-world data helped students to integrate data literacy into their disciplinary learning.

To overcome the apparent restriction imposed by context, Wilkerson and Laina (2017) relied on socio-scientific scenarios to investigate how middle school students made sense of visual representations of data. They found that students were not always able to coordinate mathematical, representational, and domain-specific resources to make sense of the visualizations. Although the study of visual representations in socio-scientific settings proved to be favorable in sense making, it is also necessary to provide participants, especially those at an early age, with tools such as dynamic software to explore and propose their own visual representations (Araujo Inastrilla, 2023).

Various efforts have been made to educate communities for public engagement as global citizens. Specifically related to climate change, the study by Boaventura et al. (2021) engaged elementary school students in citizen science projects to monitor climate change. Participants identified species (data) and stored the information on a platform for collective analysis. Although the students were authentically engaged in the tasks, data handling (data-ing) was not the protagonist of the experience. The study by Boaventura et al. (2021) is a clear example that meaningful data does not always imply successful data-ing processes.

Exploring and visualizing datasets are essential tasks for extracting insights from data. Guimarães et al. (2022) proposed classroom activities for exploring and visualizing climate change data using *R* and *RStudio*, taking advantage of large datasets produced by several scientific groups. Although each of the proposals is innovative and interesting, coding is a limitation when working with elementary school students. Data exploration and visualization with elementary school students should be supported by user-friendly software that underpins future work with software that requires much more sophisticated digital literacy skills. In this sense, integrating the study of open civic data describing global

issues with user-friendly software could support data literacy at an early age to educate global citizens.

This exploratory study attempts to overcome the limitations identified in the literature regarding the connection to context, visual representations made by others, the separation of data and data-ing, and the use of sophisticated software that requires coding. This study integrates relevant contexts, user-friendly software, and connects data with data-ing to trace the traits of global citizenship that emerge from participants' study of open civic data on global issues.

### 3 Theoretical framework

The concepts of *data* and *data-ing* are epistemologically related. While *data* refers to an object (a noun), *data-ing* refers to the processes performed on that object (a verb) to develop the object or to use the object. A similar epistemological relationship is found between the concepts of sample-sampling and model-modeling. Thus, *data*, as a model of a real-world situation (Hancock et al., 1992), refers to an object composed of information. However, for the purposes of this study, the characteristic of this information is that it is open data about phenomena of a social nature that require some processing to produce useful knowledge for decision making. In a complementary way, *data-ing* refers to the process of preparing the *data* (cleaning and formatting) for subsequent data handling, as well as the data handling process itself (exploration, visualization, searching for patterns, identifying trends, transformation, modeling, summarization, and interpretation [Hancock et al., 1992]) to produce knowledge. The production of knowledge is an essential resource in a globalized world because it overcomes the instrumental view of *data* and *data-ing* (focusing attention exclusively on the technical aspects of *data* and *data-ing*), while seeking to transform the current conditions of global issues. To illustrate this point, it is not enough to run a *data-ing* process on a dataset that reveals an increase in the amount of plastic levels in the oceans (see for example the study by Lebreton et al., 2018), but it is necessary to put this knowledge at the service of society in order to take action to solve such issues.

Global citizenship education is an initiative of UNESCO (2015, 2018) that calls for efforts to build a more inclusive, just, and peaceful world, as well as an invitation to prepare students for the challenges of the twenty-first century such as informed, critically literate, socially connected, ethical, and engaged individuals. Global citizenship “refers to a sense of belonging to a broader community and common humanity. It emphasizes political, economic, social, and cultural interdependency and interconnectedness between the local, the national and the global” (UNESCO, 2015, p. 14).

The concept of global citizenship encompasses three interrelated domains of learning: cognitive, socio-emotional, and behavioral. The cognitive domain has to do with the acquisition of *understanding*, knowledge and critical thinking. *Understanding* involves both developing data literacy skills (gathering, organizing, processing and making sense of data) and using those skills to interpret global issues (generate useful knowledge and communicate essential knowledge). *Data* and *data-ing* are critical to developing such an *understanding* and to framing how students engage with complex data in active ways, such as the movements and explorations facilitated by software that enrich accessible interpretations. The socio-emotional domain is related to a sense of *belonging* to a common humanity, shared values and responsibilities, empathy and solidarity. The idea of belonging has to do with being “part of” rather than “being inside” a community (Carrijo, 2023).

The behavioral domain refers to the actions that individuals exhibit to demonstrate their commitment to their local and global communities. This domain emphasizes the practical application of values and knowledge in everyday life, and it takes the form of activism, participation and empowerment. The behavioral domain is related to the ability to *act* effectively and responsibly at local, national and global levels. In Gutstein’s (2006) terms, it would be the ability to act with tools to “read the world” (understanding), to contribute to its transformation, to “write the world”, to create a more peaceful and sustainable world. It is essentially a form of agency. In this sense, studying the exhibition of traits of understanding, belonging, and action using open civic data examines the cognitive, socio-emotional, and behavioral domains of global citizenship, an initiative consistent with the Organisation for Economic Co-operation and Development (OECD) (2018) recommendations on the future of citizenship education. Global citizenship is a form of public engagement, a transformative force that seeks to build the knowledge, skills, values and attitudes that learners need to contribute to a more inclusive, just and peaceful world.

Elementary school students do not necessarily need to master sophisticated statistical procedures for data handling, but, like any citizen of a global village, they do need certain skills to make sense of data for public engagement. The concept of public engagement is associated with the ability to develop awareness and participate in society, and it is a concept closely related to the notion of citizenship, an emancipatory and transformative force that seeks justice and democracy (Geiger et al., 2023; Martínez-Castro & Zapata-Cardona, 2021). Unfortunately, there is no empirical evidence that shows a causal relationship between raising awareness about global issues and undertaking transformative actions (Chesky & Wolfmeyer, 2015), but public engagement is not conceivable without raising awareness about global issues.

Global citizenship requires educated citizens to contribute to justice, peace, well-being, sustainability, and inclusive practices (González-Valencia et al., 2022). However, citizens of different origins have unequal access to natural resources, health care, education, adequate nutrition, living conditions and employment opportunities. Some scholars have shown that among these disparities, those belonging to underserved communities suffer the most from social inequalities (O’Neil, 2016). Focusing this study on underserved groups is an opportunity for them to recognize their living conditions, to question global issues that affect communities differently, and to develop a global awareness that could potentially translate into a public commitment to building a new reality. It is an attempt to educate informed (literate), responsible (involved) and socially connected citizens.

## 4 Methods

This exploratory study was designed to use open civic data (data) related to a global problem to investigate the data handling (data-ing) skills that young children from underserved populations bring to bear when exploring datasets. Evidence of understanding, belonging and action was tracked in participants’ oral statements. The interest was in the idea of citizenship that emerges from children’s interaction with data, and in particular, the public engagement that this process of data handling activates. The traits of global citizenship exhibited by underserved students in the study of global issues were traced as participants handled open civic data. To do this, a learning experience, based on a global issue, was designed and the participants carried it out while they were interviewed individually.

### 4.1 Context and participants

The study was conducted with five fourth grade students (9–10 years old, three males and two females) from public schools in low-income neighborhoods of a metropolitan area in northeastern Colombia. Participants were invited through colleagues with academic ties to the schools. Inclusion criteria were age, willingness to participate and written permission from parents or guardians. The participants had received statistical training in school since the first grade and the curriculum structure followed the ministerial guidelines (Ministry of Education, 2006), which for the fourth grade focused on tabular and graphical presentation of data, possibilities of occurrence of events, collection, and analysis of data in conventional formats, and calculation and interpretation of measures of central tendency. The participants had not used software to explore datasets, and the use of computers was not a common theme in their daily school activities. School computers were used exclusively for

computer science instruction, not to support instruction in other subjects. This context provided an opportunity to study traces of global citizenship in the participants' discourse as they engaged in tasks that used data to perform data-ing processes.

## 4.2 Learning experience

A learning experience was designed to explore open civic data on a global issue (climate change). Gal (2022) calls them “burning issues” because they influence social and economic progress and affect the well-being of citizens and communities. A multivariate dataset was prepared, organized, and stored in CODAP (Common Online Data Analysis Platform) format, an interactive online visualization software, following a canonical structure with information about the Tenana River ice breakup from the Nenana Ice Classic contest. In the Nenana Ice Classic contest, competitors are challenged to predict the exact date and time that a tripod placed over the frozen Tenana River will fall when the ice breaks up in spring. This time series has been used in other studies to explore patterns of climate change (Sagarin & Micheli, 2001). The data was taken from the NASA (National Aeronautics and Space Administration) website and the dataset contains information from 1917 to 2021 regarding variables such as *year*, *month*, *day of the month*, *decimal day of the year*, and *breakup time*. The dataset (data) was presented to the participants ready to use. Using a pre-prepared dataset simplifies many of the data cleaning and data preparation tasks required in the data-ing process. This methodological choice oversimplifies the laborious task of dealing with open civic data, while reducing the tensions related to the skills required for the organization and the anxiety associated with overwhelming large datasets (Dove et al., 2023). A pre-prepared dataset allows participants to focus on the data handling tasks involved in data-ing, such as exploration, visualization, transformation, modeling, summarization, and interpretation.

In a 30-min training segment, participants were given a link to a ready-to-use 17-variable dataset of nutritional information for granola bars. The dataset was organized according to a canonical form with variables in columns (9 categorical: *brand*, *type*, *nuts*, *chocolate*, *organic* ...; and 8 numerical: *calories*, *sodium*, *fat*, *protein*, *sugar*, ...) and observations in rows (33 brands). The interviewer walked the participants through each element of the dataset to ensure that they understood the information represented by the columns and rows. To check their understanding, the interviewer asked “Can you tell me what row 3 of this dataset means? Only when comprehension was verified did the training continue. Participants were instructed to use the graphical and numerical methods provided by the software to explore the dataset in univariable, bivariate, and

multivariate analyses and to communicate their interpretations. To aid in the interpretation process, the interviewer asked questions such as “What does this graph tell you?” “What does this numerical information tell you?” For each answer given, the interviewer provided positive feedback; and when participants gave incomplete interpretations, the interviewer offered additional information to supplement and enrich the analysis. Participants were allowed to ask questions, and the interviewer offered detailed explanations about the software or about the data handling process. The training ended when the participants showed skills in using the software and showed flexibility in producing different forms of representation. The training took only 30 min because CODAP is an interactive, accessible, visual, and easy-to-use software. CODAP was chosen for this study because of its user-friendly and intuitive interface, which allows participants to engage in exploratory data analysis even if they do not have extensive experience with data analysis tools. In addition, the multiple visualization options help participants effectively understand and interpret the phenomena described by the data.

## 4.3 Data collection and analysis

Data collection consisted of individual interviews based on the participants' interaction with the learning experience. Participants were interviewed in virtual settings through the Google Meet platform while performing data handling processes (data-ing) on a dataset (data) related to the learning experience previously designed by a research team. The meetings were audio and video recorded to provide a detailed record of the participants' actions and statements during the learning experience. The interviews were conducted individually and outside the classroom, sometimes inside the school and sometimes outside. In both cases, participants were provided with a computer with Internet access for data handling (data-ing). Each interview lasted approximately 60 min; the first 30 min were spent training in the use of the software by studying a dataset in a different context.

For the learning experience segment, participants were provided with a web link with the location of the dataset containing the results of the Nenana Ice Classic contest (the database is available in the [link](#)). The Nenana Ice Classic was not a familiar scenario to the participants, but the global issue underlying the historical data of the contest was of general interest, and the participants confirmed this by being fully engaged during the interviews. Participants accessed the dataset and had full control over the device (personal computer or desktop computer) they use for data handling. Because the interview was virtual, the interviewer had no way of manipulating the participants' devices. Every decision to undertake the data handling was made exclusively by each participant. The learning experience was initiated by

asking participants “What does the dataset say?” Only when they performed specific actions (exploration, visualization, modeling ...) on the dataset that led to different forms of representation, other questions were asked: What do you want to discover with that action? What does this graph tell you? Could the dataset say something else? Is there anything else you want to add? The most common data handling tool used by participants was the software’s “graph” function. This function allowed them to explore, create visual representations, classify based on certain attributes, model data, and measure. None of the participants performed any data transformations as part of the data handling process.

The interviews were conducted in Spanish, videotaped, and later transcribed word for word to facilitate analysis. With the help of the transcripts and videos, the participants’ statements were tracked to search for the domains of global citizenship (cognitive, socio-emotional, behavioral) represented in the understanding, belonging and action, that are exhibited in the data handling of open civic data (data-ing). An analysis process adapted from Powell et al. (2003) was followed. The purpose of the analysis was to find examples within the interviews that illustrated the traits of global citizenship. First, the videos were viewed to get a general idea of the flow, and a rough video log was created to capture the basic content and sequence of each interview. A second round of visualization was then conducted, but this time with annotations to identify what Powell et al. (2003) call critical incidents: rich segments of video that can illustrate the key features of global citizenship under study. The segments were viewed again to select examples that best illustrated the domains of global citizenship and to reduce repetition. The transcription of the selected segments was reviewed and translated into English for presentation in the analysis.

## 5 Analysis and discussion

In this section, some segments of the interviews are presented in the voices of Federico and Ángel (pseudonyms). This does not mean that the voices of the other participants are less relevant, only that the selected segments are the ones that best illustrate the traits of global citizenship that suggest public engagement.

### 5.1 Cognitive domain—Understanding

Exploring and creating visual representations are some of the skills for data-ing. In the process of data-ing on the given dataset, participant Federico used the software to generate six different visual representations, of which three were univariate and three bivariate. One of the bivariate visual representations created was *month vs. day* (Fig. 1) and the following is the exchange that took

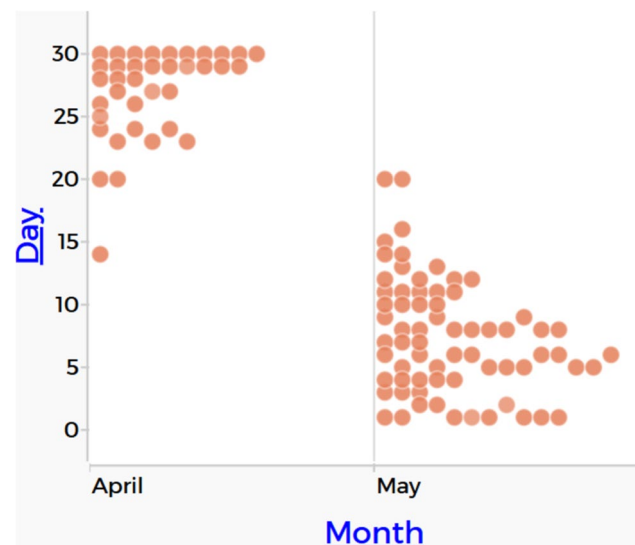


Fig. 1 Bivariate visual representation of *month vs. day*

place between the interviewer and the participant that led Federico to construct such a representation in the data-ing process.

Interviewer (16:47): What is this data telling you? What explorations do you think you could do with this data?

Federico (17:10): A graph to look at like the days. I don't know (*he dragged the “month” variable into a graph*).

Interviewer (17:20): Well, that is fine. Which would you like to explore? So, look, there it is telling you that many happened in April, and many happened in May. Why in April and May? How can you explain that?

Federico (17:45): Because they are like the months when everything is thawing.

Interviewer (18:00): Because the seasons change, right. And the weather is a little bit warmer. What you just did there, can you explain to me what you are trying to find out? (*he dragged the variable “day” and placed it on the y-axis of the graph, creating the “month” vs. “day” graph*. Figure 1). What can you see from what you just did there?

Federico (18:15): A graph to see which month has the most. What is more like, what is the one in which it defrosts easier.

By choosing this bivariate visual representation, *day* split by *month*, Federico was trying to make sense of the data in the context of the river’s thawing. He looked at the *month* with the most thaw events, while coordinating with the *day* of the month. With this move, it seems that the participant was looking for a pattern or a tendency that would allow him to better understand the phenomenon under study. According

to UNESCO (2015), people need knowledge and thinking skills to better understand the world and its complexities.

Another visual representation that Federico created in the data-ing process was a univariate graph with the distribution of the *breakup time* variable (Figs. 2 and 3) when the tripod fell due to the river thawing.

Based on the visual representations displayed in Figs. 2 and 3, the participant was encouraged to talk about patterns, trends, summaries, typical and atypical data points and interpretations, actions that are related to the data-ing process. The participants verbal explanations were stimulated by questions such as “What does this graph tell us?”, “What is the time of day when that displacement of the river or the fall of the tripod is most likely to occur?” Moving the cursor over the graph, the participant noticed that the data points concentrated between noon and three, but also noticed some outliers. A related comment was made to stimulate his talk “There is some melting in 2019, it was not long ago, what does that tell us?” The participant stated:

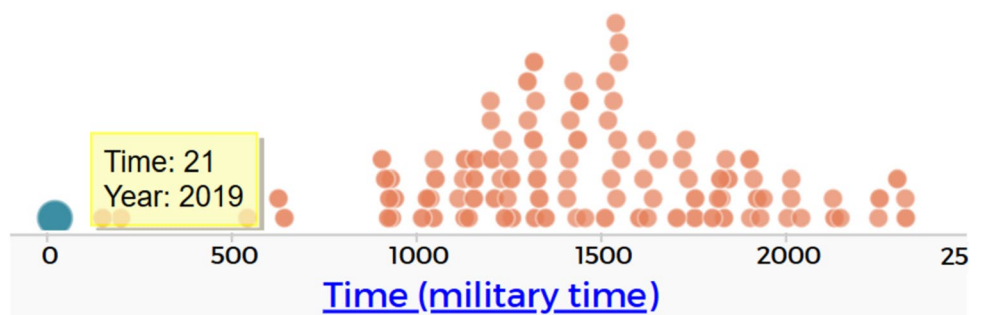
Midday is usually when there is the most sun. So, when it is sunnier there is a greater chance that the river will melt [...]. There were also some [hour] around three in the afternoon [...], and there was one that melted almost at midnight. [...] It seems that the Earth is warming (Federico, April 1, 2023).

The participants’ statement is striking for several reasons: it describes the behavior of the variable, proposes an informal statistical inference, and seems to suggest an understanding of the global issue, indicators that would

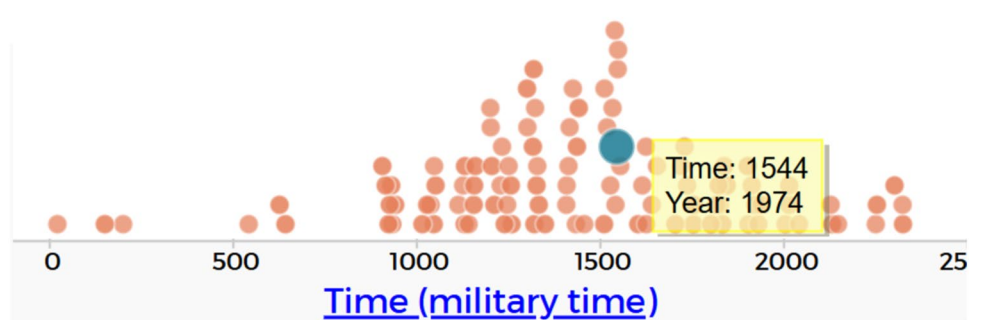
make evident the cognitive domain of global citizenship. In terms of describing the behavior of the variable, the statement suggests a coherent description of the distribution of the *breakup time* variable when the tripod falls. In the expression “when it is sunnier there is a greater chance that the river will melt” the participant seems to recognize the congregation of data points in a central tendency while using probabilistic language. Although the variable under study was the *breakup time*, the participant connected this variable with temperature. Studies using nontraditional datasets to investigate climate change patterns have also suggested an association between the *breakup time* with thermal effects “warmer climate would be expected to advance the time of breakup” (Sagarin & Micheli, 2001, p. 811). In addition, the expression “there was one that melted almost at midnight” appears to refer to an outlier that the participant identified by moving the cursor over the data points; an observation from 2019 where the breakup time occurred at 21 min past zero hours on April 14 (Fig. 2).

With respect to informal statistical inference, the expression “it seems that the Earth is warming” can be considered an informal statistical inference (in the sense understood by Makar & Rubin, 2009) since it (1) articulates a language of uncertainty, (2) considers the aggregate to propose a generalization beyond the data, and (3) uses the data as evidence. Regarding the global issue under study, the expression “it seems that the Earth is warming” is a conclusion derived from the study of the visual representation that appears to suggest that there is an understanding of a global issue, in the sense understood by Mesa (2019), who recognizes

**Fig. 2** Distribution of the *breakup time* variable (military time format) with a highlighted outlier



**Fig. 3** Distribution of the *breakup time* variable (military time format) with one of the typical times highlighted



climate change as one of the major global challenges facing humanity.

The participant Angel undertook several tasks related to the data-ing process. The participant proposed four visual representations, two of them were univariate and two were bivariate. Using a similar strategy to that used by Federico, Angel also proposed a visual representation of the *breakup* time (as in Figs. 2 and 3), and the discussion focused on the extreme observation where the tripod fall was at 21 min of the day in 2019, an observation corresponding to the 104th decimal day of the year. The participant was encouraged to talk about this representation with some input from the interviewer. The exchange followed:

Interviewer (8:33): What information does this graph, you just made, give you?

Angel (8:47): The time it [the tripod] fell? (*for a little while, the participant interacts with the software without speaking*)

Interviewer (10:34): Go to the first dot on the left of that graph [year 2019 with time 21 minutes]. Do you want to tell me what that first dot means?

Angel (10:57): That little dot, the military time and the year

Interviewer (11:00): And what year was that?

Angel (11:03): In 2019

Interviewer (11:06) And what time did the tripod fall in 2019?

Angel (11:35): Early in the morning at the 21 minutes

Interviewer (11:40) Exactly. Almost at 1:00 in the morning. Good. So, what I want us to analyze here is: if that tripod falls almost at 1:00 in the morning in 2019, what is going on there?

Angel (12:09): The ice is melting faster.

The exchange shows some evidence that the participant was making sense of the data by recognizing the changes in the times at which the tripod falls. In one of the bivariate visual representations, participant Angel coordinated the variables *year* and *decimal day of the year* when the tripod fell on the same graph (Fig. 4). The exchange with the interviewer suggests that the participant was trying to establish an association:

Interviewer (17:10) Now we are going to try to make sense of this graph. Can you point out any dot on the graph that catches your attention? whatever you want to pick out.

Angel: This one (*the participant points to the year 2019 with the 104<sup>th</sup> decimal day of the year as it is shown in Figure 4*).

Interviewer (17:27): Very good. You have seen that dot many times in the other graphs.

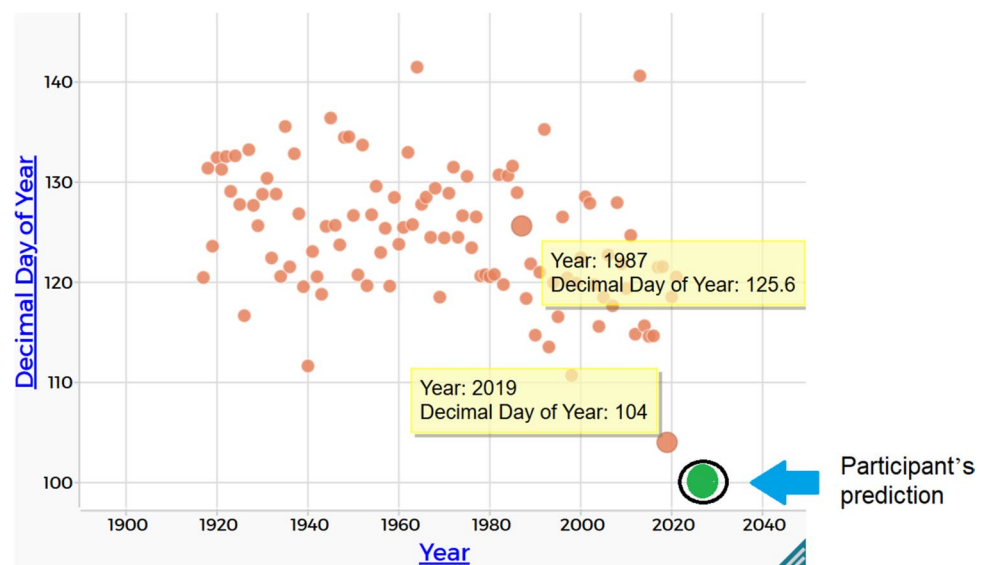
Angel (17:40): This one? (*participant again points to 2019 with 104*)

Interviewer (17:42): Yes, that dot has appeared to you several times in the other graphs. It is from the year 2019 and it is the decimal day of the year 104. See, it is the earliest day in the year that the tripod fell. Now chose another dot. Point out another dot on the graph that catches your attention (*participant points to 1987 with the 125<sup>th</sup> decimal day of the year*). That is 1987 and there the tripod fell on the 125<sup>th</sup> decimal day of the year. Do you notice any trends? Is there any particular behavior of these dots in this graph?

Angel (18:54): Yes.

Interviewer (18:57): What do you see happening there? What does the graph tell you? What do you notice?

**Fig. 4** Two-dimensional representation year versus decimal day of the year





Angel (20:05): Now the sun is getting rougher than before. The ice used to last longer, but not now because it melts faster [in the year].

Although the two variables depicted in Fig. 4 have a weak linear relationship with a fragile negative correlation coefficient (-0.4) and are derived from observational processes that could not by themselves represent causal relationships (as Engel, 2019, warns), the two-dimensional graph suggests a trend for thaw to occur earlier in the year as the years go by. Figure 4 is not an easy representation for an elementary school student to interpret because it integrates the simultaneous behavior of two variables. Hancock et al. (1992) pointed out that it is difficult for students to make sense of the associations between variables in graphical representations because they must consider the aggregate entity (rather than individual cases) that describe causal or correlational relationships (aggregate reasoning). However, in the participant's verbal statements, at the end of the last exchange, aggregate reasoning was evident when tried to describe the trend.

The participant's statement suggests establishing a relationship between two variables. In this regard, Makar et al. (2023) pointed out that literature illustrates children's ability to interpret complex relationships as associations even without having formalized the concepts. To corroborate the consistency of the interpretation given to the visual representation, the participant was asked to place a data point on the graph that would represent the prediction for 2024. The participant pointed with the cursor to the area where the point would be located (the one with the larger size and surrounded by a line in Fig. 4) and expressed: "each time the points will be located lower because the ice will melt faster" (Angel, August 28, 2023). Angel's statement seems to describe a trend over time and confirms the initial interpretation. Some researchers have reported the success of elementary students with two-dimensional graphs (Hancock et al., 1992) when there is occasional external guidance from questions. It is possible that this was the case and that the researcher's questions helped to solidify the participant's interpretation.

Both Federico, when describing the variable, proposing an informal inference, and showing signs of understanding the issue, and Angel, when making an association between variables, seem to account for the cognitive domain of global citizenship (UNESCO, 2015). This domain is related to the acquisition of knowledge, understanding and critical thinking about global issues, which are essential for the empowerment of individuals (González-Valencia et al., 2022). This generation of knowledge becomes fundamental for the development of dispositions and the establishment of actions as described in the next section.

## 5.2 Socio-emotional and Behavioral domains – Belonging and Action

The participant Federico, after recognizing the patterns of climate change observed in the visual representations produced in the data handling process, expressed a statement that seemed to reveal a sense of belonging, empathy, solidarity, and responsibility with the environmental problem under study, and at the same time he implied a transformative action that reflects a concern for a sustainable world. The participant said: "maybe we need to produce less stuff" (Federico, April 1, 2023). This statement seems to show traces of the socio-emotional and behavioral domains of global citizenship (as proposed in UNESCO, 2015).

With this statement, the participant went beyond data and data-ing to produce useful knowledge that materialized a sense of belonging even without initial familiarity with the Nenana Ice Classic contest. By saying "we need to produce less stuff", the participant includes himself as part of the global community, does not communicate a binary narrative between "us and them" and "here and there" and emphasizes the sense of belonging of "being part of" (Carrizo, 2023; González-Valencia et al., 2022). The participant no longer recognizes only that the Earth, as an abstract-isolated-remote entity, "is warming," but also recognizes the concreteness-connectedness-closeness of the global issue and the interdependence between the local and the global. An everyday action at the local level can have effects at the global level. Furthermore, the use of the pronoun "we" in the statement breaks with the sense of marginality, of not feeling "part of," that underserved communities systematically experience (Carrizo, 2023).

The participant's verbal statement seems to suggest a sense of empathy and solidarity that recognizes a finite and limited world, both in terms of the resources and regenerative capacity. In this regard, some scholars have suggested that humanity must redefine its relationship with the environment in which it lives and on which it depends (Mesa, 2019). The statement also seems to materialize an ecological awareness that recognizes the consequences of unfriendly habits with the planet and questions the ideas of unlimited growth leading to limitless consumption. Authors such as Mesa (2019) have expressed that consumption is no longer an option for a world that has exceeded its limits. Humanity has already surpassed four out of the nine planetary boundaries identified by Rockström et al. (2009). "Our consumption and production patterns have severely affected the Earth's carrying capacity" (United Nations, 2020), but it is worrying that only 50% of the population believes that climate change is anthropogenic in nature (Bikovska & Liew, 2023).

The verbal statement also seems to suggest that the participant takes an active role within the global community and assumes responsibility for a global problem

(González-Valencia et al., 2022). It proposes a form of self-regulation of the predatory actions of a consumer society in favor of environmental protection (Mesa, 2019). The statement seems to propose the search for solutions to a global problem and the generation of constructive responses and initiatives that transcend destructive practices. In a few words, the participant used the knowledge produced in data and data-ing for decision-making to propose a grounded, ethical, civic, creative, just, environmental, and sustainable gesture to “produce less stuff.” Although it is only a verbal statement and not an action as such, it suggests a willingness to take transformative action. In the statement, the participant seems to perceive himself as a change agent (Gutstein, 2006) proposing sustainable actions. The statement embodies an emancipatory meaning that integrates awareness of global issues, overcoming passive roles and acting as autonomous and transformative agents. In this sense, the statement suggests that the participant acts as a global citizen who assumes public engagement.

## 6 Conclusions

This exploratory study was based on the need for data literacy for global citizenship to advocate against social asymmetries, and the need to create opportunities for young people to engage with data and data-ing with open civic data to promote public engagement. The research question explored was: What traits of global citizenship do underserved students exhibit in studying global issues using open civic data? The findings suggest that proposing learning experiences combining data-ing with open civic data allows students from underserved populations to engage in discussions and reflections about global issues required for public engagement. Public engagement is essential for strengthening democracy and social development, but for communities that systematically suffer some form of exclusion or segregation it is an opportunity to advocate for environmental justice.

One of the purposes of data-ing is to generate useful information for decision making for public participation. Public engagement on global issues can be stimulated by studying open civic data related to relevant issues of society. As the results of this study show, data and data-ing provide opportunities for understanding, belonging and action. By engaging with data and data-ing, participants exhibited understanding by describing variable behavior, suggesting informal inferences and establishing relationships between variables; in addition, they exhibited belonging and action by proposing transformative conducts. In this sense, data and data-ing could be used as knowledge generators to educate citizens who are critical, aware, and capable of making decisions for the benefit of the global community.

Participants’ discourse, as they worked through the learning experience, appear to show traces of the cognitive, socioemotional, and behavioral domains of global citizenship. However, future research needs to focus on designing learning experiences that can reveal the behavioral domain of global citizenship in more detail and with greater reliability. This study has the limitation of not being able to directly evaluate transformative actions, because it is based exclusively on oral statements expressed by the participants and not through concrete actions.

The results of this exploratory study could guide policymakers, school leaders and teachers who want to prepare students for a global world. Education in a globalized world must address problems in an interdisciplinary way (Gimeno, 2005). It is not enough to expose students to data-ing processes. These processes must be stimulated with a purpose, to be anchored in information that has a social meaning and in the search of solutions to global issues. For students to function in an interconnected world, knowledge cannot be isolated. The study of global issues from an early age could be a tool to counteract the asymmetrical consequences of globalization.

Schooling must go beyond substantive, objective, and rational knowledge to reflect on global issues that need to be addressed. Experiences such as those presented in this study seek to provide citizens and young people from underserved communities with analytical tools for data and data-ing that enable them to participate in active citizenship. The Organisation for Economic Co-operation and Development (OECD) (2018) recommendations on the future of citizenship education point in the same direction, with the need to connect data and data-ing to explore global issues.

Exposing young children from underserved communities to data and data-ing about global issues at an early age could help them understand, reflect on, develop empathy for, build skills to talk about, and develop dispositions to address these issues. In addition, studying global issues that have a collective impact could help students from underserved communities not only to “being inside” the world, but also to feel “part of” the world as agents of change.

There are two additional limitations of this study that need to be acknowledged. First, the study involved a limited number of participants and focused on public schools. This may have significant implications for the generalizability of the study. Any generalization should be made with caution, as additional research is needed to strengthen the findings and confirm whether they are consistent in other settings. Another limitation relates to do the data collection process. This study was based on a ready-to-use dataset, as the participants were primary school students with very little experience in the data preparation process (cleaning and formatting) of the data-ing. However, it is worth asking how different the results might be if the participants were also

encouraged to carry out the data preparation process as well. Is data preparation of open civic data for public engagement a facilitator or a barrier?

**Acknowledgements** This research was carried out thanks to the support from MinCiencias and Icetex, Grant number 2023-0631.

**Funding** Open Access funding provided by Colombia Consortium.

## Declarations

**Conflict of interest** The author has no relevant financial or non-financial interests to disclose.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Araujo Inastrilla, C. R. (2023). Data visualization in the information society. *Seminars in Medical Writing and Education*, 2. <https://doi.org/10.56294/mw202325>.
- Bikovska, D., & Liew, C. L. (2023). Museums and communicating climate change-related issues on Facebook platforms. *Online Information Review*, 47(5), 974–988. <https://doi.org/10.1108/OIR-05-2022-0255>
- Boaventura, D., Neves, A. T., Santos, J., Pereira, P. C., Luís, C., Monteiro, A., Cartaxana, A., Hawkins, S. J., Caldeira, M. F., & Ponces de Carvalho, A. (2021). Promoting ocean literacy in elementary school students through investigation activities and citizen science. *Frontiers in Marine Science*. <https://doi.org/10.3389/fmars.2021.675278>
- Carrizo, M. (2023). Microexclusion and immigrant students. *Educación Matemática e Pesquisa*, 25(4), 261–283. <https://doi.org/10.23925/1983-3156.2023v25i4p261-283>
- Chen, B., Bartucz, J., Scharber, C., Rao, V., & DeLiema, D. (2023). *Towards a Framework for Justice-Oriented Data Science Education in K-12 Schools*. AERA 2023 Annual Conference.
- Chesky, N., & Wolfmeyer, M. (2015). *Philosophy of STEM Education*. Palgrave Macmillan. <https://doi.org/10.1057/9781137535467>
- Dove, G., Shanley, J., Matuk, C., & Nov, O. (2023). Open data intermediaries: Motivations, barriers and facilitators to engagement. *Proceedings of the ACM on Human-Computer Interaction*, 7, 1–22. <https://doi.org/10.1145/3579511>
- Engel, J. (2019). Cultura estadística y sociedad: ¿Qué es la Estadística Cívica? [Statistical literacy and society: What is Civic Statistics?]. En J. M. Contreras, M. M. Gea, M. M. López-Martín y E. Molina-Portillo (Eds.), *Actas del Tercer Congreso Internacional Virtual de Educación Estadística*. Universidad de Granada.
- Erwin, R. W. (2015). Data literacy: Real-world learning through problem-solving with data sets. *American Secondary Education*, 43(2) 18–26. <http://www.jstor.org/stable/43694208>
- Gal, I., Nicholson, J. & Ridgway, J. (2022). A conceptual framework for Civic Statistics and its educational applications. In J. Ridgway (Ed) *Statistics for empowerment and social engagement: Teaching Civic Statistics to develop informed citizens* (pp. 37–66). Springer. [https://doi.org/10.1007/978-3-031-20748-8\\_3](https://doi.org/10.1007/978-3-031-20748-8_3)
- Gal, I. (2022). Critical understanding of civic statistics. Engaging with important contexts, text, and opinion questions. In J. Ridgway (Ed) *Statistics for empowerment and social engagement: Teaching Civic Statistics to develop informed citizens* (pp. 323–344). Springer. [https://doi.org/10.1007/978-3-031-20748-8\\_13](https://doi.org/10.1007/978-3-031-20748-8_13)
- Geiger, V., Gal, I., & Graven, M. (2023). The connections between citizenship education and mathematics education. *ZDM Mathematics Education*, 55, 923–940. <https://doi.org/10.1007/s11858-023-01521-3>
- Gimeno, J. (2005). *La educación que aún es posible [The education that is still possible]*. Paidós.
- González-Valencia, G., Massib Sabater, M., & Santisteban Fernandez, A. (2022). Critical Global citizenship education: A study on secondary school students. *Teacher Education*. <https://doi.org/10.3389/educ.2022.867113>
- Gould, R. (2021). Toward data-scientific thinking. *Teaching Statistics*, 43, S11–S22.
- Guimarães, N., Vehkalahti, K. Campos, P., & Engel, J. (2022). Exploring Climate Change Data with R. In J. Ridgway (Ed) *Statistics for empowerment and social engagement: Teaching Civic Statistics to develop informed citizens* (pp. 267–296). Springer. [https://doi.org/10.1007/978-3-031-20748-8\\_11](https://doi.org/10.1007/978-3-031-20748-8_11)
- Gutstein, E. (2006). *Reading and writing the world with mathematics: Toward a pedagogy for social justice*. Routledge.
- Hancock, C., Kaput, J., & Goldsmith, L. (1992). Authentic inquiry with data: critical barriers to classroom implementation education development center. *Educational Psychologist*, 27(3), 337–364. [https://doi.org/10.1207/s15326985Sep2703\\_5](https://doi.org/10.1207/s15326985Sep2703_5)
- Heinzman, E. (2022). I love math only if it's coding": a case study of student experiences in the introduction to data science course. *Statistics Education Research Journal*, 21(2), 5. <https://doi.org/10.52041/serj.v21i2.43>
- Johnson, L., & Morris, P. (2010). Towards a framework for critical citizenship education. *The Curriculum Journal*, 21(1), 77–96. <https://doi.org/10.1080/09585170903560444>
- Khalil, A., Ahmed, S. G., Khattak, A. M., & Al-Qirim, N. (2020). Investigating bias in facial analysis systems: A systematic review. *IEEE Access*, 8, pp. 130751–130761. <https://doi.org/10.1109/ACCESS.2020.3006051>
- Lanzetta, M. (2021). Los desafíos del cambio climático en grandes metrópolis latinoamericanas. Apuntes y reflexiones para la adaptación [The challenges of climate change in great Latin American metropolises. Notes and reflections for its adaptation] *Revista Estado y Políticas Públicas*, 17, 19–38. [https://revistaeypp.flacso.org.ar/revista/numero-17\\_200](https://revistaeypp.flacso.org.ar/revista/numero-17_200)
- Lebreton, L., Slat, B., Ferrari, F., Sainte-Rose, B., Aitken, J., Marthouse, R., Hajbane, S., Cunsolo, S., Schwarz, A., Levivier, A., Noble, K., Debeljak, P., Schoenich-Argent, R., Brambini, R., & Reisser, J. (2018). Evidence that the Great Pacific Garbage Patch is rapidly accumulating plastic. *Scientific Reports*, 8, 4666. <https://doi.org/10.1038/s41598-018-22939-w>
- Lesser, L. M. (2007). Critical values and transforming data: teaching statistics with social justice. *Journal of Statistics Education*, 15(1), 1–21. <https://doi.org/10.1080/10691898.2007.11889454>
- Makar, K., Fry, K., & English, L. (2023). Primary students' learning about citizenship through data science. *ZDM Mathematics Education*. <https://doi.org/10.1007/s11858-022-01450-7>
- Makar, K., & Rubin, A. (2009). A framework for thinking about informal statistical inference. *Statistics Education Research Journal*, 8(1), 82–105. <https://doi.org/10.52041/serj.v8i1.457>

- Martínez-Castro, C., & Zapata-Cardona, L. (2021). El pensamiento estadístico como herramienta para la ciudadanía crítica en la formación inicial de profesores de estadística [Statistical thinking as a tool for critical citizenship in the initial preparation of statistics teachers]. *Yupana*, 13(19), 8–27. <https://doi.org/10.14409/yu.v0i13.10823>
- Mesa, M. (2019). La Educación para la Ciudadanía Global: Una apuesta por la Democracia [Global Citizenship Education: A Commitment to Democracy]. *Revista Internacional de Educación para la Justicia Social*, 8(1), 15–26. <https://doi.org/10.15366/riejs.2019.8.1.001>
- Ministry of Education (2006). *Estándares Básicos de Competencias en Lenguaje, Matemáticas, Ciencias y Ciudadanas [Basic Standards of Competencies in Language, Mathematics, Science and Citizenship]*. Revolución Educativa Colombia Aprende.
- O'Neil, C. (2016). *Weapons of math destruction. How big data increases inequalities and threatens democracy*. Crown.
- Organisation for Economic Co-operation and Development (OECD). (2018). The Future of Education and Skills, Education 2030: The Future We Want (position paper). OECD. [https://www.oecd.org/education/2030/E2030%20Position%20Paper%20\(05.04.2018\).pdf](https://www.oecd.org/education/2030/E2030%20Position%20Paper%20(05.04.2018).pdf)
- ProCivicStat Partners (2018). *Engaging civic statistics: A call for action and recommendations*. [https://iase-web.org/islp/pcs/documents/ProCivicStat\\_Report.pdf?1545118071](https://iase-web.org/islp/pcs/documents/ProCivicStat_Report.pdf?1545118071)
- Powell, A. B., Francisco, J. M., & Maher, C. A. (2003). An analytical model for studying the development of learners' mathematical ideas and reasoning using videotape data. *The Journal of Mathematical Behavior*, 22(4), 405–435. <https://doi.org/10.1016/j.jmathb.2003.09.002>
- Ridgway, J. (2022). Why engage with Civic Statistics. In J. Ridgway (Ed) *Statistics for empowerment and social engagement: Teaching Civic Statistics to develop informed citizens* (pp. 1–16). Springer. [https://doi.org/10.1007/978-3-031-20748-8\\_1](https://doi.org/10.1007/978-3-031-20748-8_1)
- Rockström, J., et al. (2009). Research Planetary Boundaries: Exploring the safe operating space for humanity. *Ecology and Society*, 14 (2). <https://www.ecologyandsociety.org/vol14/iss2/art32/>
- Rowan, R. (2018). Open-source technical communication in the classroom. In G. Y. Agboka & N. Matveeva (Eds.), *Citizenship and advocacy in technical communication: Scholarly and pedagogical perspectives* (pp. 265–284). Routledge.
- Rubin, A. (2022). Using data for good: a matter of geography. In S. A. Peters, L. Zapata-Cardona, F. Bonafini, & A. Fan (Eds.), *Bridging the Gap: Empowering & Educating Today's Learners in Statistics. Proceedings of the 11th International Conference on Teaching Statistics (ICOTS11)*, Rosario, Argentina. International Association for Statistical Education. <https://doi.org/10.52041/iase.icots11.T2I2>
- Sagarin, R., & Micheli, F. (2001). Climate change in nontraditional data sets. *Science, New Series*, 294(5543), 811. <https://doi.org/10.1126/science.1064218>
- Santamaría-Cárdaba, N., & Lourenço, M. (2021). Global citizenship education in primary school: a comparative analysis of education policy documents in Portugal and Spain. *Iberoamerican Journal of Development Studies*, 10(2), 130–158. [https://doi.org/10.26754/ojs\\_ried/ijds.585](https://doi.org/10.26754/ojs_ried/ijds.585)
- United Nations Educational, Scientific and Cultural Organization [UNESCO] (2015). *Global citizenship education: Topics and learning objectives*. UNESCO. <http://unesdoc.unesco.org/images/0023/002329/232993e.pdf>
- United Nations Educational, Scientific and Cultural Organization [UNESCO]. (2018). Preparing teachers for global citizenship education: a template - UNESCO Digital Library. Bangkok: UNESCO. <https://unesdoc.unesco.org/ark:/48223/pf0000265452>
- United Nations (2020). Harmony with Nature Website. <http://www.harmonywithnatureun.org/>
- Weiland, T. (2017). The importance of context in task selection. *Teaching Statistics*, 39(1), 20–25. <https://doi.org/10.1111/test.12116>
- Wilkerson, M. H. & Laina, V. (2017). Youth Reasoning with Interactive Data Visualizations: A Preliminary Study. *Proceedings of the 2017 Conference on Interaction Design and Children, Conference: Interaction, Design and Children*, 411–416. <https://doi.org/10.1145/3078072.3084302>
- Zapata-Cardona, L., Escobar-Franco, D. M., Santa-Ramírez, Z., García Barco, L. & Riascos Forero, Y. (2022). Promoting critical citizenship in professional development of mathematics teachers via statistical investigations. In S. A. Peters, L. Zapata-Cardona, F. Bonafini, & A. Fan (Eds.), *Bridging the Gap: Empowering & Educating Today's Learners in Statistics. Proceedings of the 11th International Conference on Teaching Statistics (ICOTS11)*, Rosario, Argentina. International Association for Statistical Education. <https://doi.org/10.52041/iase.icots11.T1C2>
- Zapata-Cardona, L. (2023). The possibilities of exploring nontraditional datasets with young children. *Teaching Statistics*, 45, S22–S29. <https://doi.org/10.1111/test.12349>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.