

Difficulties in Understanding the Arithmetic Mean: An Experience with Problem-Based Learning and ICT in Hybrid Teaching

Rubén Darío González Aranguiz, Carlos Mario Vanegas-Ortega*

Departamento de Matemática y Ciencia de la Computación, Universidad de Santiago de Chile, Santiago de Chile, Chile

*Corresponding Author

DOI: <https://doi.org/10.51584/IJRIAS.2026.11080128>

Received: 03 September 2026; Accepted: 09 September 2026; Published: 17 September 2026

ABSTRACT

Research in statistics education has repeatedly shown that school students master the computational algorithm of the arithmetic mean without appropriating its meanings. This article describes the difficulties that Chilean seventh-grade students display in understanding the arithmetic mean during a teaching sequence based on problem-based learning (PBL) with contextual data and on the use of information and communication technologies (ICT), specifically the Fathom software. The study follows a qualitative descriptive design and was carried out in a single class of twenty students at a municipal school in Santiago de Chile during the COVID-19 pandemic, under hybrid conditions in which one group of students attended in person while another followed the same lessons remotely. Because of the rotating attendance system, effective participation varied between twelve and twenty students across sessions. Data were collected through a pre-test, the written productions generated during three phases of a constructivist learning cycle, and a post-test, and were analysed by means of inductive content analysis, from which six categories emerged. The results show that participants enter and leave the sequence with a stable procedural command of the algorithm and without the interpretations of the mean as a fair share or as a balance point. Most students identify that the mean changes when an outlier is added, but attribute that change to the increase in the number of data rather than to the magnitude of the added value. Given the descriptive scope of the study, the findings do not allow an effect to be attributed to the sequence; they describe that the initial difficulties persist. The study suggests that the mediating role commonly attributed to ICT is not a property of the software but an effect of its didactic integration, and that the rules of PBL must enter the didactic contract in order to operate.

Keywords: arithmetic mean; problem-based learning; ICT; statistics education; hybrid teaching

INTRODUCTION

The study of mathematics has been a cornerstone of human development. Scientific knowledge, together with the other areas of learning, has advanced on the basis of its progress. This, however, does not seem to be matched by the way mathematics is taught in schools, where instruction still relies largely on worksheets of exercises that are remote from students' realities and that do little to develop critical reasoning.

The same can be said of statistics. Statistics education in Chile—and in Latin America and part of the Caribbean more generally—has been described as deficient, a diagnosis confirmed by the poor perception people hold of the discipline regarding what it is and what it does [1]. To this must be added the widespread view, reported by Juárez and Inzunza [16], that statistics is a subject full of formulas and laborious procedures that must be memorised in order to solve a problem, a view held by students and teachers alike.

One way of bringing statistics closer to students is through the use of ICT, which has been shown to offer several advantages: it broadens the educational offer, enables more flexible learning environments, strengthens interactive settings and scenarios, and provides access to a wide range of learning resources [13]. Caution is nevertheless required. The introduction of technological resources in educational institutions often responds

more to economic interests than to a genuine intention to renew the teaching and learning process [8]. Moreover, even though ICT is a powerful resource, it may become an obstacle for teachers. Several authors report that teachers have faced difficulties in incorporating ICT into their practice, difficulties reflected in an inadequate articulation between these technologies and the disciplines they teach, poor contextualisation and, at best, an insufficient exploitation of their potential [15]. Teachers themselves frequently attribute their limited use of computers to a lack of technological preparation [8].

What happens, then, with the teaching of statistics with ICT? Although several statistical packages exist —SPSS, R, Excel, OpenStat or ezANOVA, among others— most of them are aimed at university students in fields such as economics, statistics or mathematics, and some require a programming language, which makes them harder to handle. None of them is designed for school students, and software of that kind is scarce. As Behar and Ojeda [4] point out, most of the existing software is oriented towards data processing and is not conceived for pedagogical purposes. This makes learning statistics more tedious and complex for students and leaves aside essential aspects of its understanding.

This article therefore advocates teaching statistics through ICT and the use of real, contextualised data drawn from the perspective of problem-based learning (PBL). Although the wider research addressed measures of central tendency in general, the present report focuses on the arithmetic mean in order to illustrate how, from this perspective, statistics may move from a utilitarian tool towards forms of understanding through which students respond to situations of personal and civic interest.

The objective of this article is to describe the difficulties in understanding the arithmetic mean displayed by Chilean seventh-grade students when the content is mediated by problem-based learning with contextual data and the incorporation of ICT. Accordingly, the research question guiding the study is: what difficulties do seventh-grade students display in understanding the arithmetic mean when working through problem-based learning with the use of ICT?

Theoretical Framework

Meanings of and difficulties with the arithmetic mean

The arithmetic mean is usually presented at school as one of the simplest topics in statistics, since its computation requires two operations that students have known since their earliest years of schooling. That apparent procedural simplicity, however, conceals the conceptual complexity of the object. As Batanero [3] argues, the arithmetic mean condenses a plurality of meanings that cannot be derived from one another and that students must progressively articulate: mastering the algorithm does not entail mastering the concept, and a student who correctly states the formal definition may still be unable to decide whether the value obtained makes sense in the situation that produced it. In their analysis of secondary school textbooks, Cobo and Batanero [6] show that this reduction of the concept to its computational expression is not an accident of teaching but a feature of the material with which it is taught: textbooks present the mean almost exclusively in association with the procedure and with very little variety in the situations of use, which restricts the field of problems from which students can construct its meaning.

A first way of characterising this complexity is to identify the properties a student must recognise in order to be said to understand the mean. Strauss and Bichler [21] established a set of seven properties that has guided much of the subsequent research: the mean lies between the extreme values of the data set; the sum of the deviations of the data from it equals zero; its value is affected by every one of the data considered; it need not coincide with any of the values in the set; it may correspond to a number with no counterpart in the physical reality being described; the computation must include null values when these form part of the situation; and, finally, the mean is a value that represents the data set it summarises. The authors document that the properties of a statistical nature —particularly representativeness and the compensation of deviations— are considerably more difficult for schoolchildren than those of an arithmetic nature, and that mastering them does not automatically accompany mastery of the computation.

A second and complementary path consists in distinguishing the meanings of use that the mean acquires according to the problem being solved. Following Cobo [5] and Cobo and Batanero [6], at least four can be recognised. The first is that of fair share: the mean answers the question of how much each case would receive if the total were distributed equally, and it is the meaning that underpins expressions such as the average number of pencils per child or the average number of flights per day. The second is that of representative element of the set, that is, the value that can stand for all the data when the interest lies in characterising them with a single figure. The third is that of an estimate of an unknown quantity in the presence of measurements subject to error, a meaning that is historically prior to the others and linked to the treatment of variability. The fourth is that of balance point of the distribution, which rests on the property that deviations cancel out and makes it possible to understand why the mean shifts towards extreme values when these appear in the set. These meanings are not interchangeable: a student may mobilise fair share in a situation involving school grades and yet lack the balance-point meaning when asked to anticipate the effect of an outlier.

The difficulties documented by research are organised around that distance between procedure and meaning. Pollatsek, Lima and Well [20] showed, with university students, that automatic recourse to the algorithm leads to systematic error in situations requiring weighting, which reveals that the formula operates as a rule devoid of reference to the data set it summarises. In a longitudinal study, Watson and Moritz [22] described how the notion of average develops from a colloquial idea —what is normal, what usually happens— towards a properly statistical understanding, and how that transition is slow and not always completed during compulsory schooling. Mayén, Batanero and Díaz [18], for their part, identified persistent semiotic conflicts in students' work with measures of central tendency, associated with the difficulty of relating the computed value to the distribution that generates it. Taken together, this literature suggests that difficulties in learning the mean are not resolved by further practice of the computation, because they do not reside in the computation.

In the Chilean context, this characterisation has been taken up by recent research. Alvarado [1] states that teaching and conceiving the mean as the mathematical expression that allows its calculation prevents students from fully understanding the concept, and Gatica [11] warns about the scarce presence, in school teaching, of tasks that require working with real data and with the variability these introduce. To this must be added what Juárez and Inzunza [16] found regarding the perception of statistics as a repertoire of formulas and procedures to be memorised, a perception that becomes established among students and teachers alike. In the Chilean national curriculum the mean is a compulsory minimum content from the fifth year of primary education, which means that seventh-grade students reach its formal treatment with an established procedural command and, frequently, without any of the meanings described above.

This framework has direct consequences for the design of the sequence reported in this article. Of the four meanings identified, those of fair share and balance point were chosen for two reasons. The first connects with situations familiar to students and makes it possible to give reference to the algorithm they already master. The second articulates the most demanding statistical properties —the compensation of deviations and sensitivity to extreme values— and makes visible the relationship between the mean and the complete distribution of the data. The tasks in the sequence were designed to appeal successively to both meanings, and the Fathom software was incorporated in order to manipulate data and observe the effect of that manipulation on the statistic, in line with what Garfield and Ben-Zvi [10] describe as working with variability in the construction of statistical reasoning.

Problem-based learning and ICT in the teaching of statistics

Traditional or expository teaching has been unable to meet all the goals set for school education. This form of teaching, in which teachers transmit knowledge and students merely receive it, has resulted in a decline in the understanding of the content studied as well as in general demotivation among students.

How, then, might mathematics be taught more effectively? PBL is proposed as a novel and appealing way of addressing not only knowledge but also the enjoyment of learning, with students as the protagonists of their own learning. This may be reinforced by the skills students develop within this methodology. As Ortega et al. [19] state, such skills include critical thinking and the capacity for analysis and synthesis, so that promoting their development facilitates the learning of mathematical concepts and content and gives students useful tools to recognise, pose and solve real-life problems collaboratively.

What relationship can be established between ICT and PBL? Alzate et al. [2] note that the use of ICT can be a powerful vehicle for introducing new teaching methodologies, since today's students are deeply immersed in everything related to technology. In other words, ICT constitutes a powerful tool that catalyses the learning students progressively acquire, because computations become faster and teaching can focus on improving the understanding of concepts rather than on calculation. Technology can also provide students with the opportunity to conduct real investigations, to ask real questions of real interest [4]. This allows students to work with reliable and current information, which, as Gea et al. [12] point out, enhances interdisciplinarity in the statistics classroom and enables the learning of content not usually acquired through textbook problems, such as the effect of outliers on the computation of a statistic.

The appropriate use of statistical software can make the teaching and learning of statistics more effective, since it fosters active learning and improves students' understanding of statistical concepts and their problem-solving skills [23]. Because statistical software gives dynamism to data, it contributes to exposing students to variation when parameters are changed, as well as to the differences between models, thereby improving learning experiences [23]. By modifying data, students have the opportunity to develop argumentation based on scientific information and reasoning about the effect of variables in contextual problems. As Gea et al. [12] argue, ICT fosters students' use of verbal, symbolic, iconic and graphical language, facilitates and broadens the range of procedures available to them, makes concepts and properties ostensive —above all through simulation— and promotes reasoning and argumentation.

METHODOLOGY

Approach and design

The methodological approach of this research is qualitative. As Hernández et al. [14] put it, qualitative research provides depth of data, dispersion, interpretive richness, contextualisation of the setting, detail and unique experiences; this kind of information proved indispensable for meeting the objective, since it made it possible to answer the research question. The scope of the study is descriptive, as its purpose is to specify important properties and characteristics of the phenomenon under analysis, gathering information independently or jointly on the concepts or variables to which it refers [14].

The methodological design comprised three stages. In stage one a pre-test was administered, divided into two sections: the first with conceptual questions asking students to explain in their own words what the mean is, and the second concerned with computing it and arguing what the result indicates. The purpose was to establish students' prior knowledge of the mean and whether they were able to compute and interpret it. Stage two corresponds to the sequence of activities based on the constructivist learning cycle, understood as a theoretical and formal construction which, grounded in scientific and ideological assumptions, seeks to interpret school reality and direct it towards particular educational aims [17]. Three of the four phases of the cycle were considered: exploration activities, which involve analysing concrete and simple situations that are as close as possible to students' reality [7] and which situate students within the topic under study while activating prior knowledge [9]; activities introducing new knowledge, which lead students to explain and adjust initial models, identify new points of view, solve the problems posed and seek attributes that define concepts and the relations between prior and new ideas [7]; and structuring and synthesis activities, which help students construct knowledge as a consequence of interaction with the teacher and peers and personal adjustment [9], showing what they have learned and the changes in their points of view [7]. In stage three a post-test was administered with the same structure as the pre-test, adding a question on what happens to the mean when an outlier is incorporated.

Participants

A single class was selected in order to analyse the development of students' responses in a pre-test, during the implementation of the activities and in a post-test. Owing to the COVID-19 pandemic, students took part in hybrid mode: one group was physically present in the classroom while the other followed the same instructions online.

The class in which the intervention was carried out comprised twenty students. Because the school operated under a capacity-based rotating attendance system, and because participation in each session also depended on actual attendance and on the connectivity of those following the lesson remotely, both the number of participants and the composition of the in-person and remote groups varied from one phase to another: the same students did not attend every lesson, nor did they always attend in the same modality. For this reason the data reported in the results section do not correspond to a fixed group but to the set of students who submitted responses on each occasion. Table 1 summarises this participation.

Table 1. Students who submitted responses in each phase, by modality of participation

Phase / instrument	In person	Remote	Total
Stage 1. Pre-test	7	11	18
Stage 2. Exploration activities	11	7	18
Stage 2. Introduction of new knowledge	7	13	20
Stage 2. Structuring and synthesis	5	10	15
Stage 3. Post-test	7	5	12

Participation therefore fluctuated between twelve and twenty students, that is, between 60% and 100% of the class. The pre-test was answered by eighteen students and the post-test by twelve. In the structuring and synthesis phase only the responses actually submitted were recorded, so the figure reported for the in-person modality corresponds to that record and not necessarily to the total number of students attending. Consequently, the pre-test and post-test results do not permit individual pairing, and the comparisons presented below must be read as a characterisation of the set of responses obtained at each moment rather than as the learning trajectory of the same students.

The teaching sequence and the software

The purpose of the activities was to help students improve their understanding of measures of central tendency through everyday and simple situations and, progressively, to encourage them to solve more complex and realistic cases, placing them in a hypothetical situation in which they might work as researchers on a specific topic. Some of the advantages of PBL are thereby developed as the study of measures of central tendency advances [2]: the sequence seeks to stimulate the development of personal skills, to promote interaction between students and teacher, and to improve motivation and collaboration across disciplines.

The Fathom statistical software was chosen not only because it plots data but also because it gives them dynamism, facilitates the visualisation of the changes produced in a graph and of how variability may affect certain statistics. This makes it possible to expose students to variation when parameters are changed as well as to the differences between models, thereby improving learning experiences [23]. Figure 1 shows a representation in Fathom of the activity on the mean proposed to the students.

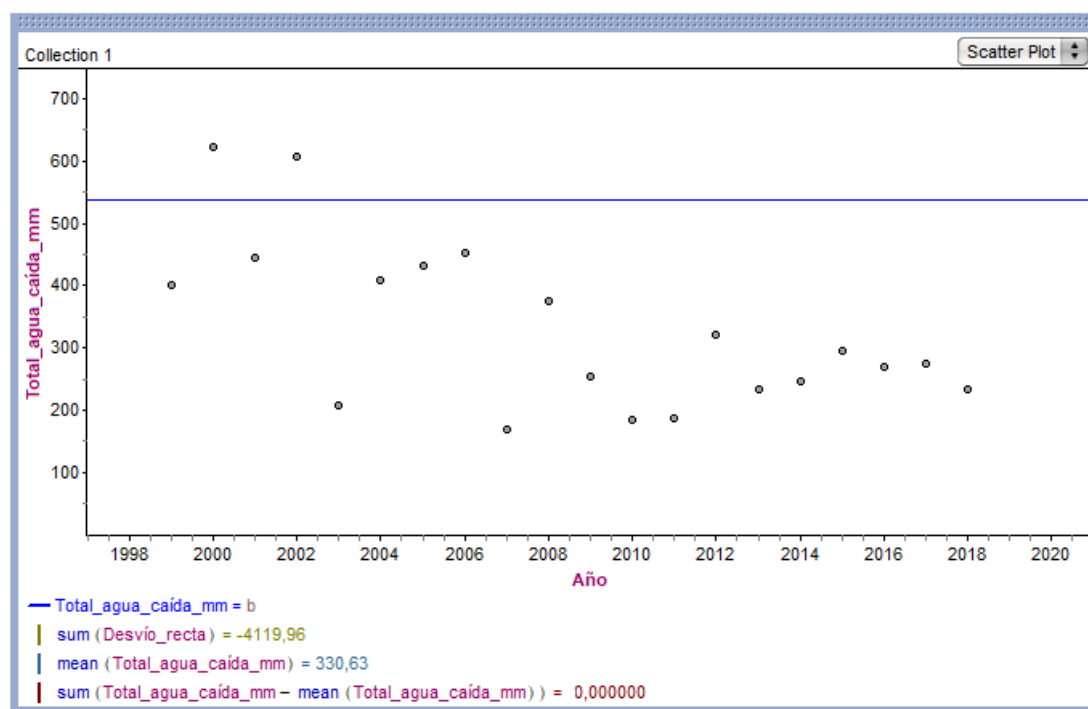


Figure 1. Representation of the data for computing the mean and the sum of deviations in Fathom

The relationship between the theoretical framework and the activities designed in the didactic sequence, together with the pre-test and post-test, lies in the use of questions that do not only involve computing measures of central tendency but also include open-response questions associated with contextual data and problems, in which students must state their ideas and write them down before developing them jointly with their peers, whether they worked remotely or in person.

Instruments and validation

The instruments (tests and lesson sequence) were validated by two doctors in statistics education, who examined the relevance of the tasks with respect to the meanings of the mean that the sequence sought to mobilise, the appropriateness of the language for the students' level, and the role assigned to the software in each phase. As a result of this review, adjustments were proposed to the wording of the open questions —so that they would not admit purely algorithmic answers—, to the simplification of the contextual statements, and to the inclusion of explicit instructions for data manipulation in Fathom, so that the use of technology would not be restricted to visualisation. All the proposed adjustments were adopted by the researchers prior to implementation. In addition, the instruments were piloted with seventh-grade students of similar characteristics to those of the participants, in a school in the Valparaíso region that was also working in hybrid mode.

Data analysis

Responses were analysed by means of inductive content analysis. Students' written productions in the pre-test, in the worksheets of each phase and in the post-test were read successively, grouping responses by similarity in the way they referred to the mean, without departing from a previously defined system of categories. Six categories emerged from this process, and these organise the presentation of the results:

- Algorithmic definition of the mean: the student characterises the arithmetic mean as the sum of the data divided by the number of data, without referring to what that value represents. This includes formal definitions reproduced from the internet.
- Imprecise definition or absence of definition: the response is omitted or resorts to formulations that do not allow the concept to be recognised (for example, “an infinite set of numbers” or “the final result of several values”).

(c) Incomplete or erroneous computation: the student begins the computation but omits one of its steps — typically the division— or makes operational and rounding errors.

(d) Contextual interpretation without sharing out: the value obtained is named in the terms of the statement (“the average of flights for the week”, “the average of pencils per child”) without making explicit the distribution that value implies.

(e) The mean as a fair share: the student refers to the mean as distributing the total equally among the cases considered.

(f) The mean as a balance point: the student recognises that the mean shifts in the presence of outliers, that it lies between the extreme values of the set, and that the sum of the deviations from it cancels out.

The first two categories account for the meanings with which students enter the sequence; the third and fourth for the procedures and readings they deploy in contextualised situations; and the last two correspond to the interpretations the didactic sequence sought to promote which, following Cobo and Batanero [6], constitute two of the meanings of use of the arithmetic mean. Coding was carried out by the research team and discrepancies were resolved by agreement, jointly reviewing the responses in dispute.

Ethical considerations

The study complied with the ethical procedures required for research with human beings and, in particular, with minors. Authorisation was obtained from the school management, informed consent from the parents or guardians and informed assent from every student, to whom the purpose of the study, the voluntary nature of their participation and their right to withdraw at any time without consequence were explained. Participation had no bearing on subject grades. All data were anonymised and the responses reproduced in the figures are presented without any element that would allow their authors to be identified, with express authorisation for their use for scientific dissemination purposes.

RESULTS AND DISCUSSION

Analysis of the pre-test responses

The analysis was carried out in two parts: the first concerning the definition of the mean and the second the computation of and argumentation about the arithmetic mean through exercises. The analysis also considers each question separately.

Concepts

When students were asked to define the mean or arithmetic mean in their own words, six of the eleven who worked online responded similarly, stating that it is “an infinite set of numbers”, an answer also given by one student who responded in person. Two students stated that the arithmetic mean is “a single number taken as the representative of a list of numbers” or “the characteristic value of a series of qualitative data”. Three further students gave answers such as “the average value found by adding the values and dividing by the number of existing values”, “what is obtained by adding all the numbers or data and dividing the result by the total number of numbers or data”, and “the mean of a set of numbers, sometimes simply called the average, is the sum of the data divided by the total number of data”. In other words, they provide the formal definition of the mean, clearly referring to obtaining it through the algorithm.

Of the seven students who took part in person, five did not answer the question about the definition of the mean. The two remaining answers were “it is the final result of several values” and “percentage based on ‘percentages’ to make a calculation”.

One of the difficulties encountered was determining whether students already had knowledge of the mean, while in other cases it is apparent that many of them copied their answers from the internet, which casts doubt on

whether they understood them fully. Although the internet has become a powerful tool for gathering information and clarifying doubts, what it provides is not always entirely correct or intelligible for students' level of knowledge.

Computation and argumentation

In this second part of the pre-test students were asked to compute the arithmetic mean and interpret the value obtained. As the following example illustrates, students had difficulties with both tasks. What is striking about these errors is that they occur with content already covered in the fifth year of primary education.

An airline pilot wishes to know the number of flights from Santiago de Chile to other regions during one week. After a few days the operator provides the following information (Table 2).

Table 2. Number of flights from Santiago de Chile to other regions during one week

Day of the week	Number of flights
Monday	23
Tuesday	34
Wednesday	25
Thursday	29
Friday	28
Saturday	34
Sunday	35

When asked to compute the mean, the seven students attending in person answered incorrectly, since they only added the number of flights during the week but did not divide the result by the number of days in a week. Five students who answered online made the same mistake. Only six students in that modality answered correctly (29.71 flights on average); some of them rounded their answers to a whole number (29 or 30). As the images in Figure 2 show, an interesting aspect to note is the error made by the twelve students who answered the question, which consisted in not dividing by the number of days in the week.

a.1) Calcula el promedio de viajes que hubo durante esta semana (2 pts)

hubo un promedio 208 vuelos esta semana

a.1) Calcula el promedio de viajes que hubo durante esta semana

Y es que hubo un promedio de 217 vuelos en total

Figure 2. Errors made by two students when computing the mean number of flights during the week

When asked what the value they had obtained meant, six of the seven in-person students answered the question and only one did not; the others state that it is “the number of flights there were during the week”, without referring to a distribution between the days and the number of flights per day. Among the students who took part remotely, eight answered similarly, stating that it is the average of flights during the week. Other answers were

of the type “it is the result of adding all the flights and then dividing them by seven”, that is, referring to the procedure or algorithm for computing the mean. As with their in-person peers, none of them refers to the mean as a distribution.

Another pre-test problem posed the following situation: Sofia looks at her grades in a subject's tests, which were 5.4, 3.3, 6.7, 4.3, 5.1, 3.3, 5.8, 4.9, 6.0 and 5.7. Of the seven in-person students, six did not compute the mean or state what that value meant. Only one gives an estimate, indicating that the grade average lies between 5.0 and 5.1, and none provides an answer about the meaning of the value of the mean, this time understood as a balance point which, in the context of this problem, refers to a grade that equalises the sum of the distances between the highest and the lowest grades. Of the eleven students taking part remotely, two do not answer; the other nine give the correct answer to the computation of the grade average (5.05). However, and just as their peers in the classroom, none of them offers an interpretation of the value of the mean. This corroborates Alvarado's [1] claim that considering the arithmetic mean from the expression that allows its computation—the sum of all the data divided by the number of data added—prevents proper reflection on the arithmetic mean and on its consideration as a representative of a data set.

In sum, the pre-test shows that students do not handle a definition of the mean that goes beyond its computational algorithm. It is worth noting that students covered the topic of the mean in the fifth year of primary education, since it is a compulsory minimum content of the Chilean curriculum at that level. Nevertheless, they do not appear to understand that the mean might be interpreted as a distribution. This is consistent with Alvarado's [1] observation that teaching and conceiving the mean as the mathematical expression that allows its computation prevents students from fully understanding the concept.

This first result allows the starting point of the intervention to be characterised precisely: the participants' difficulty is not one of computation but of meaning. Students have the algorithm at their disposal—they apply it fluently when the context is familiar and scholastic, as with grade averages—but they do not have a mathematical object to which that procedure refers. In Alvarado's [1] terms, the mean functions for them as a computational instruction rather than as a representative of the data set, which explains how they can compute it correctly and, at the same time, be unable to say what the resulting number means. To this must be added a finding specific to the hybrid modality: consultation of the internet during the pre-test produced formally correct definitions that do not correspond to any verifiable understanding. This qualifies the diagnosis of Juárez and Inzunza [16] concerning statistics as a repertoire of formulas to be memorised: in a remote environment the formula no longer needs to be memorised, only located, which shifts the didactic problem from retention towards interpretation.

A further difficulty during the pre-test concerned students taking part remotely, since they copied their answers and used the internet to respond to the questions asking them to define the arithmetic mean. Some of them provided formal definitions, but it became evident that they probably did not understand them or simply did not know what they were answering.

Analysis of the responses in the exploration phase

The analysis of students' responses considered only those questions that are most significant for the objective of this article, both because of the way they were answered and because of the difficulties they revealed. Questions in which difficulties reported in other research became apparent were also considered. Let us consider the following problem: a doctor records the height of the adolescent patients who come to her practice. The first records were (in metres): 1.67, 1.70, 1.69, 1.65, 1.64, 1.73. A new patient then arrives whose recorded height is 1.97 m. By how much has the value of the mean varied? Why is this so, if we have added one more datum?

Of the eleven students who took part in person, four leave the question blank, three state that the mean changes because more data are added, and another three argue that the mean has varied by four centimetres with respect to the previous value because there is a person who is taller than the other patients; no record of a response is available for the remaining student. Of the seven students who answered this question remotely, one provides the value of the mean using the eight patients' heights; one does not answer; two state that the mean changes

because one more datum was added; and three argue that the mean changes because the last patient was very tall.

It can be observed that a small group of participants gives the correct answer, namely that the mean changes its value (from 1.68 to 1.72) because there is a patient who is very tall. This response occurs in both modalities: in three students who worked in person and in another three who worked remotely. Both the case of the student who answered in person and that of the student who took part remotely are illustrated in Figure 3.

d.5) Posteriormente, ingresa un nuevo paciente a la consulta. La estatura que registra la doctora es de 1,97 m. ¿Cuánto ha variado el valor del promedio? ¿A qué se debe esto si hemos agregado un dato más? *

el promedio de ser 1.68 a pasado a ser 1.72, ya que lo que mide el paciente es mas de lo que el máximo demostró (1.73, ósea que a mide 0,19 cm que el otro paciente) el nuevo paciente es un dato atípico

d.5) Posteriormente, ingresa un nuevo paciente a la consulta. La estatura que registra la doctora es de 1,97 m. ¿Cuánto ha variado el valor del promedio? ¿A qué se debe esto si hemos agregado un dato más?

se debe ha que se lo agrego un digito superior y mucho mas grande (Por que era mucho mas grande que los demas)

Figure 3. Types of responses given by two students who took part in person and remotely

It may be inferred that students know the algorithm for computing the mean, since they associate it with computing their own grade averages. However, they have difficulty realising that this measure of central tendency can be interpreted as an equitable distribution, as they do not appear to have encountered this way of seeing it, even though the condition does not always hold. It also appears, at least initially, that they did not know that the arithmetic mean varies considerably in the presence of values that depart from the norm, nor that the value of the arithmetic mean lies between the extreme values of a data set.

What this phase reveals is an asymmetry between what the task offered and what students were in a position to read in it. The situation involving heights was designed precisely to make visible the effect of an outlier on the mean, which is the kind of phenomenon that, according to Gea et al. [12], technology makes it possible to address and that is not usually acquired through textbook problems. Most participants, however, merely note that the mean changed, attributing this to the increase in the number of data rather than to the magnitude of the datum added. That is, they identify the variation but not its cause. This suggests that mere exposure to a well-constructed situation is not enough: if students do not have available the idea that the mean is sensitive to extreme values, the variation observed is explained by the resource at hand—there are more data, therefore the result changes—and the task is solved without the concept at stake ever being put to the test.

Analysis of the responses in the phase introducing new knowledge

To illustrate the results of this phase, the following question is considered, related to the previous example on grades: what will happen to the average if this student's grades had been 5.2, 4.3, 5.0, 5.8, 7.0 and 6.1? And if the grades had been 1.5, 5.0, 6.1, 4.3, 5.8, 7.0? Why has the average varied if only one grade has changed? What can you conclude from this?

Of the seven students taking part in person, six submitted incomplete answers, in which four of them computed one of the two averages of the grades given, or merely wrote that the average rises or falls depending on whether

there are high or low grades. As illustrated in Figure 4, one student gives the complete answer, computing both averages and indicating the variation that occurs when there are high or low values.

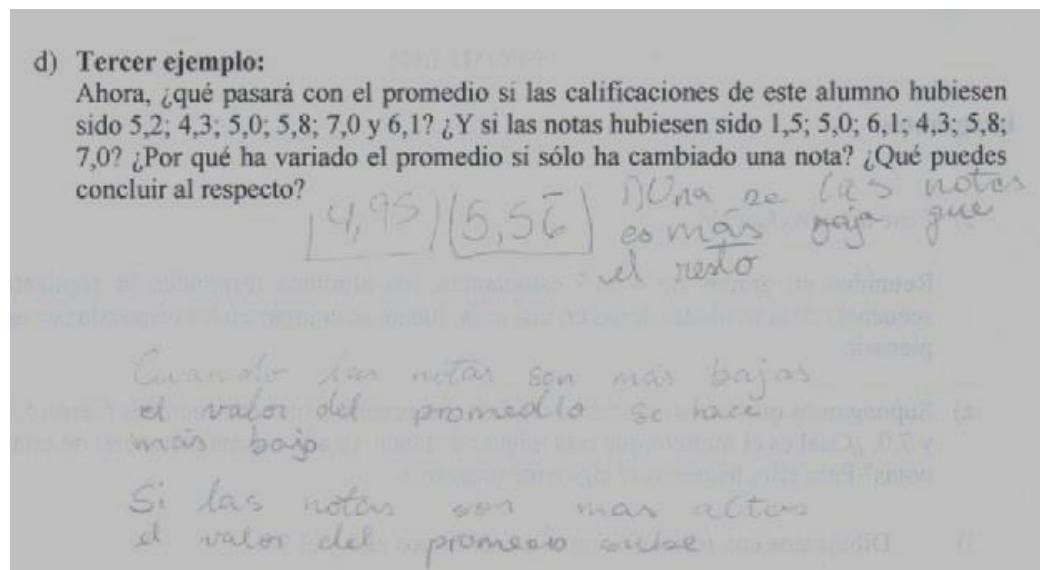


Figure 4. Answer given by a student when comparing the two grade averages and their variation

The thirteen students who took part remotely were likewise not free of difficulties: five of them did not answer, another three give inconclusive answers, and only five argue that the average changes when grades vary considerably because there is a grade that is lower than the others (Figure 5).

con la primera nota son 5,5 Y la otra nota seria 4,9 y la variación se debe a la nota 1,5 de el otro grupo el cual baja todo el promedio a el alumno

Figure 5. Conclusions given by a student on how the mean varies when one grade changes

This activity was more complex for students, particularly for the group working in person. This can be contrasted with the answers put forward by their peers working remotely, since their arguments were more accurate and conclusive and gave more reasons why the average rises or falls in the presence of higher or lower grades. Although all students know how to compute the arithmetic mean of a set of grades, despite making errors in the addition or the division, the interpretation of the mean as a balance point was not achieved.

Analysis of the responses in the structuring phase

Working in groups, students answered the following activities following the sequence set out in a worksheet, using the Fathom software: how much has the amount of rainfall in Santiago de Chile decreased? It is known that, in Santiago, global warming has led to a decrease in rainfall, which has resulted in recent years being declared the driest for the capital. But how true is this claim? Suppose you are looking for this information and wish to corroborate it. To this end, you are given the following table, which summarises the amount of rainfall per year (Table 3).

Table 3. Amount of rainfall in Santiago between 1999 and 2018

Year	Total rainfall (mm)
1999	400.4
2000	621.8
2001	443.3

2002	608.0
2003	207.2
2004	408.5
2005	431.8
2006	453.6
2007	167.7
2008	374.8
2009	254.9
2010	184.7
2011	187.4
2012	320.0
2013	232.7
2014	246.2
2015	295.0
2016	268.3
2017	273.5
2018	233.1

During 2019 only 69.5 mm of rain fell, making it the driest year in recent decades. Enter this datum into the software and see how the mean changes. By how much has it changed with respect to the value of the mean you obtained earlier? What happens to the value of the total sum of the deviations from the mean?

Five students who worked in person answer the question correctly, indicating the values of the mean rainfall with the data given in the software and adding the new value. Of this group, three argue that the sum of the deviations becomes very close to zero—it should be noted that, strictly speaking, that sum equals zero, and that the differences observed are due to the decimal rounding with which they worked.

The ten peers who took part remotely reported greater difficulties, as their answers show: four of them provide the values by which the mean changed and, of these, only two state that the sum of the deviations becomes equal or very close to zero. Another student indicates how much the value of the arithmetic mean varied before and after adding the new datum, but does not write down the values of the means with the data provided. Of this group, only one student attempts to justify why the mean changed so much, stating that it was because one more datum was added, which is an incomplete answer since it fails to mention that this datum was smaller than the others.

The structuring phase, as its name indicates, aims for students to understand, using educational software, that the sum of the deviations from the mean equals zero even when atypical data are added. Students also worked with real data that would allow them to appreciate that statistics is a tool capable of responding to other, scientific or social, fields. In this respect it is worth noting that the problem already identified by Gatica [11] emerged:

students are not used to working with real data, which frequently require high-level interpretation and reasoning, and changes in situations mean that results are not unique, with greater variability than in other areas of mathematics. This is illustrated by the arguments participants put forward in this phase: their answers were rather brief and terse, which shows that they have difficulty explaining in words when faced with open questions.

It is worth dwelling on this observation, because the argument for real data operates in two directions. The literature consulted presents them as a way of giving meaning to school statistics [4], [12], and that was the reason for incorporating the Santiago rainfall series. Nevertheless, and in line with Gatica's [11] warning, real data simultaneously increase the cognitive demand of the task: they require deciding what is relevant, tolerating variability and sustaining an interpretation that is not given by the statement. For students who have not yet stabilised the meaning of the mean, that additional demand competes with the conceptual aim of the phase instead of supporting it. The result observed—brief answers centred on the numerical value rather than on its interpretation—suggests that contextualisation with authentic data requires, as a prior condition, sustained work on the interpretations of the statistic, and cannot replace it.

Analysis of the post-test responses

Concepts

When asked to define in their own words what the mean or arithmetic mean is, of the seven students taking part in person in this final stage, two do not know what to answer and a third gives an unclear response; another two answer that it is a distribution, but only one adds that it is “equitable”; one student attempts to define it by giving “the grade average” as an example; and, finally, one states that the mean is “a way of equalising something”, without specifying what that something refers to. This last answer is striking, since the student possibly attempts to refer to the mean as a balance point, implicitly indicating the equalisation of deviations from the mean.

Of the five students who answered remotely, three state that the arithmetic mean is the sum of the data divided by the total number of data, as can be seen in Figure 6, that is, they employ the definition of the mean through its algorithm. Other answers were “the average is to share out”, without developing this idea, although it appears to indicate an equitable distribution. Finally, one student defines this measure as “the central value computed among a set of numbers”.

a) Promedio o media aritmética *

Promedio es cuando de un distinto tipo de números se suman todos y se dividen por el número de veces, ej 3 notas. sumas las 3 notas y las divides en 3

a) Promedio o media aritmética *

El promedio o media aritmética es el valor central calculado entre un conjunto de numeros

Figure 6. Two types of responses given by students who took part in remote mode

Computation and argumentation

Students were asked to solve the following problem: in order to practise the concept of the mean, a teacher asks the children to obtain the average number of coloured pencils each of them has in their pencil case. The following are the quantities stated by each of the fifteen children in the class: 4, 2, 1, 5, 10, 3, 7, 9, 2, 5, 5, 2, 7, 1, 12. How could you interpret that value?

Of the seven students taking part in person, five leave this question blank; one argues that the value could be interpreted as meaning “that they do not have many pencils”, while the other explains the algorithm for computing the mean (“I add the data and then I divide them”). In short, this group of students shows that the initial difficulties in interpreting the mean in a contextualised situation persist.

Of the five students who worked remotely, two merely state that the average number of pencils per child is 5, while a third indicates that the value obtained in the previous question is interpreted as “the average of pencils for each child”. Another student justifies the answer in an unclear way, as can be seen in Figure 7. Only one student in this group gives a coherent answer, stating that the mean is an equitable distribution.

b) ¿Cómo podrías interpretar este valor? *

AQUÍ SE REPARTIÓ ENTRE TODOS LA MISMA CANTIDAD

b) ¿Cómo podrías interpretar este valor? *

que de cada 15 estudiantes un niño tendrá 5 lápices por promedio

Figure 7. Two answers given by two students to the question of how they would interpret the average number of pencils per child

Given the responses obtained from both groups of students during the post-test, it may be stated that they show no indication of the interpretations of the mean that the sequence sought to promote —fair share and balance point— and that its characterisation through the computational algorithm prevails, just as in the pre-test. Given the descriptive nature of this study, the number of participants and the conditions under which the sequence was implemented, these results do not allow any particular effect to be attributed to the proposal, nor do they permit conclusive comparisons between the two moments; they do allow it to be described that the difficulties identified at the outset persist at the end of the intervention.

As regards the incorporation of the Fathom statistical software for working with measures of central tendency, and with the mean in particular, the responses collected do not make it possible to identify advances in the understanding of the concept or of its properties, with very few students achieving an incipient notion of the content addressed during these lessons. This result qualifies Faustino and Pérez's [8] claim that students display significant learning when they make appropriate use of ICT.

The contrast with Faustino and Pérez [8] must nevertheless be read carefully, because their claim is conditional: significant learning is associated with an appropriate use of ICT, and it is precisely that condition which the present study did not satisfy. The software was used in only one of the three phases of the sequence, with closed instructions and without students having had a previous opportunity to become familiar with the environment, so that Fathom operated as a device for displaying results rather than as an instrument of exploration. This is consistent with Behar and Ojeda's [4] observation that much of the available software is oriented towards data processing and not towards pedagogical purposes, a warning that extends to the way in which the teacher integrates it. Likewise, Zamora et al.'s [23] expectation regarding the improvement produced by exposing students to variation in parameters presupposes repeated exposure that allows the relationship between on-screen manipulation and the mathematical property represented to be constructed; a single, limited use does not generate that relationship. Consequently, the results of this study do not contradict these authors but delimit the scope of their claims: they describe what happens when the condition of appropriate and sustained use is not met.

It must be understood that ICT is only a means that contributes to improving the learning teachers expect of their students. Gutiérrez and Ojeda's argument that, since expository methods have demonstrated their ineffectiveness, one must turn to technology as a potentially better alternative [4] must therefore be qualified: with the results obtained, it becomes evident that the issue is not merely the contrast between expository teaching and the use of ICT, but that there are various factors conditioning students' mathematical learning.

Synthesis of the discussion

Taken as a whole, the results allow three tensions running through the experience to be identified and made explicit. The first is the distance between algorithm and meaning: participants enter and leave the sequence with a reasonable procedural command and without the interpretations of the mean as fair share and as balance point,

which confirms Alvarado's [1] diagnosis for this case and shows that such interpretations do not emerge spontaneously from solving contextualised problems.

The second tension concerns the place of technology. The literature reviewed attributes a catalysing function to ICT [2], [12], [23], on the understanding that it frees up computation time and makes concepts and properties ostensive. The results obtained here suggest that this function is not a property of the software but an effect of its didactic integration: without continuity of use, without prior instrumental appropriation and without tasks that force a conjecture about the data to be put to the test, visualisation is exhausted in itself. The agreement with Hernández and Cuevas's [15] warning about the inadequate articulation between technologies and the discipline taught proves, in this sense, more pertinent for explaining what occurred than the optimistic expectations gathered in the theoretical framework.

The third tension concerns PBL itself. The sequence was organised according to the phases of the constructivist learning cycle [17] and drew from PBL the contextualisation of tasks and collaborative work [19], but its rules were never made explicit to students. That is, the methodology remained on the side of the teacher's design and did not enter the didactic contract. This helps to explain why open questions were answered briefly and why working with real data was not perceived as an investigation: students responded to the sequence with the expectations of an ordinary mathematics lesson, in which a numerical result is expected rather than an argument. This observation is what grounds the main recommendation formulated below.

CONCLUSION

The results of this descriptive study suggest that the mere convergence of mathematics education perspectives with PBL and ICT is not sufficient for an improvement in the understanding of the concept of the mean to occur. Various aspects that affected the implementation of the activities must be taken into account. One of them is that PBL demands time and effort in designing the activities to be implemented, since these are expected to address everyday situations or situations of interest to students, which makes the search for examples meeting these requirements a continuous challenge. Although other skills are worked on, such as argumentation, collaborative work and respect for diverse opinions, the incorporation of PBL in this research presented a disadvantage: the slow pace of progress shown by students.

The ICT component, in turn, did not have the expected effect. The Fathom software, aimed at school and university students for the teaching of mathematics and statistics, was used; nevertheless, students did not reach the expected level of understanding of the properties of the arithmetic mean. Although the software helped them visualise the concept during the lesson, that process of visualisation did not translate into responses reflecting what had been worked on by the end of the sequence. The results of this study therefore suggest that a punctual and unsystematic use of the software is insufficient to foster statistical learning, and that continuous work is required, oriented towards students establishing relationships between the problem, the data and what the software represents.

By way of summary of these two points, the example proposed in the structuring phase concerning rainfall in the Metropolitan Region of Santiago was not appealing to the schoolchildren, even though it is a topic addressed in science subjects and constantly discussed in the media. Furthermore, the effect of changing one datum and observing how it affects the value of the mean was difficult for students to understand, since they did not grasp the concept of the mean as a balance point, which was the aim of this phase. This contrasts with Zamora et al.'s [23] claim that exposing students to variation when parameters are changed, as well as to the differences between models, improves learning experiences.

The limitations of this study include the fact that the whole intervention was carried out under unfavourable conditions: hybrid lessons, with part of the class in the classroom while another group worked online; the reduction of lesson time decided by the school, from 90 to 60 minutes per lesson; and the removal of statistics content from the national curriculum, as it was not included in the curricular prioritisation established by the Ministry of Education during the pandemic. To this must be added the variation in participation from one phase to another—between twelve and twenty students—resulting from the rotating attendance system and irregular attendance, which prevents individual pairing between pre-test and post-test responses and restricts the scope of

the comparisons made. Likewise, since this is a single class and a study of descriptive scope, the results are not generalisable and must be read as the characterisation of a case.

For future studies or teaching processes with similar characteristics, a more integral use of ICT is recommended, applying it from the outset and making explicit the topic to be worked on. It is also suggested that the intended learning outcomes be made explicit, as well as what lies behind PBL; that is, that the methodology become part of the didactic contract, making it known to students and committing them to the intention of working in this way. This could generate a different disposition towards learning, since it would allow students to acquire more thorough knowledge and to appreciate that they are capable of responding to real situations contextualised within current social and scientific demands.

Ethical Approval

The study was conducted in accordance with the ethical standards for research involving human participants and minors. Authorisation was granted by the school management. Written informed consent was obtained from the parents or legal guardians of all participating students, and informed assent was obtained from each student. Participation was voluntary, could be withdrawn at any time without consequence, and had no bearing on subject grades. All data were anonymised prior to analysis, and the student productions reproduced in the figures are published with express authorisation and contain no identifying information.

Conflict Of Interest

The authors declare that they have no conflict of interest.

Data Availability Statement

The anonymised data supporting the findings of this study (pre-test, post-test and worksheet responses) are available from the corresponding author upon reasonable request. The data are not publicly archived because they consist of written productions by minors.

REFERENCES

1. Alvarado, J. E. (2017). Propuesta de innovación para la enseñanza de la media aritmética [Master's thesis, Pontificia Universidad Católica de Valparaíso].
2. Alzate, E. J., Montes, J. W., & Escobar, R. M. (2013). Diseño de actividades mediante la metodología ABP para la enseñanza de la matemática. *Scientia et Technica*, 18(3), 542-547.
3. Batanero, C. (2000). Significado y comprensión de las medidas de posición central. *UNO. Revista de Didáctica de las Matemáticas*, 25, 41-58.
4. Behar, R., & Ojeda, M. (2016). El papel de algunas opciones tecnológicas en la educación estadística. *Heurística*, 18, 21-35.
5. Cobo, B. (2003). Significado de las medidas de posición central para los estudiantes de secundaria [Doctoral dissertation, Universidad de Granada].
6. Cobo, B., & Batanero, C. (2004). Significado de la media en los libros de texto de secundaria. *Enseñanza de las Ciencias*, 22(1), 5-18.
7. Espinoza, C., & Sánchez, I. (2014). Aprendizaje basado en problemas para enseñar y aprender estadística y probabilidad. *Paradigma*, 35(1), 103-128.
8. Faustino, A., & Pérez, S. (2013). Utilización de las TIC en la enseñanza de la estadística en la educación superior angolana. *Prisma Social*, (11), 1-31.
9. Gallego, D. E., Quiceno, Y., & Pulgarín, D. (2014). Unidades didácticas: un camino para la transformación de la enseñanza de las ciencias desde un enfoque investigativo. *Tecné, Episteme y Didaxis: TED*, (special issue), 923-930.
10. Garfield, J., & Ben-Zvi, D. (2008). Developing students' statistical reasoning: Connecting research and teaching practice. Springer.

11. Gatica Riquelme, C. (2017). Análisis didáctico para promover ideas fundamentales de estadística relacionadas a medidas de tendencia central [Master's thesis, Pontificia Universidad Católica de Valparaíso].
12. Gea, M. M., Batanero, C., López, M. M., & Contreras, J. M. (2015). Los recursos tecnológicos en la estadística bidimensional en los textos españoles de bachillerato. *Tecné, Episteme y Didaxis: TED*, 38, 113-132.
13. Gutiérrez, A., Agudelo, Ó., & Martínez, J. (2019). El uso de las TIC en los procesos de enseñanza del área de probabilidad y estadística en universitarios. In *Proceedings of the 17th LACCEI International Multi-Conference for Engineering, Education, and Technology*. LACCEI.
14. Hernández, R., Fernández, C., & Baptista, P. (2014). *Metodología de la investigación* (6th ed.). McGraw-Hill.
15. Hernández, S., & Cuevas, J. (2013). Programas informáticos de uso libre y su aplicación en la enseñanza de la estadística. *Revista Investigación Operacional*, 34, 166-174.
16. Juárez, J., & Inzunza, S. (2014). Comprensión y razonamiento de profesores de matemáticas de bachillerato sobre conceptos estadísticos básicos. *Perfiles Educativos*, 36, 14-29.
17. Marzábal, A., Rocha, A., & Toledo, B. (2015). Caracterización del desarrollo profesional de profesores de ciencias - parte 2: proceso de apropiación de un modelo didáctico basado en el ciclo de aprendizaje constructivista. *Educación Química*, 26, 212-223.
18. Mayén, S., Batanero, C., & Díaz, C. (2009). Conflictos semióticos de estudiantes mexicanos en un problema de comparación de datos ordinales. *Revista Latinoamericana de Investigación en Matemática Educativa*, 12(2), 151-178.
19. Ortega, J. M., Valencia, V. E., Becerra, M., & Durán, J. (2020). Matemática y vida cotidiana: experiencia escolar de aprendizaje basado en problemas (ABP). *Números*. *Revista de Didáctica de las Matemáticas*, 104, 103-117.
20. Pollatsek, A., Lima, S., & Well, A. D. (1981). Concept or computation: Students' understanding of the mean. *Educational Studies in Mathematics*, 12(2), 191-204.
21. Strauss, S., & Bichler, E. (1988). The development of children's concepts of the arithmetic average. *Journal for Research in Mathematics Education*, 19(1), 64-80.
22. Watson, J. M., & Moritz, J. B. (2000). The longitudinal development of understanding of average. *Mathematical Thinking and Learning*, 2(1-2), 11-50.
23. Zamora, J. A., Aguilar, E., & Guillén, H. (2021). Educación estadística: tendencias para su enseñanza y aprendizaje en educación secundaria y terciaria. *Revista Educación*, 1-19.