REDIMAT
Journal of Research in Mathematics Education

Hipatia Press
www.hipatiapress.com

Instructions for authors, subscriptions and further details:

http://redimat.hipatiapress.com

Pre-service Teachers' Common Content Knowledge Regarding the Arithmetic Mean

Juan Jesús Ortiz de Haro[1] and Vicenç Font Moll[2]

1) University of Granada, Spain.
2) University of Barcelona, Spain.

Date of publication: October 24th, 2014
Edition period: October 2014-February 2015

To cite this article: Ortiz, J.J. & Font, V. (2014). Pre-service teachers' common content knowledge regarding the arithmetic mean. *REDIMAT*, *Vol 3*(3), 192-219. doi: 10.4471/redimat.2014.51

To link this article: http://dx.doi.org/10.4471/redimat.2014.51

REDIMAT, Vol. 3 No. 3 October 2014 pp. 192-219

Pre-service Teachers' Common Content Knowledge Regarding the Arithmetic Mean

Juan J. Ortiz de Haro
University of Granada

Vicenç Font Moll
University of Barcelona

(*Received: 22 January 2014; Accepted: 30 September 2014; Published: 24 October 2014)*

Abstract

The main goal of this study is to determine the common content knowledge of a group of pre-service primary teachers regarding the arithmetic mean. The cognitive configuration tool proposed by the Onto-semiotic Approach of Cognition and Mathematics Instruction shows that the arithmetic mean can have a variety of meanings, and the application of this tool here revealed significant difficulties related to the students' understanding of this mathematical object and some of its properties. This article concludes with some educational implications for teacher training in the field of statistics.

Keywords: Arithmetic mean, teachers' knowledge, teacher education

2014 Hipatia Press
ISSN: 2014-3621
DOI: 10.4471/redimat.2014.51

REDIMAT, Vol. 3 No. 3 Octubre 2014 pp. 192-219

Conocimiento Común del Contenido de Futuros Profesores sobre la Media Aritmética

Juan J. Ortiz de Haro
University of Granada

Vicenç Font Moll
University of Barcelona

(*Recibido: 22 Enero 2014; Aceptado: 30 Septiembre 2014; Publicado: 24 Octubre 2014)*

Resumen

El objetivo principal de esta investigación es determinar el conocimiento común del contenido de un grupo de futuros profesores de educación primaria sobre la media aritmética. La herramienta configuración cognitiva propuesta por el Enfoque Ontosemiótico de la Cognición e Instrucción Matemática muestra que la media aritmética puede tener una gran variedad de significados, y su aplicación aquí ha revelado importantes dificultades relacionadas con la comprensión de los estudiantes de este objeto matemático y algunas de sus propiedades. Este trabajo concluye con algunas implicaciones educativas para la formación de profesores en el campo de la estadística

Palabras clave: Media aritmética, conocimiento de profesores, formación de profesores

2014 Hipatia Press
ISSN: 2014-3621
DOI: 10.4471/redimat.2014.51

Hipatia Press
www.hipatiapress.com

In Spain, interest in teaching statistics has been strengthened by the Royal Proposition which set out the core curriculum for primary education (Ministerio de Educación y Ciencia, 2006). This legislation includes the following content: gathering and recording data using elementary survey, observation and measurement techniques; different ways of representing information, including the graphical representation of statistical information; the arithmetic mean, mode and range; applications to familiar situations. In this context the present paper highlights the need to start studying statistical phenomena as soon as possible, to make teaching methods more active and exploratory, to foster a greater understanding of statistics as they appear in the media, and to strength both pupils' interest and their ability to evaluate statistical knowledge for the purposes of decision making. These recommendations have already been made in other curricula (e.g. National Council of Teachers of Mathematics, 2000).

According to Stohl (2005) better teaching of statistics requires better training for the teachers involved, because with no specific training they are likely to fall back on what are often erroneous beliefs and intuitions, which would then be passed on to their pupils, as was demonstrated in the study by Ortiz, Mohamed, Batanero, Serrano and Rodríguez (2006). It is important, therefore, to assess the competence of pre-service primary teachers regarding solving elementary statistical problems, especially when it comes to basic concepts such as the arithmetic mean.

In fact, in recent years we have seen a growing interest in research on the knowledge that mathematics teachers need to master in order for their teaching to be effective. However, very few studies have focused on the design of instruments to explore aspects of teachers' mathematical knowledge regarding specific topics. This paper presents the results obtained from the application of a questionnaire that was designed to explore the mathematical knowledge of pre-service primary teachers regarding the arithmetic mean. The teachers were all students at the University of Granada (Spain). The main aim of the study was to determine aspects of their common content knowledge about the arithmetic mean at the start of the course Mathematics and its Teaching. The specific focus was on the mathematical practices employed by students when solving problems concerning the arithmetic mean, as well as on certain mathematical entities which are involved in a way in these mathematical practices, which can be separated or individualized.

This article is organized into six sections. Following this introduction, section two summarizes previous research on the arithmetic mean. Section three briefly describes some of the constructs of the onto-semiotic approach, the theoretical framework used in this research. The specific methodology is described in section four. Section five presents the results of the semiotic analysis of pre-service teachers' answers to the analysed problems. Finally, section six sets out the conclusions and a number of implications regarding the statistical education of teachers

The Arithmetic Mean and Teacher Education

Research About the Notion of Arithmetic Mean

Batanero, Godino and Navas (1997) assessed the knowledge of averages among 273 pre-service primary teachers and found that they had difficulties with the treatment of invalid and atypical values when calculating the arithmetic mean, with the choice of the most appropriate measure of centre for a particular situation, and with the use of averages in comparing distributions. The authors also noted that the aforementioned difficulties remained even after the teachers received specific training. They therefore suggested that instead of algorithm-based teaching, greater emphasis should be placed on the interpretation of results and reflection upon the conditions in which statistical procedures are applied.

Research About the General Concept of Average

Callingham (1997) surveyed 100 pre-service and 36 in-service teachers regarding four problems involving averages. The results showed that teachers provide relatively good solutions to the first three questions (one about calculating the mean from a set of data and the other two involving the comparison of two data sets) using bar charts. However, they had more difficulties with the fourth problem, which required them to determine the weighted mean from a set of data. In this case, only 58% of the teachers responded correctly. Regarding the problems in which the data were presented graphically it seems that the teachers based their answers on numerical arguments rather than solely on the appearance of the data.

Begg and Edwards (1999) studied 22 in-service and 12 pre-service elementary school teachers and found that the majority of them were not familiar with the mathematical definitions of the terms mean, median and mode. In terms of their understanding of these measures, teachers were clearer about the meaning of the mean than they were about the median and mode.

Research by Leavy and O'Loughlin (2006) with 263 pre-service elementary school teachers found that while 57% of them used the mean to compare two sets of data, only 21% gave a correct answer to a problem about the weighted mean, and 88% of them were able to construct a data set that had a predetermined mean. The results also revealed that only 25% of these teachers demonstrated some kind of conceptual understanding of mean, while the remainder showed a procedural understanding. The authors concluded that in order to improve the statistical training of future teachers it was necessary to provide trainees with experiences that would increase their conceptual understanding of the mean, especially the qualitative and quantitative aspects of data representation.

Estrada (2007), in a study to assess the statistical knowledge of 367 pre-service primary teachers, observed that although more than 50% of them produced correct answers to the proposed statistical problems, the results also indicated a lack of knowledge of basic statistical concepts such as the mean, median and mode, as well as mistakes concerning the average; for example: not being aware of the effect on the mean of atypical values, not being skilled in inverting the algorithm of the mean, and confusing mean, median and mode. The findings indicate a need to improve the statistical training of pre-service teachers.

García Cruz and Garrett' (2008) contributions on 130 secondary education pupils and 97 university students, of whom 31 were studying to be primary maths teachers, showed that participants displayed different types of reasoning about the arithmetic mean, and that their answers to the proposed problems could be linked to the five levels of understanding described in the SOLO (Structure of the observed learning outcome)[1] taxonomy of Biggs and Collis (1991). A further finding was that there were no significant differences between university students and secondary education pupils in terms of the observed levels of interpretation. These results suggest that in order to address the difficulties and errors that occur when learning the arithmetic mean it is necessary to work with real-life problems and to encourage students to be more proactive in developing their knowledge.

In a study to assess the statistical and pedagogical knowledge of 55 pre-service elementary school teachers, Godino, Batanero, Roa and Wilhelmi (2008) observed that although many of them had a good idea of equiprobability, they undervalued variability. Specifically, only 29% of participants made use of the mean to compare the results obtained in real and simulated coin-tossing sequences. The authors concluded that significant changes needed to be made to initial teacher training in order to improve the statistical knowledge of pre-service teachers.

As regards research with university students, Pollatsek, Lima and Well (1981) described mistakes in calculating simple and weighted averages from a frequency table, while Mevarech (1983) reported difficulties in applying certain properties of the mean.

Research About the Mean, Median, and Mode

Groth and Bergner (2006) used the SOLO taxonomy of Biggs and Collis (1991) to classify into four categories the understanding shown by 46 pre-service teachers about mean, median and mode. Eight pre-service teachers were assigned to the unistructural/concrete symbolic level of thinking, as their responses only involved definitions of the similarities or differences between mean, median and mode. Twenty-one pre-service teachers were regarded as showing the multistructural/concrete symbolic level of thinking, as their answers suggested that these measures of centerness represent a mathematical object rather than just a procedure. The relational/concrete symbolic level of thinking was exhibited by 13 pre-service teachers whose responses indicated that these measures represent a characteristic value of the data set. Finally, three pre-service teachers reached the extended abstract level of thinking, since their answers included discussion of which measure of centre is more representative of a given data set. According to the authors, the small number of pre-service teachers who reach the highest level of thinking could be due to certain limitations in the design of the proposed tasks.

Jacobbe (2008) carried out a case study of the understanding of average shown by three pre-service elementary school teachers. The teachers were presented with three distributions (one skewed to the left, one skewed to the right and a third that was normal) and were asked to indicate which would have the smallest to largest values of the mean, median and mode. The results showed that although some of the teachers found difficult applying the algorithm of the mean in certain contexts they were able to use the shape of the distribution to determine when a given set of data would have a greater mean, median and mode than another one would. The authors conclude that although future teachers do have certain skills when it comes to solving statistical problems they would nonetheless benefit from more formal training in this regard.

The above findings highlight the need for further research in this area, since in addition there are very few studies focused on teachers' statistical content knowledge. Knowing that teachers' mathematical knowledge have an effect on their pupils' achievement (Ball, 1990) it is reasonable to assume that the observed lacks in the common content knowledge of pre-service primary teachers could prevent them from appropriately managing their pupils' mathematical knowledge about the arithmetic mean. Such lacks justify the need for specific training designed to develop the common and specialized content knowledge of pre-service primary teachers, and which is able to consider the embedded complexity in the different meanings of a mathematical object (in this case, the arithmetic mean).

Theoretical Framework

The didactic and mathematical knowledge required to teach mathematics is an issue of considerable interest. Noteworthy contributions in this regard include the considerations and recommendations of Shulman (1986) and the studies by Ball (2000), Ball, Lubiensky and Mewborn (2001), and Hill, Ball and Schilling (2008). All of them have characterized different components of knowledge that teachers must have in order to teach effectively and to facilitate their pupils' learning process. However, as Godino (2009) points out, the models of mathematical knowledge for teaching, which have emerged from research on mathematical education, are based on very broad categories. It would therefore be useful to have models that enable a more detailed analysis of each type of knowledge that is brought into play when teaching mathematics. Furthermore, a deeper understanding of this knowledge framework requires a focus on specific topics, for example, the knowledge that pre-service primary teachers need in order to teach the arithmetic mean.

Our goal here is therefore to evaluate partial aspects of the mathematical knowledge for teaching (MKT) in a group of pre-service primary teachers, drawing on the arithmetic mean. According to the model of Ball and colleagues (Ball, Lubienski, & Mewborn, 2001; Hill, Ball, & Schilling, 2008), MKT includes six types of knowledge: common content knowledge, specialized content knowledge, expanded content knowledge, knowledge of curriculum, knowledge of content and students, and knowledge of content and teaching. Our concern here is with common content knowledge, i.e. the mathematical knowledge that is typically known by competent adults and which teachers are responsible for developing in their pupils.

In order to identify the common content knowledge that is required to teach the arithmetic mean we use some of the theoretical constructs proposed by the onto-semiotic approach (OSA) to cognition and mathematics teaching (Godino, Batanero & Font, 2007; Font, Godino & Gallardo, 2013). In some studies conducted within the framework of the OSA (for example, Malaspina & Font, 2010; Godino, Font, Wilhelmi & Lurduy, 2011) with the aim of examining students' mathematical outputs, the research process begins by analysing mathematical practices and then moves towards consider the mathematical objects and processes that are activated within these practices. This article focuses on the mathematical objects that are activated during the practices involved.

Drawing on the mathematical objects that are activated in performing and evaluating the practice that enables a problem to be solved (for example, proposing and solving an arithmetic mean problem), what we can see is the use of representations (verbal, iconic, symbolic, etc.). These representations are the ostensive part of a series of concepts/definitions, propositions and procedures that are involved in the development of arguments which are used to decide whether or not the practice carried out is satisfactory. Thus, when a student performs and evaluates a mathematical practice, s/he activates a cluster of objects formed by problem situations, representations, definitions, propositions, procedures and arguments, which in the OSA is referred to as the cognitive configuration of primary mathematical objects. That is, these six types of primary entities will form 'configurations', defined as the network of objects involved and emerging from the systems of practices and the relationships established between them. These configurations can be epistemic (networks of institutional objects) or cognitive (network of personal objects).

The theoretical framework OSA (Godino, Batanero & Font, 2007) is based on elements taken from diverse disciplines such as anthropology, semiotics and ecology. It also assumes complementary elements from different theoretical models used in mathematics education to develop a unified approach to didactic phenomena that takes into account their epistemological, cognitive, socio cultural and instructional dimensions. 'Mathematical practice' in this approach is defined as any action or manifestation (linguistic or otherwise) carried out by somebody to solve mathematical problems, to communicate the solution to other people, so as to validate and generalize that solution to other contexts and problems. In the study of mathematics, more than a specific practice to solve a particular problem, we are interested in the 'systems of practices' (operative and discursive) carried out by the people involved in certain types of problem-situations. The 'system of practices' that a person carries out (personal meaning), or are shared within an institution (institutional meaning), to solve a type of problem-situations. The system of practices and the configurations are the basic theoretical tools to describe mathematical knowledge, in its double personal and institutional facets. Therefore, this kind of analysis is well suited to the identification of common content knowledge (use of procedures, representations, definitions, etc.).

Methodology

Participants

Participants were 40 pre-service primary teachers in the first year of their training at the University of Granada (Spain) and who were currently studying the course Mathematics and Its Teaching. The mathematical content of this course is the same as the one used in the curriculum for elementary schools in Spain and it covers four thematic blocks: Numbers and operations; Measurement: estimation and calculation of magnitudes; Geometry; and Dealing with information, chance and probability (Ministerio de Educación, 2006).

Generally speaking, these students have limited mathematical training. Indeed, before starting university the only statistical training they will have had is the one received at secondary school, which means that they should have studied basic statistical concepts such as data, statistical variables, frequency distributions, graphs, and measures of centerness and dispersion.

The Questionnaire

As demonstrated in the textbook analysis undertaken by Cobo and Batanero (2004) the meaning of the mathematical object 'arithmetic mean' is complex. In order to ensure that the instrument used in the present study was appropriate for pre-service teachers in terms of its item content, we decided rather than designing a representative sample of problems that can be solved with the arithmetic mean using a historical, epistemological and didactic analysis, we opted for an indirect method, namely the use of an existing questionnaire that was developed for this kind of research, thereby ensuring that the proposed problems were representative for the students concerned. The questionnaire was created by Batanero (2000) and is well suited to the assessment of pre-service teachers' common content knowledge because each of its items activates a different meaning of the arithmetic mean. The questionnaire includes five problems: the first refers to the estimation of an unknown quantity in the presence of measurement errors; the second one is about obtaining equal shares in order to achieve a uniformed distribution; the third one consists in finding an element that represents a set of given values whose distribution is approximately symmetrically, framed by a context being one of data comparison; the fourth one concerns a situation in which the task is to determine the value that is most likely to be obtained when selecting a random element from a population; and the fifth one is about the weighted arithmetic mean. The questionnaire was administered to the pre-service teachers prior to the start of the course Mathematics and Its Teaching.

In this article we discuss the problems 1 and 3. The first problem (Figure 1) is a particular example of a set of problems in which an unknown quantity has to be estimated in the context of measurement errors. In this case, where no atypical values exist, the best estimate would be the arithmetic mean of the weights obtained by the eight students, which in this case is 6.15 grams.

Eight students from a class weigh a small object using the same instrument, with the following values in grams being obtained: 6.2, 6.0, 6.0, 6.3, 6.1, 6.23, 6.15 and 6.2. What would be the best estimate of the object's real weight?

Figure 1. Problem: Estimation of an unknown quantity.

The second problem (Figure 2) is a particular example from a set of problems in which a representative element needs to be chosen for a set of given values, whose distribution is approximately symmetrical, and in the context of data comparison. In order to test the effect of the training, the average of the height jumped would be calculated prior to and after the training. This method would demonstrate if the training was effective as the first mean is 115.6 cm and the second is 120.4 cm. To represent a set of values the arithmetic mean is chosen due to its properties as 'centre of gravity' of a sample (or population) and its sense of "central location". In the case of an asymmetric distribution it would not be an appropriate choice and other measures such as the median or the mode should be used.

The following values were obtained when measuring the jump height (in cm) attained by a group of schoolchildren prior to and after training. Do you think that the training is effective?

Height jumped in cm

Pupil	Ana	Bea	Carol	Diana	Elena	Fanny	Gia	Hilda	Ines	Juana
Prior to training	115	112	107	119	115	138	126	105	104	115
After training	128	115	106	128	122	145	132	109	102	117

Figure 2. Problem. Representative element of a set of values.

Analysis of the Pre-service Teachers' Answers

The method used to study the common content knowledge of pre-service primary teachers regarding the arithmetic mean was based on the work of Malaspina and Font (2010). Thus, the first step involved examining the mathematical practices carried out by students, which in this case corresponded to reading the problem and producing a written answer. Secondly, the method considers the cognitive configuration of primary mathematical objects (Font, Godino & Gallardo, 2013) that are employed by students in their answers.

Results

This section presents the different categories of pre-service teachers' common content knowledge regarding the arithmetic mean, drawing on the semiotic analysis of their answers to each of the items. The analysis is based on the cognitive configuration of primary mathematical objects that were activated by each student when solving each of the problems. For each category a distinction is made between those students who provide an argument and those who does not, and between those who answer correctly and those who does not.

Analysis of the Problem *Estimation of an Unknown Quantity*

By analysing the cognitive configuration of primary mathematical objects in the pre-service teachers' solutions to this problem we obtained four categories of common content knowledge regarding the arithmetic mean. Category 1 corresponds to 20 pre-service teachers who performed all the necessary calculations to obtain the arithmetic mean, but who neither mentioned the term 'mean' explicitly nor justified their answer; in addition, eight of them answered incorrectly. Category 2 comprises seven pre-service teachers who carried out all the necessary calculations to obtain the arithmetic mean and explicitly used the term 'mean'; however, only one of them provided an argument and answered correctly, while of the six who didn't justify their approach only two answered correctly. Category 3 includes nine pre-service teachers who considered that the best estimate is the value within a certain range or that which is most often repeated. Of these, four teachers provided an argument but gave an incorrect answer, while the remainder failed to provide an argument and also answered incorrectly. Finally, Category 4 corresponds to two pre-service teachers 4 who answered with other arguments.

The types and frequencies of the observed cognitive configurations are set out in Table 1, which also indicates whether or not the answers given were correct and whether an argument or justification was provided. It can be seen that 27 pre-service teachers (67.5%) used the mean as the best estimate for the real weight of the object.

Table 1

Frequency of the types of common content knowledge regarding the arithmetic mean observed in relation to the problem Estimation of an unknown quantity

Type	Mathematical practice	Correct		Incorrect		No answer	Total
		A*	No A*	A*	No A*		
Uses mean implicitly	Use the arithmetic mean but who don't explicitly mention the term 'mean'.	0	12	0	8		20
Uses mean explicitly	Use the arithmetic mean and explicitly used the term 'mean'.	1	2	0	4		7
Range or repeated value	Consider that the best estimate is over a certain range or the value that is repeated most often.	0	0	4	5		9
Other	Not possible to know, or simply give a possible definition of "best estimate".	0	0	2	0		2
No answer							2
Total		1	14	6	17	2	40

A* means "Argumentation."

No A* means "No argumentation."

Fifteen pre-service teachers (37.5%) obtained the correct answer, of whom 12 used the mean implicitly but failed to provide an argument; the other three used the mean explicitly but only one of them justified the answer. Twenty-three pre-service teachers (57.5%) answered incorrectly, making the following errors: 12, despite using the arithmetic mean, made errors of calculation or when counting the number of cases, and they failed to provide an argument; nine considered that the best value is found within a certain range or is the value that repeats the most (mode), while a further two used another incorrect cognitive configuration. The six pre-service teachers who gave arguments used an incorrect justification. Finally, there were two students (5%) who didn't answer. Each of the categories are described below.

Category 1 (twenty pre-service teachers, 50%). Students who used the mean implicitly.

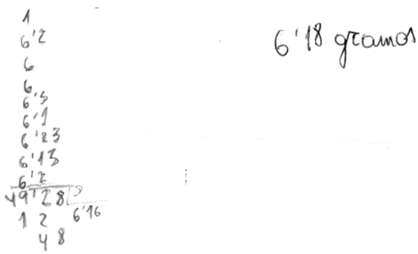
Five pre-service teachers used the arithmetic mean and the concept of estimation (or approximation) implicitly in their practice, but they did not provide an argument and their answer was incorrect. Table 2 shows an example of this kind of cognitive configuration of primary mathematical objects.

Three students used this cognitive configuration but when answering made errors of addition (1 student) or division (2 students). The other two students missed some cases (they only counted 7 values) but carried out the operations correctly.

Five pre-service teachers used the arithmetic mean and the concept of estimation (or approximation) implicitly and without argumentation, but nonetheless produced a correct answer (they calculated correctly and counted all the cases).

Five pre-service teachers used the arithmetic mean and the concept of estimation (or approximation) implicitly and without argumentation, but nonetheless produced a correct answer (they calculated correctly and counted all the cases).

Table 2
Cognitive configuration of primary mathematical objects for Student 1

Situation-problem	Problem. Estimation of an unknown quantity (A small object is weighed…)
Language 	Verbal, related to the context: grams. Symbolic: whole numbers and decimals (6, 6.2, etc.). Addition (vertical) and division bracket.
Concepts-definitions	Arithmetic mean (implicit, $\bar{x} = \bar{x}_i/n = (x_1 + x_2 + \cdots + x_n)/n$) Addition, division Estimation (implicit)
Propositions	The mean (implicit) is the best estimate of the real weight
Procedures	Addition, division Calculation of the mean
Arguments	None stated

Seven pre-service teachers using the arithmetic mean implicitly, and the concept of estimation (or approximation) explicitly, failed to provide an argument; but produced a correct answer. Another three pre-service teachers used a similar cognitive configuration but their answers were incorrect due to calculation errors. It is significant that out of the twenty students in this category, six of them did the calculations incorrectly.

Category 2 (seven pre-service teachers, 17.5%). Students who used the arithmetic mean explicitly.

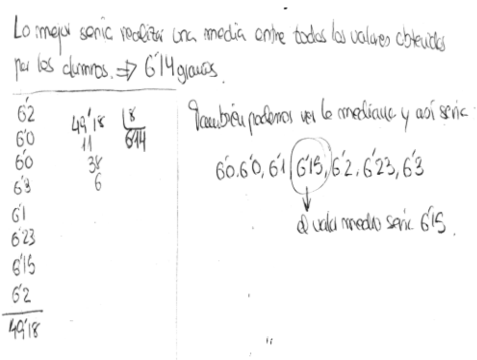
Two pre-service teachers used the arithmetic mean explicitly and the concept of estimation (or approximation) implicitly, but they did not provide an argument and their answers were incorrect. Their cognitive configurations are similar because nobody added or divided correctly. In addition one of them didn't count all the cases whereas the other did.

A similar cognitive configuration is shown by the pre-service teacher (Student 24) who used the arithmetic mean explicitly and the concept of estimation (or approximation) implicitly, failed to provide an argument but gave a correct answer (this teacher did not make the same calculation errors and counted all the cases). Only one pre-service teacher used the arithmetic mean explicitly and the concept of estimation (or approximation) implicitly, as well as providing an argument and answering correctly. Table 3 shows an example of this cognitive configuration of primary mathematical objects.

There is one pre-service teacher (Student 6) who used the arithmetic mean and the concept of estimation (or approximation) explicitly but without argumentation, and whose answer was correct: "the best estimate would be 6.15 grams". This student is one of the few who used the more formal representation of arithmetic mean, and although the plus sign is missing, he calculated correctly.

Table 3
Cognitive configuration of primary mathematical objects for Student 7

Situation-problem	*Problem. Estimation of an unknown quantity (A small object is weighed…)*
Language	Verbal, related to the context: grams. Symbolic: natural and decimal numbers (8, 6,14), vertical addition, division, mean, median, mean value



The best estimate would be the mean of all the values obtained by the pupils: → 6.14 grams
You could also determine the median like this…
the middle value would be 6.15

Concepts – definitions	Arithmetic mean $\bar{x} = \bar{x}_i/n = (x_1 + x_2 + \cdots + x_n)/n)$ Vertical addition Division bracket Median
Propositions	The best estimate is the mean 6.14 Another estimate is the median
Procedures	Addition and division Calculation of the mean Calculation of the median
Arguments	Conclusion: The mean is the best estimate: 6.14 Argument 1: The mean and the median are estimates Argument 2: The mean is the best estimate Argument 3. The mean is 6.14 Argument 4: The median is 6.15

The cognitive configuration of the two pre-service teachers who used the arithmetic mean and the concept of estimation (or approximation) explicitly, but without argumentation and whose answers are incorrect, are similar to the previous case except that one of them made an error when adding and dividing (Student 37) and the other (Student 5) when adding and by missing one of the cases.

Of the seven students who used the mean explicitly only three got the correct answer. The rest made errors in their calculations (2 students) or when counting the number of cases (2 students).

Category 3 (nine pre-service teachers, 22.5%): Students who used the concept of range or mode.

There are two pre-service teachers who used the idea of an intermediate value within a range and the concept of estimation (or approximation) explicitly in their practice, as well as providing an argument; however, their answer was incorrect: “Between 6.1 and 6.15. Estimate: 6.125” (Student 4). These two pre-service teachers (Students 4 and 11) make the same error when counting the number of cases, as they both count the value ‘6.2’ three times.

The cognitive configuration of one of the two pre-service teachers (Student 34) who used the idea of an intermediate value within a range and the concept of estimation (or approximation) explicitly in their practice, but without argumentation and whose answer was incorrect, is similar to the above example except that it lacks an argument and the concept of the midpoint of the interval estimation. The other student (Student 36) simply answered "“6.1” without writing anything else.

There are two pre-service teachers who used the idea of an intermediate value within a range and the concept of estimation (or approximation) implicitly in their practice, but who failed to provide an argument and also gave incorrect answers. The first one (Student 33) answered that “the real weight would be between 6.0 and 6.2”, whilst the second (Student 35) gave a similar answer but without making a verbal reference to the context (he didn’t mention the weight).

Two other pre-service teachers used implicitly the idea of an intermediate value within a range (and which is most often repeated) and the concept of estimation (or approximation), but despite providing an argument their answer was incorrect. The first one (Student 2) answered “6.2 because it is the intermediate value and it is also the one that repeats the most”, while the other one (Student 3) answered in a similar manner, but then stated that the mean is the value that repeats the most. Finally, one pre-service teacher (Student 32) used a similar cognitive configuration to the two previous students, but without any argumentation, answering “6”.

Category 4 (two pre-service teachers, 5%): Students who didn’t make any calculations and gave other arguments.

There are two pre-service teachers in this category, one who considers that “it is unknown as different results are obtained by all”, while the other answers “The value that is closest to the real value of the object”.

Analysis of the problem *Representative element of a set of data*

By analysing the cognitive configuration of primary mathematical objects in the pre-service teachers’ solutions to this problem we obtained three categories of common content knowledge regarding the arithmetic mean. Category 1 comprised eight pre-service teachers who used the arithmetic mean, provided an argument and gave correct answers (with one exception). Category 2 corresponds to 27 pre-service teachers who analysed the cases separately and whose arguments and answers were incorrect. Category 3 includes two pre-service teachers who did not use any calculation and whose arguments and answers are also incorrect. Based on the procedure used, Category 2 can be sub-divided into three groups: Group 1 comprises seven pre-service teachers who calculated the percentage of cases that had improved after training; Group 2 includes 16 pre-service teachers who calculated the number of cases that had improved after training; and Group 3 corresponds to four pre-service teachers who used the total deviation between the values obtained prior to and after training for each case (one of these teachers made errors of calculation).

The type and frequency of each kind of common content knowledge regarding the arithmetic mean is shown in Table 4, which also indicates whether or not the answers given were correct and whether an argument or justification was provided. It can be seen that only eight pre-service teachers (20%) used the arithmetic mean, despite it being the most efficient way of comparing the two distributions.

Table 4
Frequency of the types of common content knowledge regarding the arithmetic mean observed in relation to the problem Representative element of a set of data

Type	Mathematical practice	Correct		Incorrect		No answer	Total
		A*	No A*	A*	No A*		
Mean	Use the arithmetic mean to compare the two proposed distributions.	7	0	1	0		8
Separated cases	Use separate cases from each of the two proposed distributions to compare them.	0	0	27	0		27
Other	Do not use any calculations to compare the two distributions	0	0	2	0		2
No answer							3
Total		7	0	30	0	3	40

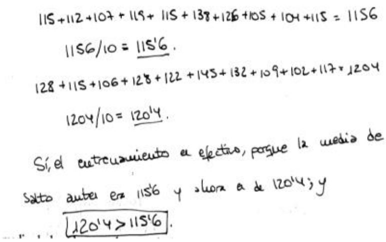
A* means “Argumentation.”
No A* means “No argumentation.”

Only seven pre-service teachers (17.5%) gave the correct answer and provided the correct argument. Thirty teachers (75%) answered incorrectly, having made the following errors: 27, who analysed the cases separately, based their arguments on procedures that do not always provide the correct solution; one did use the mean but misinterpreted it, answering that the training was not effective; and two did not use any calculations and argued erroneously. Three pre-service teachers (7.5%) did not answer at all. Each of the categories are described below.

Category 1 (eight pre-service teachers, 20%): Students who use the arithmetic mean to compare the two distributions

There are seven students who used the arithmetic mean to compare the two proposed distributions, as well as providing an argument and giving a correct answer. Table 5 shows an example of this cognitive configuration of primary mathematical objects.

Table 5
Cognitive configuration of primary mathematical objects for Student 24

Situation-problem	Problem. Representative element of a set of data (On measuring the height…)
Language  The training is effective because the mean jump height before training was 115.6 and afterwards it was 120.4; and…120.4>115.6	Verbal, related to the context: training, jump. Symbolic: whole and decimal numbers, addition (horizontal), division, fraction
Concepts – definitions	Arithmetic mean: $\bar{x} = x_i/n = (x_1 + x_2 + \cdots + x_n)/n)$ Comparison of averages Addition, Division

Propositions	The training is effective
Procedures	Calculation of the arithmetic mean Comparison of the two means
Arguments	Conclusion: The training is effective Argument: because the mean jump height was previously 115.6 and now it is 120.4; and 120.4 >115.6

Four of these pre-service teachers (Students 22, 25, 29 and 31) used the same cognitive configuration. Two other students (namely 21 and 38) stated that "in the majority of cases it is effective", despite the fact that they calculated the mean and checked that the mean of "height jumped after the training" was higher. This means that although they used the arithmetic mean to compare the two distributions they do not have a very clear concept, as they continue to base their arguments on the behaviour of the separate cases.

One pre-service teacher (Student 6) used the arithmetic mean to compare the two proposed distributions and also provided an argument, but nonetheless gave an incorrect answer. The cognitive configuration of this teacher would be similar to the previous one, as he calculated the means of each distribution but then concluded that the training "has not been effective as the mean prior to training was = 115.6 and after it was = 120.4".

Category 2 (27 pre-service teachers, 67.5%): Students who analysed the cases separately in order to compare the two distributions.

The answers given by these 27 pre-service teachers were incorrect: despite stating that the training was effective, 24 of them based their arguments on procedures that do not always provide the correct solution; the other three considered that the training was not effective. Based on the procedures used, this second category can be sub-divided into three groups:

Group 1: Students who analysed separate cases, calculating the percentage of pupils who jumped higher after training. An example of this cognitive configuration, corresponding to one such teacher who provided an argument but gave an incorrect answer: "The training is effective because 80% of the children jumped higher than before, compared to 20% whose jump height was less after training" (Student 36).

Group 2: Students who analysed separate cases, calculating the number of pupils who jumped higher after training and providing an argument, but whose answers were incorrect. This group comprised 16 pre-service teachers whose cognitive configuration are similar to the one above, except they count the number of cases rather than the percentages. Two of these teachers state that the training is not effective: 1) "because some pupils have done well whereas others have got worse" (Student 11); or 2) "because with the exception of a couple of cases the pupils jump higher after training, which means you don't have a stable measurement" (Student 20).

Group 3: Students who analysed separate cases, calculating the total deviation between the values obtained prior to and after the training for each of the pupils. This group includes four pre-service teachers: three argued that the training is effective but gave an incorrect explanation (Students 23, 27 and 40), with the latter making a calculation error in one subtraction; the fourth student considered that the training is not effective as "there are several heights and there is no variation within the same unit" (Student 28).

Category 3 (two pre-service teachers, 5%): Students who did not use any calculations to compare the two distributions.

There are two pre-service teachers in this category: one who considers that the training is effective, "as with training the body becomes ready to improve its performance" (Student 17); and another one who argues that "it is effective in some cases but not in others, because they don't do any better than they did at the outset" (Student 32).

Discussion and conclusions

The theoretical categories provided by the OSA have enabled us to conduct a detailed analysis of pre-service teachers' output, thereby capturing the complexity of the common content knowledge (representations, concepts, properties, etc.) that was activated when they were asked to solve problems involving the arithmetic mean. Indeed, the cognitive configurations of primary mathematical objects that were derived from this analysis reveal that the trainee teachers used a wide range of common content knowledge when solving these statistical problems. A total of seven categories were identified, four in relation to problem 1 (Estimation of an unknown quantity) and three in relation to problem 2 (Representative element in a set of data). Although the tasks and the questionnaire used were taken from previous research rather than designed specifically for this study, the onto-semiotic analysis (Malaspina & Font, 2010) of the pre-service teachers' answers to the problems set is much more detailed than previously published analyses.

With regard to problem 1, a third of pre-service primary teachers did not recognize the arithmetic mean as the best estimate of the real weight of the object, either failing to give an answer or using other incorrect cognitive configurations, such that they solved the problem incorrectly. Furthermore, a high percentage of them (77.5%) provided no justification for their answer. The percentage of correct answers (37.5%) is lower than that obtained by Estrada (2007) and Batanero et al. (1997), although in these two cases the problem used was multiple choice and included an atypical value.

With respect to problem 2, 80% of the pre-service teachers did not take the mean into account in order to compare the two distributions, this percentage being much higher than was reported by Godino et al. (2008) with an equivalent problem. Furthermore, they made use of other incorrect cognitive configurations, or even failed to answer, and some of them made calculation errors. A high proportion of them (75%) did, however, provide an argument, albeit an incorrect one, for their answer. The percentage of correct answers was only 17.5%, well below the figure obtained by Leavy and O'Loughlin (2006), Estrada (2007) and Batanero et al. (1997), although in the latter two cases the problem was multiple choice and the information was presented in graphical form.

The observed errors were similar to those described by Estrada (2007), Leavy and O'Loughlin (2006) and Batanero et al. (1997). The most noteworthy errors were, with respect to the first problem, the idea that the best estimate lies within a certain range or is the mode, and, with respect to the second problem, the method based on analysing the cases separately. These errors could be due to a lack of statistical training and the fact that, in general, the pre-service teachers had received an education based on decontextualized situations in which it was only necessary to apply the algorithm of the arithmetic mean. At all events, they also made calculation errors, similar to those described by Pollatsek et al. (1981) with university students.

As regards the mathematical understanding of the arithmetic mean shown by these pre-service primary teachers, it can be concluded that those of them who did apply this concept to the two analysed problems show greater understanding than do those who neither used or related the concept to these problems. This is because the former have established a semiotic function between the problem and the mathematical object known as the arithmetic mean, thereby enabling them to recognize the proposed problem as an extra-mathematical situation which falls under the domain of the aforementioned object. This is a noteworthy finding when one considers that it is never made explicit in either problem that the mean should be used. It is also consistent with the conclusion of Leavy and O'Loughlin (2006), who consider that an indicator of the conceptual understanding of the arithmetic mean is the ability to recognize situations in which the mean is an appropriate measure.

Another important finding of the present study is the notable difference in the percentage of students who chose the mean as the most useful measure to solve the first problem (67.5%) compared to those who opted to use the mean for the second problem (20%). This indicates that a given student's level of conceptual understanding is not fixed but, rather, depends on the context. These results are consistent with those obtained by Jacobbe (2007, 2008), who found that pre-service teachers have difficulties when applying the algorithm of the mean in different contexts.

Our findings also reflect those of other authors (e.g. García Cruz et al., 2008) who have similarly found that many of these errors remain through to university, thereby highlighting the need to improve the basic statistical education of pre-service primary teachers, who will otherwise find it problematic to teach a topic which they themselves find so difficult.

These insufficiencies in teachers' understanding justify the need for specific training designed to develop their common and specialized content knowledge. Such training should consider the complexity embedded in the different meanings of a mathematical object (in this case, the arithmetic mean), since teachers' specialized knowledge must enable them to determine whether the range of problems they set their pupils are representative of the overall meaning of a given mathematical object.

Notes

[1] It is a model for the study of students' development throughout the learning process, based on a set of tasks limited to a given domain. Five levels can be differentiated: prestructural, unistructural, multistructural, transitional and the relational.

References

Ball, D.L (1990). The mathematical understanding that prospective teachers bring to teacher education. *Elementary School Journal, 90*(4), 449-466.

Ball, D.L. (2000). Bridging practices: Intertwining content and pedagogy in teaching and learning to teach. *Journal of Teacher Education, 51*(3), 241-247. doi: 10.1177/0022487100051003013

Ball, D.L., Lubienski, S.T., & Mewborn, D.S. (2001). Research on teaching mathematics: The unsolved problem of teachers' mathematical knowledge. In V. Richardson (Ed.), *Handbook of research on teaching* (4th ed., pp. 433-456). Washington, DC: American Educational Research Association.

Batanero, C. (2000). Significado y comprensión de las medidas de posición central [Meaning and understanding of measures of centre]. *Uno, 25*, 41-58.

Batanero, C., Godino, J., & Navas, F. (1997). Concepciones de maestros de primaria en formación sobre los promedios [Concepts of averages among trainee primary school teachers]. In H. Salmerón (Ed.), *VII Jornadas LOGSE: Evaluación Educativa* (pp. 304-310). Granada, Spain: University of Granada.

Begg, A., & Edwards, R. (1999). Teachers' ideas about teaching statistics. *Proceedings of the 1999 Combined Conference of the Australian Association for Research in Education and the New Zealand Association for Research in Education*. Melbourne: Australian Association for Research in Education. Retrieved from: http://www.aare.edu.au/99pap/

Biggs, J. B., & Collis, K. F. (1991). Multimodal Learning and the Quality of Intelligent Behavior. In H. A. H. Rowe (Ed.), *Intelligence: Reconceptualization and measurement* (pp. 57-76). Hillsdale, NJ: Erlbaum.

Callingham, R. (1997). Teachers' multimodal functioning in relation to the concept of average. *Mathematics Education Research Journal, 9*(2), 205–224. doi: 10.1007/BF03217311

Cobo, B., & Batanero, C. (2004). Significados de la media en los libros de texto de secundaria [Meanings of the mean in secondary education textbooks]. *Enseñanza de las Ciencias, 22*(1), 5-18.

Estrada, A. (2007). Evaluación del conocimiento estadístico en la formación inicial del profesorado [Evaluating statistical knowledge in initial teacher training]. *Uno, 45*, 78-97.

Font, V., Godino, J. D. & Gallardo, J. (2013). The emergence of objects from mathematical practices. *Educational Studies in Mathematics, 82*, 97 - 124. doi: 10.1007/s10649-012-9411-0

García Cruz, J. A. & Garrett, A. J. (2008). Understanding the Arithmetic Mean: A Study with Secondary and University Students. *Journal of the Korea Society of Mathematical Education Series D: Research in Mathematical Education, 12*(1), 49-66.

Godino, J.D. (2009). Categorías de análisis de los conocimientos del profesor de matemáticas [Categories for analysing the knowledge of mathematics teachers]. *Unión, 20,* 13-31.

Godino, J. D., Batanero, C., & Font, V. (2007). The Onto-Semiotic Approach to Research in Mathematics Education. *ZDM-The International Journal on Mathematics Education, 39*(1-2), 127-135. doi: 10.1007/s11858-006-0004-1

Godino, J. D., Batanero, C., Roa, R., & Wilhelmi, M. (2008). Assessing and developing pedagogical content and statistical knowledge of primary school teachers trough project work. In: C. Batanero, G. Burrill, C. Reading & A. Rossman (Eds.), *Proceedings of the ICMI Study 18 and 2008 IASE Round Table Conference*. Monterrey, Mexico.

Godino, J. D., Font, V., Wilhelmi, M. R., & Lurduy, O. (2011). Why is the learning of elementary arithmetic concepts difficult? *Educational Studies in Mathematics*, 77 (2), 247-265. doi: 10.1007/s10649-010-9278-x

Groth, R. E., & Bergner, J. A. (2006). Preservice elementary teachers' conceptual and procedural knowledge of mean, median, and mode. *Mathematical Thinking and Learning*, *8*, 37–63. doi:10.1207/s15327833mtl0801_3

Hill, H. C., Ball, D. L., & Schlling, S. G. (2008). Unpacking pedagogical content knowledge of students. *Journal for Research in Mathematics Education, 39,* 372-400.

Jacobbe, T. (2008). Elementary school teachers' understanding of the mean and median. In C. Batanero, G. Burrill, C. Reading & A. Rossman (Eds.), *Proceedings of the ICMI Study 18 and 2008 IASE Round Table Conference*. Monterrey, México.

Leavy, A., & O'Loughlin, N. (2006). Preservice teachers understanding of the mean moving beyond the arithmetic average. *Journal of Mathematics Teacher Education, 9,* 53-90. doi: 10.1007/s10857-006-9003-y

Malaspina, U. & Font, V. (2010). The role of intuition in the solving of optimization problems. *Educational Studies in Mathematics*, *75*(1), 107-130. doi: 10.1007/s10649-010-9243-8

Mevarech, Z. R. (1983). A deep structure model of students' statistical misconceptions. *Educational Studies in Mathematics*, *14*, 415-429. doi: 10.1007/BF00368237

Ministerio de Educación y Ciencia (2006). *Real Decreto 1513/2006, de 7 de diciembre, por el que se establecen las enseñanzas mínimas de la Educación Primaria* [Royal Decree 1513/2006, of 7 December, setting out the core curriculum for primary education in Spain].

National Council of Teachers of Mathematics (2000). *Principles and Standards for School Mathematics*. Reston, VA: Author.

Ortiz, J. J., Mohamed, N., Batanero, C., Serrano, L., & Rodríguez, J. (2006). Comparación de probabilidades en maestros en formación [Trainee teachers and the comparison of probabilities]. In P. Bolea, M. J. González & M. Moreno (Eds.), *Actas del X Simposio de la Sociedad Española de Investigación en Educación Matemática* (pp. 268-276). Huesca, Spain: SEIEM.

Pollatsek, A., Lima, S., & Well, A. D. (1981). Concept or computation: Students' understanding of the mean. *Educational Studies in Mathematics*, *12*, 191-204. doi: 10.1007/BF00305621

Shulman, L.S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher, 15(2),* 4-14.
Stohl, H. (2005). Probability in teacher education and development. In: G. A. Jones (Ed.). *Exploring probability in schools. Challenges for teaching and learning* (pp. 345-366). New York: Springer.

Juan Jesús Ortiz de Haro is professor of mathematics education, in the College of Education, University of Granada, Spain.

Vicenç Font Moll is professor of Mathematics Education, in the Department of Sciences and Mathematics Education, University of Barcelona, Spain.

Contact Address: Direct correspondence concerning this article, should be addressed to the author. Postal address: Campus Universitario de Melilla – Universidad de Granada; Calle Santander, 1; 52005 Melilla (Spain). **Email**: jortiz@ugr.es