ASSESSING SECONDARY SCHOOL STUDENT'S UNDERSTANDING OF AVERAGES

Carmen Batanero, Belén Cobo Merino and Carmen Díaz
University of Granada, Spain, batanero@ugr.es

In this paper we describe results from a questionnaire given to a sample of secondary school students to assess the personal meaning they attribute to mean median and mode. The questionnaire is made up of 9 open- ended tasks (26 sub items) where students provide detailed reasoning to their responses. Comparative results from two samples of 14 year-olds (n=168) and 16 year-olds (n=144) and multivariate analysis for the combined sample will be analysed.

1 INTRODUCTION

The teaching of averages has been included for many years in the mathematics curriculum for secondary school. However, there is a greater emphasis on the teaching of statistics in recent curricula (e.g., M.E.C., 1992; N.C.T.M. 2000) and exploratory data analysis activities are suggested. This is why more research on students' reasoning when solving statistical open- ended tasks is needed.

Theoretical framework

Our research is based on a theoretical model on the meaning/ understanding of mathematical concepts (Godino & Batanero, 1994; 1997), where the authors distinguish five interrelated components in the meaning of the concept, each of which should be specifically dealt with in organising instruction or in assessing learning. These components are described below:

1. *The field of problems from which the concept has emerged*: One such problem in the case of the "mean" is finding the best estimation of an unknown quantity X when several different measurements x_1, x_2,... x_n, of the quantity are available. We take the mean as the best estimator because it produces a minimum error and compensates positive and negative deviations. Another different situation is looking for an element $\bar{x}$, *representative* of a set of given values, the distribution of which is approximately symmetrical; in this case, we take the mean, because it is the "centre of gravity" of the distribution. Other problems are finding a fair amount to be shared out in order to achieve a uniform distribution for a salary or other numerical variables, or guessing the value that will most probably be obtained in selecting a value from a symmetrical random variable (expected value).

2. *The representations of the concept*; to solve the problems we need ostensive representations, since the concept is an abstract entity. For example we use the words "mean", "average", "expected value", the symbol $\bar{x}$, the graphical representations, such as the centre of a histogram.

3. *The procedures and algorithms* to deal with it, to solve related problems or to compute its values, such as adding the quantities x_1, x_2,... x_n, and dividing by the

number of data; computing a weighted average, computing the mean from a table, from a graph or from a data set with calculators or computers.

4. *The definitions of the concept, its properties and relationships* to other concepts, such as the fact that the mean of a set of integer data can be a non integer number, that can be influenced by extreme values; the relative position of mean, median and mode in asymmetrical distributions.
5. *The arguments and proofs* we use to convince others of the validity of our solutions to the problems or the truth of the properties related to the concepts.

It is also important to notice that different levels of abstraction and difficulty can be considered in each of the five components defined above, and that, thus, the *meaning of the mean* is very different at different institutions.

In primary school or for the ordinary citizen, a simple definition of the mean would be sufficient, using a simple notation, avoiding algebraic formulae; restricting the calculus to simple data. A statistical literate citizen (Gal, 2002) would also need to understand the use of means in the mass media or in the business world (e.g. to understand stock market, prices, employment and other economic indicators that make use of weighted means). In scientific or professional work, or at university level, however, a more complex meaning of the mean would be needed.

Much more research is still needed to clarify the fundamental components in the meaning of each specific statistics/ mathematics concept as well as the adequate level of abstraction in which each component should be taught, since students might have difficulties in all the different components of the meaning of a concept. When entering an institution such as a school or a University the *personal* meaning that a subject attributes to a specific concept might be different to the meaning of the concept in that institution, so that we distinguish between institutional and personal meanings.

Previous research

There has been an extensive previous research on the understanding of averages, although, in general, they have focussed on only some isolated elements of the meaning of the concept. For example, Pollatsek, Lima, & Well (1981) described University students' errors in computing weighted averages and found students who do not recognise the problems of finding an expected value as a problem of averages; Mevarech (1983) observed that students tend to attribute non-existent algebraic properties to the arithmetic mean, such as null element or associative property. Intuitive understanding of arithmetic mean properties in 8-12 year-olds children is investigated by Strauss & Bichler (1988). As regards the understanding of the algorithm, while the majority of 12-13 year-olds are able to compute averages, according to Cai (1995), only some of them can invert the algorithm to compute an unknown value from a given average. Other students produce mistakes in computing mean, median and mode (Carvalho, 2001). Gattusso & Mary (1998) study the effect of context and representations on the difficulty of computing averages. Watson &

Moritz (2000) students confused the words mean, median and mode. That research also suggested that students can reach different levels of understanding for averages at the same age, although there is a progression in the level of understanding with age and instruction. Reading (2002) describes profiles for understanding of averages at different levels as a part or her profiles of statistical understanding.

2 METHOD

In Spain, new curricula introduce central measures at the first year of secondary school level (13-14 years old students). A revision of the topic is made at the fourth year (16-17 years-old students). This is why we focused on these two groups of students. Two samples of 14 year-olds (n=168) and 16 year-olds (n=144) were given a questionnaire. This sample included 155 boys and 157 girls from five schools in Granada, including different social and economic backgrounds.

Table 1. Elements of meaning assessed in the different items

	Elements of meaning assessed in the item	Item							
		1	2	3	4	5	7	8	9
Definition	Definition of mean	X	X	X	X	X	X	X	X
	Definition of mode								X
	Definition of median								X
Properties	Mean does not preserve the numerical set	X							
	Mean might not coincide with any data point	X						X	
	In computing the mean all the data are relevant					X	X		
	Mean is not an internal operation	X						X	
	Zero values affect the mean						X		
	Mean of the sum of two variables		X						
	Mean is commutative						X		
	Mean is not associative		X						
	Mean is a representative value	X						X	
	Sum of deviations to the mean			X	X				
	Mean and median only coincide in symmetrical distributions					X			
	Median is robust; mean is not					X			
Problems	Finding the best representative value			X			X		
	Finding a fair share	X		X			X		
	Guessing a probable value	X							
	Estimating an unknown quantity, from repeated measures								
Procedures	Computing mean from raw data	X	X		X	X	X	X	X
	Computing weighted average		X						
	Estimating mean from a graph								X
	Estimating mode from a graph								X
	Inverting the mean algorithm	X			X		X		
	Building a distribution with a given average						X		
	Computing median from raw data					X			
	Estimating the median from a graph								X

The questionnaire was made of 16 open- ended questions, 9 of which were common for the two groups of students and are presented in the Appendix. Its building was based on a previous theoretical analysis of the concept, as well as in an empirical analysis of secondary school textbooks (Cobo, 1998, 2001).

Item 1 and 2 were adapted from Watson & Moritz (2000); item 3, 4 from Tormo (1993); item 7 from Gattusso (1996); item 8 from Konold and Garfield (1992); item 9 from Zawojeswki (1986); item 5 was new for this research. Students answered individually the questionnaire and were encouraged to give detailed explanation to their answers. Due to restriction in length, in this paper we are mainly presenting the quantitative results for the part of the questionnaire that was common to the two groups (9 items). Detailed results about the qualitative analysis were presented at the conference.

3 RESULTS AND DISCUSSION

In Table 2 percentages of correct responses and standard deviation by group are compared. We obtained a reliability coefficient Alpha=0.78, which coincides with generalizability coefficient for items (GI=0.78) and a generalizability coefficient for students GS=0.98 of data for the 9 items and combined sample. This suggests than our results are much more generalizable to other students (with the same items) that to other items (with the same students).

Table 2. Percentage of correct responses and standard deviation by course

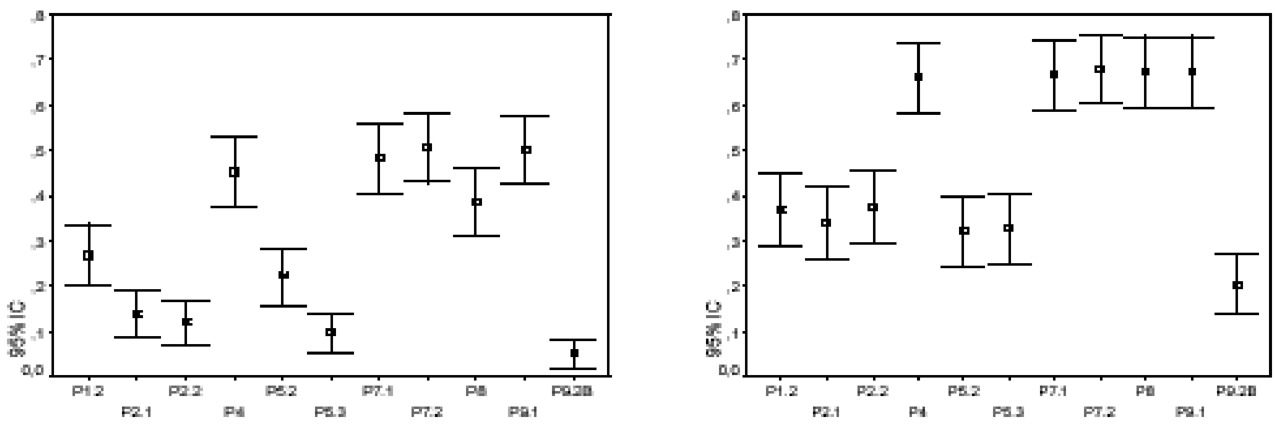
Item	Age 14 (n=168)		Age 17 (n=144)	
	Percentage	Std. Deviation	Percentage	Std. Deviation
P1.a	.63	.48	.69	.46
P1.b	.27	.44	.37	.48
P2.a	.14	.34	.34	.48
P2.b	.12	.32	.38	.49
P2.c	.36	.48	.33	.47
P3	.46	.50	.49	.50
P4	.45	.50	.66	.48
P5.a	.38	.49	.38	.49
P5.b	.22	.42	.32	.47
P5.c	.09	.29	.33	.47
P7.a	.48	.50	.67	.47
P7.b	.51	.50	.68	.47
P7.c	.53	.50	.61	.49
P8	.39	.49	.67	.47
P9.a	.50	.50	.67	.47
P9.b	.21	.41	.26	.44
P9.c	.04	.21	.20	.40

A multivariate anova to test the significance of the differences (dependent variable: vector of percentages of correct responses in the items) was significant (Wilk’s Lambda= .731; F=6.258; p-value <.001) for course and school (random factor; Wilk’s Lambda= .510; F=3.127; p-value <.001) but not for gender or interactions. This suggests that understanding of averages increased with instruction (see confidence intervals for items with significant differences in Figure 1), although results were not homogeneous in the different schools or as regards different items.

The greatest improvement was produced in item 8, which refers to abstract properties of averages (Mean might not coincide with any data point; Mean is not an internal operation; Mean is a representative value), Item 4 (Sum of deviations to the mean; Finding a fair share; Inverting the mean algorithm), Item 7.a (Inverting the mean algorithm; Building a distribution with a given average) and 7.b (Mean is commutative), Item, 9.a (Estimating mean from a graph), for which a wide percentage of students gave correct responses after teaching. There was also significant improvement in item 2.b and 2.a (Mean is not associative; Computing weighted averages), Item 5.c (Median is robust; mean is not) 5.b (Computing median from raw data), and 9.c (Estimating mode from a graph) although these items still remained difficult after instruction.

There was no difference in 2.c (Mean of the sum of two variables), 5.a (Finding the best representative value), and Item 3 (Finding the best representative value; Finding a fair share; Sum of deviations to the mean), although in these two items, more than 50% of the students in both groups provided correct responses. Results were quite poor in Items 7.c (In computing the mean all the data are relevant; Zero values affect the mean), Item 9.b (Estimating the median from a graph) and did not improve at all with teaching.

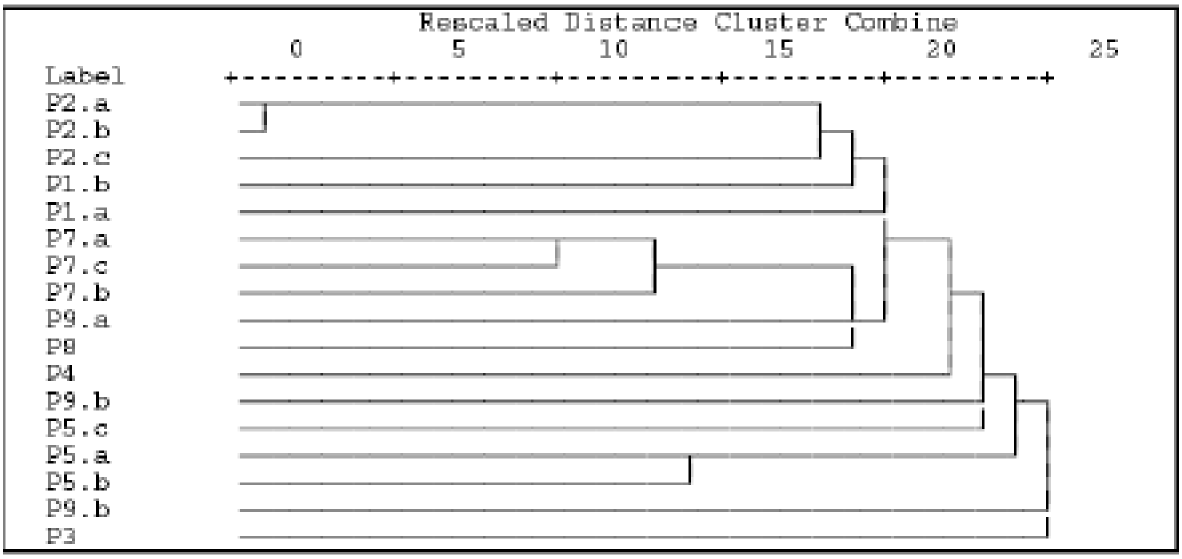
Figure 1. 95% confidence interval for items with significant difference by course



14 years-old 17 years-old

A cluster analysis of response to items (correct response=1, incorrect response=0) was made using the correlation coefficient as similitude measure and nearest point as linkage method (see Figure 2). Results show a multidimensional structure, which agree with our theoretical framework, where students can understand some components of the meaning for a mathematical object (e.g, average) and not others. There are several different clusters, usually grouping only subitems in the same item (e.g. 2.a, 2.b, 2.c; 1.a, 1.b; 7.a, 7.b, 7.c, 5.a, 5.b). Other items remain isolated; even for subitems in the same item (e.g. in item 9).

Figure 2. Dendrogram for cluster analysis



4 CONCLUSION

Some recent research is trying to describe levels of understanding statistical concepts, in particular for the case of averages (Watson & Moritz, 2000). This research is supported by neo-piagetian frameworks such as Biggs and Collis (1982, 1991), where empirical responses by students are used to classify them in discrete states along an unidimensional continuum. While we recognise the relevance of this research, which has used a very large sample of students and its interest in providing teachers with a criteria to organise instruction about a given topic (such as averages) along different curricular levels, we remark that the items used in Watson & Moritz were not based on a previous epistemological analysis and did not assessed the whole meaning of averages.

Our results are based on a more varied type of items than those used in previous research and suggest a more complex non linear structure of student's understanding. Even with a moderate sample of students they are highly generalizable to other students and multivariate statistical analysis (which was not used by other researchers) suggests the possibility of a multifactorial structure of understanding. These results are also confirmed by factor analysis and by the qualitative analyses of responses, which show that different students use a variety of elements of meaning correctly and incorrectly to solve a same task. All of this support our theoretical model and our systemic view of mathematical objects. In this model understanding a concept is a continuous constructive process where students progressively acquire and relate the different elements of the meaning of the concept. This understanding emerges from the student's meaningful practices linked to repeated solution of problems that are specific to that concept. It is through repeated activity of solving significant problems related to the concept that the student progressively acquires and widens his/her understanding.

Finally, these results suggest the interest to continue research on the meaning and understanding of statistical objects with larger samples of students and with different items which take into account the complex nature of mathematics and mathematical activity.

REFERENCES

Biggs, J. B. & Collis, K. F. (1982). Evaluating the quality of learning: The SOLO taxonomy. Academic Press: New York.

Cai, J. (1995). Beyond the computational algorithm. Students'unsderstanding of the arithmetic average concept. In L. Meira (Ed.). Proceedinsg of the 19th PME Conference (v.3, pp. 144-151). Universidade Federal de Pernambuco, Recife, Brazil.

Carvalho, C. (2001). Interaçao entre pares. Contributos para a promoçao do desenvolvimiento lógico e do desempenho estatístico no 7º ano de escolaridade. PhD. University of Lisbon.

Cobo, B. (1998). Estadísticos de orden en la enseñanza secundaria. Master's Thesis. University of Granada.

Cobo, B. (2001). Problemas y algoritmos relacionados con la media en los libros de texto de secundaria. Jornadas Europeas de Enseñanza y Difusión de la Estadística. Palma de Mallorca: Instituto Balear de Estadística.

Gal, I (2002). Adult's statistical literacy. Meanings, components, responsibilities. International Statistical Review, 70(1), 1-25.

Gattusso, L. & Mary, C. (1998). Development of the concept of weighted average among high-school students. In L. Pereira-Mendoza, C. Seu Keu, T. Wee Kee y W.K. Wong, Proceedings of the Fifth International Conference on Teaching Statistics (pp. 685-691). Singapore: International Association for Statistical Education.

Godino, J. D. & Batanero, C. (1994). Significado personal e institucional de los objetos matemáticos. Recherches en Didactiques des Mathématiques, 14(3), 325-355.

Godino, J. D. & Batanero, C. (1997) Clarifying the meaning of mathematical objects as a priority area of research in Mathematics Education. In A. Sierpinska, & J. Kilpatrick (Eds.), Mathematics Education as a Research Domain: A Search for Identity (pp. 177-195). Dordrecht: Kluwer.

M.E.C. (1992). Decretos de Enseñanza Secundaria Obligatoria. Madrid: Ministerio de Educación y Ciencia.

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

N.C.T.M. (2000). Principles and standards for school mathematics. Reston, VA; N.C.T.M. http://standards.nctm.org/

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

Reading, C. (2002). Profiles for statistics understanding. In B. Phillips (Ed.), Proceedings of the Sixth International Conference on Teaching of Statistics. Cape Town: IASE. CD ROM.

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.

Tormo, C. (1993). Estudio sobre cuatro propiedades de la media aritmética en alumnos de 12 a 15 años. Master's thesis. University of Valencia.

Watson, J. M., & Moritz, J. B. (2000). The longitudinal development of understanding of average. Mathematical Thinking and Learning, v1 (2/3), 11-50.

APPENDIX. QUESTIONNAIRE

Item 1. Let's say that the average number of children for 10 Andalusian families is 1.2. a) Explain this sentence in your own words; b) If the Garcías have 4 children and the Pérez have 1 child, show how many children the other 8 families might have. Explain your answer.

Item 2. Maria and Pedro spend an average of 8 hours each weekend practising sports. Another 8 students spend an average of 8 hours each weekend practising sports. a) What is the average number of hours the 10 students spend every weekend practising sports? b) Maria and Pedro also spend 1 hour on average each weekend listening to music and the other 8 students spend 3 hours each weekend. What is the average number of hours the 10 students spend every weekend listening to music? c) What is the average number of hours the 10 students spend every weekend on these two activities?

Item 3. Four friends met to prepare a dinner. Each of them brought some flour to make a pizza. As they want to make four equal sized pizzas, those bringing more flour gave a part to those bringing less. Is the total amount of flour given equal, higher or smaller than the amount of flour received? Why do you think this?

Item 4. We take six numbers from which the highest is number 5. By adding the six numbers and dividing by six the result is four. Is this possible? Why?

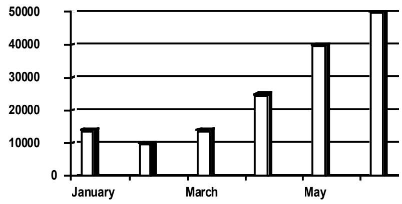
Item 5. The weights in kilograms for 9 children are 15, 25, 17, 19, 16, 26, 18, 19, 24. a) Which is the weight for the median child? b) Which is the median if we include another child who wighs 43 Kg.? c) Is it adequate to use the arithmetic mean to represent the weight of the 10 children? Why?

Item 7. Lucía, Juan and Pablo go to a party. Each of them takes some candies. The average number of candies is 11. a) How many candies did each of them take? Lucía__ Juan___ Pablo__ b) Is this the only possibility? Why? c) Another boy went to the party, but he takes no candies. What is now the average number of candies for the four children? Why?

Item 8. A small object was weighed on the same scale separately by nine students in a science class. The weights (in grams) recorded by each student are shown below.

6.2 6.0 6.0 15.3 6.1 6.3 6.2 6.15 6.2

The students want to determine as accurately as they can the actual weight of this object. Which method which would you recommend they use?



Item 9. This diagram shows the sales of sandwiches in a shop in six months the past year. Give an estimation for each of the following averages of the number of sandwiches sold each month: a) mean; b) median; c) mode.