PROSPECTIVE ELEMENTARY SCHOOL TEACHERS' INTERPRETATION OF CENTRAL TENDENCY MEASURES DURING A STATISTICAL INVESTIGATION

Raquel Santos

Investigation Unit of the Institute of Education, University of Lisbon

João Pedro da Ponte

Institute of Education, University of Lisbon

The introduction of statistics as a topic of the elementary school curriculum makes it necessary to reinforce teachers' preparation in this area as well as to understand prospective elementary school teachers' statistical content knowledge. This paper aims to understand the meanings that prospective teachers give to central tendency measures during a statistical investigation. We observe innovative interpretations as well as interpretations based on the context of each question and showing real understanding of concepts. But we also notice that most groups do not interpret central tendency measures, just analyzing data by reading graphs and tables. For the future, we suggest that prospective teachers must work with tasks requiring the interpretation of different measures to understand the differences among them.

Keywords: Statistical investigation, Mean, Median, Mode, Teacher education.

INTRODUCTION

The Portuguese official curriculum documents for kindergarten and primary levels (ME, 1997, 2007) give emphasis to statistics. This topic has great importance nowadays since society depends more and more on results from statistical studies (Groth, 2006). With industry, medicine and other society sectors recurring to data to make decisions, statistics should be a focal part of the new information era (Wild & Pfannkuch, 1999). Statistical literacy, regarded as the "ability to interpret, critically evaluate, and communicate about statistical information and messages" (Gal, 2002, p. 1), is essential in the education of every citizen and naturally in the education of prospective teachers. Therefore, it is important to know how prospective teachers interpret and communicate statistical information.

Statistical investigations allow students to become active in the learning process. During these projects, students pick a theme of their interest, define goals, select instruments to collect data, choose samples, collect, analyze and interpret data to answer the proposed questions (Batanero & Godino, 2005). Additionally, during an investigation, they perform every step of the PPDAC cycle (Problem, Plan, Data, Analysis, Conclusions) described by Wild and Pfannkuch (1999), in an environment inciting meaningful learning (Ponte, 2007). Furthermore, students are able to appreciate the importance and the difficulty of the statistical work and the interest of statistics in solving real life problems (Batanero & Godino, 2005). Moreover, teaching through statistical investigations allows the identification of students'

difficulties in their mathematical knowledge and, sometimes, even to detect concepts and ideas that seemed well consolidated but are not (Ponte, 2007).

This article shows part of an investigation that aims to understand prospective teachers' statistical and didactical knowledge. Our specific aim is to analyze the ways prospective teachers (for Pre-K to grade 6) interpret central tendency measures (mode, mean and median), based on their reports of statistical investigations.

CONCEPTUAL FRAMEWORK

Teachers are a key element in the educational process (Ponte, 1994). They need to know in depth the content they teach, as this is true even for teachers of early years (Ma, 1999). A solid mathematical knowledge is essential to promote a learning environment where students want and can learn mathematics, and this must be addressed since preservice teacher education. Heaton and Mickelson (2002) state that teachers' statistical knowledge encompasses the ability to conduct statistical investigations, describe information using different methods and form conclusions. To Mulekar (2007), prospective teachers must have a deep comprehension of statistical concepts in order to give coherent meaning to results. Central tendency measures (mode, mean and median) have a particular interest as they are frequently found in daily life (Groth, 2006). To Groth, understanding of these measures is an important component of statistical literacy. Nevertheless, according to Jacobbe (2008), even grade 1 exemplary teachers do not have conceptual knowledge of the two most basic statistical concepts – mode and mean.

In Portugal, the mode appears in the official curriculum documents in grades 3 and 4 (ME, 2007) and teachers tend to consider it an easy concept to understand. However, some studies suggest a more complex picture. For example, Fernandes (2009), researching difficulties and errors in statistics from prospective teachers (for grades 1 to 6), refers gaps on the comprehension of this concept, especially when they select the biggest frequency instead of the corresponding value of the variable. In a study with 40 prospective teachers (for Pre-K to grade 6) in their undergraduate program it was frequent to find answers like "the mode is 9, since there is a bigger number of students that see television" (Martins, Pires & Barros, 2009, p. 7). In the interpretation of this measure it is recurrent to associate the mode to the biggest number on the results table, the biggest absolute frequency, the value that appears more times and the biggest frequency category or interval (Martins et al., 2009). Some of these interpretations reveal confusion, since the value that appears more times could be seen as the number on the absolute frequencies column that is repeated more often and not as the value of the variable that is repeated more times.

The concept of mean is introduced in grades 5 and 6 (ME, 2007). Research about this concept is vast, since it is very used in statistical studies. Leavy and O'Loughlin (2006) indicate that there are two types of understanding – conceptual and procedural: "Computationally, the arithmetic average is the score around which deviations in one direction exactly equal deviations in another direction" (Leavy &

O'Loughlin, 2006, p. 55); conceptually, the mean may be seen as a balance point or center of gravity, representing the data set. To the authors, interpretations of the mean as the fair share (the value that represents the data set as if all data were equal), or as the balance point (where higher values compensate lower values) show conceptual understanding of the concept. When computing the mean, a frequent mistake is to determine the mean of the absolute frequencies in qualitative variables (Martins et al., 2009). Students use several interpretations of the mean, just restating the algorithm (Chatzivasileiou, Michalis & Tsaliki, 2010; Fernandes & Barros, 2005), indicating it as the "sum of all results divided by the existing values" or "sum of numbers" (Martins et al., 2009). Others associate the mean to the notion of balance, the average value, the value that balance the highest and lowest values (Martins et al., 2009), the fair share value, the typical expected value (Chatzivasileiou et al., 2010; Konold & Pollatsek, 2004), the location measure (a close but not exact value) (Chatzivasileiou et al., 2010), a signal in noise (where the mean of different observations is a close approximation to the actual value, ignoring the errors) or a data reduction value (a value to reduce the complexity of all data) (Konold & Pollatsek, 2004). As incorrect interpretations, there are answers based on the maximum value, the minimum value, a specific value, the median and the mode (Chatzivasileiou et al., 2010; Leavy & O'Loughlin, 2006). On the latter cases, students lack recognition of the data set as a whole and tend to focus on individual values (Chatzivasileiou et al., 2010).

Concerning the median, concept of grades 7, 8 and 9 (ME, 2007), the scenario is also problematic. Research shows that there are difficulties on the understanding and interpretation of this concept on grade 12 students (Fernandes, 2009) as well as on prospective teachers (grades 1 to 6) (Fernandes & Barros, 2005). Computing it, several prospective teachers indicate the central value of the absolute frequencies, others confuse this concept with the mode, and the most frequent mistake is to determine the central value without ordering the data (Martins et al., 2009). Even some grade 1 experienced teachers calculate the mean when they are asked for the median (Jacobbe, 2008). Interpretations of this concept include associations to the point where the number of values above equals the number of values below (Konold & Pollatsek, 2004), to the central value (although not always on the most correct way) as well as to "the value in the middle", the value that "divide in half", the value that is "somewhere in the middle", the second quartile (without more explanations), the value that "divide the sample in half and balance big values with small ones", the "average number of all results" and the "mean value" (Martins et al., 2009). The last three interpretations show some confusion between the concepts of median and mean.

METHODOLOGY

Participants in this study are prospective primary (grades 1 to 6) and kindergarten teachers in an undergraduate program of a Portuguese school of education that took a course on Discrete Mathematics, Statistics and Probabilities during the second semester of their 2nd year of studies, in 2010/11. This is the only course in the teacher education program dedicated to the development of the statistical knowledge. During

the course, prospective teachers worked statistical concepts through exploration and discovery, and strong emphasis was given to their interpretation in real contexts in different kinds of tasks. Additionally, they worked with the software Excel to organize data and calculate statistical measures. Prospective teachers also carried out statistical investigations, individually or in groups (of 2 or 3 students), in themes chosen by them, thinking on its possible use with their future students. During the statistical investigation, prospective teachers were asked to register the whole process in a written report and to present it to the class at the end of the semester, including data organization, analysis and interpretation (including all of what they have learned in class), a conclusion regarding the theme and a reflection about the work done. Towards the end of the work each group received feedback from the teacher with questions to help them to reflect about what was done and never received simple corrections of it. Besides that, after the presentation they received a final grade for all the process and final product. In this article we analyze these written reports in order to be able to understand their understanding of central tendency measures.

The 36 prospective teachers that attended the course were organized in 16 groups (6 groups attended day classes – D – and 10 groups after work classes – PL). 21 out of the 36 authorized the participation in this study. They were organized in 12 groups and chose themes such as recycling, food, water habits at home and routines. They undertook statistical investigations through questionnaires with 12 to 25 questions, with about 70% of the questions involving qualitative variables. When working the data, all groups used Excel, making use of the corresponding functions to determine the mean and the median. Regarding the mode, they realized it was more convenient to not use the Excel function "mode", since this does not work when analyzing qualitative variables and also when the variable has more than one mode. Since participants used Excel to determine the measures, the focus was on the interpretation of such measures in the context of each question. This means that, for the same concept, one group may give different interpretations, depending on the question they are analyzing. Analysis of data was exploratory and involved the categorization of interpretations observed in these reports. Categorization was made according to definitions and cases of doubt were sorted out between the two authors. Codes are used to identify the written reports of groups (G1 through G10) and the class (D/PL).

RESULTS

Mode

There are a large number of interpretations for the mode concept in the written reports of statistical investigations (Table 1). Four groups associated the mode to something that happens "more times" like happened in previous studies (Martins et al., 2009). Example of that is the analysis of the question "What is your profession?" where the group stated "the mode of the profession group is 'intellectual and scientific professions', since it is the profession group that is repeated more times in the sample" (G4D, p. 11). Interpretations made by five groups also associated mode to an answer that appears a large number of times, but made reference to "frequency"

or related words, like “the number of accompanying people more frequent is 1 (mode)” (G3D, p. 23).

Appropriate interpretations	Number of groups
Value associated to the highest frequency	5
The majority of the sample… / most…	4
Value referred/repeated/verified/chosen/appears more times	4
Value associated to more or to a greater number of respondents	4
Other interpretation using “more” and “bigger” in context	2
Predominant value	1
Satisfactory value	1

Table 1: Appropriate interpretations of the mode

Four groups used expressions like “the majority” or similar that imply a correct idea: “the mode is 3, therefore the majority of students already went 3 times to the theatre” during the analysis of the question “how many times did you go to the theatre?” (G6D, p. 28). Also related to the size of the sample was the category of interpretations where the mode was associated to more or to a greater number of respondents. In that category were interpretations like "the greater number of respondents is from the feminine gender (mode)" (G3D, p. 6). A similar idea seems to be meant by one group that referred something that is predominant (“it is the feminine gender that predominates” – G6D, p. 23) and by another group that indicated “here we obtained a satisfactory answer of ‘yes’, being the mode of this qualitative variable ‘yes’” (G8PL, p. 8). Interpretations of the mode made by two groups were connected to the context, making use of the expressions “more” or “bigger”: “the television is the information method more used by the respondents (...). Hence, it can be concluded that the mode is television” (G2D, p. 17).

Table 2 summarizes the interpretation problems made by prospective teachers. Two groups did interpretations also associated to the size of the sample, but the expressions chosen were not as efficient as in the previous cases. For example, in the analysis of the question: “Since when do you recycle?”, one group used the expression “we noted that 36% of the people started to recycle between 2006 and 2008, where, because this is the mode interval, we conclude that the greater part of the sample started to recycle by that time” (G2D, p. 9). This response is problematic due to the fact that if something is happening to a large part of the sample that does not mean that that “part” is the greater and, consequently, is the mode.

Additionally, one group, in the analysis of the question “Where do you eat breakfast?”, refers “the mode (…) is home, since is the higher value in the graph and on the table” (G1PL, p. 29). This statement, encountered also in previous studies

(Martins et al., 2009), can be problematic when the variable is not qualitative, since in the quantitative case the highest value on the table could be from the variable and not from the frequency.

Problematic interpretations	Number of groups
A great part of the sample…	2
Highest value	1
Confusion	1
Without interpretation (quantitative variables)	7
Without interpretation (qualitative variables)	9

Table 2: Problematic interpretations of the mode

One group that made the following statement: "the ages to which we made more questionnaires were young people aged 19" (G1D, p. 8). We observe some confusion but it is unclear if it is in the comprehension of the mode or in the phrase construction in Portuguese. This statement conveys the idea that the questionnaires were made to ages and not to people, which can be confusing to who reads the reports.

Also important is the fact that in the case of qualitative variables, nine groups did not make any reference to the concept of mode and to what this measure represents and means when analyzing the data. The same happened to seven groups regarding quantitative variables. This means that, especially in the cases of variables where the mode is the only statistical measure that can be determined, some participants make only readings from graphs and tables.

Mean

Table 3 summarizes the adequate interpretations of the concept of mean. One group argued the following expression "the average age is 3.4, which represents the age balance" (G1PL, p. 11). This response demonstrates that the group understands that the mean can be seen as a balance point (Leavy & O'Loughlin, 2006), but it is unclear the real understanding of the prospective teachers of the meaning of this measure. This interpretation may reveal conceptual comprehension (Leavy & O'Loughlin, 2006), but this is not evident, in this case.

Appropriate interpretations	Number of groups
Value that represents the equilibrium (balance point)	1
Fair share model	2

Table 3: Appropriate interpretations of the mean

Two groups used the fair share model when interpreting the mean, like the example of the analysis of the questions “How many favorite games do you have?” where the group wrote "it means that if every child had the same amount of favorite games, each one had 13, 33 games" (G2PL, p. 21). This type of interpretation also found by others researchers (Chatzivasileiou et al., 2010) reveals conceptual understanding of the mean concept as suggest by Leavy e O’Loughlin (2006).

Problematic interpretations of the mean are associated with misconceptions regarding the distinction between this concept and other statistical measures (Table 4). A group confused mean and spread measures, using the mean (and the median) to indicate whether the data was spread or not: “the values of the mean is 7.6 and of the median is 8, then we can conclude that the values are not very spread” (G6D, p. 28). We note that, in some quantitative variables, all groups made statements like “the mean is…” or “the sample, in average…” that do not represent an interpretation of the mean.

Problematic interpretations	Number of groups
Confusion between the mean and spread measures	1
Without interpretation	12

Table 4: Problematic interpretations of the mean

Median

There were several adequate interpretations of the concept of median (Table 5).

Appropriate interpretations	Number of groups
Value that divides the ordered data	1
Up to the median there is 50% of the sample...	1
50% of the respondents… or less…	3
50% of the respondents… at maximum	1

Table 5: Appropriate interpretations of the median

We observed the following interpretation of one group: “up to the median there is 50% of the sample, and after the median, there also is” (G5D, p. 13). This shows that these participants comprehend that the median divides the sample in half, 50% to each side of the median. Nevertheless, this can be problematic because of the non-inclusion of the median on the second half of the sample (“after the median”).

Another group made the interpretation "the median is 7.5, which is what divides the ordered data" (G4PL, p. 11), revealing that the group understands that the median is a number that divides the data when ordered, but do not reveal a real understanding of this concept in context.

Four groups make appropriate interpretations showing understanding of this concept in the context of the data, like: “50% of the people expect to return to the fair, in the same year, at maximum 2 times” (G3D, p. 31) or “50% of the children have 4 or less games” (G2PL, p. 21). This type of interpretation discloses understanding of the meaning of median in the context of the variable at study, making use of expressions that, in reality, make sense on data interpretation.

Nonetheless, not all interpretations were appropriate (Table 6). Three of the interpretations show that the median is a central value that divides data in 50% on each side, however without revealing the real comprehension of the concept. For example, one group wrote: “the Q2 is the median, which mean, the value that is in the center, equivalent to 50%” (G1D, p. 11). This interpretation shows confusion as it states that the median is equivalent to 50% of the data, which is incorrect. In the following example a group tried to make a statement like the one previously indicated as appropriate and very useful in context, however got confused, maybe because they do not really understand the concept: “50% of the people at maximum come to the fair with 3 accompanying persons” (G3D, p. 23).

Problematic interpretations	Number of groups
Value that is in the center	1
Value equivalent to 50%	1
Value to a maximum of 50% of respondents	1
Confusion between median and spread measures	1
Without interpretation	12

Table 6: Problematic interpretations of the median

Like in the case of the mean, one group used the median to make assumptions about the data spread, showing some confusion between the median and other statistical measures. Moreover, it should be noted that none of the groups of prospective teachers did the interpretation of the median concept in all quantitative variables that they studied.

CONCLUSIONS AND IMPLICATIONS

The analysis of the written reports from the twelve groups highlights gaps on prospective teachers’ understanding of central tendency measures. Even when analyzing contextualized data, they interpret the measures in a way that demonstrates some confusion about their meaning, sometimes not distinguishing among different statistical measures. Additionally, they use interpretations not mentioned in previous researches maybe probably because they had to write a report with the analysis of several variables and try to not repeat the same interpretation all over. Example of that are the interpretations of the mode like something that happens to “the majority”

and to "the greater number" of respondents (appropriate interpretations) or as a "great part" or "higher value" (problematic interpretations). Moreover, some groups gave appropriate meanings, like indicating the mean as a balance point (Leavy & O'Loughlin, 2006) or referring to the median as the center of the ordered data, but without showing a real comprehension of the meaning of the concepts. Additionally, we observe concrete and contextual interpretations like in the case of the mode as the value related to the majority, the mean as the fair share model and the median in where 50% of the sample is associated to this value or less. We suggest that interpretations related to the context of data are those that demonstrate stronger understanding of the concepts. Furthermore, an interpretation may be categorized as appropriate or not, depending both on the type of variable (quantitative or qualitative) and on the associated information (for example, if it is stated the relative frequency in the case of mentioning "great part"). Lastly, most or all groups do not make interpretations of central tendency measures, giving only readings of graphs and tables.

These results are important since they show a large variety of interpretations of central tendency measures, some appropriate and others not, that go beyond those mentioned in previous researches (e.g., Chatzivasileiou et al., 2010; Fernandes & Barros, 2005; Martins et al., 2009). Prospective teachers should experience and discuss all of them, in order to realize which ones better convey the information provided by data. Additionally, it is essential to make prospective teachers to compare and contrast the three central tendency measures to recognize their differences. Finally, this study shows the importance of emphasizing concrete interpretations of contextualized data, to consolidate the understanding of central tendency measures and to learn to distinguish them and grasp their utility.

ACKNOWLEDGEMENT

This work is financed by national funds through FCT – Organization to Science and Technology under the project *Developing Statistical Literacy* (contract PTDC/CPE-CED/117933/2010).

REFERENCES

Batanero, C., & Godino, J. D. (2005). Perspectivas de la educación estadística como área de investigación. In R. Luengo (Ed.). *Líneas de investigación en Didáctica de las Matemáticas* (pp. 203-226). Badajoz: Universidad de Extremadura.

Chatzivasileiou, E., Michalis, J., & Tsaliki, C. (2010). Elementary school students' understanding of concept of arithmetic mean. In C. Reading (Ed.), *Proceedings of the 8th International Conference on Teaching Statistics*. Ljubljana, Slovenia.

Fernandes, J.A. (2009). Ensino e aprendizagem da Estatística: Realidades e desafios. *Proceedings of the XIX EIEM: Números e Estatística.* Vila Real, Portugal.

Fernandes, J. A., & Barros, P. M. (2005). Dificuldades em estocástica de uma futura professora do 1º e 2º ciclos do ensino básico. *Revista Portuguesa de Educação, 18*(1), 117-150.

Gal, I. (2002). Adults' statistical literacy: Meanings, components, responsabilities. *International Statistical Review, 70*(1), 1-25.

Groth, R. E. (2006). An exploration of students' statistical thinking. *Teaching Statistics, 28*(1), 17-21.

Heaton, R. M., & Mickelson, W. T. (2002). The learning and teaching of statistical investigation in teaching and teacher education. *Journal of Mathematics Teacher Education, 5*(1), 35-59.

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 Conf*. Monterrey, Mexico.

Konold, C. & Pollatsek, A. (2004). Conceptualizing an average as a stable feature of a noisy process. In D. Ben-Zvi & J. Garfield (Eds.), *The Challenge of Developing Statistical Literacy, Reasoning and Thinking* (pp. 169-199). Dordrecht: Kluwer Academic Publishers.

Leavy, A., & O'Loughlin, N. (2006). Preservice teacher understanding of the mean: Moving beyond the arithmetic average. *Journal of Mathematics Teacher Education, 9*(1), 53-90.

Ma, L. (1999). *Knowing and teaching elementary mathematics: Teachers' understanding of fundamental mathematics in China and the United States*. Mahwah, NJ: Erlbaum.

Martins, C., Pires, M. V., & Barros, P. M. (2009). Conhecimento estatístico: Um estudo com futuros professores. *Proceedings of the EIEM: Números e Estatística*. Vila Real, Portugal.

ME (1997). *Orientações curriculares para a educação pré-escolar*. Lisboa: DEB.

ME (2007). *Programa de Matemática do ensino básico*. Lisboa: DGIDC.

Mulekar, M. (2007). Preparing teachers of statistics in the United States. *Proceedings of the 56th Session of International Statistical Institute*. Lisboa.

Ponte, J. P. (1994). Mathematics teachers' professional knowledge. In J. P. Ponte & J. F. Matos (Eds.), *Proc. PME XVIII* (Vol. I, pp. 195-210). Lisboa, Portugal.

Ponte, J. P. (2007). Investigations and explorations in the mathematics classroom. *ZDM, 39*, 419-430.

Wild, C. J., & Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. *International Statistical Review, 67*(3), 223-265.