THE VARIATION CONCEPT: A STUDY WITH SECONDARY SCHOOL MATHEMATICS TEACHERS

Ms. Cláudia Borim da Silva
Universidade São Judas Tadeu, Brazil
Dra. Cileda de Queiroz e Silva Coutinho
Pontifícia Universidade Católica de São Paulo, Brazil
dasilvm@uol.com.br

This research paper investigates how teachers working in Basic Education apply concepts related to the idea of variation. The researchers analysed the variation concept of ten mathematics teachers, using activities derived from a research methodology conventionally known as Didactic Engineering. Qualitative analysis of data revealed two different manners of understanding variation: the range and high value of variance or standard deviation as an indicator of large variation of data. However, the participants could not construct the concept of variation around the mean. They did not know how to analyse the values of variance or standard deviation but they knew how to compute these variation measures.

INTRODUCTION

The importance of statistics within the Brazilian national curriculum derives from the educational guidelines contained in the 'National Curriculum Parameters' (PCN's), which were published in 1998. The PCN's recommended that from the age of seven to fourteen, mathematics teaching should include statistics, probability and combinatory. A complementary document to the PCN's (the PCN+) recommended that between the age of fourteen and eighteen, mathematics teaching should include average, variance and standard deviation (Brasil, 2002).

According to Moore (1997), Watson and Kelly (2002) and Wild and Pfannkuch (1999), measures of variation help to build statistical thinking, despite the fact that the task of employing these measures in teaching and learning is complicated (Bem-Zvi and Garfield, 2004). The objective of this research was to identify the invariant operators that were used by the mathematical teachers when they were solving statistical problems. We are using concept according to Vergnaud (1996), which has three components: a set of problems which the concept gains significance, a set of concept representations and a set of invariants operators used to participant in a fixed manner (in the same comprehend of invariants operators of Piaget).

METHOD

This research paper was developed within the program for postgraduate studies in Mathematical Education at the Catholic University in São Paulo. The experimental stage was developed within the project for continuing education titled "Mathematical Thought in Elementary Education", which was coordinated by Prof. Dr. Saddo Ag Almouloud.

The subject of the research was ten mathematics teachers working in basic education. These teachers had already been educated in some areas of basic mathematics, such as Geometry, Fractions and Functions and were seeking education in Statistics. Activities were developed during the first semester of 2005 and the research was carried out in eighteen three hour meetings held at weekly intervals. Furthermore, a virtual forum was used to interact with the ten teachers.

The research methodology adopted the principles of 'Didactic Engineering' (Artigue, 1995). The first study was historical epistemology, literature review and a didactic study of the concept of variation titled previous analysis. This phase allowed us to review the didactics variables that were pertinent to the organisation of the activities.

One of the didactics variables chosen was to give form to the way in which the study approached its theme. In other words, we opted for a discussion of the statistical concepts using research performed by the teachers participating in the research. The organisation of the activities of our 'Didactic Engineering' took the following form:

First stage: Application of a conceptual map using the word statistics to identify the opinions of the ten teachers regarding statistics;

Second stage: From the results of the first stage, three texts were introduced to a discussion of the development of statistical thinking. Towards this end, the participants discussed: the PCN's and the PCN+, a text about the historical relationship between statistics and mathematics and a text that provides an example of how statistics teaching uses research (Ponte, Brocardo and Oliveira, 2003).

Third stage: The participants were motivated to work with real data. Hence, towards this end, the researchers discussed methodology to facilitate the efforts of the participants in data collection. During this stage, the participants created a questionnaire to collect data from 110 teachers.

Stage Four: After data collection, the participants organized and analysed the data collected. During this stage the participants discussed the frequency distribution, measures of central tendency and variation. During this part of the research exercise, the concept of variation was at the forefront of the conversation.

The development of these four stages were recorded on audio tape and also through the participation of observers (one for every group of participants). The following phase of our methodology 'Didactic Engineering', involved the organization and analysis of the collected data (audio tapes, written material provided by the observers and material produced by the participants themselves during their sessions). That phase prepared us for the last stage of the methodology, which involved the internal validation of the research through a comparison of a priori and a posteriori analyses.

THE INVARIANT RELATIONSHIPS WITH THE VARIATION CONCEPT

The conceptual map is a schematic resource used to represent significant concepts within a set of sentences (Novak and Gowin, 1999 in Manrique, 2003). Using the conceptual map, the teachers identify statistics as a tool for the analysis of facts that could help society to understand its reality. On the other hand, such information could be used by certain people to manipulate society.

However, the conceptual map did not enable the researchers to identify the perception of the participants in relation to the idea of variation and in stage four we confirmed the absence of this concept. In the last stage, the activity requested by the participants was to describe to someone from another planet the characteristics of the people that responded to the questionnaire (this activity emulates the activity described by Ponte, Brocardo and Oliveira, 2003 and referred to by the researchers in the second stage above). It was initially requested that the participants analyse two variables, i.e. age and length of tenure in the profession, from the set of quantitative variables. These two variables were to serve as the basis for a discussion of variation. As a result of the large response rate to the questionnaire, each group of participants decided to calculate the mean of groups of data.

As a result of their calculations, one of the groups of participants found that the mean age was 38.9 years. They recorded this finding as follows: "The Mean age of the teachers is approximately 40 years." The following discussion enabled us to infer that the term approximately is used for numerical approximation and not for the concept of variation around the mean. Hence, the participants had not learnt the concept of the mean as a value around which the data varies, but only as the result of an algorithmic procedure.

Researcher: "How are we going to analyse the Mean? What does the number that you found mean? Imagine that you're a person from another planet."

Michael: "A lot…"

Susan: "The majority of the teachers are this age."

Rachel: "You will arrive in the school and the teachers will have this characteristic"

Researcher: "How do we tell someone from another planet that he shouldn't believe that?"

Rachel: "Why is only this number very vague. It doesn't give any idea of what it really is."

Michael: "I'm going to look at the graph (referring to the Histogram)."

There is a moment of silence, some participants are thinking, others are trying to find a response I the statistics text book.

Researcher: "Let's suppose that someone is coming to Brazil with the intention of distributing money to poor families and wants to know the number of children in each family. Someone says that they have 2.3 children. How are we going to help this person interpret this information?"

Susan: "So, the teachers are between 19 and 60 years old and have a Mean of 40 years, the majority are 40 years old, a few are 19 years old and a small number are 60 years old. The teachers are going to be around 40, some 35, 43, 42…"

At the moment that the participant Susan refers to an age between 19 and 60, it can be inferred that she is using the variation concept to approximate the total range of the data. Hence, she is using the minimum age and the maximum age as parameters for her evaluation of the data.

Besides the invariant related to the variation concept, another invariant that can be identified was related to the mean. In fact, books and didactic materials conventionally use symmetrical distributions, which help the learner to understand the mean as the value of the group that appears with the most frequency, i.e. the mode.

Continuing the discussion of the arithmetical Mean, the researcher said, "*Half of the teachers can be 20 years old and the other half be 60 years old and the Mean be 40 years*?"

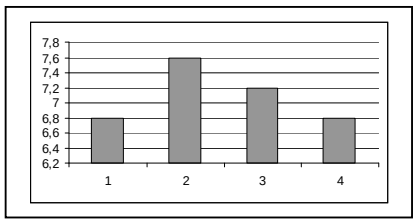
In this way, the participants realized that these measures, by themselves, would not be enough to describe the behaviour of a variable. They decided that if they want to analyse, they must research other measures in statistics text books. They found measures of variation and when the researcher asked what those measures meant, two participants responded:

Michael: "This is what I have to say: when this number here, 2 (referring to the standard deviation of the variable: length of tenure in the profession), is closer to zero it means that the data is more homogenous. If it is further from zero, I don't know how far from zero, you have to do another analysis. Many people with…"

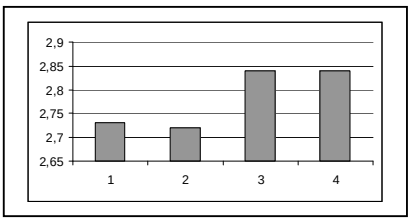
John: "The more the variance, the more dispersed are the data."

Although the participant, John, had understood the existence of the variation, he did not succeed in identifying a reference for it: mean. In this case, we describe the active invariant as follows: *if the value of the variance or of the standard deviation is high, then the variation of the data is large*. The use of that invariant, which does not indicate a relationship between the mean and the standard deviation, is illustrated in the graphs created by participant Michael (see below).

These graphs describe one of the variables of the instrument that the participants used to collect the data: the degree of importance of the statistic (score on a scale varying from 0 to 10) to day by day of the teacher, importance to the area in which he graduated, importance to his discipline and importance to the student (identified in the following graphs, in the respective bars 1, 2, 3, 4).



Graph 1: Arithmetic mean of the importance degree of the statistics



Graph 2: Standard deviation of the importance degree of the statistics

However, the participant, Michael, had identified homogeneity in the data when the value of the standard deviation was near zero, he did not associate this variation as an interval around the mean. This was made clear when he said that he did not know how far from zero the standard deviation would have to move to indicate a large variation. His lack of understanding was shown again when he created two separate graphs for the Mean and the standard deviation.

To sum up, we can infer that both participants are using the same invariant and understand the variation among the data, but do not understand the interval around the Mean. This

intuitive concept of variation among the data was already identified by Loosen, Lioen and Lacante (1985) and is described as unalikeability by Perry and Kader (2005).

DISCUSSION

In this investigation, it was possible to observe that the variation between the data could be intuitively constructed during a discussion of the arithmetic mean. The total range is a simple measure of variation, which appears naturally and does not break the thought process.

However, the concept and the significance of the standard deviation do not appear naturally in this work. The participants understood that a there was a variation in the data, but they could not identify that this variation was around the mean. This demonstrates no relationship between the mean and the standard deviation in the use of the invariant identifiers. The participants seemed to associate the invariant identifier with the terms *high standard deviation* or *low standard deviation*, but without relating high or low to the distribution of the values around the Mean.

It is interesting to observe that the participants of the study are mathematics teachers. They had already studied statistics during their graduate education and had heard of the term standard deviation, but it would not have been necessary to teach this theme.

The fact that the participants worked with real data, collected by them, is not sufficient for them to be able to construct the concept of variation around the mean. We conclude that it is necessary to continue the work with these participants, in respect of the formation of statistical thinking.

REFERENCES

Artigue, M. (1995). Ingeniería didáctica. In P. Gómez (Ed.), *Ingeniería Didáctica en Educación Matemática*, (pp. 33-59). Bogotá, Colombia: Iberoamérica.

Ben-Zvi, D. and Garfield, J. (2004). Research on reasoning about variation: A forward. *Statistics Education Research Journal*, 3(2), 4-6, http://www.stat.auckland.ac.nz/serj.

Brasil, Ministério da Educação. Secretaria de Educação Média e Tecnológica. (2002). *PCN+ Ensino Médio: Orientações Educacionais Complementares aos Parâmetros Curriculares Nacionais*. Brasília: Ministério da Educação.

Loosen, F., Lioen, M. and Lacante M. (1985). The standard deviation: Some drawbacks of an intuitive approach. *Teaching Statistics*, 7(1), 2-5.

Manrique, A. L. (2003). Processo de formação de professores em geometria: Mudanças em concepções e práticas. *Tese de Doutorado*. São Paulo: Pontifícia Universidade Católica de São Paulo.

Moore, D. S. (1997). New pedagogy and new content: The case in statistics. *International Statistical Review*, 65(2), 123-137.

Perry, M. and Kader, G. (2005). Variation as unlikeability. *Teaching Statistics*, 27(2), 58-60.

Ponte, J. P., Brocardo, J. and Oliveira, H. (2003). *Investigações Matemáticas na Sala de Aula*. Belo Horizonte: Autêntica, cap. V.

Vergnaud, G. (1996). A teoria dos campos conceituais. In J. Brun (Ed.), *Didáctica das Matemáticas*, (pp. 155-189). Lisboa: Instituto Piaget.

Watson, J. and Kelly, B. (2002). Can grade 3 students learn about variation? In B. Phillips (Ed.), *Proceedings of the Sixth International Conference on Teaching of Statistics*, Cape Town. Voorburg, The Netherlands: International Statistical Institute.

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