ACQUISITION OF NOTIONS OF STATISTICAL VARIATION BY IN-SERVICE TEACHERS

Ernesto Sánchez and Jaime García
Departamento de Matemática Educativa Cinvestav, México
esanchez@cinvestav.mx

A study with six middle school teachers on the notion of variation in the task of prediction is hereby presented. The SOLO hierarchy in which notions like randomness, structure and variation are included is applied. Variation related activities used in this study arise from the questions used in research on statistical variation but were also adapted to include computer simulation. The notion of no singular event emerged as a key for evaluation of the transition from multi-structural to relational thinking.

INTRODUCTION

There is increasing interest in statistics education into students' understanding of variability in different contexts such as sampling and probability, and researchers have called for greater emphasis on developing students' conceptions of variability in the school mathematics curriculum. Unfortunately, there is little research on teachers' conceptions of variability, particularly in the probability context (Canada, 2005; Aisling, 2006). The goal of this paper is to describe a difficulty for teachers in managing variation and show how it can be overcome.

We studied six middle school teachers who teach students between 12 and 15 years of age, all of the teachers in a Masters degree program in mathematics education. Apart from other regular courses or subjects , these teachers undertake a probability and statistics related project, in the course of the three year Masters degree program. The goal of the project is to strengthen teachers' knowledge of probability and statistics, identify problems relating to the teaching of the subject, and suggest didactic sequences and materials for its effective learning. The project is being supervised by the authors of this paper and three other collaborators. One topic in the project is statistical variation in a chance setting. Some tasks used in previous works (Watson, Kelly, Callingham & Shaughnessy, 2003) in addition to technology, facilitated the design of the activities that opens discussion about important ideas in the statistics curriculum such as randomness, centers, probability and distribution.

Variation is an omnipresent notion in statistics (Moore, 1990), and therefore it cannot be reduced to just an algorithm, process or scheme. Everyone perceives the idea of variation as something that changes but working with variation mathematically is not easy. Wild and Pfannkuch (1999) suggest that statisticians model variation for the purpose of prediction, explanation, or control. In this study a task of prediction is considered. Two questions then arise: What are the elemental notions that govern the use of variation for prediction? What are the problems being faced by teachers in the construction of those notions?

CONCEPTUAL FRAMEWORK

Randomness, structure and variation, three concepts related to statistical variability, are hereby distinguished. Randomness is considered in a wider dimension to include notions like chance, uncertainty, disorder, chaos, error, spread and deviations. Structure includes notions like regularity, centers, tendencies and distributions, while variation is a measurement that combines randomness and structure.

A hierarchy based on the Structure of Observed Learning Outcome (SOLO) model is used in this work (Biggs & Collis, 1991). Someone is located at the uni-structural level if his justification for a prediction is dominated by the consideration of randomness (anything can happen) or if it is dominated by the structure (the average value will occur). Someone is in the multi-structural level if in the justification for a prediction considers both randomness and structure but independently or relates them in an unsuitable way. Finally, someone is in relational level if the justification for a prediction combines randomness with structure. Within the context of the tasks used, providing nonsingular events as an answer to prediction tasks turns out to be a key element in identifying someone's transition from the multi-structural level

© ICMI/IASE 2008. In C. Batanero, G. Burrill, C. Reading & A. Rossman (Eds.), *Joint ICMI/IASE Study: Teaching Statistics in School Mathematics. Challenges for Teaching and Teacher Education. Proceedings of the ICMI Study 18 and 2008 IASE Round Table Conference.*

to relational level. Shaughnessy and Ciancetta (2002) identified the tendency to conceive singular events as a difficulty in variation tasks: "Many of the tasks used in research ask students to give a single number for an answer" (p. 1), stating further that, "…the teaching of statistics tends to emphasize single outcome responses and to focus on centers rather than on spreads" (p. 1).

The writers believe there is a tendency on the part of students to think of events as single, because the general concept of event is quite abstract, and its appropriate use in the answers to questions implies an advance in the understanding of variability, in such a way that the consideration of nonsingular events as an answer to prediction tasks represents the beginning of a relational thinking about variation. This change represents an effort by the students to retreat from certainty. This is a crucial step in the development of statistical thinking.

METHODOLOGY

Participants

Six teachers of public middle schools, whose pupils range between 12 and 13 years of age, were studied. The teachers who have had many years of teaching experience are taking retraining or updating professional courses in The Center for Research and Advanced Studies of National Polytechnic Institute of Mexico. Some basic information about the teachers, with fictitious names, is shown in Table 1. Only Alonso teaches all the topics related to the subject: "Presentation and Processing of Data & Probability", which is a topic in the official syllabus of the Mexican Ministry of Education. While Genaro has taught parts of topics, others have not.

Table 1. Data of participants

Name	Years as active teacher	Years teaching the same grade	Grades of Students	Qualification/Academia level
Alonso	14	2	2nd & 3rd	Major in Education
Genaro	14	2	2nd & 3rd	Geology Engineer
Gimeno	6	2	3rd	Major in Education
Lorenzo	8	1	1st	1st degree in Civil Engineering
Mónica	10		1st	1st degree in Business Administration
Weber	11		1st	1st degree in Communication & Electronics

Instruments

Two sets of pre and post questionnaires were designed: The pre questionnaire contains five individual diagnostic questionnaires and five computer based activities; two of which were carried out with Probability Explorer and three with Fathom. There were four post questionnaires; they included the teachers' response to the pre-questionnaire and the results of the computer-based activities. Interviews were then conducted to conclude the activities. Due to limited available space, only a report of the activities related to throwing the dice 60 times is presented here:

> Activity A (partial): Imagine a die is thrown 60 times; fill in the table below the number of times that you think each of the numbered faces will appear. Please explain the reason/s for giving your answer (Here a 2×2 table indicating the die face numbers and the cell to be filled with the frequencies is given).

Post Questionnaire

The post questionnaire dealt with the responses of the teachers in the pre questionnaire and the results obtained in the activities. Therefore a post questionnaire was designed specifically for each of the five teachers. Below, as an instance, we present the post questionnaire given to Alonso.

Post questionnaire for Activity A (Alonso)

R: Remember: Imagine a die is thrown 60 times; fill in the table below the number of times that you think each of the numbers will appear. The following (Fig 1, left side) was the answer you gave to the above question or in the previous activity. What are your reasons for giving this answer?

A: I gave that answer because each face has the same probability of occurring.

R: Please find the result using Probability Explorer to simulate a set of possible answers to the same question and complete the table below with the results obtained.

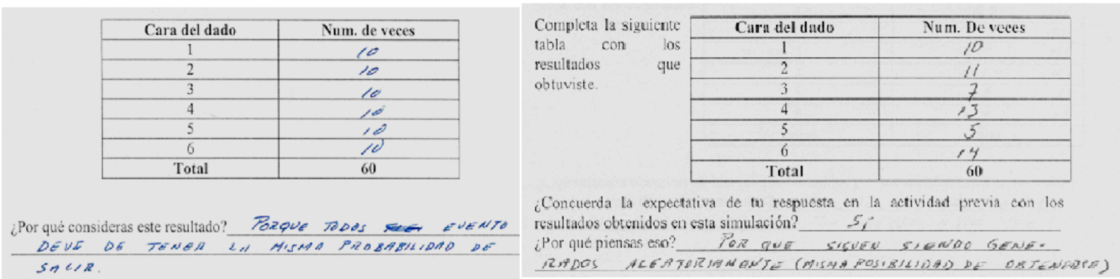
R: Does your answer from the previous activity tally with results obtained in this simulation (Figure 1, right side)?

A: Yes.

R: Why did you think so?

A: Because they were randomly generated (the same probability of getting them).

R: Why are the two results not coherent? If you were asked to fill in the table again, what are your answers likely be? (The table to be filled again). What is/are your reason/s for giving this answer? Do you think probability is contradictory? Explain.



Cara del dado	Num. de veces
1	10
2	10
3	10
4	10
5	10
6	10
Total	60

¿Por qué consideras este resultado? PORQUE TODOS ~~...~~ EVENTO DEVE DE TENER LA MISMA PROBABILIDAD DE SALIR.

Completa la siguiente tabla con los resultados que obtuviste.

Cara del dado	Num. De veces
1	10
2	11
3	7
4	13
5	5
6	14
Total	60

¿Concuerda la expectativa de tu respuesta en la actividad previa con los resultados obtenidos en esta simulación? SI

¿Por qué piensas eso? POR QUE SIGUEN SIENDO GENERADOS ALEATORIAMENTE (MISMA POSIBILIDAD DE OBTENERLOS)

Figure 1. Example of response to pre-(left) and post-questionnaire (right)

Procedures for Application of Instruments

The instruments mentioned above were applied in four stages. First, the teachers answered the pre questionnaire. Second, they carried out a guided activity, which involved the use of a computer to solve or to answer the pre questionnaire through a statistical simulation. Then the teachers answered the post questionnaire. The post questionnaire (which was computer-based) provided an opportunity to compare and contrast the results of the pre questionnaire and the activity; any conflict forced the teachers to reflect on their understanding of the underlying statistical notions at play versus expected. They needed to find the justifications or explanations for obtaining different results (before and after the use of computer) and evaluate the statistics variability. Last, the participants were interviewed for the purpose of obtaining additional information about their reasoning. The interview, which lasted for about an hour, was filmed. All this procedure, including the interview, was repeated in six sessions, each of four hours. In each session, the pre questionnaire was applied followed by the guided activity. The post questionnaire for each activity was applied in the subsequent session.

Procedures for Data Analysis.

Characteristics of four levels of answers, in line with the SOLO model of Biggs and Collis (1991), were defined to analyze the participants' results and their justifications or reasons for the answers in the following way:

- Pre-structural level: Here we included inconsistent responses, for example, "the sum of the frequencies is not 60". No pertinent or incoherent reasons were given as justification (e.g. Because God has created things in this way).
- Uni-structural level: The results in this category were based on only one governing factor; randomness or structure. Somebody was said to be randomness centered when

they explicitly or implicitly said that any event might occur. For example the person could say that six, for instance, does not occur when a die is thrown 60 times. When asked why he/she thinks so, the answer was that it is by chance. On the other hand we considered a person was structure centered if the answer was a uniform distribution of 10 for each of the six faces of the die.

- Multi-structural level: To locate a participant in this level, both structure and randomness were considered even though the participant could not draw any relation between the two. For example, he/she could supply different numbers that are approximately 10 and whose sum was 60, i.e., "(10, 11, 7, 13, 8, 11)". Even though the numbers supplied were less probable than "(10, 10, 10, 10, 10)" we interpreted that response as an attempt to express some variability around 10. Another common answer at this level, which is a bit more complex, is saying that each face of a die can occur "± 10 times" or "around 10".
- Relational level: At this level a range of possibilities for each face of a die was given and the probability of a result to fall within the given range was estimated.

RESULTS

The results and the reasons given by the six participating teachers to the question in Activity A are given below. We can observe that at the beginning (Table 2) their responses were centered on the structure of the situation, where the governing distribution for the experiment is uniform. During the working sessions, the participants noted the variation in each of their simulations, and their explanations were based on the fact that it was a chance phenomenon (Table 3), without recourse to structure. After the experiment, the majority of the participating teachers proposed results with uniform distributions with just a little variation (Table 4).

Table 2. Responses before activity A

Teacher	Alonso	Genaro	Gimeno
Proposed sequence	10, 10, 10, 10, 10, 10	10, 10, 10, 10, 10, 10	10, 10, 10, 10, 10, 10
Reason why	Because all events need to have equal probability	Because all the events have a probability of $\frac{1}{6}$, i.e. the 60 times for each case is $\frac{10}{60}$	Because all the numbers have the same probability
Teacher	**Lorenzo**	**Mónica**	**Weber**
Proposed sequence	5, 15, 5, 15, 5, 15	10, 10, 10, 10, 10, 10	10, 10, 10, 10, 10, 10
Reason why	It is probable that any series of 3 numbers is repeated more.	Because the each face of the dice have the same probability	Because each face of the dice has the same probability.

In comparing the results proposed by the teachers (Table 2) with the simulated results (Table 3), they were confronted with the questions: "Do you consider both results equal?" and "Why?" The answers are presented in Table 3. We can see that Alonso, Gimeno and Lorenzo considered that the results of the simulation were the same as that provided in their first answer "(10, 10, 10, 10, 10)", although in fact the results are different. They may have had a good idea of the kind of results that can occur, but they were not able to express it. In contrast, the other teachers realized that their answers did not match because they really expected the sequence 10, 10…

Table 3. Responses during the activity A

Teacher	Alonso	Genaro	Gimeno
Simulated result	10, 11, 7, 13, 5, 14	9, 7, 11, 9, 7, 11	7, 17, 10, 7, 8, 11
Does the result obtained tally with the one you earlier supplied?	Yes, because they were generated randomly (having the same probability of occurring)	It's not definite; I observed that as the number of times the dice is thrown increases, the number of times that each of the faces appears tends to be equal.	Yes, I think it is a question of chance. Now two occurs 6 times not because it has the greatest probability. Simply, in this exercise it falls more times.
Teacher	**Lorenzo**	**Mónica**	**Weber**
Simulated result	10, 12, 9, 10, 6, 13	6, 11, 10, 10, 11, 12	13, 12, 6, 13, 13, 3
Does the result obtained tally with the one you earlier supplied?	Yes, because the simulation is conserved and it is by chance. Presently "6" predominates	No; the probability for each face is not the same.	No; for the simple reason that I proposed the same probability without the activity. With this the possible results can be compared.

In Table 4 we present the answers to the activity: "Imagine a die is thrown 1000 times; fill in the table below the number of times that you think each of the numbers will appear". The teachers, Alonso, Genaro and Lorenzo, gave a very narrow variation; although they recognized that there is variation, they supposed that it reduces as the number of repetitions increases. This made them suggest an almost uniform distribution. On the other hand Weber gave too wide a variation. While Weber perhaps was more randomness centered, the other teachers oscillated between randomness and structure but just a little. The teachers tend not to respond with singular events in the activity of predicting the number of heads in 10 throws.

Table 4. Responses and range of the results proposed for 1000 throws of a dice

Alonso	Genaro	Gimeno	Lorenzo	Mónica	Weber
165, 167, 167, 168, 164, 169	164, 166, 165, 167, 168, 170	160, 173, 170, 173, 168, 156	167, 166, 167, 166, 167, 167	180, 170, 160, 170, 160, 160	50, 70, 350, 400, 40, 60
Range = 5	Range = 6	Range = 17	Range = 1	Range = 20	Range = 350

When asked to predict events with probabilities of 0, 0.5 and almost 1 (Table 5), the teachers realized that the singular events no longer worked for them. Events that include a range of results were proposed.

Table 5. Events proposed with probabilities 0, 0.5 and almost 1

Probability	Alonso	Genaro	Gimeno
Almost Zero	That the event is 10 or 1	Obtaining, 0 or 10 heads when throwing a coin 10 times	10
P= 0.5	That the event is 5 and 6	Obtaining 4 or 5 or 6 heads when throwing a coin 10 times	4, 6 and 8
Almost 1	That the event is between 2 and 8	---	All the events from 0 to 10

Some of the teachers' comments in the interview indicated that they were aware that thinking about single events was an obstacle. In the interview the teachers were asked to guess the result of drawing out a ball with replacement 30 times from an urn containing three balls labeled A, B and C. Genaro remarked: "Well, at the beginning I believed I could get 7 As, 12 Bs, and 11 Cs, but it is very difficult to obtain this precise event. If we consider now obtaining between 7 and 12 As, between 7 and 12 Bs, between 7 and 12 Cs, this event would appear more easily".

The other teachers also made comments expressing the likelihood of offering as answers no-singular events after computing the probabilities that the result would be that event, thereby helping to make precise predictions.

CONCLUSIONS

We observed that the teachers studied had some limitations in understanding the complex relations that exist between randomness and structure in variation. One of their difficulties was their inability to express their sense of variation mathematically. Their answer at the first encounter with the prediction task was almost determinism: (10, 10, 10, 10, 10, 10). After the simulation they included variability in their answers, but they used the same language of certainty (for example: "164, 166, 165, 167, 168, 170").

Giving their predictions and the results obtained by simulations, the teachers were able to learn that a definite answer is always incorrect, and that instead, the solution must be a range of values; this raises the opportunity to combine randomness and structure without returning to the uni-structural level governed by one or another factor. The transition from a multistructural to relational level began when the teachers realized they must propose a range of values (no-singular events) as an answer. This notion must be completed with the calculation of the probability of getting a result within that range. In the course of the activities, a great difficulty faced by the teachers was a misconception of the law of large numbers. They believed that the distribution of frequencies becomes uniform as the number of repetitions increases, ignoring the fact that the law of large numbers refers to relative frequencies and not absolute frequencies. The computer software and the task offered fundamental assistance to overcome this difficulty.

REFERENCES

Aisling L. (2006). Using data comparisons to support a focus on distribution: Examining preservice teachers' understandings of distribution when engaged in statistical inquiry. *Statistics Education Research Journal*, *5*(2). Online: www.stat.auckland.ac.nz/serj/.

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 N. J.: Lawrence Erlbaum Associates.

Canada, D. (2006). Elementary pre-service teachers' conceptions of variation in a probability context. *Statistics Education Research Journal*, *5*(1), 36-63. Online: www.stat.auckland.ac.nz/serj/.

Moore, D. S. (1990). Uncertainty. In L. A. Steen (Ed.), *On the shoulders of giants* (pp. 95-137). Washington D. C. USA: National Academies Press

Shaughnessy, J. M., & Ciancetta, M. (2002). Students' understanding of variability in a probability environment. In B. Phillips (Ed.), *Proceedings of the Sixth International Congress on Teaching Statistics*. [CD-ROM]. Cape Town, South Africa: International Association for Statistical Education .

Watson, J. M., Kelly, B. A., Callingham, R. A., & Shaughnessy, M. (2003). The measurement of school students' understanding of statistical variation. *International Journal of Mathematical Education in Science and Technology, 3*4(1), 1 –29.

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