6-2015

Levels of Reasoning of Middle School Students about Data Dispersion in Risk Contexts

Ernesto Sánchez

Antonio Orta

Follow this and additional works at: <https://scholarworks.umt.edu/tme>

Part of the [Mathematics Commons](#)

Let us know how access to this document benefits you.

Recommended Citation

Sánchez, Ernesto and Orta, Antonio (2015) "Levels of Reasoning of Middle School Students about Data Dispersion in Risk Contexts," *The Mathematics Enthusiast*. Vol. 12 : No. 1 , Article 23.

DOI: <https://doi.org/10.54870/1551-3440.1348>

Available at: <https://scholarworks.umt.edu/tme/vol12/iss1/23>

This Article is brought to you for free and open access by ScholarWorks at University of Montana. It has been accepted for inclusion in The Mathematics Enthusiast by an authorized editor of ScholarWorks at University of Montana. For more information, please contact scholarworks@mso.umt.edu.

Levels of Reasoning of Middle School Students about Data Dispersion in Risk Contexts

Ernesto Sánchez

Departamento de Matemática Educativa, Cinvestav-IPN, México

Antonio Orta

Escuela Nacional para Maestras de Jardines de Niños, México

Abstract: The aim of this research study is to explore students' reasoning concerning variation when they compare groups and have to interpret dispersion in terms of risk. In particular, we analyze in this paper the responses to two problems from a questionnaire administered to 82 ninth-grade students. The problems consist of choosing between two and three groups of data by comparing them. The first one composed of losses and winnings coming from a hypothetical game; the second is about medical treatments. The results show the difficulty students had in interpreting variation in a risk context. Although they identify the data group with more variation, this is not enough for interpreting the variation in terms of risk and making a rational decision. The psychological categories of risk-seeking and risk-aversion are used to explain the behavior of students who choose one group or another when they identify correctly the risk in each situation. As a conclusion, it is suggested that more risk context situations should be studied.

Keywords: variation, dispersion, variability, risk, middle school students.

Introduction

Variation is the underlying reason for the existence of statistics (Watson, 2006, p. 217) since variability is in everywhere and therefore in data. Moore (1990) emphasized the omnipresence of variation and the importance of modeling and measure variation in statistics; Wild & Pfannkuch (1999) included the perception of variation as part of the fundamental types of statistical reasoning. Burrill & Biehler (2011) proposed a list of seven fundamental statistical ideas in which variation is the second after data. But this growing recognition of the importance of variation in basic education is relatively recent. Just 18 years ago Shaughnessy (1997) wondered where the educational research on variability was and called to the community of statistical educators to investigate on statistical variation. As a consequence several researchers began to explore scenarios that would allow students to display their understanding about variation.

As a fundamental idea it is supposed that variation “can be taught effectively in some intellectually honest form to any child at any stage of development” (Heitele, 1975, p. 187), so research to improve instruction in variation can be located at any level. In a growing number of studies on students' thinking about variation, contexts and problems have been proposed to encourage students from different scholar levels to perceive, describe or/and measure variation in data. For example, variability in sampling (Watson & Moritz, 2000), chance (Watson & Kelly, 2004), repeated measurement, and natural variation in the plants growth (Lehrer & Schauble, 2007, Petrosino, Lehrer & Schauble, 2003), weather (Reading, 2004), all these situations have been explored for develop statistical thinking of students at different levels. Risk situation could provide another context to investigate variability. In this paper we propose a way to formulate some decision making problems in risk contexts where dispersion is relevant and we explore the reasoning of middle school students in front of these problems.

Conceptual framework

The concepts that constitute our conceptual framework try to provide a comprehensive understanding of the study. The first is that of tasks as a key element in teaching and learning. In statistics the context of the tasks has an important role for understanding statistical ideas; therefore the second concept of our framework is risk context in its relation to variation. Reasoning as a way to close to students' thinking is our third concept. Finally, we include SOLO model as an instrument to organize the reasoning of students.

Tasks

An important part of research in mathematics and statistics education is to seek high-level tasks that promote the capacity to think, reason, and solve problems related to the fundamental ideas of the study area. The tasks should also encourage to the students to engage with the concept to be learned. In statistical unlike mathematics, reasoning must articulate abstract mathematical ideas with real situations through the data, i.e. statistical reasoning is intimately linked with the contexts. The formulation of tasks that need to collect data is not sufficient to emerge themselves abstract ideas (concepts) that must be learned. Nor is it enough to organize students to work with tasks that only require the manipulation of concepts and properties independently to situations. Finding a balance between these extremes depends largely on the choice of good problems. These are the means by which the actions of the teacher are transformed in student engagement in reflection and action.

Problems on decision making under uncertainty are common in statistics; this kind of problems has been widely used to promote, and also to analyze, some important aspects of the statistical reasoning of people. On the other side, tasks on comparing two groups are frequently used to engage students in reasoning about data since in many statistical studies is necessary compare groups. We are proposing two problems on decision making where comparing groups of data is required and dispersion is significant, in addition their solution implies some risk preferences.

Risk context

The interpretation of dispersion depends on the situation from which the data come. One kind of elemental problems where variation could emerge can be formulated in risk context. When the uncertainty present in a process implies any threat to the effect of a result, it is called a risk. These situations appear when there are potential and unwanted results that, as a consequence, lead to losses or damages. Defining risk means to specify both the valuable and unwanted results in a way that reflects the value attributed to them. Analyses of risk situations offer information for decision-making. The theory of decision making under situations of risk has two aspects. On one hand, it defines abstract rules for what people should do; and, on the other, it studies what people really do when facing the risk.

The theory of decision making in situations of risk is a broad theory, but what is necessary for our research are very basic elements. These are restricted to the concepts of prospect, risk aversion and risk seeking as they have been defined by Kahneman and Tversky (2000).

Consider the following problem:

Problem: Choose between	
A: p_0 chance to win X_0	B: p_1 chance to win X_1
q_0 chance to win Y_0	q_1 chance to win Y_1

The overall utility of each game is:

$$U_A = p_0 \times X_0 + q_0 \times Y_0$$

$$U_B = p_1 \times X_1 + q_1 \times Y_1$$

According to classical theory of utility, comparing U_A to U_B the most suitable game is determined.

These decision making problems are generalized as a choice between prospects or gambles in the following way. A prospect $(x_1, p_1; x_2, p_2; \dots; x_n, p_n)$ is a contract that yields outcome x_i with probability p_i , where $p_1 + p_2 + \dots + p_n = 1$. In Prospect Theory (Kahneman & Tversky, 2000) a utility function for outcomes $u(x_i)$ is postulated. In our case we suppose that such function is the identity $u(x_i) = x_i$ since the outcomes are given in the monetary value that the gamblers win (or in time that the patients live with each treatment). Then the overall utility of the prospect is the expectation:

$$U(x_1, p_1; \dots; x_n, p_n) = p_1 x_1 + p_2 x_2 + \dots + p_n x_n$$

Consider a situation where a sample $x_1, x_2 \dots; x_n$ of outcomes of a game is given and their corresponding probabilities are unknown. Due to lack of information may be reasonable to assume that the probability of each outcome is $1/n$. In such case the overall utility is the arithmetical mean:

$$U\left(x_1, \frac{1}{n}; \dots; x_n, \frac{1}{n}\right) = \frac{x_1}{n} + \dots + \frac{x_n}{n} = \bar{x}$$

Now it is possible to formulate problems of decision-making as follows:

Problem. The gains of realizations of n times the game A and m the game B are:

Game A: $x_1, x_2 \dots; x_n$

Game B: $y_1, y_2 \dots; y_m$

Which of the two games would you choose to play in?

If the games are thought as prospects, according to classical theory of utility, solution is obtained by comparing $\bar{x}$ and $\bar{y}$. (This procedure of solution is not easily understood by students).

However, when the expected gains coincide $\bar{x} = \bar{y}$ not necessarily the most appropriate decision is to choose any of the two games since the dispersion of the values in each set can be significant for the decision maker. Indeed, the comparison of the dispersion of each set gives an account of the difference in risk terms. Consider for example the following problem:

Problem. The gains of 2 realizations of two games are:

Game A: $-1, 3$

Game B: $1, 1$

Which of the two games would you choose to play in?

The overall utility of each game is 1 but for people the games are not equivalent if games are thought as prospects:

A:	50% chance to win 3 50% chance to loss 1	B:	1 for sure
----	---	----	------------

Frequently people prefer game B because they do not like to take risks. In the research of how people answer this kind of problems psychologists have elaborated the concepts of risk aversion and risk seeking:

The preference for a sure gain is an instance of risk aversion. In general, a preference for a sure outcome over a gamble that has higher or equal expectation is called risk aversion, and the rejection of a sure thing in favor of a gamble of lower or equal expectation is called risk seeking (Kahneman & Tversky, 2000, p. 2)

It is worth noting that in a game, the dispersion of gains (including losses) can be considered a measure of risk. With this idea, an extension of these attitudes toward risk can be developed to explain some of the behaviors observed in the responses of students to decision-making problems studied in this work. Consider a problem of making decision on two data sets whose arithmetic means is the same but with different dispersion, and suppose the data belong to a variable that carries risk. Let's say that a preference is motivated by risk aversion when an option whose data have less dispersion over another whose data have greater dispersion is preferred. The decision is motivated by risk seeking when the option whose data have greater dispersion is chosen.

Reasoning

The statistics education community has distinguished three overlapping areas of statistics to organize and analyze the objectives, activities and results of statistical learning: statistical literacy, reasoning and thinking. This study is located in the area of statistical reasoning. The purpose of the research on statistical reasoning is to understand how people reason with statistical ideas (Garfield & Ben-Zvi, 2008) in order to propose features to create learning scenarios. It is worth then, to make some remarks on the idea of reasoning in general. For this we take some lessons from *inferentialism* (Brandom, 2000); Bakker & Derry, (2011) point out three lessons:

- Concepts should be primarily understood in terms of their role in reasoning and inferences
- To understand concept use we should privilege holism over atomism
- Privileging an inferentialist approach to education over representationalist one.

The first implies that “to talk about concepts is to talk about roles in reasoning” (Brandom, 2000, p. 11); concepts are understood in social practices of asking and giving reasons.

The second lesson tells us that “One cannot have any concepts unless one has many concepts. For the content of each concept is articulated by its inferential relations to other concepts. Concepts, then, must come in packages.” (Brandom, 2000, p. 15-16). The third lesson is the recommendation of Bakker and Derry for using an inferentialist approach for analyzing educational process.

When students try to justify their responses, elements that they think are important to the situation are revealed; in particular, the data they choose, operations made with these and knowledge and beliefs on which they rest for doing that, are important in reasoning. Unfortunately, the answers given by the students are usually not as explicit as to give us information on these three aspects; anyway, we will try to identify their main features of reasoning from their responses. For this task we will rely on our interpretation of the SOLO model.

SOLO Model

The Biggs and Collis (1982, 1991) *Structure of Observed Learning Outcomes* (SOLO) model has been used by many researchers to identify levels of student reasoning on different concepts. The SOLO model is based on the assumption that development can be represented in hierarchical structures. In a similar way to the stage theory of development of Piaget, five modes of representation are postulated in SOLO model: Sensorimotor (from birth), Ikonc (from around 18 months), Concrete symbolic (from around six years), Formal (from around 14 years) and Post-formal (from around 20 years). “Modes,

then, are levels of abstraction, progressing from concrete actions to abstract concepts and principles, which form the basis of the developmental stages”(Biggs & Collis, 1991, p. 62).

Within each mode and related to a task in a conceptual web, levels of reasoning that progress from incompetence to expertise can be identified. At the *Prestructural* level (P) responses only show that students engage to task but do not use any relevant aspect to its solution. In the *Unistructural* level responses that have one relevant aspect to the task solution are classified. The responses at *Multistructural* level present more than one relevant aspect but without integrate them. For the *Relational* level responses integrate in a coherent way more than one relevant aspect to the task. Finally, at the Extended Abstract (EA) level responses would show a higher abstract response. A hierarchy that describes levels of increasing complexity in a students' reasoning is obtained.

Some relevant aspects of the solution to the task are chosen according to an a priori analysis, however, is in the process of analyzing data that emerge unanticipated patterns which reveal peculiar forms of reasoning of students. It should be clarified that in the data analysis, the responses of the students are classified in different levels but are not intended to classify students. Finally we subscribe the Watson commentary about that SOLO “[...] is a useful model because it stresses that it is observed in students' responses that is analyzed, not what the observer thinks the student might have meant” (Watson, 2006, p. 13).

Method

The participants were 82 students, the teacher in charge of both groups, and the researchers who are the authors of this work. Students (aged 14 to 16) belonging to two different ninth grade groups in a private school in Mexico City (In Mexico 9th grade is the last year of middle school). Two problems were designed to explore the reasoning of the students; the first is related to a game where the risk is represented by possible losses, the second is related to life time after follow different treatments.

Problem 1

In a fair, the attendees are invited to participate in one of two games, but not in both. In order to know which game to play, John observes, takes note and sorts the results of 10 people playing each game. The losses (-) or cash prizes (+) obtained by the 20 people are shown in the following lists:

Game 1:

15	-21	-4	50	-2	11	13	-25	16	-4
----	-----	----	----	----	----	----	-----	----	----

Game 2:

120	-120	60	-24	-21	133	-81	96	-132	18
-----	------	----	-----	-----	-----	-----	----	------	----

a) If you had the possibility of playing only one of the two games, which one would you choose? Why?

b) In which of the two games do the data have more variability? Why?

Problem 2

Consider you must advice a person who suffers from a severe, incurable and deathly illness, which may be treated with a drug that may extend the patient's life for several years. It is possible to choose between three different treatments. People show side effects to the medication; while in some cases the drugs have the desired results, in some others the effects may be more favorable or more adverse. The following lists show the number of years the patients have lived after being treated with one of the different options; each number in the list corresponds to the time in years a patient has survived with the respective treatment. The graphs corresponding to the treatments are shown after.

Time in years (Treatment 1)	Time in years (Treatment 2)	Time in years (Treatment 3)
5.2	6.8	6.8
5.6	6.9	6.8
6.5	6.9	6.9
6.5	7.0	7.0
7.0	7.0	7.0
7.0	7.0	7.1
7.0	7.1	7.1
7.8	7.1	7.1
8.7	7.2	7.2
9.1	7.4	7.4

a) What kind of treatment would you prefer (1, 2 ó 3)? Why?

b) In which of the treatments there is more variability? Explain your answer

An activity was designed and developed during two teaching sessions of 50 minutes each. The students solved some problems where they have to analyze data and calculate means, ranges and mean deviation but without learned more widely their meanings. The two problems were administered before and after the activities; however the analysis presented below is made taken together all data. Some remarks about the differences between frequencies of responses of pre-test and post-test are mentioned at the end.

Results

In this section we present examples of responses to questions 1a, 1b, and 2a described above, in order to show the kind of answers that were classified at each SOLO level. At the end of the examples of each question a table with the frequencies of responses of each level is presented.

Examples of Responses to Question 1a

If you had the possibility of playing only one of the two games, which one would you choose? Why?

Prestructural level. In the Figure 1 response, the student chose game 2, “because you win more”, but there is no evidence of how the data are used, although one can assume that the answer is motivated by something the student perceived in them. The answer is circular because the questions: What game would you choose?, and: In what game you win more?, are equivalent. This kind of responses provides no progress in understanding the situation.

a) Si tienes la posibilidad de participar en un solo juego ¿Cuál juego elegirías?

El juego 2

¿Por qué? porque segan? más

Figure 1. Prestructural level answer.

Unistructural level. In the response of the Figure 2, the student chooses Game 1 “because you can also lose as in game two but fewer and you risk less”, that is he thinks that in game 2 he loses less and provides an indication that he compares the loss of both games (probably the minimum of each set or the sum of losses) and also alludes to risk. Indeed, it is possible that the way in which the student approaches the problem is influenced by risk aversion, since he skews his attention toward the losses, ignoring the information that provides positive gains.

a) Si tienes la posibilidad de participar en un solo juego ¿Cuál juego elegirías?

Juego 1

¿Por qué? porque al igual que el dos tiene perdidas, pero en menor cantidad + no te arriesgas tanto

Figure 2. Unistructural level answer.

In the response of Figure 3, the student chose Game 1 arguing “because you win more and if you lose, do not lose too much”. In addition he wrote in the margin of each data list the result of the sum of the values, but with an error in the first sum. Although we considered that responses, in which students

add and compare the gains of each game, are Multi-structural, in this case, the answer is located in Uni-structural, taking into account that the student made a mistake in the sum.

Juego 1:	15	-21	-4	50	-2	11	13	-25	16	-4 = 57
Juego 2:	120	-120	60	-24	-21	133	-81	96	-132	18 = 49

a) Si tienes la posibilidad de participar en un solo juego ¿Cuál juego elegirías?

en el juego 1

¿Por qué? porque ganas mas y si pierdes no pierdes mucho

Figure 3. Unistructural level answer (mistake in the sum).

Multistructural level. In response of the Figure 4 the student chooses Game 2, and argues “If you win you get more money but if you lose, lose more”. There are features in his worksheet showing that he added and compared the values of each data list. However, he does not draw the conclusion that both games yield the same profit, instead he put attention on the extreme values, since effectively in Game 2 it is possible win a lot, but also lose a lot. This response reflects risk seeking. The response is classified in Multistructural because the student considered two relevant aspects of the problem: a pre-figuration of the arithmetic mean and the range, but without relate them in a convenient way.

Juego 1:	15	-21	-4	50	-2	11	13	-25	16	-4 49
Juego 2:	120	-120	60	-24	-21	133	-81	96	-132	18 49

a) Si tienes la posibilidad de participar en un solo juego ¿Cuál juego elegirías?

En el 2º

¿Por qué? si ganas es mas dinero aunque si pierdes, pierdes mas.

Figure 4. Multistructural level answer (risk seeking).

The response of Figure 5, student chose the game 1, “because you lose less money or the quantities are lower”. This response is similar to the previous because the student adds and realizes that the results of both list are equal, but like before, does not draw the conclusion that in both games, on average, the same profit is obtained. He also pays attention to extreme values, but now chooses Game 1, probably due to risk aversion.

Juego 1:	15	-21	-4	50	-2	11	13	-25	16	-4 49
Juego 2:	120	-120	60	-24	-21	133	-81	96	-132	18 49

a) Si tienes la posibilidad de participar en un solo juego ¿Cuál juego elegirías?

El primero

¿Por qué? Por qe pierdes menos dinero o los contenidos son menores

Figure 5. Multistructural level answer (risk aversion).

Relational level. There was no response classified in this level. They would be responses in which the students realize that both games have the same utility but are different respect to risk, and then they made a decision according his risk preferences.

Percentages by level:

Level	Pre-test frequency	Post-test frequency
Pre-structural	77 (89 %)	54 (70 %)
Uni-structural	6 (7 %)	4 (5 %)
Multi-structural	3 (3 %)	19 (25 %)
No response	1 (1 %)	
Total	87	77

Table 1. Number of responses by level

With respect to Table 1, frequency of pre-structural responses decreased in post-test respect to pre-test while that of the multi-structural responses increased. The main reason is that there was greater use of the procedure for adding the amounts of each list, probably as a result of the students solved problems that used the arithmetic mean in the episodes of instruction. However, in the answers to post-test, students only added the amounts of each list without actually calculating the arithmetic mean.

Examples of Responses to Question 1b.

In which of the two games do the data have more variability? Why?

Prestructural level. We found answers in which the student makes a choice of a game but the argument offered would lead to the election of the other or such argument does not correspond with the choice made. We have classified these responses in pre-structural. For example, in Figure 6 the student chose the game 2 because "there are all different numbers"; however this is not true.

b) ¿En cuál de los dos juegos hay más *dispersión* de los datos? Juego 2

¿Por qué? hay numeros todas diferentes

Figure 6. Prestructural level answer.

In Figure 7 the student chooses Game 1 because "there is a greater difference [in the game 1] between the highest datum and lowest datum [than in the game 2]", but actually meets this property the game 2 and not 1. Other responses classified in prestructural give vague arguments as "because there are very low data and very high data" or "because there are very large quantities"

b) ¿En cuál de los dos juegos hay más *dispersión* de los datos? en el primero

¿Por qué? Porque existe mayor diferencia entre el dato mayor y el menor

Figure 7. Prestructural level answer.

Unistructural level. In the Unistructural level we have classified responses, whose argument provides evidence that the student observed some trait in data related to dispersion and his choice is consistent. For example, in response of Figure 8, the student chooses Game 2, "because there are greater differences between the amounts". In this response the differences compared are not specified, but the fact of having chosen the game whose data are more scattered, indicates that probably the range of each game, or the sum of the differences between successive data, was considered.

b) ¿En cuál de los dos juegos hay más *dispersión* de los datos? en el segundo

¿Por qué? porque se ve más la diferencia entre las cantidades

Figure 8. Unistructural level answer.

Multistructural level. The responses in which game two was chosen as the most dispersed and the decision was argued by comparing the ranges of both data sets were classified as Multistructural. In Figure 9 the student chose the game 2 and argued “the difference between the maximum and minimum is different in each game” and in his worksheet wrote $133 + 132 = 265$ indicating that he considered at least the range of game 2 data.

Juego 1:
15 -21 -4 50 -2 11 13 -25 16 -4

Juego 2:
120 -120 60 -24 -21 133 -81 96 -132 18

b) ¿En cuál de los dos juegos hay más *dispersión* de los datos? en el 2

¿Por qué? la diferencia del menor dato al mayor dato en cada juego es diferente ~~por lo tanto en el 2 es de 265~~

133
132
265

Figure 9. Multistructural level answer

Relational level. There is no response classified in this level.

Level	Pre-test frequency	Post-test frequency
Prestructural	74 (85 %)	48 (62 %)
Unistructural	9 (11 %)	14 (18 %)
Multistructural	2 (2 %)	13 (17 %)
No response	2 (2 %)	2 (3 %)
Total	87	77

Table 2. Frequencies by level

Examples of Responses to Question 2a.

What kind of treatment would you prefer (1, 2 or 3)? Why?

Prestructural level. The arguments of the pre-structural responses do not provide evidence that students have taken into account the data or how the data were used to make the decision. For example, in one response where treatment X was chosen, the student argues "there is more time and quality of life". In a class of responses where treatment 1 was chosen, the argument was "because you can live more" but without clarifying the relationship with the data. In another class of responses where treatment 2 was chosen, the arguments alluded to stability, regularity or control, for example, "because it is more stable and it can be said that is more effective"

Unistruktural level. In the Unistruktural level have been classified responses where is taken into account only a relevant datum and the treatment chosen is consistent with the argument. A kind of answers in this level are based on consideration of an extreme value, for example, in Figure 10 a student chooses treatment 2 and argues "because there are likely to live 6.8 years or more", that is, compares the minimum values of each treatment and selects the greatest of them.

a) ¿Qué tipo de tratamiento preferirías (1, 2 ó 3)? 2
 ¿Por qué? porque vives años buenos años y hay mucha probabilidad de que te salga 6.8 para arriba

Figure 10. Unistruktural level answer.

In the Unistruktural level we have also included responses in which the decision is based on the observation of the modes, and the largest of them is chosen, for example, a student chooses treatment 3 "because 3 patients reach to live 7.1 years with this treatment". Finally, another type of response included at this level consists of adding the times of each treatment, but with errors on the sum, so that students do not realize of the equality of results. Then they choose the treatment in which data accumulate the greatest amount. For example, in Figure 11 is shown that the student got 71, 70.4 and 70.4 as a result of summing the data of Treatments 1, 2 and 3 respectively. Consequently he chose treatment 1, "Because with this treatment you live a longer time"

Tiempo en años (Tratamiento 1)	Tiempo en años (Tratamiento 2)	Tiempo en años (Tratamiento 3)
5.2	6.8	6.8
5.6	6.9	6.8
6.5	6.9	6.9
6.5	7.0	7.0
7.0	7.0	7.0
7.0	7.0	7.1
7.0	7.1	7.1
7.8	7.1	7.1
8.7	7.2	7.2
9.1	7.4	7.4

71 70.4 70.4

a) ¿Qué tipo de tratamiento preferirías (1, 2 ó 3)? El 1.
 ¿Por qué? Porque has más tiempo de año de vida

Figure 11. Unistruktural level answer (mistake in the sum).

Multistruktural level. In the multi-structural level responses were divided into two types. The first type consists of the answers that consider the two extreme values of each dataset. For example, in Figure 12 is shown that the student chose the treatment 3 "Because maybe I will not live nine years but I have secured from 6.8 to 7.4". Although not mentioned, it appears that the student perceives the risk involved in the first treatment (the possibility to live only 5.2 years) because it gives up the opportunity to live 9, "ensuring" live at least 6.8 years

a) ¿Qué tipo de tratamiento preferirías (1, 2 ó 3)? Tratamiento 03
 ¿Por qué? Porque tal vez no viva 09 años más
pero así tengo asegurados de 6.8 a 7.4

Figure 12. Multistructural level answer (considering extreme values of each dataset).

In the second type of responses, the values of each data set are summed and the results are compared, getting to the conclusion that is the same follow any of the three treatments. In Figure 13, the student said that any of the three treatments can be chosen "because I added the data from each treatment and I got 70.4 and then all are better".

Tiempo en años (Tratamiento 1)	Tiempo en años (Tratamiento 2)	Tiempo en años (Tratamiento 3)
5.2	6.8	6.8
5.6	6.9	6.8
6.5	6.9	6.9
6.5	7.0	7.0
7.0	7.0	7.0
7.0	7.0	7.1
7.0	7.1	7.1
7.8	7.1	7.1
8.7	7.2	7.2
9.1	7.4	7.4
70.4	70.4	70.4

a) ¿Qué tipo de tratamiento preferirías (1, 2 ó 3)? cualquiera de los 3
 ¿Por qué? porque sume cada número de
cada tratamiento y me salió 70.4
y entonces todos son mejores

Figure 13. Multistructural level answer (the values of each data set are summed).

Relational level. In the relational level we include responses that supported its decision taking into account center and dispersion of each dataset. Only two cases were included in this level. In Figure 14, treatment 3 was chosen "because it is more likely to live 7.1 years or less, but the results are not so far apart and are more likely to live from 6.8 to 7.2 years".

a) ¿Qué tipo de tratamiento preferirías (1, 2 ó 3)? El tratamiento 3
 ¿Por qué? Es más seguro que viva 7.1 o menos pero ~~no está~~ los resultados
no estén tan alejados y es más seguro que viva de 6.8 a
7.2 años

Figure 14. Relational level answer.

Level	Pre-test frequency	Post-test frequency
Prestructural	56 (64 %)	47 (61 %)
Unistructural	21 (24 %)	12 (15 %)
Multistructural	10 (11 %)	16 (21 %)
Relational		2 (3%)
Total	87	77

Table 3. Frequencies by level

Conclusions

On the Table 4, we can see that there was in average a decreasing of 15% in Prestructural level from the pre-test to post-test. Several students answered in post-test using data while in pre-test they gave circular arguments. The frequencies of responses classified in Unistructural level in average decreased in 1.34%, although the frequencies to the question 1b increased from 9 to 14. This is because some students improved their perception of dispersion in data, but they were unable to relate this notion with choosing an option in the other problems. An increase in average of 15.67% from pretest to posttest was registered in the frequencies of responses classified at Multistructural level. Most of new responses classified at this level were based on the procedure of seeking the game or treatment where the sum of the data is greater than the sum of the data of the other game; in most of these the conclusion was that the games were equivalent. In the rest of responses in this level students took into account the maximum and the minimum of each set, then they perceived the risk and choosing according to their risk attitudes (risk aversion or risk seeking); in these cases no mention of mean was done. In only two responses to question 2 of the posttest, classified in Relational, students took into account the center (mode) of the sets and observed that one of these involved greater risk; for this, students compared the maxima and minima.

Level	Question 1a		Question 1b		Question 2a	
	Pre-test frequency	Post-test frequency	Pre-test frequency	Post-test frequency	Pre-test frequency	Post-test frequency
Pre-structural	77 (89 %)	54 (70 %)	74 (85 %)	48 (62 %)	56 (64 %)	47 (61 %)
Uni-structural	6 (7 %)	4 (5 %)	9 (11 %)	14 (18 %)	21 (24 %)	12 (15 %)
Multi-structural	3 (3 %)	19 (25 %)	2 (2 %)	13 (17 %)	10 (11 %)	16 (21 %)
Relational	-	-	-	-	-	2 (3%)
No response	1 (1 %)		2 (2 %)	2 (3 %)		
Total	87	77	87	77	87	77

Table 4. Frequencies by level

Our main result is a hierarchy SOLO which describes the patterns of reasoning emerging from the students' efforts to respond to problems. In Table 5 such hierarchy is presented.

Hierarchy SOLO for problems

In general, predominate Prestructural responses; students whose answers are classified at this level understand what is asked them and they make a choice but fail to use the data to support their preferences. However, there are students who see in a single value of each data set (maximum, minimum or mode) a key to make a decision. These responses have been classified in the Unistructural level; they prefigure the valid scheme of solution. The value chosen is one that students consider a representative of the set. In responses of Multistructural level, a step forward towards the solution scheme is given, since more than one value of each data is considered. Two main strategies were identified: 1) compare the sum of values of each set data, 2) take into account the maxima and minima. Each of these strategies is an early or primitive form of the two main statistical tools of the case: the mean and dispersion. The

second strategy led some students to perceive the risk. Finally, In Relational level responses, both strategies are used and the decision is made according of attitude towards risk. In these responses a scheme of solution is complete.

Level	Description
Prestructural	One option is chosen but without justification or with a circular response as “because is the better”.
Unistructural	The maximum, minimum or mode of each data list is observed and compared
Multistructural	The sums of the data lists, or mean, are compared or both the maximum and minimum are considered and, in this case, risk is perceived.
Relational	The sums or means of the data lists are compared and range (or another measure of dispersion) are considered; the final decision is influenced by preferences about risk

Table 5. SOLO hierarchy

As a final commentary we would like to comment that the indeterminacy of the answers to many of the questions that emerge in Probability and Statistics may be a major cause of the frequent distrust and even rejection toward the discipline by students. In general, an indeterminate response is not considered a satisfactory one; however, good answers that are found in probability and statistics are in some way indeterminate due to the random nature of the phenomena modeled. In the problems that we have reviewed may be not convincing that the choice of a game or treatment is not completely determined by the behavior of the data, but also depends on the solver attitude towards risk. This relativity disturbs those who believe that science must give absolute and conclusive answers to the problems that arise on it. Relativity of the responses may obscure the main point which is that the analysis of the ranges, and more generally the analysis of variation, provides information about the risks involved and therefore helps to make rational decisions. The use in teaching of problems as those treated in this study can help the students to construct schemes for assessing the results of the statistical analysis and help them to retreat from certainty in a profitably way.

References

- Bakker, A. & Derry, J. (2011). Lessons from inferentialism for statistics education. *Mathematical Thinking and Learning*, 13, 5-26.
- Biggs, J. B. and Collis, K. (1982) *Evaluating the Quality of Learning: the SOLO taxonomy*. New York: Academic Press.
- Biggs, J. & Collis, K. (1991). Multimodal learning and the quality of intelligent behaviour. In H. Rowe (Ed.), *Intelligence, Reconceptualization and Measurement* (pp.57–76). New Jersey: Laurence Erlbaum Assoc.
- Brandom, R. (2000). *Articulating Reasons: An Introduction to Inferentialism*. Cambridge MA: Harvard University Press.
- Burrill, G., & Biehler, R. (2011). Fundamental statistical ideas in the school curriculum and in training teachers. In C. Batanero, G. Burrill, & C. Reading (Eds.), *Teaching statistics in school mathematics Challenges for teaching and teacher education: A joint ICMI/IASE Study* (pp. 57-69). New York: Springer.
- Garfield, J., & Ben-Zvi, D. (2008). *Developing students' statistical reasoning: Connecting research and teaching practice*. New York: Springer.

- Heitele, D. (1975). An epistemological view on fundamental stochastic ideas. *Educational Studies in Mathematics*, 6, 187-205.
- Kahneman, D. & Tversky, A., (2000). Choices, values, and frames. In D. Kahneman & A. Tversky (Eds.), *Choices, Values, and Frames*. Cambridge: Russell Sage Foundation.
- Lehrer, R., & Schauble, L. (2007). Contrasting emerging conceptions of distribution in contexts of error and natural variation. In M. Lovett & P. Shah (Eds.), *Thinking with data* (pp.149-176). Mahwah, NJ: Lawrence Erlbaum Associates.
- Moore, D. (1990). Uncertainty. In L. A. Steen (Ed.), *On the shoulders of giants: New approaches to numeracy* (pp. 95-137). Washington, DC: National Academy Press.
- Petrosino, A. J., Lehrer, R., & Schauble, L. (2003). Structuring error and experimental variation as distribution in the fourth grade. *Mathematical Thinking and Learning*, 5(2&3), 131-156.
- Reading C. (2004). "Student Description of Variation While Working with Weather Data", *Statistics Education Research Journal*, Vol. 3(2), 84-105.
- Shaughnessy, J.M. (1997). Missed opportunities in research on the teaching and learning of data and chance. In F. Biddulph and K. Carr (Eds.), *People in mathematics education* (Vol. 1, pp. 6-22). Waikato, New Zealand: Mathematics Education Research Group of Australasia.
- Watson, J.M., & Moritz, J.B. (2000). Developing concepts of sampling. *Journal for Research in Mathematics Education*, 31, 44-70
- Watson, J. M., & Kelly, B. A. (2004). Statistical variation in a chance setting: A two-year study. *Educational Studies in Mathematics*, 57, 121-144.
- Watson, J. M. (2006). *Statistical literacy at school: Growth and goals*. Mahwah, NJ: Lawrence Erlbaum.
- Wild, D. J., & Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. *International Statistical Review*, 67(3), 223-265.