RANDOM WALKS IN TEACHING PROBABILITY AT THE HIGH SCHOOL

Hugo Mael Hernandez Trevethan[1], Verônica Yumi Kataoka[2] and Marcelo Silva de Oliveira[3]
[1]Sciences and Humanities School - National Autonomous University of Mexico, Mexico
[2]Bandeirante University of São Paulo, Brazil
[3]Federal University of Lavras, Brazil
animal_estocastico@hotmail.com

Since there's a need of supporting Probability teaching, it is important to present to teachers different didactical activities, to support their practice. In this context, the "Colegio de Ciencias y Humanidades (CCH)" in Mexico, offered a course for a pedagogical formation and even conceptual support to their recent-hired Stats teachers. In 2009, two professors were selected for being guided by a researcher and responsible of the course, to apply in their regular classes a didactical sequence for Probability teaching, named Mônica's random walks. This work aims to present and discuss this didactical sequence applied to high school students in CCH. This activity allows the students to work with diverse probability concepts, such as, difference between random and deterministic experiments, theoretical and frequentist probability, and is a recommended activity to help teachers in improving the level of literacy probability of their students.

INTRODUCTION

In order to do a better understanding of the many random phenomena in daily life, we need to learn, even at elementary school, some probability concepts. In addition, probability is used in the context of statistical inference, by helping in hypothesis verification, and therefore, in making decisions. However, according to Kataoka et al. (2008), not all probability concepts are easily understood at first, as many of them are rather abstract, requiring caution and skill from the teacher during the process of teaching - learning, in order to provide pupils with the proper development of their probability reasoning.

On the other hand, Peck and Gould (2005), Ainley and Monteiro (2008) mentioned pedagogical difficulties for probability and statistics teaching by mathematics teachers in basic school courses, since many of them sometimes had a basic education in probability and statistics but are not always prepared to teach such contents. In many cases, they did not get any courses related to the didactic in statistics.

In this context, the Colegio de Ciencias y Humanidades (CCH, Mexico), as an education institution, is concerned about teachers' in-service education. In Mexico, most of the high-school teachers receive professional education different from a major in education; this means that most of the mathematics teachers are mathematicians, engineers, or actuaries. Most of the in-service teachers' needs relate to pedagogical education, including conceptual support. For statistics teachers, this last point is more evident than in other areas of knowledge, since most of them have a very weak understanding in this science.

As a part of developing this understanding, CCH gives to their recently hired statistics teachers a 40-hours course named "Formation course for recently hired teachers in Statistics and Probability". The teachers in the course were between 25 and 30 years old, with no more than three years of experience in teaching CCH. The course passes through all the subjects contained in the scholar program of statistics and probability, which includes descriptive statistics, bivariate data, probability, distributions, sample distributions and inference. It also tries to show didactical sequences for all those topics.

In 2009, this course was conducted under the guidance of two researchers, both experienced teachers in statistics at CCH (one of them, the first author of this paper, who will be referenced here as researcher-1). For each of the lectures, every topic was seen under the idea of problem solving and simulation – both physical and virtual. Specifically, in probability, within a two-hour session, there was applied a didactical sequence, named Mônica's random walks.

During the application of the didactical sequence in the course, both researchers selected two professors to later apply the same activity it in their regular classes, to their students. The objective was to verify if the sequence could, in fact, support those teachers in their teaching of basic probability concepts, and later during introduction to the binomial distribution to verify if this

In C. Reading (Ed.), *Data and context in statistics education: Towards an evidence-based society. Proceedings of the Eighth International Conference on Teaching Statistics (ICOTS8, July, 2010), Ljubljana, Slovenia*. Voorburg, The Netherlands: International Statistical Institute. www.stat.auckland.ac.nz/~iase/publications.php

formation course is really working as an important part of those teachers' understanding in didactics. In order to do so, researcher-1 went through four stages with teachers: application of the activity in teachers' course (stage 1); application of the activity at teachers' classroom (stage 2); institutionalization of the probabilistic concepts seen in the activity with students (stage 3); and the final stage (stage 4) that is about to happen along the first semester in 2010, which is introducing binomial distribution at the classroom.

In this context, this work aims to relate just the second stage of this general study, by presenting and discussing the didactical sequence "Mônica's random walks" applied with the high school students "Colegio de Ciencias y Humanidades" in Mexico. It is worth indicating that application of the activity occurred also at researcher-1's classroom, in order to show to teachers the use of the sequence *in situ*, and to discuss, with more information in hand, results of both teachers' groups, and for them to have the chance of being observers at researcher-1's classroom along stages three and four.

MÔNICA'S RANDOM WALKS

According to Batanero and Godino (2002), the construction of probability concepts should start from the understanding of three basic notions: perception of chance; the idea of random experience; and a notion of probability. Likewise, Lopes (2003) state that it is desirable that the teacher would approach such concepts through activities in which pupils make experiences and observe the events, promoting their intuitive expression of chance and uncertainty, constructing, from the results, mathematical methods for the study of those phenomena.

Based on all those recommendations, the didactical sequence "Mônica's random walks" was chosen, which allows to work with probability concepts such as events, sample space, single event probability, to explore the difference between random and deterministic experiments, to approach probability from relative frequencies, to calculate theoretical probability from a tree diagram, and to compare observed and expected patterns.

This activity was adapted by Cazorla and Santana (2006) to be used in elementary school, from Fernandez and Fernandez (1999), whom proposed it to teach the binomial distribution to major students.

The sequence was analyzed by Gusmão and Cazorla (2009) by applying it to mathematics teachers, in lectures of teachers' formation, based on the Ontosemiotic theory (Godino, 2002). This analysis showed sequence's variability on teaching basic concepts on probability, so it pointed out different semiotic conflicts, due, mainly, to the lack of previous knowledge in teachers who were facing for the first time some of these concepts. This brought out a major intervention of the researcher and teacher of the discipline.

The full activity has four different sections and 23 questions. In each section, questions must be answered based on a previous action. In the first section, students have to read this story:

"Mônica (the girl on the left corner down) and their friends live in the same quarter. Mônica´s house is four blocks away to the houses of Horácio (dinosaur), Cebolinha (three-haired kid), Magali (girl at center), Cascão (the kid with with braces) and Bidu (dog), according to Figure 1. Mônica used to visit her friends along weekdays on a pre-set order: Monday, Horácio; Tuesday, Cebolinha; Wednesday, Magali; Thursday, Cascão; and Friday, Bidu. For turning those visits more thrilling, the group decided that randomness should pick the friend to be visited by Mônica. To do so, when leaving home and on each crossing, Mônica has to toss a coin; when the head occurs (H), she will walk a block North, and when the tail occurs (T), she will walk a block East. Each toss represents a block of her route. Mônica has to toss the coin four times in order to reach some friend's house".

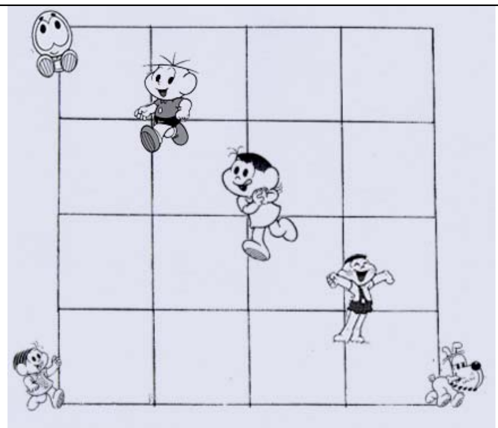

Figure 1. Map of Mônica's random walks

After reading the story, students answer the some questions without tossing the coin. At section II, students carry out the experiment; at section III, they have to construct the tree diagram;

and at section IV, they have to compare theoretical and frequency probabilities, and between the random and deterministic experiment.

METHOD

Participants

Two in-service teachers in the course were selected to apply the sequence at their classes along the second semester of 2009; they were both teachers with a major in mathematics and two-year experience in CCH. Also, a third class of high-school students was selected to solve the strategy. This group was attended by researcher 1 of the course.

The activity was solved by the 91 senior high-school pupils, aged between 16 and 19 years old, and with just a very elementary background in probability.

Procedures and data analysis

The activity was applied in a two-hour session with the 91 students separated in 31 small groups (2 to 5 students, with some groups up to 7 students), as a hands-on activity, without any teacher's intervention.

As previously stated, the complete sequence has 23 questions, but in this paper just 17 questions will be discussed, since the other six questions are instructions about the execution of the experiment, recording results in tables, and constructing a tree diagram. There are 4 questions in the first section, 5 in the second, 3 in the third, and 5 in the fourth.

The students' answers for each question were categorized in different hierarchic levels of knowledge, varying from a non-existing knowledge about the question up to a full knowledge level demanded to solve the question, summarized by Table 1, with exception to question III.P1.

Table 1. Description of categories used to systematize students' answers

Section/Problem	Categories		
	0	1	2
I.P2, I.P4, III.P2, IV.P5	a) No answer b) Wrong answer	Right answer	----------------
II.P3, II.P4, III.P3, IV.P2	a) No answer b) Answer without justification or with inconsistent justification	Answer with consistent justification	-----------------
I.P1, I.P3, IIP.1, II.P2, II.P5, IV.P1, IV.P3, IV.P4	a) No answer b) Wrong answer c) Right answer without justification or with inconsistent justification	Right answer with informal justification	Right answer with formal justification

We developed our categorization based on the work previously presented by Díaz and Batanero (2009), whom classified a wrong answer as 0, partially wrong as 1 and totally correct as 2; and upon the categorization of Hernandez et al. (2009), who used 0 (blank), 1 (wrong or incomplete), 2 (almost right answers) and 3 (right answer). We also adapted that categorization to fit the needs raised from the nature of the responses in our study.

RESULTS

Categorized answers from the 31 student groups to the 7 questions may be observed in Table 2. On question I.P1, we asked if there was any difference between the old and the new way of Mônica visiting her friends, and by observing results in Table 2, we verified that most of the students (67.7%) answered correctly with a formal justification (category 2). We consider a formal justification when the students realize the old way is an example of a deterministic experiment, and the new way is an example of a random experiment, like: "Mônica used to have a pre-set order for visiting her friends, now is indicated by randomness for everyone". An example of a correct answer with informal justification (category 1) is: "In the old way, it was a pre-set day for visiting a friend, it was stated in the new way it is not a sure thing that they will meet, and that she will visit a different friend each day in which some visits could repeat".

Table 2. Percentage development of the students for the 7 presented problems in the instrument

Categories	I				II		III
	P1	P2	P3	P4	P3	P4	P3
0	3.2	22.6	12.9	71.0	58.1	16.1	16.1
1	29.1	77.4	87.1	29.0	41.9	83.9	83.9
2	67.7	---	---	---	---	---	----

The objective in question I.P2 was to present the concepts of events and sample space, since students had to verify, which are the possible outcomes of tossing a coin. 22.6% of the groups did not answer correctly (category 0), and as an example we have: "4^2 = 16 {tttt, ttth, ttht, thtt, tthh, thhh, thth, thtt, hhht, hhhh, hhth, hthh, hhtt, httt, htht, hthh}". This group did not realize what we were asking when tossing one coin.

The objective in question I.P3 was to identify students' concepts in probability and to discuss ideas such as probability, likely, randomness, and classical probability. Most of them (87.1%) answered correctly (category 1). As two examples of an incorrect answer (category 0) we have: "2 out of 4 = 1/2" or "It has only two options of 4 times that the coin is tossed", since we were asking about "chances" of having a head or tail by tossing a coin.

Question I.P4 is repeated in II.P3, II.P4 and III.P3: "Does every friend have the same chance to be visited?" The difference among them was just the moment in the instrument in which this question was asked: I.P4 was asked before the random experiment; II.P3 before systematization of experiment results in a table; II.P5 after the construction of the table; and III.P3 after construction the tree diagram. The objective was to check out if the student would need the experimentation or tree diagram to realize that probabilities were not equal. It is worth to point out that despite analyzing the four answers jointly, in questions II.P3, II.P5 and III.P3, students had to justify their answer in addition to answering yes or no; this explains why answers were categorized differently for I.P4 (Table 1).

We observed at the beginning that 22 groups (71%) answered to I.P4 that chances of being visited by Mônica were equal for each friend. Even without asking for a justification, from our experience using the sequence with other classes, we suppose that most of the students considered probabilities as equal because: coin tosses were at random and independent; distance to any friend's house was the same (four blocks); and probabilities for head and tail are equal.

After the experimentation, at II.P3 we observed that 9 groups changed their minds, and from those 9 groups who answered correctly to I.P4, at the time they justified why probabilities were not equal, 5 groups could not do so and their answers were categorized as 0.

In II.P4, with the table, 13 groups changed their minds and justified correctly; adding this to those answers in II.P3, we had 83.9% of groups realizing that probabilities are not equal and that Magali has greater probability. In III.P3, after constructing the tree diagram, an additional change happened. We noticed that most of the students needed to carry out experiments to get to the right answer; this shows that the concepts of equal probability and single events are a strong part of the way the student understands the concept of probability. We assumed that the students did not aim to identify that in this case the events were composed.

Results of the other 9 questions in the instrument may be observed in Table 3.

In question II.P1, students, after experimentation, should answer who has greater probability of being visited, Magali or Horácio. In this question we considered a correct answer with informal justification (category 1), when the students based their answer just on the experiment, what happened with 25.8% of the groups. And a correct answer (54.8% of the groups) with formal justification was considered when some theoretical aspect was involved, such as: "Magali, since it is less likely to have the same side on the coin four times".

In question II.P2, we asked if there is any friend with a chance of not being visited by Mônica, and just like II.P1 we considered a correct answer with informal justification when it was justified on the experiment, as was the case of 45.2% of the groups, with answers like: "Yes. The combination (TTTT) for visiting Bidu never appeared".

Table 3. Percentage development of the students for the 9 presented problems in the instrument

Categories	Section/Problem								
	II			III	IV				
	P1	P2	P5	P2	P1	P2	P3	P4	P5
0	19.4	16.1	48.4	38.7	38.7	45.2	41.9	16.1	25.8
1	25.8	45.2	35.5	61.3	48.4	54.8	41.9	58.1	74.2
2	54.8	38.7	16.1	---	12.9	---	16.1	25.8	---

The objective in question II.P5 was to demonstrate differences among the experimental results of the groups. So, we asked the students to trace the bar graph for their observed frequencies for the visits and to compare with their classmates, in order for them to answer if the results were equal. Most of the groups realized that the results could be different because of the variability involved in experimentation. Nevertheless, just 16.1% of them were able to explain this situation accurately. Among those whom answered incorrectly, we noticed some answers stating that results may be the same since for all of them Magali has the highest probability.

After constructing the tree diagram, we asked the students to answer III.P1 involving how many paths does Mônica has in total. This question was not categorized since just one group failed on it by answering “62”. The rest of the groups simply answered 16, or also presented some mathematical calculation based on combinatory (just one group).

At question IIIP.2, we were asking the students to find out a relation on the paths heading to each friend’s house. 61.3% of the students found out correctly that the number of paths depended on the number of heads (or tails) in the four tosses. In addition, some of the groups reported that that they did not understand the question.

The question IVP.1 asked the students to compare both ways of calculating probabilities, frequency and theoretical. Just 12.9% of the students aimed an answer with a formal justification (category 2). Students classified as category 1 (48.9%), justified difference between both, saying that theoretical probability is more "exact", maybe limited by the high variability found in the groups’ experiments, since in II.P5 they compared results.

In IV.P2 we asked which of two ways for calculating probabilities was better and why. Like in IV.P1, justification on category 0 (45.2% of the groups) was based on exactitude. When we asked students in question IV.P3 to trace the bar graph for theoretical probability and to compare their results with their classmates for saying if outcomes were equal or not, again many groups (41.9%) justified that the graphs were equal by using the phrase “more exact”. Just 25.8% of the groups aimed to answer correctly with a formal justification, like: “They are equals, since the results come from the tree diagram”.

In question IV.P4 we asked if they find this new way of Mônica visiting her friends was equal. Most of the groups (58.1%) realized that this new form was not equal (many justifications were about having some friends with greater probabilities than others), but just 12% gave a formal justification. We’ve considered a formal justification when students based their arguments on the idea of “possibilities” or “probabilities”.

In question IVP.5, students had to propose a new way for Mônica visiting her friends, in case they found the new way not equal. Answers to this question were divided in two categories, 0 and 1, but category 1 was divided as 1.1 and 1.2, given that both of them are right answers, one based on a deterministic proposal, and the other based on a random proposal with a uniform distribution. At category 1.1 we have 10 groups out of 31, and at category 1.2 we have 13 groups. Examples of this are: “The former method, one day each”, and “Putting papers with the names in one bag and taking out one for each day”. A wrong answer would be: “We think this one is the only possible way”.

After application of the activities and systematization of the results, researcher-1 assessed students’ answers jointly with teachers; the following results were important in the analysis:

- Even when the activity was applied without teachers’ intervention and knowing that students had not worked deeply with probability in the scholar semester, it was possible to identify which previous knowledge’s related to the probability concepts in the activity the pupils had, and if those were coherent with a possible background in probability previously experienced in elementary and high school.

• Identifying what were the main misconceptions.

Every given discussion about each group's responses will be used to plan the intervention that teachers will have at the moment of institutionalize probabilistic concepts (3rd stage), and later on when introducing the concept of Binomial Distribution, along the first semester 2010 (4th stage).

FINAL CONSIDERATIONS

In a global analysis, we observed that there was no difference among results of the three classes; students were highly motivated and relaxed along the experiment (tossing a coin), aimed to answer all the questions and realized the importance of some probability concepts. And despite many of them presenting many informal justifications, we consider that, generally, students understood the main objective on the activity, namely to present differences between deterministic and random experiment, and between theoretical and frequency probabilities as well.

In addition, we consider that the activity can actually contribute to the teaching of basic concepts in probability, and therefore, to help students in improving their level of probabilistic literacy.

REFERENCES

Ainley, J., & Monteiro, C. (2008, June). Comparing Curricular Approaches for Statistics in Primary School in England and Brazil: A Focus on Graphing. *Proc: JOINT ICMI/IASE STUDY - Teaching Statistics in School Mathematics: Challenges for teaching and teacher education*. Monterrey, México.

Batanero, C., & Godino, J. (2002). *Estadística y su didáctica para maestro*s: Proyecto Edumat maestros (Stochastics and its didactics for teachers: Edumat-Teachers project). Granada, Universidad de Granada. (in Spanish)

Cazorla, I. M., & Santana, E. R. dos S. (2006). *Tratamento da Informação para o Ensino Fundamental e Médio (1rd. ed., v.1)*. Itabuna, Bahia, Brasil: Via Litterarum (in Portguese).

Diaz, C., & Batanero, C. (2009). University students' knowledge and biases in conditional probability reasoning. *International Eletronic Journal Mathematics Education*, v. 4, 131-162.

Dirección General del Colegio de Ciencias y Humanidades (2004). Programa Estadística y Probabilidad I y II. CCH-UNAM; México (in Spanish).

Fernandez, D. W. X., & Fernandez, D. X. (1999). O prazer de aprender probabilidade através de jogos: descobrindo a distribuição Binomial. *Anais da Conferencia Internacional "Experiências e Expectativas do Ensino de Estatística – Desafios para o Século XXI*. Florianópolis, SC (in Portguese).

Godino, J. D. (2002). Un enfoque ontológico y semiótico de la cognición matemática. *Recherches en didactique des Mathématiques*. 14(3), 325-355. (in Spanish)

Gusmão, T., & Cazorla, I. (2009). Uma análise semióticados passeios aleatórios da Mônica: atividade para ensinar conceitos básicos de probabilidade. *Anais do IV Seminário Internacional de Pesquisa em Educação Matemática.* SBEM, Taguatinga (in Portguese).

Hernandez, H. M. T., Silva, C. B., Cazorla, I. & Kataoka, V. Y. (2009). (In)Formal aspects of Independent Events. *Proceedings 57th Session Intern. Statistical Institute*. Intern. Stat. Institute, Durban.

Kataoka, V. Y., Souza, A. A., Oliveira, A. C. S., Fernandes, F. M. O., Paranaíba, P.F., & Oliveira, M.S. (2008). Probability Teaching in Brazilian Basic Education: Evaluation and Intervention. In*: 11th International Congress on Mathematical Education*, Monterrey.

Lopes, C. A. E. (2003). *O Conhecimento Profissional dos professores e suas relações com Estatística e Probabilidade na Educação Infantil*. PhD thesis. Univ. Est. Campinas, Campinas, Brasil . (in Portguese).

Peck. R., & Gould, R. (2005). Preparing secondary teachers to teach statistics: A distance education model. *Proceedings 55th Session Intern. Statistical Institute*. Intern. Stat. Institute, Sydney.