TEACHING PROBABILITY WITH THE SUPPORT OF THE R STATISTICAL SOFTWARE

ROBSON DOS SANTOS FERREIRA
Universidade Bandeirante Anhanguera, Brazil
robson1santos@yahoo.com.br

VERÔNICA YUMI KATAOKA
Universidade Estadual de Santa Cruz, Brazil
veronicayumi@terra.com.br

MONICA KARRER
Universidade Bandeirante Anhanguera, Brazil
mkarrer@uol.com.br

ABSTRACT

The objective of this paper is to discuss aspects of high school students' learning of probability in a context where they are supported by the statistical software R. We report on the application of a teaching experiment, constructed using the perspective of Gal's probabilistic literacy and Papert's constructionism. The results show improvement in students' learning of basic concepts, such as: random experiment, estimation of probabilities, and calculation of probabilities using a tree diagram. The use of R allowed students to extend their reasoning beyond that developed from paper-and-pencil approaches, since it made it possible for them to work with a larger number of simulations, and go beyond the standard equiprobability assumption in coin tosses.

Keywords: *Statistics education research; Probabilistic literacy; Constructionism; Design experiment, R software*

1. INTRODUCTION

The process of learning probabilistic concepts requires skills from students that go beyond traditional school tasks because, according to Borovcnik and Kapadia (2009), conceptual errors in probability can affect important personal decisions related to daily life. These authors also consider that knowledge in this area is an essential tool for the understanding of inferential statistics. Thus, concepts of probability must be addressed to enable students to understand and apply them critically in different contexts – to develop their literacy in probability (Gal, 2005). To achieve these goals, Coutinho (2001) and Batanero and Godino (2002) recommend that development of ideas of probability should be based on understanding of three basic notions: perception of chance. the idea of randomness. and the interpretation of probability.

Batanero and Godino (2002) point out that we should provide students with conditions to observe the unpredictability of each result separately, in a random experiment, highlighting the variability in small samples by comparing the results obtained by each student. However, the student must become aware of the convergence phenomenon (law of large numbers), looking at the accumulation of results from the whole class and then comparing the reliability of small and large samples.

In this way, the use of computer resources constitutes an important tool for increasing the sample size for random experimentation in the classroom. According to Mills (2002), several researchers have suggested the use of computers in probability teaching, as a way to increase

Statistics Education Research Journal, 13(2), 132-147, http://iase-web.org/Publications.php?p=SERJ

students' ability to understand abstract or difficult concepts. Batanero (2007) emphasizes that the introduction of computers in schools made it possible to carry out simulations that may help students solve simple probability problems, a feature that is not possible using physical experimentation. Lane and Peres (2006) discuss research results that confirm the benefit of random experimentation and computer simulations in the development of active learning, allowing students to construct knowledge.

In Brazil, the National Curricular Parameters (Brasil, 1998) indicate that skills for setting up experiments and simulations for estimating and confirming probabilities should be started in elementary school. According to Borovcnik and Kapadia (2009), simulation is an optimum strategy and, when combined with the use of technology, helps to reduce technical calculations and provides a better focus on the concepts.

Pratt and Kapadia (2010) point out that researchers on probabilistic cognition generally agree that students should be provided with opportunities to work with random generators as well as computer simulators, as this allows them to explore two complementary kinds of mathematical activities: theoretical and empirical. Theoretical activities include classic combinatorial analysis resulting from experiences with random generators, and empirical activities include frequentist experimentation, by hand or by computer.

For Lane and Peres (2006), the use of simulation itself doesn't guarantee an active learning, since the students may be put in the position of passive observers, with lower assimilation of concepts. DelMas, Garfield, and Chance (1999) observed that students performed better when, as well as viewing the computational simulation, they were given a written test about the concepts involved. This may be because the test forced the students to confront discrepancies between expected and observed results. The computational resources should be used in an environment that provides for greater student autonomy in the construction of their own knowledge. Therefore, we chose to utilize the perspective of constructionism which, according to Papert (1980), is characterized by student initiative and control over the environment. In constructionism, learning is understood as the personal construction of knowledge.

Considering these characteristics of constructionism and the probabilistic literacy model proposed by Gal (2005), the present study aims to investigate the learning process of probabilistic concepts by students in the last year of high school. The study involves the teaching experiment 'Passeios Aleatórios da Carlinha' ('Carlinha's Random Walks'), using a paper-and-pencil environment supported by the software R. The results presented here are part of Ferreira's (2011) master's degree research in mathematics education.

This led us to the following research question: "What aspects can be observed when integrating students' learning processes into a computational environment – R software – in order to work with probabilistic concepts, within the perspective of the Gal's (2005) probabilistic literacy model and Papert's (1980) constructionism?"

2. THEORETICAL AND METHODOLOGICAL PERSPECTIVES

Valente (1994) points out that Papert's (1980) constructionism perspective stands out by providing a construction of knowledge based on the practical implementation of an action that produces a palpable product, of personal interest to those who produced it. Inside this environment, according to Papert, the mental construction must be supported by concrete constructions favoring the development of abstraction: there is a dialectical movement between the concrete and the abstract. Under this perspective, the students' control of the process helps their learning, and the teacher's role is to propose new challenges to stimulate them, respecting their cognitive levels of development. The emphasis is on students' learning processes rather than on the product they present.

We believe that in this research, the constructionist approach made important contributions to students' development of probabilistic literacy. According to Gal (2005), a student capable of critically reading and interpreting probabilistic information, and making decisions based on this, may be considered as literate in probability. Gal proposed a model for developing probabilistic literacy composed of two elements: the cognitive knowledge and the dispositional aspects. Here we focused only on cognitive knowledge, comprising five blocks: big ideas; figuring probabilities; language; context; and critical questions.

In order to develop the teaching experiment, we used the didactic activity 'Carlinha's Random Walks' (Cazorla, Kataoka, & Nagamine, 2010); this is an adapted version of 'Monica's Random Walks' (Cazorla & Santana, 2006). Earlier researchers have analyzed 'Monica's Random Walks' from different perspectives, but only in a paper-and-pencil environment. For example, Gusmão and Cazorla (2009) carried out the activity with 29 mathematics teachers, using onto-semiotic theory (Godino, 2002). They concluded that the sequence was feasible for teaching basic probability concepts, but noted the presence of several semiotic conflicts, due mainly to teachers' poor prior knowledge, as they experienced some of these concepts for the first time. Nagamine, Henriques, and Cazorla (2010) evaluated the same activity using anthropological teaching theory (Chevallards, 1992) – more specifically, the praxeological aspect – concluding that evidence from the technique and technology enabled identification of conflicts in some task requirements, leading to an improvement of the activity. Hernandez, Kataoka, and Oliveira (2010) carried out the activity with a group of 91 students in their last year at Sciences and Humanities High School (CCH), Mexico, who had not studied probability during that particular school year. In a qualitative analysis of the answers, they concluded that the students understood the difference between deterministic and random experiments, as well as between theoretical and frequentist probability, and that the activity was feasible for investigating topics in probability during that school year. Note also the version called 'Jefferson's Random Walks' developed for use with blind students (see Vita & Kataoka, 2014, in this volume).

An initial design was based on the story named 'Carlinha's Random Walks', with a key question repeated in most of the activities, asking if all the friends had the same chance of getting a visit. Different strategies were proposed in order to reflect on this question: simulating data with R software, constructing a table of frequencies, constructing the tree diagram, and obtaining the probability distribution. We aimed to investigate the steps students may take in reflecting on the key question using these different strategies.

3. DESCRIPTION OF THE TEACHING EXPERIMENT

In our research, we chose to work with a small-scale version of design experiment methodology, with seven third-year high school students, organized into two pairs and a trio, called D1, D2 (the trio) and D3. Although this limits the generalizability, it has the advantage of promoting deeper analysis of data, and bringing to light specific aspects that couldn't be observed in large-scale research. We consider the elements for analysis to be: students' written responses, and computer screencaptures obtained with Camtasia Studio® software. Prior to the development of the teaching experiment, students carried out an activity to familiarize themselves with R software.

The teaching experiment was divided into five sections with all the tasks and questions based on the following story:

Carlinha used to visit her friends in the week, in a particular order: on Monday, Luiz; on Tuesday, Felipe; on Wednesday, Fernanda; on Thursday, Alex; and on Friday, Paula. To make the meetings more exciting, the group decided that they would choose by chance the friend that Carlinha was going to visit. In order to do this, Carlinha should flip a coin on leaving her house and at every intersection. If she got 'heads' (C), she would walk one block to the north; if she got 'tails' (X), she would walk one block to the east. Each flip represented a block on her route. Carlinha would flip the coin four times in order to get to her friends' houses, as the diagram below shows.

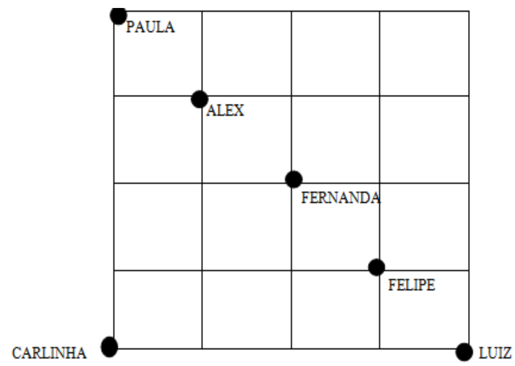

Figure 1. Illustration of Paths (Ferreira, 2011, p. 87).

3.1. SECTION I: THE STORY AND CONTEXT

This section, comprised of five tasks, aims to investigate the intuitive conceptions of probability, as well as the differences between a deterministic situation and a random experiment. These initial tasks were carried out using paper and pencil.

By only reading the story, without carrying out the simulation, answer the following:

(1) What is the difference between the old way Carlinha used to visit her friends and the new way?
(2) What are the possible outcomes when flipping a coin?
(3) What is the chance of getting "heads"? And what is the chance of getting chance "tails"? Why do you think that?
(4) Do all her friends have the same chance of being visited? If no, what are the chances? If yes, what are the chances? Why do you think that?
(5) Imagine that you have flipped the coin four times. How would you take notes about the imaginary results?

It was expected that the probabilistic concepts were known by the students in their third year at a São Paulo State High School. According to the official curriculum (Secretaria da Educação, São Paulo, 2011), this content should be addressed in the second quarter of the second year, so we assumed that these students had already acquired this knowledge. In addition, we thought that, regardless of the school environment, the word 'chance' is frequently used in people's daily lives.

We observed that in task (1) all groups identified the differences between a random experiment and a deterministic situation. In Task (2), only the students in group D2 did not understand the question, and so did not identify the two possible outcomes: 'heads' or 'tails'. In task (3), all groups justified that when flipping a 'fair' coin, the chance of getting 'heads' (or 'tails') was 50%. In task (4), all groups reported that the friends did not have the same chance of being visited and D1 and D3 justified this by saying that a friend could be visited more than once a week, depending on the results of the coin tosses. D2 reported (incorrectly) that if the first toss gave 'heads,' then Paula could not be visited. The explanations given by the groups confirmed the results of Gusmão and Cazorla (2009). Hernandez, Kataoka, and Oliveira (2010) pointed out that students' explanations could be based on frequentist or belief views of probability theory, but we cannot tell which at this stage.

In task (5), answers such as XXCC, or X, X, C, C, were expected, and such records would demonstrate that the students understood the statement. No group used such schemes to imagine the possible results when flipping the coin. D1 said that they could get: 'heads', 'heads', 'tails', 'heads', and so Alex would be the friend visited. D2 imagined three 'heads' and one 'tails', and that way Alex would be visited. D3 answered 'tails', 'heads', 'heads', 'heads', and with this result, Felipe would be visited. According to Nagamine, Henriques and Cazorla (2010), the answers given by D1 and D3 are acceptable, but show a misunderstanding of the statement. On the other hand, D2's answer shows a record that did not specify the order of occurrence of the events.

By the end of this section, students showed a higher level of autonomy in relation to the teacher-researcher. The sequence of tasks turned them into active participants in the discussion of the results. The students felt themselves to be participants in an open study, rather than a closed task with a specific correct or incorrect answer. This led us to observe from this very first section an important characteristic foreseen in constructionism: the integration of the participants into the proposed task, and their need to feel part of the process of constructing their own knowledge. In terms of the cognitive elements of Gal's (2005) probabilistic literacy model, we noted that the students were already able to work with big ideas, context and language.

3.2. SECTION II: THE SIMULATION

The tasks in this section consist of a simulation activity followed by seven questions: (1), (5) and (7) were carried out using R, and the others using paper and pencil, but with the support of the results from R. Note that the students did not flip the coin, but simulated the flipping. This section aimed to make students reflect on the answers they gave in Section I, using the experimental results they obtained. It was expected that in each task the students could create their own understanding of subjective probability and that this could be observed in the later results they obtained during the teaching experiment.

(1) Carlinha's Problem: In order for Carlinha to visit a friend, you (each pair of participants) will have to simulate the coin-flipping four times, which we will call the 'experiment'. If, as a result, you have 'heads' (0), Carlinha will walk one block north. If you have 'tails' (1), she will walk one block east. To represent this simulation in the R software, we will use the following language: to represent 'heads' we will use the digit 0, and to represent 'tails' we will use the digit 1. You will simulate this experiment 30 times. For example, if you have the following, 'heads', 'heads', 'tails', 'heads', that is, (0, 0, 1, 0), you should assign her visit to Alex. Fill in Table 1.

Table 1. Simulation results

Experiment	Sequence	Friend
1		
2		
. . .		
30		

(2) Who is probably the most visited, Paula or Fernanda? Why?
(3) Is there a chance that Carlinha will not be visiting any friend? ( ) No ( ) Yes. Why?
(4) Once you have performed the simulation, would you change your opinion in response to the following question: "Do all friends have the same chance of being visited?" Think about your answer considering question 4 of section I. ( ) No ( ) Yes. Why?
(5) Can you summarize the section II results obtained using R in a table that represents the data? This table is called a 'frequency distribution table'(FDT, Table 2).

Table 2. Distribution of the number of times Carlinha visited each friend

Friend	Number of times visited (hi)	Probability (pi)	Percentage
Luiz			
Felipe			
Fernanda			
Alex			
Paula			
Total	30	1,00	100,00

(6) Once you have organized the results in the FDT, would you change your opinion regarding the question: "Do all friends have the same chance of being visited?" Think about your answer by considering question 3 of section II. ( ) No ( ) Yes. Why?

(7) Choose any other pair of students, and build a graph that compares their results with yours. Are they equal? ( ) Yes ( ) No. What do you think about that?

All groups performed task (1) without any difficulty, and in task (2) identified Fernanda as the most-visited friend. In task (3), D2 and D3 found that there was a chance that one friend of the group will not be visited, based on their simulation results, while D1 obtained a result closer to the theoretical probability. In task (4), groups D2 and D3 maintained their previous opinion that friends had different chances of being visited, but were uncertain about the reasons, whereas D1 maintained that friends had the same chance of being visited and gave an explanation based on their own simulation results. Task (5) was carried out by all groups, and in task (6) they all reported that they had not changed their previous opinion that friends had different chances of being visited. Nevertheless, D1 and D2 changed their explanation, influenced by their simulations, while D3 stayed with their previous justification. Finally, in task (7), all groups reported that the graphics were different, whereas D2 and D3 found out that, despite the differences, there were friends who were more likely to be visited.

One goal of this block of tasks was achieved when students realized that they could use the tools from R to carry out the required tasks. This illustrates an important feature of constructionism: students need to have access to the necessary tools and recognize them as being useful for solving the proposed tasks.

It was also remarkable that the simulation provoked new reflections on the earlier results, demonstrating significant development of the idea of probability. The simulation results in task (1) led students to realize that some friends were going to be visited more than others, thus intuitively deconstructing the idea that the visits were random. Students noticed that it was possible to say whether any friend would have more chance of being visited than another, and in task (6) they identified which were more likely to be visited, confirming their results during the comparison of charts in task (7).

Of the cognitive elements in the probabilistic literacy model, students developed their ability to work with big ideas, figuring probabilities, language, and critical questions. The element of critical questions was particularly important, demonstrating the growing engagement of participants with the tasks.

3.3. SECTION III: THE TREE DIAGRAM ('TREE OF POSSIBILITIES')

This section consists of five tasks: task (5) could be done in R, and the other four tasks in a paper-and-pencil environment. The aim of this section is to discuss the concepts of theoretical probability from the frequentist perspective.

(1) Complete the 'tree of possibilities', indicating the sequence randomly selected for the number of 'heads' and the friend visited (Figure 2). Note that each branch opens in two new branches (one for 'heads' and the other 'tails') for each draw:

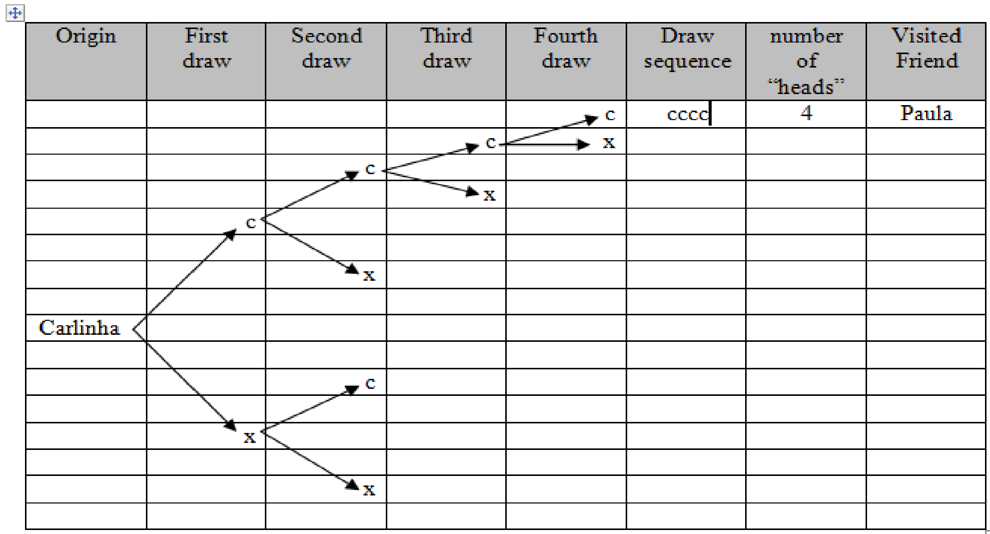

Figure 2. The Tree Diagram (or 'tree of possibilities').

(2) Now, find out how many ways are there altogether.
(3) Find out if is there a common feature of all the paths leading to each of her friends: Paula, Alex, Fernanda, Felipe and Luiz.
(4) After analyzing how many paths lead Carlinha to each friend's house, would you change your opinion for the following question: "Do all friends have the same chance of being visited?" Think about your answer considering question 5, in Section II. ( ) No ( )Yes. Why?
(5) After analyzing and systematizing the results of the tree diagram, fill in Table 3:

Table 3. Distribution of the probability that Carlinha visits each friend

Friend	Number of paths	Number of paths/total number of paths (fractions)	Probability (p)*
Luiz			
Felipe			
Fernanda			
Alex			
Paula			
Total			

(*) realize the division to express the results in decimal form

Task (1) was designed so that the students could visualize, through the tree diagram, the 16 mutually exclusive paths (task 2). We expected that students would realize that each of these paths had the same probability, 1/16, and therefore that the probability of visiting each friend would depend on the number of paths leading to that friend. We also expected that the students would see that which friend is visited depends only on the number of 'heads' and 'tails' and not on their order. In the study carried out by Hernandez, Kataoka, and Oliveira (2010), almost two-thirds of groups (61%) noticed this fact.

Group D1 made the tree but had difficulty when finding the friends visited. The teacher-researcher intervened, helping the group with an example sequence and showing which friend was visited and this enabled the group to complete the task.

D2 performed the task, but made a mistake by attributing five visits to Felipe and five to Fernanda. We noticed that they incorrectly wrote a sequence as XCXX rather than XCCX, and so misidentified a friend as Felipe instead of Fernanda (see Figure 3). Despite this mistake, the group had no difficulty making the tree, once they had seen the relation between the sequence and the friend to be visited, as requested in task (3).

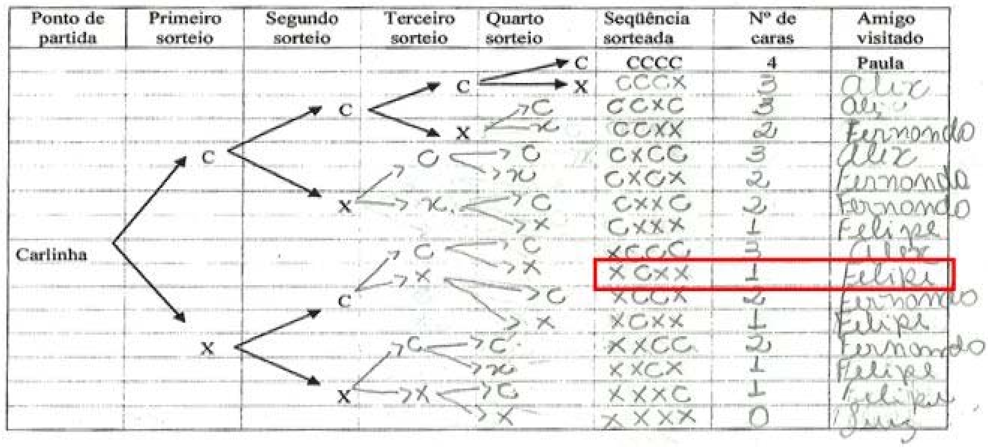

Figure 3. Results of D2 for Task 1 (Ferreira, 2011, p. 113).

D3 had more difficulty with the task. Initially, the students did not understand how to complete the tree, requiring the teacher-researcher's intervention. He showed the group the same example used with group D1. Later, D3 also had trouble composing the sequences. Again, the teacher-researcher intervened and explained, using the previous example, how the sequences were formed, and finally, still with some difficulty, the students completed their tree.

Note that questions about tree diagrams were explored in the second year of high school, so the students should have been familiar with this kind of representation. It is probable that the difficulties the students presented could be attributed to the fact that, from the determination of the sequence, they still had to relate the results to the number of 'heads' and to the name of the friend that was going to be visited.

In task (2), all groups identified that there was a total of 16 paths. In task (3), D1 and D2 were able to make a connection between the number of visits and the number of 'heads', helped by looking at the last two columns of the table (number of 'heads' and name of friend). As mentioned, D3 had difficulty making the tree, and did not notice this connection; however, they did notice that the number of 'heads' was four for Paula and zero for Luiz since there was only one path to each of these friends. For the other friends, they made only qualitative statements, reporting, for example, that Alex had a smaller chance of being visited than Fernanda because there were fewer paths that led to Alex's name. Task (4) presented more difficulties, and the members of D3 were the only students who began to realize that they needed to change their explanation about whether all friends had the same chance of being visited. We realized this when analyzing their explanation: "Fernanda has more chances of being visited, because she has more possible paths." ("A possibilidade de sorteio é maior da Fernanda por ter mais caminhos a serem saídos.") This response demonstrates an improvement in the group's concept of probability: using the tree, they deconstructed their explanation that the visits occurred only at random. The improvement in D3's observations made us reinforce the need for knowledge construction analyses throughout the development of this experiment.

Although D3 had earlier difficulties, it did not prevent their improvement during this sequence of tasks, or when they rethought their answers to each new question. Groups D1 and D2 did not change their minds or their explanations: they believed that the tree did not bring anything new, and once again, the visits were still attributed only to randomness when flipping the coin. It may be that students focused on making the tree diagram rather than the probabilistic concepts involved. We noticed that the idea of theoretical probability became clearer only after the discussions in later sections.

Task (5) was carried out by all groups without difficulty, though D2 maintained the mistake made in task (1). The teacher intervened, questioning the group about the number of visits each friend would have, and asked to the group to look again at task (1). The group checked and were confident about their results so the teacher-researcher chose not to point out the error and the group continued to work on the tasks. Only through later comparison with students from other groups did the group notice this mistake.

Overall, it was noticed that the participants were now fully involved with the problem, able to face each new task as an important step towards solving a problem as a coherent group. This led us to identify an important feature of constructionism since, according to Maltempi (2004), learning should occur through an active process in which learners contribute to the development of the tasks. As predicted in constructionism, this block of tasks contributed to providing learners with a position of control over their own knowledge building.

We believe that there were important contributions to the development of probabilistic literacy (Gal, 2005) through the reflections proposed in this section. We think that all five elements (big ideas, figuring probabilities, language, context, and critical questions) highlighted by Gal (2005) were strengthened. The tree diagram made students reflect, even if only informally, on the elements of randomness, prediction, and uncertainty, and encouraged them to think about probabilistic calculations in different ways.

It was also important that students began to assign a more contextualized meaning to the term ‘probability’ so that the issue no longer had a specific ‘school’ character but was treated as an everyday idea. This was helped by the use of a context that students considered ‘less formal’ than usual problems in their mathematics classes. Finally, the most rewarding feature of the tasks in this section was the increasing use of critical questions. Based on the results obtained from the tree diagram, students made the connection between number of paths and probability, showing recognition that some friends have higher chances of being visited than others.

3.4. SECTION IV: THE DECISION

This section discusses the differences between results from theoretical probability and relative frequency, and whether Carlinha’s way of visiting her friends was fair. The tasks were carried out with paper-and-pencil, except for the graphing in task (4).

(1) Fill in Table 4 with the results of Tables 2 and 3:
(2) What is the difference between these two ways of assigning probabilities?
(3) Analyzing the results, which of these two ways of assigning probabilities is the most appropriate for you? Why?
(4) Make a bar graph that compares the experimental results with theoretical probability. Are they the same? ( ) Yes ( ) No. What can you conclude?
(5) Do you think the *new* probability distribution of Carlinha’s walks is fair to all her friends? ( ) Yes ( ) No. Why?
(6) If you think that distribution it is unfair, could you indicate another way to decide the friends to be visited by Carlinha?

Table 4. Comparative table of assignment of probabilities

Friend	Relative frequency (hi)	Probability (pi)
Luiz		
Felipe		
Fernanda		
Alex		
Paula		
Total		

None of the groups had difficulty in performing task (1). In task (2), we expected that students would realize that the value of the relative frequency was an estimation of the theoretical probability, and that it could vary from sample to sample. We now share group answers that stood out during analysis. D1 stated: "The first time we did a simulation, the second time we specified if it was going to be heads or tails." ("Na primeira nós fizemos uma simulação, a segunda a gente atribui que ia ser cara ou coroa.") This demonstrates that they understood that in the first case the results were related to the experimentation, and in the second case the result occurred independently of the experiments. D2 said: "One is probability made from the experiment and the other from theory." ("Uma é probabilidade feita do experimento e a outra da teórica."; they used the term "feita", "made" or "done", when referring to the experimentation). D3 reported that "The practical has an exact result, but the theoretical is based on a calculation" ("Que a prática é a que tem resultado exato, o téorico, cálculo"). Here, the students were getting results directly from the experimentation, but they were not referring to the probability of visiting in a general way, but to the result obtained in that particular experiment. Moreover, they understood that the theoretical approach provided a result that would never be exact, because it did not depend on the experimentation: for example, the fact that a particular friend had a higher probability than others did not guarantee that he or she would be visited more often in practice.

At the end of this task, the teacher-researcher led a discussion with the groups in order to name the second way to assign probability. The groups suggested that there should be a way to assign probability without flipping the coin, and D3 named it as "theoretical probability", a label that was used from then on by all the groups.

In task (3), it was noted that only one of the three groups chose theoretical probability. The groups gave the following reasons: "Theoretical, because you have a basis for assigning each heads or tails." ("Teórica, porque você tem uma base que para cada uma você vai atribuir cara ou coroa." D1); "The experiment, because it is showing us that the probability from an experiment is more concrete than theoretical." ("A do experimento, pois ela é a que nos mostra probabilidade a partir de um experimento uma coisa concreta a teórica não." D2); and "Experimentation because you have more chance of getting the exact number in real conditions" ("A pratica porque você tem mais chance de adquirir o número exato em prova viva." D3).

In task (4), the students realized that the graph of experimental results had the same trend as the graph of the theoretical probability results. This is consistent with the observations of Cazorla, Gusmão, and Kataoka (2011), and suggests that this task constitutes an important tool for further observation of the convergence phenomenon (law of large numbers), when the number of simulations is increased.

With respect to task (5), the studies of Hernandez, Kataoka, and Oliveira (2010), and Gusmão and Cazorla (2009) suggest that students may think that Carlinha's new way of making visits is unfair, since friends who live further from central point were disadvantaged. Indeed, all groups responded that the new way of visiting is not fair, because some friends can be visited more than others. In task (6), D1 said that the best way to visit everyone would be the initial way – a different friend every day of the week. D2 proposed that the names of all friends should be written on paper and selected at random, with friends who have already been visited excluded from the new selection. Only D3 did not suggest any other way to determine the visits.

We now highlight some important items related to constructionism. We observed significant progress once there was clear evidence of the five dimensions – pragmatics, syntonic, syntactic, semantic, and social – that form the basis of constructionism (Maltempi, 2004).

Maltempi states that the pragmatic dimension is linked to the feeling of learning something that has immediate use. We noticed that students were already totally integrated into the sequence of tasks from their earlier answers, and they had the understanding to give clear answers to new questions. Therefore, we considered that the pragmatic dimension had been addressed. As we predicted earlier, we saw that the students had already had contact with the concept of probability in other periods of their school life. However, they had not been encouraged to reflect effectively about its true meaning in more visible situations. According to Maltempi (2004), when working in the constructionist perspective, we must establish space in the 'syntonic' dimension. Rather than the dissociated learning practiced in the traditional classroom, we need to construct contextualized projects, tuned to situations that students regard as important. We noticed that our questions aroused the students' interest since, rather than discussing abstract differences between experimental and theoretical probability, the questions provoked reflection on the concept of probability using context that was part of the students' world.

According to Maltempi (2004), in the syntactic dimension, learners should be able to access easily the basic elements that compose the learning environment, and progress in handling these according to their needs and cognitive development. In this section, we noticed that the students were already manipulating the tools of R very easily. We had the initial hypothesis that, on the basis of the introductory activities to familiarize them with R, students would be able to use the software when needed to carry out the tasks and achieve the objectives for each task. Students had difficulty with this in the early sections, but by now their familiarity with the computational tool was now evident and facilitated the execution and understanding of the tasks.

In the same vein, we noticed that the tasks carried out using R had a meaning in the context of the activity. Each time students were asked to look at a previous task or to respond to previous questions, it was possible to see evidence of a new view imbued with new meanings. This leads us to think that the semantic dimension is also being contemplated, providing the learners with an ability to manipulate the elements that carry meaning and to make sense to them, in line with Maltempi (2004).

We noticed that the questions in this section made the students reflect on the meaning of probability in everyday situations and reflect on the 'justice' in various situations that are dependent on probability. We believe that this initiates a reflection on social areas, signaling that this is also a dimension to be considered. Maltempi (2004) claims that there should be an integration of activity with personal relationships and the culture of the learner's environment, using the computer as a powerful social tool to achieve this integration. In this section, the tasks made it possible for the students to work with three elements of the cognitive component of the probabilistic literacy model (Gal, 2005): figuring probabilities, language, and critical questions.

3.5. SECTION V: OTHER EXPLORATIONS

This final section consists of eight questions aimed at provoking discussion of the concept of unequal probabilities. This idea is not addressed much in basic education, so we considered it important to include in order to reinforce the idea that not all events are equiprobable.

Section V.1.

(1) Perform a simulation with 12,000 experiments. What do you notice when comparing the results of this simulation with a simulation of 30 experiments? And with the theoretical probabilities?

(2) Using the same approach of the flipping coin, what could you do to change the fact that Fernanda was the friend visited most often?

Section V.2.

(3) If we used a coin with a probability of 0.6 to get 'heads', who is (are) the friend(s) visited most often?

(4) What if the probability of getting 'heads' were 0.8: who would be the most visited friend(s)?

(5) What if the probability of getting 'heads' were 0.1: who would be the most visited friend(s)?

Section V.3

(6) Considering that we are simulating the flip of a coin, how would you describe the coins in the proposed tasks 3–5? And what about the coin mentioned in the earlier tasks, with success probability of 0.5?

(7) Now, try to exchange the success probability for 0.8. What are your conclusions?

(8) Considering the 12,000 simulation experiment performed in task 1, change the success probability to 0.6. What do you observe? In this case, who is (are) the friend(s) most visited?

All groups carried out the simulation of the 12,000 experiments in R. Representative results are shown in Figure 4.

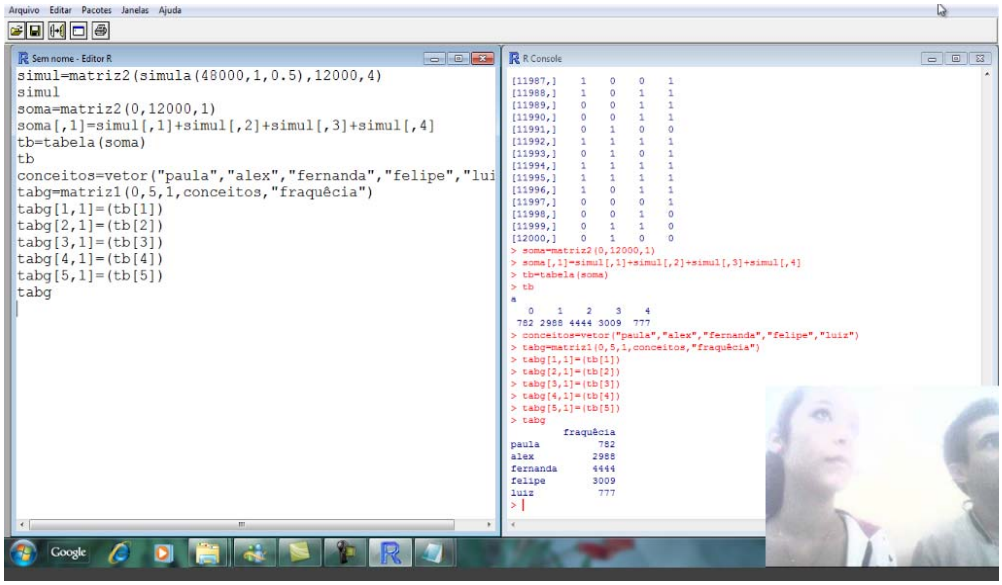

Figure 4. Simulation results from group D3 (Ferreira, 2011, p. 125)

When comparing the results with task (1)'s simulation of thirty experiments, all groups said that there were no major differences, and D1 and D3 stated that only the actual sequence of numbers changes. Comparing them with the theoretical probabilities, they stated that the latter gave more certainty: "Theoretically, you are assured of the probability." ("A teórica você tem a certeza da probabilidade." D1); "With theoretical probability, Paula and Luiz are least visited." ("Com a probabilidade teórica Paula e Luiz menos visitados." D2); "Theoretically, you are sure of the probability." ("Na teórica você tem a certeza da probabilidade." D3).

For task (2), D1 and D3 suggested exchanging the order of the friends' names on the board. D2 suggested flipping the coin but discarding any sequence that leads to Fernanda and flipping the coin again. From the results, we can infer that the students did not realize at this stage that changing the probability of 'heads' could change the results and lead to the desired solution. They seemed to have the fixed idea that the outcomes were equiprobable since coins have two faces – 'heads' and 'tails'. For task (3), we expected that Fernanda and Alex would be identified as the most visited friends, based intuitively on the fact that Fernanda needs two 'heads' (50%) and Alex needs three 'heads' (75%) for a visit, those probabilities being closest to the 0.6 or 60% probability of 'heads' for the coin. We did not expect that students would use the concept of binomial distribution to carry out a calculation, but we believed that the task could encourage important initial reflections for future formalization of the binomial concept.

All three groups identified Alex as one of the most-visited friends. D2 also included Fernanda, whereas the two other groups included Paula but not Fernanda—possibly because Fernanda required 50% 'heads' while the coin had a greater (60%) chance of 'heads' and so suggested those friends who needed a higher proportion.

In task (4) groups D1 and D3 identified Alex and Paula as the most-visited friends, while D2 indicated only Paula. In task (5), we expected students to pick Luiz as the most-visited friend, since he was the only one who did not need any 'heads' to be visited. However, all three groups picked Felipe as the most visited friend, contradicting our expected answer. The answers seem to have been influenced by poor understanding of probability. Students were able to identify that Fernanda was no longer the most-visited friend, but they related the fact that Felipe needed only one 'head' with the coin's 0.1 probability of getting a 'head'.

In these questions in the second subsection, we noted our failure to ask students to justify their answers, a fact that hampered our analysis of what led them to their decisions. The previous sections had already been used and reviewed by other researchers and we made only a few changes to integrate the use of R. Section V, however, was developed specifically for this study and will be improved in the next iteration.

In task (6), all groups labeled the coins as 'honest' and 'flawed', because these terms are often used in game scenarios, independent of the school environment. We also noticed that we mistakenly placed task (7) before (instead of after) task (8), leading to confusion in the answers. In task (8), D2 acknowledged that Fernanda was still the most-visited friend, while D1 and D3 said that the most-visited friends would be Paula and Alex. When transposing the answers, the pairs D1 and D3 may have confused the names Paula and Fernanda, since they performed the simulation correctly (see Figure 5).

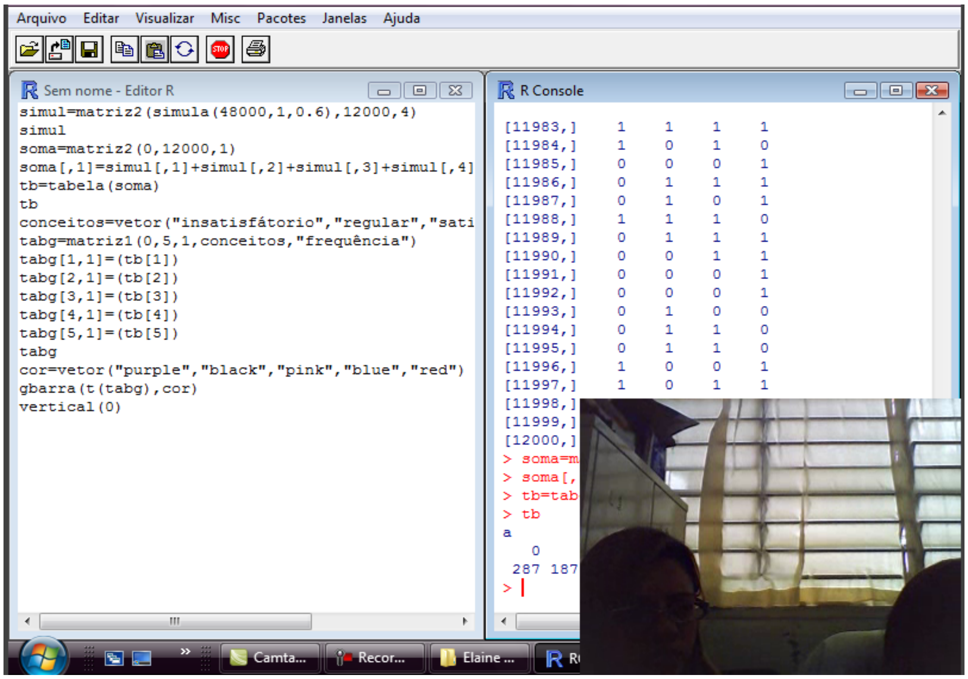

Figure 5. R script from group D1 for task (8) (Ferreira, 2011, p. 128)

In this section, we noticed that the use of R software became an important tool. Having already simulated the results from 30 trials, students started to reflect on the relationship between these results and the theoretical probabilities. However, using R made it easy for students to visualize the results of the experiment with 12 000 trials. The task took on an important role, contributing to students' understanding of the phenomenon of convergence, now with a much larger number of trials. Batanero (2001) commented about the care that must be taken when working with random experimentation to prevent the 'Law of Large Numbers' being transformed into belief in a 'Law of Small Numbers'.

We believe that this section led students to new reflections on the fairness or bias of a coin, by utilizing the software tool that made it possible to change the probability of 'heads'. Biased coins are sometimes used in classroom exercises, although those coins are likely abstractions as the students are not usually in contact with a concrete representation or simulation software. The tasks in this section enabled reflection by means of a concrete realization in the context of the activity, which, once again, reinforced the characteristics of constructionism. These tasks enabled exploration of all five elements of the cognitive component of Gal's (2005) model of probabilistic literacy: big ideas, figuring probabilities, language, context, and critical questions.

4. FINAL CONSIDERATIONS

We return to the research question: “What aspects can be observed when integrating students’ learning process into a computational environment – R software – in order to work with probabilistic concepts, within the perspective of Gal’s (2005) probabilistic literacy model and Papert’s (1980) constructionism?” We infer that students work with more autonomy to build their own knowledge when linked with the use of computer resources, and that this became an important tool for building probabilistic knowledge by the end of the experiment. We saw that the work led to important discussions about simulated versus theoretical results, and about the phenomenon of convergence – the law of large numbers, among others, as proposed in the initial objectives. Furthermore, making use of experimental methodology allowed the students to participate actively in the process and thus learn in a constructionist environment.

There are still many challenges to overcome in order to develop such learning materials for the school environment, to allow students to work with autonomy with significant use of the computer and to construct their probabilistic knowledge. Such a teaching approach is rarely used in basic education, and faces barriers imposed by the school dynamic. Our choice to work with the ‘Carlinha’s Random Walks’ teaching experiment was fundamental to achieving the desired goals. As an experiment already explored by other researchers in different perspectives and with different audiences, it had a solid foundation on which we could make the adjustments required.

The use of R lived up to our expectations. As an open-source software, it is freely available for schools and does not require any specialized hardware. Students were able to become familiar with the language quickly, and soon overcame initial difficulties. Working with R provided students with a sense of control in being able to manipulate the actions, and not to work only with pre-specified commands. There were two other key points about using the software to develop the concept of probability. Firstly, the simulation of 12 000 experiments enhanced the visualization of convergence of probability, and secondly, the possibility of working with coins with different degrees of bias extended reflection beyond the standard equiprobable situation, with which students are generally quite familiar.

The basic question “Do all friends have the same chance of being visited?” had an important role, as it was repeated in various sections of the experiment. We noticed that these repetitions corresponded with a significant maturation of students’ justifications and understandings of concepts of probability. The work with the tree diagram also played an important role in this maturation, since it provoked students’ important reflections on theoretical probability as an alternative to their previous frequentist reasoning.

We believe that these aspects all contributed to the development of probabilistic literacy, since all five elements of Gal’s (2005) model were addressed and worked with, formally or informally. Even though this work was carried out with students in the last year of Brazilian high school, we identified gaps in students’ understanding of probabilistic concepts. This leads us to suggest that further discussion of the pedagogy of probability are needed in order to achieve the recommendations of the National Curriculum Parameters for elementary school.

Despite all the difficulties encountered – in terms of understanding the concepts of probability, the level of autonomy of the students, and the unusual approach for school study – we believe that we provided the students with an opportunity to reflect on and develop concepts of probability through a different perspective using computer software. The results seem to indicate that this type of experiment can constitute an important educational resource for teachers working with probabilistic concepts in basic education, and contribute to the increase of students’ probabilistic literacy.

REFERENCES

Batanero, C., & Diaz, C. (2007). Probabilidad, grado de creencia y proceso de aprendizaje [Probability, degrees of belief and the learning process]. *XIII Jornadas Nacionales de Enseñanza y Aprendizaje de las Matemáticas.* Granada.

[Online: http://www.ugr.es/~batanero/ARTICULOS/PonenciaJAEM.pdf]

Batanero, C., & Godino, J. (2002). *Estocástica y su Didáctica para Maestros* [*Stochastics and its Teaching for Teachers*]. Proyecto Edumat-Maestros, Granada, Universidad de Granada. [Online: http://www.ugr.es/~jgodino/edumat-maestros/manual/6_Estocastica.pdf]

Borovcnik, M., & Kapadia, R (2009). Research and developments in probability education. *International Electronic Journal of Mathematics*, *4*(3). [Online: http://www.iejme.com/032009/IEJME_p00_introd_E.pdf]

Brasil: Ministério da Educação (1998b). Secretaria de Educação Especial. *Parâmetros Curriculares Nacionais*: *Matemática* [*National Curriculum Parameters: Mathematics*]. Brasília: MEC/SEF.

Cazorla, I., Gusmão, T., & Kataoka, V. Y. (2011). Validação de uma sequência didática de probabilidade a partir da análise da prática de professores, soba ótica do enfoque ontosemiótico [Validation of a didactic sequence for probability based on the analysis of the practice of teachers, from an onto-semiotic focus]. *Bolema*, *24*(39), 537–560.

Cazorla, I., Kataoka, V. Y., & Nagamine, C. M. L. (2010). *Os Passeios Aleatórios da Carlinha* [*Carlinha's Random Walks*]. Tutorial AVALE. [Online: http://ambiente.educacao.ba.gov.br/conteudos/download/1622.pdf]

Cazorla, I. M. & Santana, E. R. S. (2006). *Tratamento da Informaçã para o Ensino Fundamental e Médio* [*Information Processing for Elementary and Secondary Education*]. Itabuna, Bahia, Brazil: Via Litterarum.

Chevallard, Y. (1992). Concepts fondamentaux de la didactique: perspectives apportées par une approche anthropologique [Fundamental concepts of didactics: perspectives provided by an anthropological approach]. *Recherches en Didactique des Mathématiques*, *12*(1), 73–112.

Coutinho, C. (2001). *Introduction aux Situations Aléatoires dès le Collège: de la modélisation à la simulation d'expériences de Bernoulli dans l'environnement informatique Cabri-géomètre II*. [*Introduction to Random Situations from College: modeling Bernouilli simulation experiments in the IT environment of Cabri Geometry II*.] Doctoral dissertation, University J. Fourier, Grenoble, France.

delMas, R., Garfield, J., & Chance, B. (1999). A model of classroom research in action: developing simulation activities to improve students' statistical reasoning. *Journal of Statistics Education*, *7*(3). [Online: www.amstat.org/publications/jse/secure/v7n3/delmas.cfm]

Ferreira, R. S. (2011). *Ensino de Probabilidade com o Uso do Programa Estatístico R numa Perspectiva Construcionista* [*Teaching Probability using the Statistical Program R in the Constructionist Perspective*]. Master's thesis in mathematics education, Bandeirante University of São Paulo, São Paulo.

Gal, I. (2005). Towards "probability literacy" for all citizens: Building blocks and instructional dilemmas. In G. A. Jones (Ed.), *Exploring probability in school: Challenges for teaching and learning* (pp. 39–63). New York: Springer.

Godino, J. D. (1995). ¿Qué aportan los ordenadores a la enseñanza y aprendizaje de la estatística? [What do computers do in teaching statistics?]. *UNO*, 5, 45–56.

Godino, J. D. (2002). Un enfoque ontológico y semiótico de la cognición matemática [An ontological and semiotic approach to mathematical cognition]. *Recherches en Didactiques des Mathematiques*, *22*(2/3), 237–284.

Gusmão, T., & Cazorla, I. (2009). A semiotic analysis of random walks of Monica: an activity to teach basic concepts of probability. *Proceedings of the IV International Seminar for Research in Mathematics Education*. SBEM: Taguatinga – DF.

Hernandez, T. H. M., Kataoka, V. Y., & Oliveira, M. S. (2010). Random walks in teaching probability at the high school. In C. Reading (Ed.), *Proceedings of the 8th International Conference on Teaching Statistics* (*ICOTS8, July 2010), Ljubljana, Sloevnia.* Voorburg, The Netherlands: International Statistical Institute. [Online: http://iase-web.org/documents/papers/icots8/ICOTS8_2B1_HERNANDEZ.pdf]

Lane, D. M., & Peres, S. C. (2006). Interactive simulations in the teaching of statistics: promise and pitfalls. In A. Rossman & B. Chance (Eds.), *Proceedings of the Seventh International Conference on Teaching Statistics.* Salvador, Brazil: IASE.

[Online: http://iase-web.org/documents/papers/icots7/7D1_LANE.pdf]

Maltempi, M. (2004). Construcionismo: pano de fundo para pesquisas em informática aplicada à educação matemática [Constructionism: backdrop for research in informatics applied to mathematics education]. In M. A. Bicudo & M. C. Borba (Eds.), *EducaçaoMatemática: Pesquisa em movimento* [*Mathematics Education: Research in motion*] (pp. 264–282). São Paulo: Cortez.

Mills, J. (2002). Using computer simulation methods to teach statistics: A review of the literature. *Journal of Statistics Education*, *10*(1).
[Online: http://www.amstat.org/publications/jse/v10n1/mills.html]

Nagamine, C. L, Henriques, A., & Cazorla, I. (2010). Análise a priori dos passeios aleatórios da Monica [A priori analysis of Monica's random walks]. *Proceedings: X National Meeting of Mathematics Education*, Salvador, Bahia, Brazil.

Papert, S. (1980). *Mindstorms: Children, computers and powerful ideas*. New York: Basic Books.

Pratt, D., & Kapadia, R. (2010). Shaping the experience of young and naïve probabilists. *International Electronic Journal of Mathematics*, *4*(3), 323–338.
[Online: http://www.iejme.com/032009/IEJME_p23_Pratt_E.pdf]

Secretaria da Educação, São Paulo (2011). *Currículo do Estado de São Paulo: Matemática e suas tecnologias* [*Curriculum of the State of São Paulo*: *Mathematics and its technologies*] (Coord. Maria Inês Fini). São Paulo: SEE.

Valente, J. A. (1994). Por quê o computador na educação? [Why use the computer in education?] In J. A. Valente (Org.), *Computadores e Conhecimento: Repensando a Educação* [*Computers and Knowledge*: *Rethinking education*].(pp. 24–44). Campinas: Graphic Unicamp.

Vita, A. C., & Kataoka, V. Y. (2014). Blind students' learning of probability through the use of a tactile model. *Statistics Education Research Journal*, *13*(2), 148–163.
[Online: http://iase-web.org/documents/SERJ/SERJ13(2)_Vita.pdf]

ROBSON DOS SANTOS FERREIRA
Rua Miguel João Jorge, 380, Jardim São José, Campinas, São Paulo, Brasil.
Apartamento 32 Bloco 09
CEP: 13051-172