International Electronic Journal of Mathematics Education
2023, 18(4), em0749
e-ISSN: 1306-3030
https://www.iejme.com

MODESTUM

Research Article

OPEN ACCESS

A study on mathematics students' probabilistic intuition for decision-making in high school

Alejandro Ramírez-Contreras [1*], Leopoldo Zúñiga-Silva [1], Ezequiel Ojeda-Gómez [1]

[1] Instituto Tecnológico y de Estudios Superiores de Monterrey, Monterrey, MEXICO
*Corresponding Author: alex.ram.contreras@tec.mx

Citation: Ramírez-Contreras, A., Zúñiga-Silva, L., & Ojeda-Gómez, E. (2023). A study on mathematics students' probabilistic intuition for decision-making in high school. *International Electronic Journal of Mathematics Education, 18*(4), em0749. https://doi.org/10.29333/iejme/13586

ARTICLE INFO

Received: 01 Jun. 2022
Accepted: 03 Aug. 2023

ABSTRACT

This paper reports on an exploratory study about probabilistic intuition in learning mathematics for decision-making. The analysis was carried out on a group of high school students in relation to their probabilistic intuition in problem-solving, after performing playful learning activities on a simulation platform specifically designed for this study. The analysis used a mixed qualitative-quantitative method. The results showed that the simulation activities facilitated the students' probabilistic intuition and boosted their performance in solving problems, particularly avoiding decisions based on previous results (when unnecessary) and identifying situations, where binomial distribution can be used.

Keywords: didactic simulator, educational innovation, higher education, probabilistic intuition, Tec21

INTRODUCTION

Probability and decision making appear in our childhood, usually associated with games of chance. They are introduced in some primary education activities; however, their formal study does not occur until middle school or high school. Probability learning for decision-making is difficult for mathematics students, mainly because they can rarely understand the contexts in which problems arise and confront the difficulties encountered in learning probability concepts (Memnun et al., 2019). Moreover, teaching probability is usually very theoretical and axiomatic. Students feel that it is not very useful and has few applications because it is not given appropriate emphasis in problem-solving. In addition, some problems require knowledge and skills beyond common sense to solve; thus, the students get the idea that it is very complicated to learn and, consequently, lose interest in its study. They feel that what is taught is irrelevant to their interests, leading them to make mistakes and unsuccessfully face the problems they are proposed to solve (Ruiz, 2018).

Considering that current education focuses on developing competencies, skills such as probabilistic intuition become primarily relevant. Diccionario de la Lengua Española [Dictionary of the Spanish Language] (2021) defines intuition as "the faculty of understanding things instantaneously, without the need for reasoning." It concerns a student's belief about a phenomenon that influences how they act upon it. In particular, probabilistic intuition concerns a student's perception of a situation or problem that involves calculating probabilities. However, it is crucial to consider that, in the school environment, these intuitions are usually erroneous (Alvarado et al., 2018). This situation needs to be addressed because probabilistic intuition is essential in learning processes. According to Zamora (2014), developing probabilistic intuition in students and the natural ability to face a problem and make decisions appropriately is necessary.

One practical way to meet this need is to use technologies in strategies that allow students to visualize and understand probability problems, motivating and interesting them (Frassia & Serpe, 2017). Specifically, in this research, we designed a platform incorporating the simulation of random, repeatable experiments as a didactic resource so the student could understand the context of a problem through the interactive modeling of didactic situations. This platform is described below in the methodological framework.

Like some other competencies, probabilistic intuition is not easy to observe and analyze and involves several elements. For this study, a set of problems based on previous research is proposed. We intended that the student manifests the necessary concepts and shows evidence of adequate probabilistic reasoning in the situations involved. To solve these problems, students must be familiar with concepts such as independence, biases, binomial distribution, and combinatorial reasoning. These concepts are discussed in the light of intuition and probabilistic reasoning, with each problem containing elements related to them.

Thus, this research objective was to explore the effects of using a simulation platform on the intuition and probabilistic reasoning and decision-making of high school students studying mathematics.

THEORETICAL FRAMEWORK

Generally, decision-making in the face of a stochastic phenomenon is based on our intuition; in some phenomena, empirical intuition is sufficient to choose the most appropriate option, but in others, experience is necessary to arrive at the best solution.

Teaching probability is usually very theoretical, based mainly on the axioms proposed by Kolmogorov. However, it is necessary to note that education has been at a turning point for a few years, moving from traditional models to one focused on competencies. For this reason, it is necessary to base probability courses on problem-solving that develops the necessary skills for the best decision in the face of a stochastic phenomenon.

This situation demands focusing probability teaching on methods that allow students to develop competencies, particularly playful methods. One of the main foci of recent research has been probabilistic reasoning through simulation (Alvarado et al., 2018; Batanero, 2006; Martins, 2018; Tintle et al., 2015). In addition, introducing simulation in specific problems has been proposed. For example, Martins (2018) proposes using simulation methods to approach the birthday problem, using a particular platform applied to contextualized situations, such as a football game. Rosado et al. (2017) propose using gamification and adaptive learning in real problems to develop intuition and abstraction in the student. Alvarado et al. (2018) highlight the importance of intuition in engineering students and how it influences erroneous results in contextualized situations. Tintle et al. (2015) highlight simulation as a tool to combat "anti-statistical" reasoning.

However, most research focuses on higher education (engineering, mainly). Therefore, coinciding with Batanero (2006), it is necessary to apply new methodologies at the pre-university level that awaken probabilistic reasoning in the student because this level generally does not delve into probability. Such methodologies should include concrete experiments or simulations in real environments. Developing probabilistic intuition at early educational levels is vital because not doing this results in low performance at the higher levels (Batanero et al., 2021; Zamora, 2014). Without probabilistic intuition, the students are guided only by common sense, and if it is inadequate or wrong, they cannot achieve correct results (Alvarado et al., 2018; Batanero, 2006). This problem justifies the need to implement methods that help students develop probabilistic intuition and motivation at the pre-university level.

Probabilistic Reasoning and Intuition

Philosophically, Diccionario de la Lengua Española [Dictionary of the Spanish Language] (2021) describes intuition as the intimate and instantaneous perception of an idea or a truth that appears as evident to those with it. It is usually a product of mental development; however, intuition in the face of certain phenomena needs to be developed intentionally. Zamora (2014) discusses intuition in probability according to Fischbein, who distinguishes two types of intuition: a) primary intuition, which is acquired mainly through experience, and b) secondary intuition, which is developed through education, usually in school.

Students often base their decisions on intuition, but their intuition is usually inadequate when chance intervenes. However, coinciding with Fischbein and Schnarch and cited in Alvarado et al. (2018), it is not advisable to ignore students' confidence in their intuition, which must be developed in the school classroom so that they can use it in decision-making throughout their lives. Attention should be given to the development of probabilistic intuition so that decisions are based on probabilistic processes and methods, leaving aside the "and/or believe ... ", "I think...", and "it seems to me ...", among others.

Gamification and Interactive Computational Methods

In recent years, thanks to the needs and new challenges in education, new methodologies have emerged for the teaching-learning process and the development of students' competencies (Escobar et al., 2022). Little by little, gamification is assuming more relevance among new education methodologies, and although it is not necessarily related to technology and technological platforms, the combination of these has a more profound impact on students (Ramiro-Miranda et al., 2022). Marín-Díaz (2015) describes it as a trend based mainly on combining learning with games or playful activities, innovating the methods applied in the classroom.

According to Lukác and Gavala (2019), teaching probability, like other areas of mathematics, should incorporate visual and interactive methodologies that the student can manipulate to develop the ability to solve problems and achieve greater motivation towards learning.

Mezhennaya and Pugachev (2018) showed the impact an interactive computational system had on teaching probability to a group of students. The results were very positive; the platform facilitated students' learning. Also, Aizikovitsh-Udi and Radakovic (2012) noted the use of platforms in education and the good results achieved by using them to link theory and practice.

Gamification does not necessarily require an interactive platform; however, these two elements in recent years have been successfully combined and implemented in education.

Problem-solving through simulation

Borovcnik and Kapadia (2014), cited in Alvarado et al. (2018), highlight three main perspectives in the teaching of probability:

(1) Laplace's theory, which is based on the student's prior knowledge and a set of simple, equiprobable events (sample space);

(2) the frequentist theory, which is determined by frequency (counts), and

(3) the subjective theory, which is based on the student's personal belief.

The concept of simulation is based on the first two perspectives since a pleasant interface for the student is intended to obtain the sample space of a random experiment to immediately take a random sample from that random space and measure whether or not it meets the condition of the desired event. This process is repeated a considerable number of times, which provides a relative frequency. Finally, the frequentist probability is calculated by dividing the complying cases by the total repetitions.

Martins (2018) highlighted that using simulation in modern probability teaching has a significant impact on learning, and, in addition, students enjoy it. Similarly, Koparan and Yilmaz (2015) studied the impact of simulation on the basic teaching of probability, particularly discrete probability with finite sample space.

Although it has been found that simulation favors the teaching-learning process of probability, in particular, developing competencies and knowledge, it has been little implemented for problem-solving.

METHODOLOGICAL FRAMEWORK

The following is information regarding the method, the participants in the study, the data collection instrument, the simulation platform designed, and the process prior to applying the instrument.

Research Focus

We conducted a pilot study using a mixed, quantitative-qualitative approach to study intuition and probabilistic reasoning.

Quantitative method

The students' responses on the instrument used were analyzed, their answers graded, the relationship between correct and incorrect answers determined, the mean and the mode calculated, and a correlation analysis among the questions was performed.

Qualitative method

A detailed analysis was carried out in which the intuition and reasoning that the students manifested in the proposed problems were observed and discussed. Each response was analyzed considering both the analytical procedures and the intuitive and reasoning elements through observations and inferences based on the theoretical elements underlying this study.

Participants and Subject of Study

The study population was 49 students in the sixth semester preparatory (high school) level of the PrepaTec Zacatecas Campus in the Tecnologico de Monterrey university system in Mexico. Of these 49 students, 27 were women, 22 men, and all were aged between 17 and 18 with a standard deviation of 0.4. The school subject of the study was "mathematics for decision making." Not all students of the institution study this subject because it is designed for students with profiles not focused on engineering. The subject is students' first approach to the formal study of probability; they had only studied frequentist probability before this research. Topics covered in the subject include permutation and combination, discrete random variables, and binomial distribution. These themes were reinforced by implementing a probabilistic simulation platform (own design), described later.

Data Collection Instrument

Below are the problems given to the students on the instrument and the explanation of each:

Problem 1

Guadalajara and Zacatecas are in the basketball finals. The tournament ends when either team wins six games. At the end of the first seven games, Zacatecas has won four and Guadalajara three. How many more games are needed to guarantee a winner? Upon resuming play, what is the probability that Guadalajara will win the series? Suppose the probability of winning is 0.5 in each of the games.

Problem 2

If you throw six coins in the air one at a time and in the first five, your result was "heads", what is the probability of getting a "heads" on the sixth throw?

Problem 3

In a room, there are two people. You knock on the door, and a woman comes out. What is the probability that there are two women in the room?

a. ½
b. ⅓
c. Other

Problem 4

The probability that a person favors a particular reform is 0.5. Which of the following statements is most likely?

a. In a sample of 100 people, 50 are in favor of reform.
b. In a sample of 10 people, 5 are in favor of reform.
c. They are equally probable.

Problem 5

A group of friends usually meet on the first day of the month to celebrate that month's birthdays. When asked, seven answered that they would have their birthdays this month. What is the likelihood that at least two will be on the same day?

Tip: It can be helpful to work with a plugin.

Problem 6

There are eight white balls and nine red ones in an urn. If two balls are removed randomly without replacement, which of the following options is more likely?

a. Probably you will get a red ball followed by a white ball.

b. Probably you will get a white ball and a red ball.

c. Probably you will get a red one given that the first one was white.

d. They are all just as likely.

Explanation of the instrument problems

Problem 1 is an adaptation of the problem proposed by Pacioli in 1494 in one of the first printed books on mathematics (Abramovich & Nikitin, 2017). The problem is known as "the problem of division of stakes" and can be solved using the binomial distribution. However, you first must determine the number of games necessary to guarantee a winner. It also addresses the importance of the order in which games can be won. An alternative solution can be using the product rule, independence, and considering all possible cases.

Problem 2 is based on "the player's fallacy," which has been addressed in several articles and different ways. Arévalo et. al. (2015), in their article "intuitions and probability," mention some variations of this fallacy and how people get carried away by the previous immediate results. They describe, for example, the first variation that gave rise to this fallacy in the casinos, where a person obtains black sequentially several times and then believes it is increasingly likely the next will be red. They also mention the example of Don Quixote, where it was said that all the bad things happening to him were a sign that good luck would be forthcoming. Another classic example came from Edgar Allan Poe. In one of his books, he stated that, if by throwing a die repeatedly, five consecutive deuces were obtained, the probability of obtaining a 2 when throwing the die next would be less than one-sixth. A further example of this situation is when a sports commentator claims that if a team has won a series of games, defeat is getting closer, or vice versa.

Vásquez Ortiz and Alsina (2017) analyzed the probabilistic knowledge of a group of primary education teachers. One of the items examined was related to the player's fallacy and the unconscious bias produced by the number of times a result has appeared. Batanero et al. (2018) emphasize how the result of previous experiments influences intuition about the result of the next experiment when they are independent. Also, Rodríguez-Alveal et. al. (2018) analyzed how a sequence of values with a string of equal characters, such as "CCCSS" and "CSCSC," can influence the decision to understand whether or not it is random. The objective of this problem in the present study is that students can comprehend the independence of experiments and not bias the probability by personal beliefs.

Problem 3 derives from Gardner's problem in 1959. It is known as "the two-child problem" or "Smiths' children's problem." The intention is to differentiate between independence and conditional probability. This problem has had many variants. For example, in Batanero et al. (2014), some of these variants were mentioned and used to improve students' intuitions, highlighting that they are usually incorrect. In Alvarado et al. (2018), probabilistic intuitions in engineering students were studied, and a questionnaire was proposed. One of the items was related to this problem.

Problem 4 relates to the importance of sample size and how it influences the calculation of probabilities. It is intended that the student connects it with binomial distribution and can calculate the probabilities. This problem is an adaptation of the one proposed by Tversky and Kahneman in 1974, where a problem of babies born in a small hospital and a large hospital was utilized to refer to sample size. Alvarado et al. (2018) mentioned an adaptation of this problem, which was proposed to analyze intuition in students' binomial and combinatorial reasoning.

Problem 5 is an adaptation of the "birthday problem." It has been studied with different approaches. In all studies, there is agreement on the intuition that the students manifest. Lesser and Glickman (2009) address this problem and call it a "magic trick", as students expect approximately 183 people to be required to reach a 50% chance that at least two will have birthdays on the same day when in fact, only 23 are required. It spotlights the difficulties of the students' intuitions. Martins (2018) analyzed this problem through simulation, showing that this playful approach helped students face the difficulties of their intuition. Maxara and Biehler (2006) analyzed it with a new simulation focus for introductory probability courses, emphasizing the development of competencies.

Problem 6 can be approached with a hypergeometric distribution. The intention is to observe whether the student identifies when order is relevant or not and consider several cases that meet a single condition (compound event) and consider the one, where there is a dependence on the conditional probability. This differentiates between the use of binomial and hypergeometric distributions.

Simulation Platform

The development of the simulation platform was not intended for problem-solving. Instead, it was intended that students develop probabilistic intuition by visualizing and manipulating simulations of random experiments, and with this, they would be

Figure 1. Initial interface (Source: Authors' own elaboration)

Figure 2. Simulation problems (Source: Authors' own elaboration)

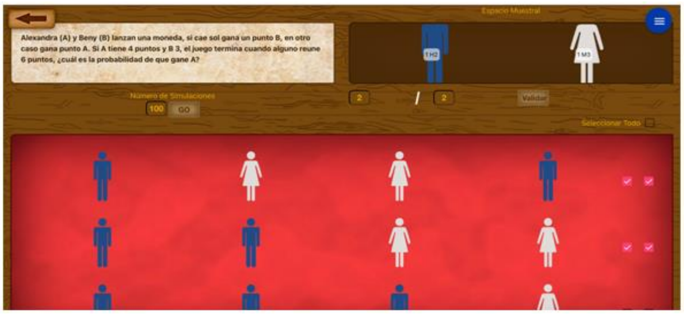

Figure 3. Example of simulation for problem one (Source: Authors' own elaboration)

able to solve the proposed problems. The platform's development was based on gamification methods to get the student to conceive it as a game (**Figure 1**) and advance at different levels to promote motivation. In addition, the platform was designed with an interface that was easy for students to manipulate.

The simulator has four worlds (levels), and the student must advance from one level to another. The student must solve a certain number of problems using simulation to move forward. For example, a problem appears once they enter one of the worlds, as shown in **Figure 2**.

In the simulator interface, the students can see the problem, the possible results of each element, and the number of simulations (this can be edited to obtain a more accurate probability). They must then select which elements meet the request (cases in favor) and the total cases (the student can select all or one at a time when it is a conditional probability).

The platform aimed to develop intuition and probabilistic reasoning through simulation, visualization elements, and gamification. This platform was implemented as didactic support in the probability theory classes, as described below.

Using the simulator in class sessions

The activities carried out with the students related to each problem of the data collection instrument are described below.

Regarding problem 1: This problem's activities were performed in the class sessions with two approaches. First, a simulation chain was run to identify how many games were necessary to guarantee a winner. The second approach simulated chains with the necessary repetitions to guarantee a winner. Once the number of games needed was identified, the goal was to show several ways to get a winner and analyze how many favored the desired player. **Figure 3** shows an example of the simulation.

Regarding problem 2: The activities were two comparative cases. The first addressed the simple probability of obtaining a result without previous results. In the second, a chain of simulations was shown, where a certain number of events had already happened, and the probability was estimated. Students could observe that regardless of the previous results, the outcome of the next event was not affected.

Figure 4. Example of simulation for problem type three (Source: Authors' own elaboration)

Figure 5. Example of simulation for problem type five (Source: Authors' own elaboration)

Figure 6. Example of simulation for problem type six (Source: Authors' own elaboration)

Regarding problem 3: During the simulation, the students performed runs to observe the elements of the sample space and notice the difference between one case and another. Similarly, the favored cases and the total cases could be analyzed. **Figure 4** shows an example.

Regarding problem 4: In this activity, simulations were run and analyzed in the sample space. For example, if a coin is thrown into the air twice, what is the probability of getting heads? The cases in which this statement was fulfilled were analyzed, for example, two out of four, then the probability of obtaining two heads was analyzed if we throw a coin four times, and so on. It is worth mentioning that obtaining half of the coin tosses was not the only case analyzed, but it was the most popular because it represented hope and was the most likely (**Figure 5**).

Regarding problem 5: In the beginning, the simulation of the related activities was somewhat complicated because the number of sample elements was large. For example, the original "*birthday problem*" has 365 days. To have a probability of 0.5, sample sizes of approximately 22 are required, which implies a tedious counting process that causes students to lose interest. Thus, we redesigned the simulation to adapt it to problems of birthdays occurring on the same day of the week or the same day of the month.

Regarding problem 6: During the simulation, the students analyzed several examples to observe the difference in several cases and distinguish which was more frequent and, consequently, more likely. The simulation helped students identify when there is a specific order and when there is a more general event (without order). **Figure 6** illustrates two related examples.

Pre-Trial Process and Application of the Instrument

The students went through the process of teaching and learning probability topics during a semester. The classes were designed with a school model that favored the development of intuition and probabilistic reasoning. The simulation platform explicitly designed for this purpose was an essential didactic resource. At the end of the experience, the data collection instrument was applied to analyze the students' intuition and reasoning.

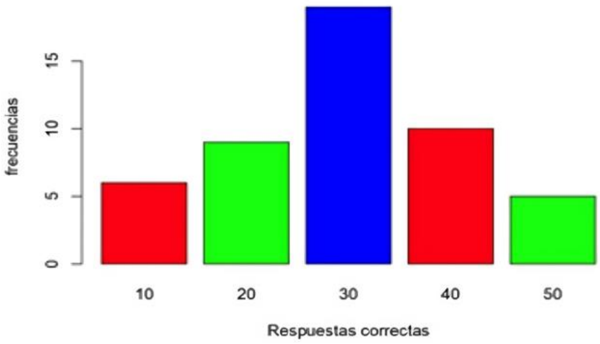

Figure 7. Histogram (Source: Authors' own elaboration)

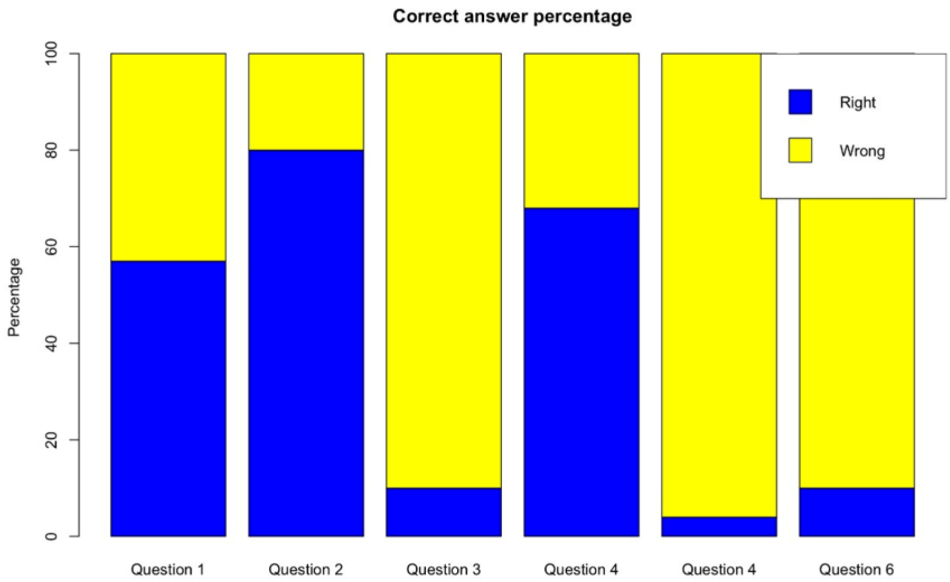

Figure 8. Correct & incorrect answers percentages (Source: Authors' own elaboration)

The evaluation instrument asked the students to answer the problem and explain the procedure and the solution. We designed the instrument to obtain information about what they were thinking when solving the problems and, thus, have elements that allowed exploring how their intuition and reasoning manifested.

RESULTS

The six problems already mentioned were used to obtain results. For the quantitative part, each was assigned ten points, considering answers and procedures. In the qualitative part, the elements of intuition and reasoning were considered.

Quantitative Analysis

The mean points per student was 29.8; i.e., on average, the students solved three problems correctly. The most frequent score was 30. No student solved all the problems, but all students answered at least one problem.

As seen in the bar graph (**Figure 7**), the results are concentrated around the mode and the mean in symmetrical behavior.

Figure 8 shows that item 3, item 5, and item 6 had the most incorrect answers, and 1, 2, and 4 received the most correct ones.

Although the mean and mode were solved correctly on three questions, most frequently in item 1 and item 2, there was no correlation between the questions solved correctly and incorrectly. This can be seen in **Figure 9**, which shows the correlation between the questions. Item 1 and item 2 are the most correlated, but only with 32%.

Qualitative Analysis

In addition to the quantitative analysis, a qualitative analysis involved examining all the students' procedures and arguments in their responses to understand the intuition they manifested.

Problem 3 and problem 5 had the most incorrect answers, and problems 1, 2, and 4 had correct answers. Then we analyzed the data obtained in each of the problems.

Figure 9. Correlation between questions (Source: Authors' own elaboration)

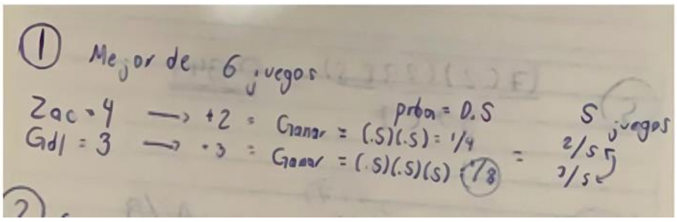

Figure 10. Student 2 answer, item one (Source: Field study)

Analysis of problem 1

Notably, in this problem some students did not achieve the solution, but they did identify the necessary number of games to guarantee a winner. According to some authors such as Abramovich and Nikitin (2017), that is the fundamental part of the problem, and it is where the probabilistic intuition manifests itself. In addition, 47% used a binomial model to solve the problem, managing to apply one of the fundamental topics included in the subject's curriculum.

Below are some responses from students.

Student 1 answer:

> "Assuming there are three more games, they can be distributed in a certain way that no team has won six. However, if there are four more games, one of the teams will win six. To calculate the probability that Guadalajara wins, you first have to see the probability of tying the series, and once that happens, multiply by 0.5."

Student 2 answer (**Figure 10**):

> "At least two games are needed if Zacatecas wins two in a row, making it the tournament winner; a maximum of 11 games (7+4) because if in 12 matches, 6-6 are won (by the probability of 0.5), then 11 are enough for there to be a winner (6-5). The probability comes from comparing how many victories each team lacks, multiplying the probabilities of this happening next."

Student 3's answer is given in **Figure 11** while student 4's answer is given in **Figure 12**.

Student 5 answer (**Figure 13**):

> "Because if Zacatecas wins, they will lack two victories; if there are three more games, there is the probability it wins one and Guadalajara two, and there would be no tournament winner, but if there are four games to play because Zacatecas wins two, there is a winner, and if Guadalajara wins three, equally, there is a winner."

Student 6 answer:

> "The sum of the number of games that the two teams need to win minus one since it only takes one team to reach the 6 games required.

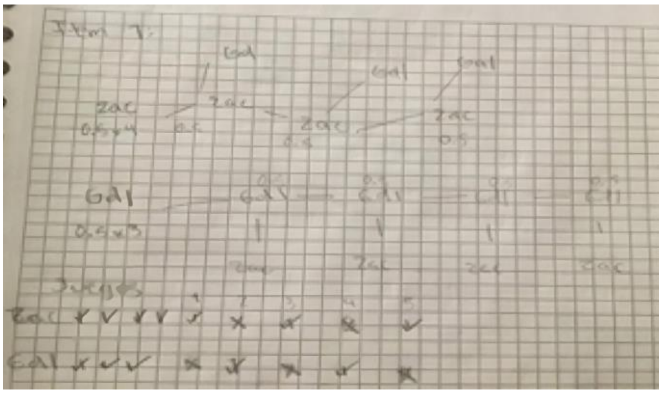

Figure 11. Student 3 answer, item one (Source: Field study)

$$P(7)=(7C7)(0.5)^7(0.5)^0$$

Figure 12. Student 4 answer, item one (Source: Field study)

Procedimiento(cálculos)
zac-4 gdl-3

4C3(0.5^3)(0.5)=0.25

Figure 13. Student 5 answer, item one (Source: Field study)

Procedimiento(cálculos)

$P(x=6)= 6C6\ (0.5)^{-6}(0.5)^{-0}$

= 0.015625

Figure 14. Student 1 answer, item two (Source: Field study)

> The results of the games were sought and multiplied by the probability of victory that Guadalajara will win, elevated to the number of times that their victories were needed."

It can be observed that student 1 identifies very well how many games are necessary to guarantee a winner, considering that with three more games, the series can be tied at five, and with one more game, a winner is guaranteed. For calculating probabilities, apply the same strategy: find the probability of tying the series at five and then multiply by the probability of winning one more game. Student 2 correctly identifies the number of games needed to guarantee a winner, arguing that, with 11 games, one team can win five and the other six, which is the smallest difference between the games. However, to calculate the probability, only consider one case per team: when Zacatecas wins the next two or when Guadalajara wins the next three. This gives evidence that the student does not identify that there are more cases in which a winner can be obtained; their solution was only based on the fastest way to win the series. Student 3 did not identify the minimum number of games needed. However, the strategy he applied is fascinating: he made a tree diagram to illustrate the solution. It is a suitable method, considering that there are few games, but the student did not get the correct answer because he did not complete the diagram. Student 4 used a binomial distribution but calculated the probability of one team winning 7 consecutive games, not explaining his argument. The evidence of student 5 and student 6 shows that they solved the problem as expected and adequately justified their answers, even explaining each of the elements of the binomial probability distribution.

We can infer that the work done by the students with the simulator positively influenced their ability to determine the number of games needed to guarantee a winner, even though, in the end, some of them failed to find the right solution.

Analysis of problem 2

In this problem, which received the highest percentage of correct answers, most students identified that they were independent events, using phrases such as "completely independent" "does not depend," "previous tosses do not affect," "already occurred," among others. The students who did not attain the expected result performed an operation but did not notice that they were independent events. This problem is presented to compare "logic" or probabilistic intuition versus common sense or intuition. It occurs very frequently in everyday situations, where a decision has to be made. Some authors such as Arévalo et al. (2015), Batanero et al. (2018), Rodríguez-Alveal et al. (2018), and Vásquez-Ortiz and Alsina (2017) agree that most people get distracted psychologically by results obtained previously and make their decisions accordingly, which leads to incorrect results.

We present some student answers to problem 2 below:

Student 1 answer (**Figure 14**):

> "We have to make a binomial distribution. The probability is 1/2, because there are only two options."

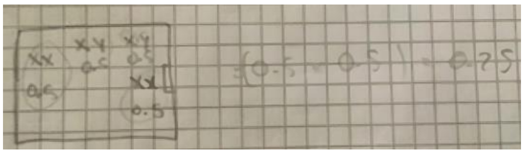

Figure 15. Student 3 answer, item three (Source: Field study)

Student 2 answer:

> "We have several chances of getting five heads from six coins; heads and tails have the same probability, so we multiply those odds by the probability of getting a heads on the last toss."

This was one of the problems in which more correct answers were obtained. Also, the students identified that it was an independent event. It is striking that the most common mistake of the 20% of students who did not reach the correct answer was failing to consider the independence between events and the conditional. The results of the first tosses were already known, calculating the probability of obtaining six heads. However, given the number of correct answers and the arguments offered, we can infer that the activities in the simulator related to this problem helped the students develop probabilistic intuition to avoid making decisions based on previous results.

Analysis of problem 3

In this problem, it is noteworthy that, approximately, one third of the students assigned an order, referring to the person who left, as "the first". If that were the case, the answer would be 0.5; however, this problem does not imply a specific order, so that assumption is incorrect. Also noteworthy is that about 60% of the students mentioned that independence between the events and that the person inside only has two options, and, thus, the problem is "the same as the previous one," alluding to problem 2. Although some research studies such as Bar-Hillel and Falk (1982) and Nickerson and Zenger (2004) have recognized a certain degree of ambiguity in problems of this type (called "paradoxes"), they help develop intuition and can be solved theoretically by various methods. The results in our research are similar to those mentioned in Batanero et al. (2014). In addition, we can establish that the main difficulty is that the problem speaks of people, and the students assume that, due to this fact, they are independent of each other. However, the fact that it is "at least one" should be highlighted, and it is not known which of the two, so both cases should be considered. Herein lies the principal error that students make. Below are some of the answers.

Student 1 answer:

> "Here, it is not saying what is the probability that it is a woman GIVEN that the one who came out was a woman, then I understand that it is independent. As it can be a woman or a man, the probability is 1/2."

Student 2 answer:

> "We have the possibility that a woman or a man will come out first, but there is a possibility that a man or a woman has stayed inside.
>
> Woman-woman, woman-man, man-woman, man-man.
>
> We took the first two options, so only one meets the condition."

Student 3 answer (**Figure 15**):

> "While you really cannot know exactly if there are two women in the room because of today's gender identities, the question lies in what is the probability that there are two women in the room, not that there is another woman in the room, so multiplying probability that it is a woman and the probability that it is another woman (0.5×0.5) gives us a result of 0.25."

Student 4 answer:

> "We know that one of the people in the room is a woman, the second is either a woman or a man; there are only two possible outcomes: either two women or a woman and a man."

In this problem, generally, similar answers were obtained as the previous problem in terms of reasoning (although the number of errors in the solution was much higher). Notably, the students affirmed that it was an independent event without noticing the conditional "it is known that at least one is a woman." Despite this, some interesting results were obtained. For example, student 1 identified that conditional probability should be used: he uses the phrase "given that." Student 2 described the sample space but referred to the person who came out as the "first person" without considering the order, leading to a wrong answer. Student 3 was the one who was closest to the correct answer. He described the sample space, "xx, xy, yx," as the one, where there is at least one woman, and, in addition, highlighted the one of interest, which is "xx", that is, two women. However, in the end, he only calculated the probability of having two women and did not divide it over the conditional. What student 4 offered was an example of the most common response. This suggests that the simulator activities did not help most students recognize situations in which conditional elements were established.

$$P(100) = (100C5)(0.5)^{50}(0.5)^{50}$$

$$P(10) = (10C5)(0.5)^{5}(0.5)^{5}$$

Figure 16. Student 3 answer, item four (Source: Field study)

Procedimiento(cálculos)

(10C5)(0.5)^5 (0.5)^5 =0.25

(100C50)(0.5)^50 (0.5)^50=0.08

Figure 17. Student 4 answer, item four (Source: Field study)

Analysis of problem 4

Concerning this problem, although it was one of the most correctly answered, only half of the students did the calculations using the binomial distribution. The rest justified with phrases such as *"the smaller the number of options, the greater the probability," "smaller sample is easier to work with," "in both, there is half, but there are fewer ways to take 5 from a sample of 10 than 50 from 100,"* among others. Of those who answered incorrectly, one replied that, with a larger sample, there is more likelihood to get half of the successes. The other students agreed that the probability was equal, allowing themselves to be carried away by the rule of "favorable cases among total cases" and other similar rules of probability to argue that they are just as likely. According to Alvarado et al. (2018), the essence of this type of problem lies in going beyond the basic rule of probability and not getting distracted by common sense. It is also important to note that combinatorial reasoning applies, where there are more ways and probabilities of obtaining a result.

Here are some answers:

Student 1 answer:

> "In both groups, it can be expected that half is favored because of the proposed probability of 0.5."

Student 2 answer:

> "I think they are just as likely, but if we perform the experiment, paragraph 1 is more likely by the law of large numbers: the larger the sample, the closer it gets to the theoretical probability."

Student 3 answer (**Figure 16**):

> "The fewer people, the higher probability."

Student 4 answer (**Figure 17**):

> "Although in both, it is half, there are fewer ways to take five from a sample of 10 than 50 out of 100, so the probability is higher."

As already mentioned, this problem was one of the ones that had the most correct answers, and a considerable number of the students used the expected procedure of binomial distribution. Some interesting answers were also obtained. For example, student 1 alludes to mathematical expectation, although it is true that it is expected to obtain half of the successes in both cases. This is also the most likely result; if we compare them, the probability changes. Student 2 speaks incorrectly of the law of large numbers, which is unrelated to this topic. What student 3 and student 4 offer are examples of expected answers because in both cases, the probabilities are calculated by binomial distribution; however, the contrast of arguments is interesting: student 3 only offers a common argument. On the other hand, student 4 speaks of ways to accommodate, manifesting combinatorial reasoning and understanding of the binomial model.

Similar to the results in problem 1 and problem 2 and considering the number of correct answers and the arguments offered, we can infer that the simulator activities related to this problem also boosted probabilistic intuition in the students to identify situations, where the binomial distribution can be used.

Analysis of problem 5

In this problem, approximately 40% tried to use the rule of "cases in favor out of total cases", using the number of people as the cases in favor and the number of days of the month as total cases. It is important to note that this problem is complex for students. For this reason, the purpose was not to achieve the solution theoretically but to approach it differently, thanks to work done on the simulation platform. According to Martins (2018) and Maxara and Biehler (2006), simulation helps develop or strengthen the skills needed to solve these problems. Students can go beyond common sense.

Here are some particular cases:

Student 1 answer (**Figure 18**):

> "There are 30 days in the month; that is why I divided 21 by 30."

Procedimiento(cálculos)
7C2=21
21/30= 7/10= 0.7

Figure 18. Student 1 answer, item five (Source: Field study)

(7C0)(29/30)^7=0.79
(7C1)(1/30)(29/30)^6= 0.19
0.79+0.19+=0.98

Figure 19. Student 2 answer, item five (Source: Field study)

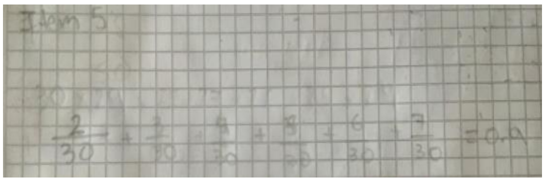

Figure 20. Student 3 answer, item five (Source: Field study)

Procedimiento(cálculos)
(7C2)(23C5)/(30C7)

Figure 21. Student 4 answer, item five (Source: Field study)

Procedimiento(cálculos)
P(A')= 30*29*28...24/30^7=0.4691
P(A)= 1- 0.4691= 0.53

Figure 22. Student 5 answer, item five (Source: Field study)

Student 2 answer (**Figure 19**):

> "We remove the probability that zero and one have the same birthday to be able to take out the complement."

Student 3 answer (**Figure 20**):

> "Since the problem tells you at least two people, you have to sum the odds of it being two, three, four, five, six, seven people in the 30 days of each month."

Student answer 4 (**Figure 21**):

> "From the 30 days of the month, we take away seven, which are the days that can be birthdays. It is taken out with 7C2; and two because it is the number of people we want to know who coincide on the same day. We multiply 7C2 by the remainder of the days and celebrated, 23C5, divided by the total of everything, 30C7."

Student 5 answer (**Figure 22**):

> "I did a complement procedure, where P(a)=two people have birthdays the same day, considering that a month is 30 days and P(A')=no person has a birthday the same day. Then there are seven fractions with a denominator of 30, and the numerator goes down one by one."

This problem was one of the most complicated for students, mainly due to the type of calculations needed. However, we can observe that answers based on intuition appear; for example, just over 30% of the students concluded that the probability that at least two have birthdays on the same day is greater than 0.5. Interesting answers appear. Student 1 performs calculations in which two students can be selected from a group of seven and divides it by the number of days of the month. It can be inferred that this student is only considering the case that exactly two people have birthdays on the same day, and at least two are required. The calculation of this probability is incorrect.

For his part, student 2 identifies that he should use the complement. He tries to calculate the probability that zero students have birthdays on the same day and that a student will have a birthday on the same day but tries to use a binomial model, which is not suitable for this problem. Student 3 does not use a complement but identifies that several cases are birthdays and, therefore, calculates each of these probabilities; however, when calculating each, he does it incorrectly. Student 4 attempts to use a hypergeometric distribution, which should not be applied to this model. Finally, student 5 is an example of the expected way to solve the problem, also arguing the use of the complement.

It can be observed that the use of the simulator helped some students regarding their probabilistic intuition in some aspects of the solution process, but it was not sufficient to reach the correct solution.

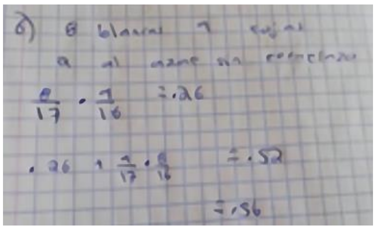

Figure 23. Student 2 answer, item six (Source: Field study)

Procedimiento(cálculos)

$$P\ (R\ dado\ que\ B) = \frac{P(B \cap R)}{P(B)}$$

$$\frac{(\frac{[illegible]}{17})(\frac{[illegible]}{16})}{8/17} = 0.5625$$

Figure 24. Student 3 answer, item six (Source: Field study)

Procedimiento(cálculos)

(9/17)(8/16)=0.26

(8/17)(9/16)(2)=0.53

(9/17)(8/16)=0.26

Figure 25. Student 4 answer, item six (Source: Field study)

Analysis of problem 6

Only 10% of the students solved this problem correctly. About 40% incorrectly calculated one of the probabilities (conditional probability), which was the students' main mistake. It is worth mentioning that some students did not even try to solve it.

Here are some answers.

Student 1 answer:

> "This option is more probable as it is easier for an event to succeed when it does not carry an order."

Student 2 answer (**Figure 23**):

> "I did the calculations to know the probabilities of each event, but at first glance, the last one also seemed the most likely because it had the least conditionals."

Student 3 answer (**Figure 24**):

> "There is a very minimal difference between the probability of the first and third subsections, but I guess this one came out bigger because of the way the balls exited."

Student 4 answer (**Figure 25**):

> "The probability is higher in the second option since it does not have a specific event order compared to the other two options."

This last problem showed that the main difficulty faced by the students was calculating conditional probability. Student 1 explained the reasoning in which the order is relevant to calculating probabilities, and the more ways that must be considered, the more the probability increases. Student 2 and student 3 performed the calculation of the probabilities expectedly. In particular, student 2 argued that subsection (c) is more likely since it had fewer conditions and, therefore, higher probability. For his part, student 3 performs the calculations in more detail using mathematical notation. Finally, student 4's answer is an example of the most common mistake. This occurs in the conditional probability since only the probability that the second event is red is sought without considering the probability of what already happened. In general, students calculated both probabilities; they made no distinction between a conditional and a joint probability (intersection).

Similar to problem 3, we observe that a significant percentage of the students failed to discern situations, where conditional elements must be considered, specifically, conditional probability calculations, after using the simulator in the classes.

In summary, the main positive or partial effects of students using the simulator are the following:

Positive effects:

1. Their ability to determine the number of games needed to guarantee a winner improved, although some failed to find the right solution, as in problem 1.
2. Probabilistic intuition was favored in students regarding avoiding making decisions based on previous results, as in problem 2.

3. The simulator activities related to problem 4 also favored the students' probabilistic intuition by identifying situations using the binomial distribution.

Partial or insufficient effects:

4. It was impossible to get most students to recognize or discern situations, where conditional elements must be considered, as in problem 3 and problem 6.
5. Using the simulator helped some students use probabilistic intuition in some aspects of the solution process, such as calculating probabilities of individual events or using the complement. Still, it was insufficient to arrive at the correct solution, as observed in problem 5.

CONCLUSIONS

The use of simulation in probability courses has been studied at various educational levels, highlighting its virtues. This study explored how simulation influences a group of mathematics students for decision-making and developing probabilistic intuition.

In the analysis carried out, we could observe and verify the evidence of elements of probabilistic intuition in the students' proposed arguments, differing from common sense that people typically use. The students generally showed reasoning that led them to the right solution; however, not all could arrive at correct answers when the problems were more complex. Despite this, we can say that the simulation platform positively impacted the students' probabilistic intuition. The results showed the importance of implementing this type of didactic strategy, considering the previous results, and leveraging the areas of opportunity to improve learning in future experiences.

The research's main objective was to perform an exploratory study on using a simulator to impact students' probabilistic intuition when solving problems. In addition, it motivated the students to learn this topic and boosted their perception of the importance of studying probability and using it in decision making.

However, the students presented significant difficulties with some conditional probability problems because they could not use the known information; instead, they took the situation proposed to them as a new or different event. This result suggests that the didactic activities designed for this aspect were not practical enough.

Another vital aspect to consider is the link between the solution pathways offered by the activities supported by a simulator and the theoretical, analytical procedures. While it is true that, in this study, the impact on the development of intuition had positive results in some probabilistic concepts of students, it does not seem to be enough. Students must be able to use their probabilistic intuition to solve problems using knowledge, mathematical notations, and processes, which lead them to the correct answers and not remain incorrect somewhere in the solution processes. The designed simulator strengthened the students' probabilistic intuition in some ways, such as avoiding making decisions based on previous results or identifying situations, where the binomial distribution should be used. Using the simulator led to good ideas about how to address particular problems. However, it is necessary to look for alternatives that link this with the ability to use resources and theoretical mathematical procedures to reach correct solutions.

It can be concluded that using simulation as a technological didactic tool helps students develop their probabilistic intuition; however, there is still much to be done in this field.

Author contributions: All authors have sufficiently contributed to the study and agreed with the results and conclusions.

Funding: The authors acknowledge the financial and technical support of the Writing Lab, Institute for the Future of Education, Tecnologico de Monterrey, Mexico, in the production of this work. The authors appreciate the financial support of Novus Grant with PEP No. PHHT011-19ZZ00003, Institute for the Future of Education, Tecnologico de Monterrey, Mexico, in the production of this work.

Ethical statement: The authors stated that ethical review and approval was not required for the study on participants in accordance with the local legislation and institutional requirements. Consent was obtained from the participants before the research was conducted. Participants were informed about the research, anonymity and their rights about withdrawing from the study without adverse repercussions.

Declaration of interest: No conflict of interest is declared by authors.

Data sharing statement: Data supporting the findings and conclusions are available upon request from the corresponding author.

REFERENCES

Abramovich, S., & Nikitin, Y. Y. (2017). Teaching classic probability problems with modern digital tools. *Computers in the Schools, 34*(4), 318-336. https://doi.org/10.1080/07380569.2017.1384687

Aizikovitsh-Udi, E., & Radakovic, N. (2012). Teaching probability by using GeoGebra dynamic tool and implementing critical thinking skills. *Procedia-Social and Behavioral Sciences, 46*, 4943-4947. https://doi.org/10.1016/j.sbspro.2012.06.364

Alvarado, H., Estrella, S., Retamal, L., & Galindo, M. (2018). Intuiciones probabilísticas en estudiantes de ingeniería: Implicaciones para la enseñanza de la probabilidad [Probabilistic insights in engineering students: Implications for the teaching of probability]. *Revista Latinoamericana de Investigación en Matemática Educativa* [*Latin American Journal of Research in Educational Mathematics*], *21*(2), 131-156. https://doi.org/10.12802/relime.18.2121

Arévalo, M. P., Fernández, J. M., & Sanjuán, E. L. (2015). Intuición y probabilidad [Intuition and probability]. *LI Olimpiada Matemática Española* [*LI Spanish Mathematical Olympiad*], *2015*, 83.

Bar-Hillel, M., & Falk, R. (1982). Some teasers concerning conditional probabilities. *Cognition, 11*(2), 109-122. https://doi.org/10.1016/0010-0277(82)90021-X

Batanero, C. (2006). Razonamiento probabilístico en la vida cotidiana: Un desafío educativo. Investigación en el aula de matemáticas [Probabilistic reasoning in everyday life: An educational challenge. Research in the mathematics classroom]. In *Proceedings of Statistics and Chance* (pp. 1-17).

Batanero, C., Begué, N., & Gea, M. M. (2018). ¿Cómo desarrollar el sentido del muestreo en los estudiantes? [How to develop the sense of sampling in students?] In *Proceedings of the 3rd Colombian Meeting of Stochastic Education*.

Batanero, C., Cañadas, G., Arteaga, P., & Contreras, J. M. (2014). La paradoja del niño o niña: Aplicaciones para la clase de probabilidad [The boy or girl paradox: Applications for the class of probability]. *Revista Digital Matemática* [*Mathematics Digital Magazine*], *14*(1), 1-13. https://doi.org/10.18845/rdmei.v14i1.1562

Batanero, C., Hernández-Solís, L. A., Gea, M. M., & Álvarez-Arroyo, R. (2021). Comparing probabilities in urns: A study with primary school students. *Uniciencia, 35*(2), 129-143. https://doi.org/10.15359/ru.35-2.9

Diccionario de la Lengua Española [Dictionary of the Spanish Language]. (2021). *Real Academia Española, 23.ª ed., versión 23.4 en línea* [*Royal Spanish Academy, 23rd ed., version 23.4 online*]. https://dle.rae.es

Escobar, B. R. P., Carhautocto, G. R., Salazar, C. A. H., & Sánchez, W. A. C. (2022). Competencias transversales en el contexto educativo universitario: Un pensamiento crítico desde los principios de gamificación [Transversal competences in the university educational context: Critical thinking from the principles of gamification]. *Revista Prisma Social* [*Social Prism Magazine*], *38*, 158-178.

Frassia, M. G., & Serpe, A. (2017). Computer-based activity's development for probability education in high school. In *Proceedings of the International Conference on New Horizon in Education*.

Koparan, T., & Yilmaz, G. K. (2015). The effect of simulation-based learning on prospective teachers' inference skills in teaching probability. *Universal Journal of Educational Research, 3*(11), 775-786. https://doi.org/10.13189/ujer.2015.031101

Lesser, L. M., & Glickman, M. E. (2009). Using magic in the teaching of probability and statistics. *Model Assisted Statistics and Applications, 4*(4), 265-274. https://doi.org/10.3233/MAS-2009-0137

Lukác, S., & Gavala, T. (2019). Interactive learning environment supporting visualization in the teaching of probability. *ICTE Journal, 8*(1), 48-60. https://doi.org/10.2478/ijicte-2019-0005

Marín-Díaz, V. (2015). La gamificación educativa. Una alternativa para la enseñanza creative [Educational gamification. An alternative for creative teaching]. *Digital Education Review, 27*.

Martins, R. M. D. C. (2018). Learning the principles of simulation using the birthday problem. *Teaching Statistics, 40*(3), 108-111. https://doi.org/10.1111/test.12164

Maxara, C., & Biehler, R. (2006). Students' probabilistic simulation and modeling competence after a computer-intensive elementary course in statistics and probability. In *Proceeding of the 7th International Conference on the Teaching of Statistics* (pp. 1-6).

Memnun, D. S., Ozbilen, O., & Dinc, E. (2019). A qualitative research on the difficulties and failures about probability concepts of high school students. *Journal of Educational Issues, 5*(1), 1-19. https://doi.org/10.5296/jei.v5i1.14146

Mezhennaya, N. M., & Pugachev, O. V. (2018). On the results of using interactive education methods in teaching Probability Theory. *Problems of Education in the 21st Century*, *76*(5), 678. https://doi.org/10.33225/pec/18.76.678

Nickerson, J. A., & Zenger, T. R. (2004). A knowledge-based theory of the firm. The problem-solving perspective. *Organization Science, 15*(6), 617-632. https://doi.org/10.1287/orsc.1040.0093

Ramiro-Miranda, A., López Valencia, M. L., & Serra, S. V. (2022). Percepciones de estudiantes de fonoaudiología sobre el aprendizaje basado en simulación y gamificación [Speech-language pathology students' perceptions of simulation-based learning and gamification]. *Educación Médica Superior* [*Higher Medical Education*], *36*(2).

Rodríguez-Alveal, F., Díaz-Levicoy, D., & Vásquez-Ortiz, C. (2018). Evaluación de la alfabetización probabilística del profesorado en formación y en activo [Assessment of the probabilistic literacy of teachers in training and active]. *Estudios Pedagógicos (Valdivia)* [*Pedagogical Studies (Valdivia)*], *44*(1), 135-156. https://doi.org/10.4067/S0718-07052018000100135

Rosado, R., Castillo, R., Yescas, F., & Santiago, R. (2017). Probabilidad y estadística con gamificación y aprendizaje adaptativo [Probability and statistics with gamification and adaptive learning]. *Proyecto Novus 2016* [*Novus Project 2016*].

Ruiz, E. F. (2018). Empleo de aplicaciones tecnológicas en el tratamiento de temas de probabilidad y estadística. Dificultades presentadas por los estudiantes en la formulación de planteamientos correctos [Use of technological applications in the treatment of probability and statistics issues. Difficulties presented by students in the formulation of correct approaches]. *Revista Iberoamericana Para la Investigación y el Desarrollo Educativo* [*Iberoamerican Journal for Educational Research and Development*], *8*(16), 216-245. https://doi.org/10.23913/ride.v8i16.339

Tintle, N., Chance, B., Cobb, G., Roy, S., Swanson, T., & VanderStoep, J. (2015). Combating anti-statistical thinking using simulation-based methods throughout the undergraduate curriculum. *The American Statistician, 69*(4), 362-370. https://doi.org/10.1080/00031305.2015.1081619

Vásquez-Ortiz, C., & Alsina, Á. (2017). Aproximación al conocimiento común del contenido para enseñar probabilidad desde el modelo del conocimiento didáctico-matemático [Approach to common content knowledge to teach probability from the didactic-mathematical knowledge model]. *Educación Matemática* [*Mathematics Education*], *29*(3), 79-108. https://doi.org/10.24844/EM2903.03

Zamora, I. (2014). *Desarrollo de la intuición probabilística en educación primaria* [*Development of probabilistic intuition in primary education*] [Master's thesis, Universidad de Valladolid].