An exploratory study of the difficulties related to the probabilistic modeling process

Abdennabi Lakrim[1], Mohamed Chergui[2], Bouazza El Wahbi[1]
[1]Department of Mathematics, Laboratory LAGA, Faculty of Sciences, University Ibn Tofail, Kenitra, Morocco
[2]Department of Mathematics, EREAM Team, Regional Center of Training and Education Professions, Kenitra, Morocco

Article Info

Article history:

Received Mar 30, 2024
Revised Sep 5, 2024
Accepted Sep 19, 2024

Keywords:

Didactics of mathematics
Difficulties in modeling
Modeling
Probabilistic modeling cycle
Probabilities

ABSTRACT

This study is a contribution to efforts to promote practices for dealing with the difficulties encountered by learners in probabilistic modeling situations. We attempt to elucidate as precisely as possible the types of difficulties that secondary school students face in the process of modeling with probability tools. By referring to a large and extensive literature on the specific difficulties of modeling, we have been able to establish a typology of possible difficulties that may emerge during the implementation of the modeling process by secondary school students. We determined sub-categories for the probabilistic modeling stages. Based on this typology, we constructed a test and we administered it to a random sample of secondary school students. The analysis of the responses via a grid developed for this purpose, we were able to deduce that the sub-categories of difficulties we identified are both positively correlated. With regard to student performance, some particular difficulties were encountered in the implementation of the modeling stages. We also observed that the choice of the appropriate probabilistic model is hampered by a lack of conceptual understanding of the probabilistic situation, while the stage of probabilistic calculations is mainly affected by the failure to choose the correct mathematical methods and techniques.

Corresponding Author:

Mohamed Chergui
Department of Mathematics, EREAM Team, Regional Center of Training and Education Professions
Avenue, Abdel aziz Boutaleb, Kenitra, Morocco
Email: chergui_m@yahoo.fr

1. INTRODUCTION

Probability is of great importance in the teaching of mathematics, offering a significant opportunity to develop several skills for learners, in particular those that are mathematical as affirmed by Brase *et al.* [1]. More specifically, probability promotes the development of language that gives students the ability to explain phenomena and gives them the resources they need to make critical decisions. For this last point, we refer the interested reader to the study on insights into decision-making using choice probability undertaken by Crapse and Basso [2] and the references therein. However, teaching and learning probability is not exempt from certain difficulties that are sometimes inherent to this concept, and which are perceived by both students and teachers. This last statement is the result of numerous studies carried out on this topic, as reported in [3] or pointed out recently by Batanero and Álvarez-Arroyo [4]. For example, students, during their first exposure to probability in high school, may find the subject particularly difficult due to its paradoxes and the abstract nature of its fundamental concepts. In their study of secondary school students' learning difficulties based on their level of probability understanding, Anggara *et al*. [5] revealed that students have difficulties in describing the sample, constructing the events of an experiment, developing mathematical models, and understanding the principles of evaluating and interpreting experimental results. These authors were able to deduce that students have epistemological obstacles. By this category of obstacles we mean, as defined by Vergnaud [6], any difficulty linked to learners' pre-existing conceptions, which may be valid in some contexts and inappropriate in others. In this context, Ang and Shahrill [7] described several misconceptions about probability and developed the following classification of misconceptions:

- Equiprobability bias: the student tends to assume that random events are equally probable in nature. He regards the chances of different outcomes as equally probable events;
- Beliefs: student thinks that the outcome of an event depends on a force beyond their control. Sometimes this force is luck or wishes;
- Human control: people generalize the behavior of random generators such as dice, coins and spinning tops. They believe that the results depend on how these different devices are thrown or manipulated.

Identifying students' difficulties in learning probability is of great usefulness, as it enables teachers to prepare adapted learning processes and overcome potential obstacles [5]. Understanding these obstacles is an essential step to teachers to search for appropriate methods and attenuate the potential difficulties encountered by students during their learning process [5]. It is necessary to recognize that these obstacles may arise from pre-existing conceptions and differences between learning contexts, requiring a well-considered didactic approach to promote a thorough understanding of the concept of probability [6]. Girard has identified a variety of difficulties that frequently confront the teaching of probability at different educational levels, encompassing epistemological, didactic, mathematical, psychological and modeling-related obstacles [8].

In a study conducted among a group of high school students in the United States, Garfield and Ahlgren [9] identified three main challenges faced by the majority when learning probability. Firstly, students may have difficulty conceptualizing probability, struggling to understand this abstract notion as a measure of chance or frequency relative to the occurrence of a certain event. They may also experience difficulties in perceiving random events, struggling to recognize and interpret them correctly. Finally, probabilistic reasoning can be problematic, as students may have difficulty applying the rules of probability calculation, estimating probabilities or interpreting results in a meaningful way in a given context as asserted by Saldanha and Liu in [10].

In his study to assess students' understanding of probability concepts and their associated vocabulary, Green [11] observed that students have difficulty mastering and using the vocabulary associated with probability, such as "certain", "probable", and that systematic teaching methods are needed to develop a deeper and more complex understanding of the notion of probability for students. The role of language as a source of difficulties in learning probability was also emphasized by the authors in [12]. Aware of this problem, Post and Prediger [13] have conducted research to identify the conditions for successful work on multiple semiotic representations to improve secondary school students' conceptual understanding of conditional probabilities.

Research carried out by Godino *et al*. [14] has identified several difficulties encountered by teachers when teaching probability. These include a limited understanding of the different nuances of probability and the concept of hazard, as well as reasoning errors such as associating hazard with equiprobable results. In addition, some teachers find it difficult to relate their mathematical knowledge to concrete probability problems.

In summary, teaching probability in high school presents a number of complex challenges for students, including the introduction of concepts linked to the use of set vocabulary, the logic of events, and the application of different probability knowledge. It follows, then, that it is necessary to seek effective approaches to overcoming this widely reported state of dissatisfaction. In a move towards pedagogical reform, which has also focused on paradigms, several studies have called for the integration of mathematical modeling as a learning objective and as a method for teaching mathematics [15]. This has been particularly true for the teaching of probability, which plays a central role in prediction, decision-making and many other areas of application [16]. This proposal is based, in part, on the observation that many students find this particular area of mathematics abstract and challenging to understand [17]. This is precisely where modeling comes into its own as a promising and powerful teaching method [18]. Indeed, modeling offers a pedagogical framework that involves creating visual representations, real-life simulations and hands-on experiments to help students understand probabilistic concepts. Moreover, in the modeling process, which will be described later, it becomes possible for students to see the practical application of probability through the involvement of real data [19].

In the context of teaching and learning probability, the implementation of modeling makes it possible to visualize and manipulate probabilistic situations in a concrete way. To simulate games of chance, such as throwing dice or drawing cards, for example, we can use computer tools that enable students to acquire an intuitive and practical understanding of probabilistic concepts such as conditional probability and the independence of events [20]. Simulations and interactive models enable students to manipulate variables and see how they affect probabilities, thus reinforcing learning. Students can consolidate their understanding by schematizing the situations studied with Venn diagrams and probability trees. These are visual representation tools that help students to represent and solve complex probability problems [21].

It follows, then, that by using concrete models, teachers can make probabilistic concepts more accessible to students and help them develop crucial mathematical and critical thinking skills. Therefore, modeling is proving to be a valuable pedagogical tool for teaching and learning probability, providing significant advantages in terms of understanding probabilistic concepts. However, it is essential to consider the challenges associated with this approach in order to conceive effective teaching programs and efficient pedagogical scenarios. Motivated by these issues, we set out in this article to explore the difficulties that secondary school students may encounter in the probabilistic modeling process. In connection with this problematic, we define our research questions as: i) what types of difficulties encountered by high school students when implementing probabilistic modeling process? and ii) what are the main traits that characterize the types of difficulties encountered by students?

The main objective of this work is to explicitly identify the sub-steps of probabilistic modeling that pose difficulties for students. Consequently, the added value of this research is to provide practicing teachers with the ability to make appropriate adjustments to the teaching process and to the learning of probability. It should also be emphasized that this work is a contribution to filling the gap in studies that have looked at the difficulties in implementing modeling using probabilistic tools.

2. THEORITICAL FRAMEWORK

Didacticians have conceptualized mathematical modeling in various manners, considering it to be a crucial step linking the concrete world to the world of mathematics. A wide range of studies in various fields support this view, as shown in [22] and [23], for example. Two distinct definitions of mathematical modeling emerge, each having its roots in distinct pedagogical approaches.

In the French literature, the definition is based on the necessity to define a mathematical model as a central activity. Abassian *et al.* [24] describe modeling as a symbolic representation of certain aspects of the real world. This perspective has its place in the context of problem-based learning, an approach used to highlight the applicability of mathematics. Indeed, mathematical skills can be effectively acquired using appropriate mathematical or non-mathematical frameworks. This proves invaluable in improving the skills and didactic thinking of pre-service teachers, as pointed out by Güner and Erbay [25].

As part of the process of formalizing the notion of modeling, we should mention that some authors [26] refers to an intermediate model between the real situation and the mathematical model, called a pseudo-concrete model, which represents a first level of abstraction from reality. For the German authors, mathematical modeling is a cyclical process consisting of stages linking real-life situations and mathematics in both directions. In connection with this framework, we provide in Table 1 the sub-competencies characterized by Greefrath and Vorhölter [27] in accordance with the modeling cycle developed by Blum and Leiß [28].

To improve probabilistic problem solving through modeling, some authors, as stated by Pérez and Delgadillo in [19], inspired by the work of Blum and Leiß [28] and Shaughnessy [29], have developed a modeling cycle adapted to this domain. This modeling process consists of seven distinct and complementary stages, each playing a crucial role in the creation and validation of the model. The first step consists in simplifying the real experiment in linguistic terms, with the aim of transforming the real experiment into a random situation that is easier to model. The main target is to adapt this experience to usual situations such as random draw patterns or dice throws. This stage is immediately followed by the structuring of the model, where it is necessary to choose or build a model that corresponds to the simplification carried out. This means mobilizing all the data to match the situation studied to a mathematical model. Next, mathematical activities are undertaken. This involves formalizing the model developed in the previous stage, enabling the construction of a precise probabilistic model. For example, models such as Bernoulli's can be used. Next comes the purely mathematical work involved in solving the probabilistic problem associated with the mathematical model constructed. This requires the probabilistic calculation of events and random variables, for example. The results obtained are then subjected to an initial validation. This is performed by checking that the results do not conflict with the properties of probability.

After this stage, the results are submitted for final validation, taking into account the real-life context and the underlying logic of the problem posed. We check whether the results are logically coherent with the actual situation studied. This is the interpretation stage, followed by a final reformulation of the answers. Once the results have been validated and interpreted satisfactorily, the answers are reformulated using clear, precise language, to present the conclusions drawn from the model in a way that is understandable and useful for solving the original problem.

In conclusion, we can consider that modeling involves a non-mathematical element at the beginning and at the end of the process. The modeling process can be described schematically by distinguishing three stages: the development of a model based on reality, the operation of the model itself within the framework of mathematics, and the comparison of the model's results with reality. An important stage in modeling is when students move from their own intuitive mathematical strategies to formal working methods. This insight into the notion of modeling and its stages enabled us to design the methodological framework described in the next section.

Table 1. Sub categories of modeling

Sub-competency	Explanation
Understanding	Students represent the problematic situation and form their special mental model. This allows them to acquire an understanding of the issue.
Simplifying	Students distinguish between important and irrelevant data about a realistic situation.
Mathematizing	Students convert simplified real-life situations into equations, figures, diagrams, and functions., thus forming a mathematical model.
Working mathematically	Students employ some heuristic approaches and use their mathematical background to solve the mathematical problem.
Interpreting	The students transfer the results deduced from the model to the real context and thus obtain tangible results.
Validating	Students examine the appropriateness of the actual findings in the situation model.
Exposing	Students match the answers found in the model with the actual data and thus develop an answer to the main question.

3. METHODOLOGY

To answer the questions posed in this study, and based on the literature review presented above, we will use a test to explore the difficulties experienced during the various stages of probability modeling. The results obtained will be analyzed both quantitatively and qualitatively. Quantitative methods aim to produce objective information that can be clearly expressed by means of statistics and figures, while qualitative methods attempt to provide more in-depth information on the phenomena under study. This complementarity between the two approaches ensures greater rigor in the study undertaken

3.1. Data collection

The course on probability is included in the Moroccan secondary school curriculum [30]. At the beginning of the course, enumeration tools are introduced to be used in the Laplacian definition of the probability of an event. This is followed by an inventory of concepts that will be conveyed in the course, such as random experiment, eventuality and event. After the presentation of certain properties of probabilistic calculus, the program stipulates the study of certain particular situations such as the independence of events and random experiments, conditional probability and the binomial model. Our survey will be carried out by administering a test consisting of 10 questions, each conceived in the light of the skills targeted by the educational guidelines for probability courses. Table 2 lists the questions and skills required to produce answers. The test was administered in May 2023 to 56 final year students (17-18 years old) in the experimental sciences discipline after having followed the calculation course on probability. Test participants come from various secondary schools in the Rabat-Sale-Kenitra Regional Academy for Education and Training.

Table 2. Items correspondent test

	Question statement	Required abilities
Q1	An urn contains indistinguishable balls distributed as follows: 4 red, 3 black and 5 green. 3 balls are drawn successively from the urn. a) What is the number of possible draws? b) What is the number of draws containing 3 green balls?	Using of the appropriate enumeration model
Q2	Give an interpretation for each of the following numbers: A_7^3 ; $\binom{8}{2}$; 5^4.	
Q3	An unrigged die with numbers from 1 to 6 is rolled once. One student considers that there is more chance of getting an even number than a multiple of 3. Do you share this student's conclusion?	Using probability calculus
Q4	The distribution of the number of students in a secondary school is 240, 90 and 70 for the Science, Literature and Technical disciplines respectively. We randomly select a student from this secondary school. What is the probability that the chosen student is from the technical discipline?	
Q5	Three bulbs from a set of 20 are randomly selected at the same time, 7 of which are defective. Calculate the probability of having at least one defective bulb.	
Q6	In a population, 35% of individuals are smokers, and of these, 75% are infected with disease M. Of the remaining individuals, 30% are infected with disease M. An individual is randomly selected from this population. What is the probability of this individual being infected with disease M?	Using conditional probability
Q7	An urn contains 10 indiscernible balls, three of which are green and numbered 0-1-1, four of which are red and numbered 1-1-2-2, and three of which are black and numbered 0-0-2. Two balls are drawn successively from the urn, without being returned. We are interested in the sum of the numbers on the drawn balls. Calculate the probability of each possible value for this sum.	The use of a random variable
Q8	A player participates in a game in which a balanced six-sided die numbered 0-0-0-1-1-2 is rolled once. If the die shows 0, the player must pay 2 Dirhams, and 1 Dirham if the die shows 1. The player wins 2 dirhams if the die shows 2. Let us denote by X the random variable that associates with each number obtained the value paid or won by the player. Determine the probability distribution of the random variable X.	
Q9	The probability distribution of a random variable X equal to the number of green balls drawn from a bag is given by P(X=1) = 1/7, P(X=2) =2/7 and P(X=3) = p. Determine the value of p and determine the average number of green balls that can be drawn from the bag.	The use and interpretation of the mathematical expectation of a random variable.
Q10	In a factory, the probability of a resistor being inoperative is equal to 0.02. A sample of 30 resistors is taken at random. What is the probability that exactly 7 resistors will be inoperative?	Using the binomial distribution in a probabilistic situation.

3.2. Results analysis tools

On the basis of the conceptual framework outlined in the previous section, we have developed the grid shown in Table 3, which presents the corresponding difficulties at each stage of probabilistic modeling, along with explanations for identifying these difficulties. Difficulties noted in the students' responses to the test questions are classified according to the sub-categories presented in Table 3. Each sub-category of difficulties corresponding to each stage has been coded to simplify processing of the data collected. An enumeration of the numbers in each sub-category is carried out in order to perform a statistical study. For this purpose, we have chosen to process the data using SPSS and EXCEL software.

Table 3. Difficulties related to the probabilistic modeling process

Modeling steps	Code	Difficulties corresponding to each stage	Description of difficulties
B: Choosing the appropriate probabilistic model	B_1	Lack of conceptual understanding of the probabilistic situation.	– Confusion between certain concepts. – Misunderstanding of some concepts. – Difficulties due to language formulation of the situation
	B_2	Not realizing the dependencies between variables.	– Difficulty in identifying elements describing the probabilistic situation studied (colors, numbers, and objects drawn or left in an urn.)
	B_3	Inability to identify conditioning situations.	– For example, link stochastic independence to chronological order.
C: Probabilistic calculations	C_1	Failure to identify events correctly.	– A lack of accurate perception of events.
	C_2	Inability to establish possible relationships between events (compatibility, independence, ...)	– For example, confusing the intersection of events with conditional probability.
	C_3	Failure to choose correct mathematical methods and techniques.	– The student may be confused between the use of combinatorial tools and the law of total probability.
	C_4	Not understanding the conditions under which certain properties are used.	– For example, calculate the probability of the intersection and union of two events by the product or the sum of their probabilities respectively.
D: Mathematical processing	D_1	Failure to use appropriate algorithms and resolution strategies.	– Lack of initiatives that can lead to expected results, such as the use of diagrams and probability trees.
	D_2	Technical errors	– Errors in algebraic calculation or in the application of certain algebraic properties.
	D_3	Not solving the model due to excessive complexity.	– This is the case, for example, when a random experience is made up of several phases, or involves several objects or actors.
E: Mathematical validation	E_1	Failure to identify the correct meaning of aspects of mathematical results.	– This is the case, for example, with the inability to interpret the mathematical expectation of a random variable or the values of its distribution function.
	E_2	Failure to assess the consistency of results obtained in mathematical processing.	– This is the case, for example, when failing to recognize that events are contrary or form a complete system.

4. RESULTS AND DISSCUSSION

We begin this section by presenting the statistics for students who actually answered the various test questions. For a clear comparison of the numbers obtained, we convert the data in Table 4 into a graph (Figure 1). As an initial comment on these data, we would like to quote the following:

- The number of students who answered the test questions can be considered low in global terms.
- The number of answers provided for the first two questions on counting is higher than for the questions on probability.
- The number of answers to certain questions is too low, as is the case for Q6, Q7 and Q8.
- The number of students who answered question Q4 is not compatible with the number of students who answered question Q5, even though it concerns the same probabilistic tool.

To gain a better understanding of these preliminary remarks, and to make proper use of the results obtained to provide answers to our research questions, we will carry out an analysis of the data collected at each modelling stage. This will be done first with a univariate analysis, followed by a cross-sectional analysis. The numbers of occurrences of each of the difficulties related to the probabilistic modeling process are given in Table 5. For a more comprehensive view of the quantitative results obtained, we present the following illustrations and discussions for each of the four stages concerned by our study. The results of choosing the appropriate probabilistic model are plotted in Figure 2.

Table 4. Number of students responding per question

Question	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10
Number of students responding	34	32	18	16	28	7	8	6	12	7

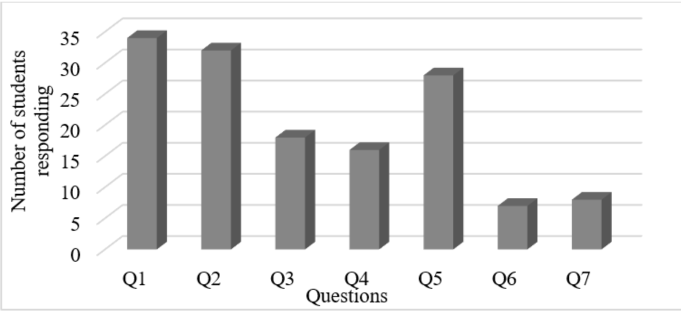

Figure 1. Number of students responding

Table 5. Summary of difficulties statistics

Difficulties	B1	B2	B3	C1	C2	C3	C4	D1	D2	D3	E1	E2
Number	145	27	1	8	1	32	25	30	14	12	7	19

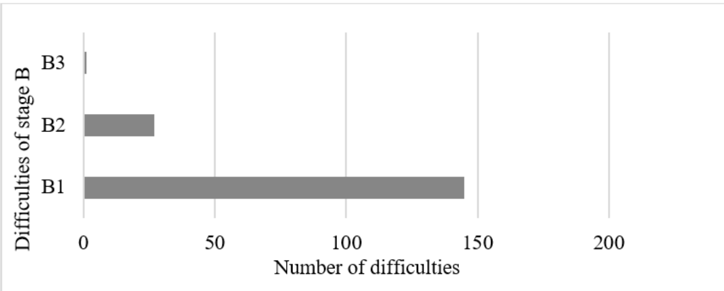

Figure 2. Redundancy of difficulties related to choosing the appropriate probabilistic model

It is very clear that difficulties linked to inability to identify conditioning situations are almost absent, unlike those due to lack of conceptual understanding of the probabilistic situation which are dominant. Particularly in the answers to questions Q1 and Q2, it was noted that students had difficulty identifying the mathematical models of enumeration appropriate to each situation. The answers produced for question Q2 reveal, on the one hand, a poor conceptualization of enumeration tools and, on the other, the importance of teaching probability through modeling situations, by virtue of the dialectic it offers between the abstract mathematical framework and the applicability of mathematical knowledge. Another source to which we should pay close attention is the complexity of the language used to formulate statements in probabilistic situations. In the productions of the tested students, this language factor took on several aspects, as described below:

- Difficulties due to misunderstanding of terms and concepts. Difficulties were noted in the use of notations, particularly those used in enumeration. Specific probability terms carry particular meanings for students. For example, the expression “we randomly select” is interpreted as equiprobability. In question Q8, several students deduce from the data “a balanced six-sided die” that the probability of each of the three numbers 0, 1 and 2 is equal to 1/6.
- Difficulties of vocabulary complexity: generally, when solving a probability problem, the student is faced with language emanating from three registers. The first is totally linked to a real-life context, and can create cultural problems for the student or even the teacher. The second register is that of enumeration. For example, students have great difficulty distinguishing between a combination and an arrangement. The last is that of probabilities, which involves terms that may have specific connotations for students, such as conditional probability, which is often confused with the probability of the intersection of two events.

– Difficulties arising from not realizing the importance of probabilistic concepts in formulating answers simply and clearly. To illustrate this point, we give the example of students who were unable to represent the different events referred to in question Q7.

The second difficulty that was clearly identified in the students' answers was the lack of realizing the dependencies between variables. This was observed in the confusion between the values of a random variable and their corresponding probabilities. In question Q9, a large proportion of students answered for the probability of the value 3 by 3/7. In connection with this type of difficulty, it was noted that in question Q6, some students concluded that the probability of an individual being infected with disease M is 30%+70%. Here again, there was a conceptual confusion between percentages and probabilities. These two examples show that students are unable to assign the right status to data in a probabilistic situation. For the probabilistic calculations stage, as can be clearly seen in Figure 3, difficulties related to failure to choose correct mathematical methods and techniques are the most frequent, followed by those due to not understanding the conditions under which certain properties are used.

In question Q10, some students answered that the probability that exactly 7 resistors will be inoperative is 7/10. They are inappropriately applying the Laplacian approach based on counting the eventualities of a random experiment. Clearly, the conditions for applying this approach are not available to the student, who is instead put in the situation of applying the binomial model. The failure to choose the right probabilistic tool was also observed in the students' answers to question Q6. Indeed, the correct answer requires the use of the total probability formula, but some students proceeded to sum certain percentage provided in the given data.

Concerning modeling stage C, difficulty subcategories C3 and C4 were successively followed by types C1 and C2. An example of failure to identify events correctly is concretized in the students' erroneous answers to question Q3. Indeed, a large number of students (82%) failed to provide an answer to this question, and many of the others failed to explain the events whose probabilities are involved in the comparison. To clarify the C2-type difficulties observed, we refer to the answers produced by students on question Q9, where several students were unable to realize the relationship between the events involved in the question statement. In fact, they were asked to determine the probability P(X=3). This can be done directly by noticing that the events (X=1), (X=2) and (X=3) form a complete system of events, and therefore the sum of their probabilities is equal to 1.

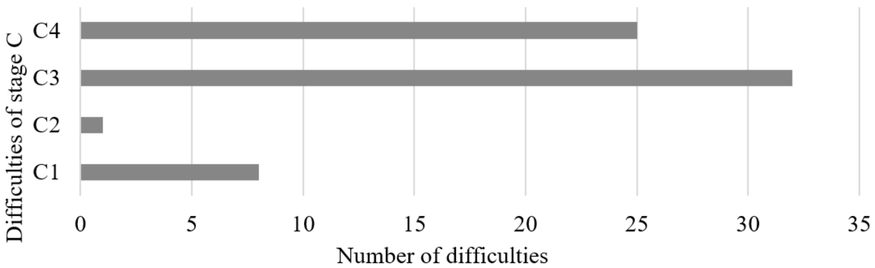

Figure 3. Redundancy of difficulties related to probabilistic calculations

In the mathematical processing modeling stage, the failure to use appropriate algorithms and resolution strategies is the most frequent type of difficulty among the students taking part in the test. This is perfectly aligned with the very low rate (around 14%) of students who answered questions Q6 and Q7. This may be justified by a lack of initiatives that can lead to expected results. Indeed, in both questions, it is possible to proceed by the representation of a probability tree which allows the explicit description of events and the determination of their probabilities. The other two types, D2 and D3, are almost equal and do not take on large values (Figure 4). The comparatively small number of technical errors does not allow us to conclude that the students' algebraic calculation skills are well developed, since calculative tasks in probability calculation are generally straightforward.

Inability to solve the model due to excessive complexity can be seen as the main factor explaining the low number of students who answered question Q8. Indeed, the determination of the probability distribution of the random variable X in the proposed situation is not a simple task, but is made up of several steps, starting with an inventory of the events arising from the random experiment to reach the possible values of the random variable, and ending with the calculation of the probability associated with each value. For the last stage, which concerns the mathematical validation of the results obtained by implementing the mathematical model, it turns out that difficulties due to failure to assess the consistency of results obtained in mathematical processing are quite prevalent among students, compared with those linked to failure to identify the correct meaning of aspects of mathematical results, as illustrated in Figure 5.

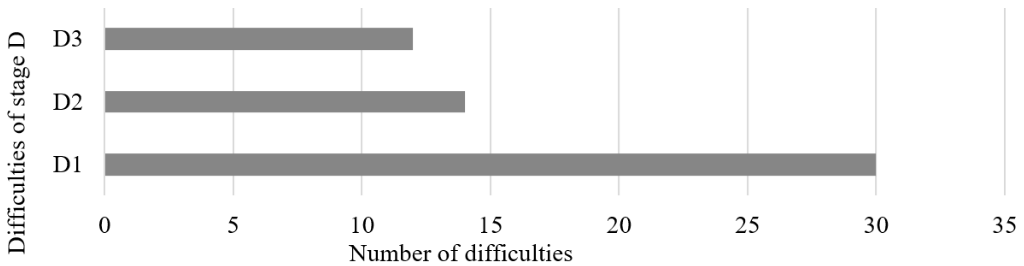

Figure 4. Redundancy of difficulties related to mathematical processing

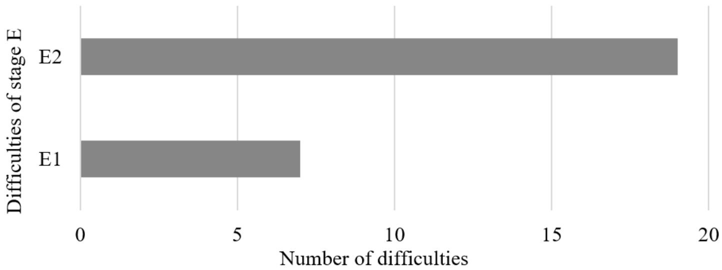

Figure 5. Redundancy of difficulties related to validation

Validation problems in probability problems are not at all strange, given the nature of this field of mathematics, which combines formal thinking and empiricism. In probability, the student does not always find the tools to elaborate rigorous proofs, as in algebra or geometry, where certain types of reasoning are generally applicable. The error made in question Q9 in determining the value of the probability p could be avoided by realizing that the sum of the three probabilities is equal to 1. Similarly, in question Q8, assigning the value 1/6 to the probability of the value 0 obtained by throwing the die was not submitted to any validation criteria. The cross-tabulation of the variables for the probability modelling stages revealed some significant correlations, at the 0.05 level (two-tailed), as shown by the values presented in the Table 6 which were computed using SPSS software. It should also be noted that the determinant of the correlation matrix gave the value 0.287, which allows us to reject that there is no correlation significantly different from 0 between the variables.

The first thing to note from Table 6 is that all correlations are positive. Most correlations are moderate. For example, sub-category C3 is moderately correlated with all other sub-categories except E2. On the other hand, for some sub-categories, the correlation is too weak. This latter situation concerns, for example, the sub-category pairs (B1, C2), (B1, D1), (B2, D2), (B3, D3) and (D1, E1). These results seem logical in the sense that mastery of one probability modeling sub-skill can only benefit performance in the other sub-categories.

Table 6. Correlation matrix

	B1	B2	B3	C1	C2	C3	C4	D1	D2	D3	E1	E2
B1												
B2	.118											
B3	.108	.195										
C1	.389	.451	.289									
C2	.090	.275	.047	.111								
C3	.262	.240	.287	.158	.267							
C4	.379	.245	.201	.142	.211	.366						
D1	.016	.126	.218	.414	.230	.385	.096					
D2	.473	.027	.284	.091	.437	.344	.281	.412				
D3	.352	.468	.001	.235	.109	.400	.414	.066	.500			
E1	.228	.181	.355	.111	.127	.500	.470	.053	.410	.156		
E2	.405	.139	.280	.107	.140	.058	.111	.347	.347	.340	.172	

5. CONCLUSION

Probability is crucial to real-life phenomena, and teaching it is a challenge. The importance of modeling in teaching and learning, particularly for learning probability, is underlined by many studies. However, its implementation is difficult due to conceptual obstacles and pedagogical difficulties linked to the management of modeling situations in the classroom. In the present work, we focus on the difficulties encountered by secondary school students in the various stages of probabilistic modeling. This has the advantage of making students' failures in probabilistic modeling explicit and, consequently, offering the opportunity to implement relevant regulations. We therefore carried out a literature review that enabled us to develop a typology of difficulties corresponding to each modelling stage.

Based on this typology, we examined the work of 56 secondary school students on a test involving probability situations. The examination consisted in classifying the errors identified according to 12 possible subcategories. Considerable number of students were unable to answer several questions on the test. This can be explained by an inability to apply probability knowledge. This inability is accentuated when students are confronted with complex situations. In terms of categories of difficulty, the students tested had more difficulty in choosing the appropriate probabilistic model. This was sometimes manifested in a misunderstanding of the data involved in the situation. In the second stage of modelling, namely probabilistic calculations, failure to choose correct mathematical methods and techniques dominated, followed by not understanding the conditions under which certain properties are used. In the third step of working mathematically, failure to use appropriate algorithms and resolution strategies was the most common, while technical errors were infrequent. For the validation of the mathematical model used, a failure to assess the consistency of results obtained in mathematical processing was clearly observed. It is important to point out that the cross-tabulation of the results obtained from the students' responses showed a positive correlation between different types of difficulty. This encourages us to continue this work in the future with a multivariate study to determine the main factors that explain the types of difficulties identified in the student probability model. It will also be an opportunity to go beyond the limitations of the present study, such as sample size and school level.

REFERENCES

[1] G. L. Brase, S. Martinie, and C. Castillo-Garsow, "Intuitive conceptions of probability and the development of basic math skills," in *Probabilistic Thinking. Advances in Mathematics Education*, 2014, pp. 161–194, doi: 10.1007/978-94-007-7155-0_10.

[2] T. B. Crapse and M. A. Basso, "Insights into decision making using choice probability," *Journal of Neurophysiology*, vol. 114, no. 6, pp. 3039–3049, Dec. 2015, doi: 10.1152/jn.00335.2015.

[3] S. Lukáč and T. Gavala, "Interactive learning environment supporting visualization in the teaching of probability," *International Journal of Information and Communication Technologies in Education*, vol. 8, no. 1, pp. 48–60, Mar. 2019, doi: 10.2478/ijicte-2019-0005.

[4] C. Batanero and R. Álvarez-Arroyo, "Teaching and learning of probability," *ZDM – Mathematics Education*, vol. 56, no. 1, pp. 5–17, Feb. 2024, doi: 10.1007/s11858-023-01511-5.

[5] B. Anggara, N. Priatna, and D. Juandi, "Learning difficulties of senior high school students based on probability understanding levels," *Journal of Physics: Conference Series*, vol. 1013, p. 012116, May 2018, doi: 10.1088/1742-6596/1013/1/012116.

[6] G. Vergnaud, "A comprehensive theory of representation for mathematics education," *The Journal of Mathematical Behavior*, vol. 17, no. 2, pp. 167–181, 1998, doi: 10.1016/S0364-0213(99)80057-3.

[7] L. H. Ang and M. Shahrill, "Identifying students' specific misconceptions in learning probability," *International Journal of Probability and Statistics*, vol. 3, no. 2, pp. 23–29, 2014.

[8] B. Chaput, JC. Girard, M. Henry, "Frequentist approach: modelling and simulation in statistics and probability teaching," In *Teaching Statistics in School Mathematics-Challenges for Teaching and Teacher Education, New ICMI Study Series*, vol. 14, Springer, 2011, doi: 10.1007/978-94-007-1131-0_12.

[9] J. Garfield and A. Ahlgren, "Difficulties in learning basic concepts in probability and statistics: implications for research," *Journal for Research in Mathematics Education*, vol. 19, no. 1, p. 44, Jan. 1988, doi: 10.2307/749110.

[10] L. Saldanha and Y. Liu, "Challenges of developing coherent probabilistic reasoning: rethinking randomness and probability from a stochastic perspective," in *Probabilistic Thinking, Advances in Mathematics Education,* Springer, 2014, pp. 367–396, doi: 10.1007/978-94-007-7155-0_20.

[11] D. R. Green, *Probability concepts in school pupils aged 11–16 years*. 1982.

[12] A. M. Nacarato and R. C. Grando, "The role of language as a source of difficulties in learning probability," *Statistics Education Research Journal*, vol. 13, no. 2, pp. 93–103, 2014.

[13] M. Post and S. Prediger, "Connecting representations for understanding probabilities requires unfolding concepts in meaning-related language," in *Twelfth Congress of the European Society for Research in Mathematics Education (CERME12)*, 2022, no. 15.

[14] J. D. Godino, C. Batanero, and M. Burgos, "Theory of didactical suitability: an enlarged view of the quality of mathematics instruction," *Eurasia Journal of Mathematics, Science and Technology Education*, vol. 19, no. 6, p. em2270, Jun. 2023, doi: 10.29333/ejmste/13187.

[15] W. Blum and M. Niss, "Applied mathematical problem solving, modelling, applications, and links to other subjects ? State, trends and issues in mathematics instruction," *Educational Studies in Mathematics*, vol. 22, no. 1, pp. 37–68, Feb. 1991, doi: 10.1007/BF00302716.

[16] A. Leung, "Boundary crossing pedagogy in STEM education," *International Journal of STEM Education*, vol. 7, no. 1, p. 15, Dec. 2020, doi: 10.1186/s40594-020-00212-9.

[17] B. D. Lazić, J. B. Knežević, and S. M. Maričić, "The influence of project-based learning on student achievement in elementary mathematics education," *South African Journal of Education*, vol. 41, no. 3, pp. 1–10, Aug. 2021, doi: 10.15700/saje.v41n3a1909.

[18] P. E. Spector, C. C. Rosen, H. A. Richardson, L. J. Williams, and R. E. Johnson, "A new perspective on method variance: a measure-centric approach," *Journal of Management*, vol. 45, no. 3, pp. 855–880, Mar. 2019, doi: 10.1177/0149206316687295.

[19] K. M. Pérez and E. M. Delgadillo, "The mathematical work of chilean secondary teachers in the domain of probability, " *Canadian Journal of Science, Mathematics and Technology Education*, vol. 22, pp. 558–575, 2022, doi: 10.1007/s42330-022-00235-0.

[20] T. Koparan, "The impact of a game and simulation-based probability learning environment on the achievement and attitudes of prospective teachers," *International Journal of Mathematical Education in Science and Technology*, vol. 53, no. 9, pp. 2319–2337, Sep. 2022, doi: 10.1080/0020739X.2020.1868592.

[21] J. D. Vermunt, "Understanding, measuring and improving simulation-based learning in higher education: student and teacher learning perspectives," *Learning and Instruction*, vol. 86, p. 101773, Aug. 2023, doi: 10.1016/j.learninstruc.2023.101773.

[22] H. Janan, H. Yaagoubi, and M. T. Abouchadi, "3D mathematical thermal analyses of PA12 powder bed on the SLS process by two numerical methods," *International Journal on "Technical and Physical Problems of Engineering*, vol. 14, no. 2, pp. 74–83, 2022.

[23] H. Al-Araji, A. Al-Zughaibi, and E. Hussein, "Testing two types of magneto-rheological (mr) damper models with quarter car suspension system response," *Technical and Physical Problems of Engineering*, vol. 1, no. 54, pp. 52–61, 2023.

[24] A. Abassian, F. Safi, S. Bush, and J. Bostic, "Five different perspectives on mathematical modeling in mathematics education," *Investigations in Mathematics Learning*, vol. 12, no. 1, pp. 53–65, 2019, doi: 10.1080/19477503.2019.1595360.

[25] P. Güner and H. N. Erbay, "Prospective mathematics teachers' thinking styles and problem-solving skills," *Thinking Skills and Creativity*, vol. 40, 2021, doi: 10.1016/j.tsc.2021.100827.

[26] P. V. Hernández, J. Huincahue, P. Cumsille and A. Nechache, "Articulating the Blomhøj modelling cycle and the mathemat working spaces: analysis of a task in higher education," *Acta Scientiae*, vol. 24, no. 7, pp. 62-91, 2022, doi: 10.17648/acta.scientiae.7135.

[27] G. Greefrath and K. Vorhölter, *Teaching and learning mathematical modelling*. Cham: Springer International Publishing, 2016, doi: 10.1007/978-3-319-45004-9.

[28] W. Blum and D. Leiß, "How do students and teachers deal with modelling problems?," in *Mathematical Modelling*, Elsevier, 2007, pp. 222–231, doi: 10.1533/9780857099419.5.221.

[29] J. M. Shaughnessy, "Misconceptions of probability: an experiment with a small-group, activity-based, model building approach to introductory probability at the college level," *Educational Studies in Mathematics*, vol. 8, pp. 295–316, 1977, doi: 10.1007/BF00385927.

[30] MEN, "General educational guidelines and mathematics program for the qualifying secondary cycle," *Ministry of National Education*, 2024. [Online]. Avalilable: men.gov.ma/Ar/Documents/Examens_2024/EXAM NATIONAL BAC 2024/CDR Adapté EXAM NAT BAC- MATH_ FILIERE SM Op Ang 2024.pdf.

BIOGRAPHIES OF AUTHORS

Abdennabi Lakrim has been teaching vacation courses in mathematics didactics at the Kénitra Faculty of Science since 2020. She received a master's degree in teaching and training professions from the Faculty of Sciences, University Ibn Tofail, Kenitra, Morocco in 2018. Currently, he is a Ph.D. student in Mathematics Didactics, and Experimental Sciences at the same faculty. His research focuses on teaching and learning probabilities. He works as a teacher of mathematics in secondary school, in Kenitra, Morocco. He can be contacted at email: abdennabi.lakrim@uit.ac.ma.

Mohamed Chergui received his Ph.D. degree from the Faculty of Sciences, Moulay Ismail, Meknés, Morocco in 2009 in Applied Functional Analysis. He is currently full Professor at Regional Center of Training and Education Professions (CRMEF) in Kenitra, Morocco. His researches are focused on Mathematics and Didactics. He is an author of more than 30 refereed scientific articles. He can be contacted at email: chergui_m@yahoo.fr.

Bouazza El Wahbi is a distinguished university professor, the head chief of Department of Mathematics at Faculty of sciences at University Ibn Tofail I Kenitra. He received his Ph.D. from the University at University Mohamed V in Rabat. He conducts research on many fields of mathematics and on teaching mathematics in secondary and high schools. Particularly, he is interested by analyzing classroom practices. These studies have been published in more than 35 refereed articles. He has directed 12 Ph-d thesis projects. He can be contacted at email: bouazza.elwahbi@uit.ac.ma.