Journal on Mathematics Education
Volume 16, No. 2, 2025, pp. 603-632

Teaching conditional probability in grade 12 using realistic mathematics education theory

Quoc Ai Nguyen[1], Nam Hong Dao[2,*], Tuong Anh Hoa[1], Ngoc-Giang Nguyen[3]

[1]Faculty of Mathematics and Applications, SaiGon University, Ho Chi Minh City, Vietnam
[2]University of Medicine and Pharmacy at Ho Chi Minh City, Viet Nam
[3]Faculty of Data Sciences in Business, Ho Chi Minh University of Banking, Ho Chi Minh City, Vietnam
*Correspondence: daohongnam@ump.edu.vn

Received: 26 December 2024 | Revised: 25 June 2025 | Accepted: 28 June 2025 | Published Online: 29 June 2025

Abstract

The perspective on applying mathematics to solve practical problems is clearly expressed in Vietnam's Mathematics General Education Curriculum in 2018. Therefore, Realistic Mathematics Education (RME) is suitable for the goals of mathematics education in Vietnam and can become one of the fundamental theories for designing mathematics teaching models in high schools. This study presents an experimental result of teaching conditional probability based on RME and following a five-step process we develop, including problem setting, experience, rediscovering conditional probability, forming conditional probability, application. Worksheets containing problems were distributed to 42 students to solve, and their work was collected immediately afterwards. The works will be analyzed to identify emergent models constructed by students. The new result of the paper is from the study of De Lange, Jupri and Drijvers, we propose a mathematization process for teaching conditional probability in Vietnam. The data analysis method is a qualitative method, through observing the students' problem-solving process. Experimental results show that students actively participated in the process of reinventing conditional probability through solving learning tasks.

Keywords: Bayes' Theorem, Conditional Probability, Model For, Model Of, Realistic Mathematics Education

How to Cite: Nguyen, Q. A., Dao, N. H., Hoa, T. A., & Nguyen, N.-G. (2025). Teaching conditional probability in grade 12 using realistic mathematics education theory. *Journal on Mathematics Education, 16*(2), 603–632. https://doi.org/10.22342/jme.v16i2.pp603-632

Realistic Mathematics Education (RME) is a pedagogical approach that emphasizes students' active construction of mathematical understanding through engagement with meaningful, real-world problems. Within this framework, students develop their understanding by exploring and solving contextually grounded problems that resonate with their experiences, while teachers provide scaffolding to support the reinvention of mathematical ideas (Freudenthal, 2002). A defining feature of RME is the prominent role of rich, "realistic" situations in the learning process. These situations serve both as the starting point for developing mathematical concepts, tools, and procedures, and as a context in which students later apply and formalize their knowledge, gradually moving toward greater abstraction and generalization (Van den Heuvel-Panhuizen & Drijvers, 2020).

According to Van den Heuvel-Panhuizen (2000), the contemporary form of RME is strongly shaped by Freudenthal's philosophical view that mathematics must be meaningfully connected to reality, closely linked to learners lived experiences, and relevant to society in order to be of human value. Rather than treating mathematics as a static body of knowledge to be transmitted, Freudenthal conceptualized it as a human activity. From this perspective, mathematics education should not focus on mathematics as a closed formal system, but rather on the process of mathematization—the dynamic activity of organizing reality through mathematical reasoning (Freudenthal, 1968).

Treffers (1991) identified five foundational principles of RME: (1) the use of meaningful contexts; (2) progression through levels of understanding and the use of models; (3) opportunities for reflection and engagement in special tasks; (4) the integration of social interaction; and (5) collaborative structuring of knowledge. These principles were later refined by Van den Heuvel-Panhuizen (2000; 2005), who articulated them under six headings: the activity principle, the reality principle, the level principle, the intertwinement principle, the interaction principle, and the guidance principle (Inci et al., 2023; Van den Heuvel-Panhuizen, 2000). Finally, the goals of RME align closely with the aims of Vietnam's 2018 Mathematics General Education Curriculum (VMGEC), issued by the Ministry of Education and Training. The curriculum emphasizes creating opportunities for students to apply mathematics in solving real-world problems, fostering connections between mathematical ideas, between mathematics and daily life, and between mathematics and other academic disciplines and educational activities.

Since 2020, Realistic Mathematics Education (RME) has attracted increasing attention from mathematics education researchers in Vietnam. Several studies have examined the historical development, conceptual foundations, and characteristics of RME, as well as the challenges associated with its implementation in the Vietnamese educational context. Other works have explored the application of RME to the teaching of specific mathematical topics at the high school level. In general, these studies suggest that RME is both feasible and beneficial for mathematics education in Vietnam, aligning closely with the vision articulated in the 2018 Mathematics General Education Curriculum (VMGEC) issued by the Ministry of Education and Training, which emphasizes connecting mathematics to real life and fostering students' mathematical competencies. Key contributions include those by Da (2022), Le et al. (2021), Loc et al. (2020), Nguyen (2020), Nguyen et al. (2020), and Tran and Tran (2021).

Nguyen (2020) provides a theoretical overview of RME, tracing its historical development and comparing the notions of mathematization and mathematical modeling. While the study presents the five core principles of RME, it lacks practical illustrations and does not offer examples to clarify the theoretical distinctions between mathematization and modeling. In contrast, Da (2022) offers a more applied perspective. This study includes diagrams of horizontal and vertical mathematization, outlines a teaching process, and identifies five characteristics of RME. Notably, the author provides an example of applying RME to the teaching of calculus. However, the discussion remains limited in scope and similarly does not address the distinction between mathematization and modeling in sufficient depth.

The work by Loc et al. (2020) proposes a lesson design framework based on RME, including a seven-step instructional process and a well-developed experimental activity. Although the study was conducted prior to the implementation of the 2018 VMGEC and relies heavily on previously established findings, it offers an original contribution in the form of a structured design methodology. Nevertheless, the study focuses solely on short-term learning outcomes and does not investigate long-term knowledge retention. Furthermore, Tran and Tran (2021) explore the potential of RME for teaching probability and statistics in medical and pharmaceutical universities. While the study discusses RME's historical background, educational objectives, and its core principles, it remains largely conceptual. The study highlights the relevance of mathematical modeling but does not present concrete applications of RME in instructional practice, nor does it report on any empirical teaching experiments.

Nguyen et al. (2020) proposed a framework for evaluating the development of RME in Vietnam by examining its integration into educational policy and practice. Their framework consists of four key components: (1) national vision and strategy; (2) curriculum and instructional materials; (3) assessments and examinations; and (4) teacher professional development. However, while the framework is useful at the systemic level, the study does not explore the practical implementation of RME in teaching specific mathematical topics. Furthermore, the article lacks a discussion of quantitative methodologies or empirical analysis to support its claims.

Similarly, the study by Le et al. (2021) provides a general overview of RME-related research and offers statistical data concerning practical issues encountered in the implementation of the 2018 VMGEC. While informative, the study shares a similar limitation with Nguyen (2020) in that it does not investigate the specific application of RME in classroom teaching or curriculum design for particular mathematical content areas.

In the present study, we build on these previous theoretical insights by operationalizing RME principles in the context of teaching conditional probability. Specifically, we adopt three key design heuristics proposed by Gravemeijer (1994): guided reinvention (or progressive mathematization), didactical phenomenology, and the use of self-developed models. Furthermore, Freudenthal (1968) conceptualized mathematics as the mathematization of reality—a process whereby learners actively organize real-world phenomena through mathematical reasoning. He emphasized that mathematics should not be taught as a finished system of abstract rules, but rather as a dynamic human activity. Later, Freudenthal (1971) elaborated on this idea, noting that the act of mathematizing is not simply a matter of translating a problem into a symbolic language; instead, symbolization itself can emerge organically through the organization of the subject matter (Gravemeijer & Terwel, 2000).

De Lange (2006) explains that the process of mathematization unfolds through the following phases: (1) beginning with a real-world problem; (2) identifying and structuring the problem using relevant mathematical concepts; (3) progressively abstracting the problem from its real-world context; (4) solving the reformulated mathematical problem; and (5) interpreting the mathematical solution in relation to the original context. The first three stages constitute the transition from real-world experience to formal mathematical representation. Building on Freudenthal's work, Treffers (1987) distinguished between horizontal mathematization—the process of transforming a contextual situation into a mathematical problem—and vertical mathematization—the internal development and refinement of mathematical ideas. Vertical mathematization involves progressing to higher levels of abstraction, often through problems that allow for multiple levels of solution strategies (Gravemeijer & Terwel, 2000). Freudenthal (2002) emphasized the complementary and equally important roles of both horizontal and vertical mathematization. While horizontal mathematization helps students connect mathematics to reality, vertical mathematization fosters deeper mathematical thinking through internal reorganization within the mathematical system. As Van den Heuvel-Panhuizen and Drijvers (2014; 2020) caution, overemphasizing real-world contexts can risk undermining students' development of abstract mathematical reasoning.

The principle of guided reinvention, as formulated by Freudenthal (2002), Gravemeijer (1994), and Gravemeijer and Terwel (2000), places emphasis on the learning process rather than on the discovery of entirely new knowledge. The aim is to enable students to experience mathematics as something they build and understand personally—knowledge that feels owned and meaningful. Historical developments in mathematics can serve as a heuristic guide for this process. More frequently, however, students' informal strategies are seen as anticipatory of formal procedures and thus provide the foundation for reinvention. In such cases, mathematizing these informal strategies becomes the essence of the reinvention process. Consequently, instructional tasks are designed to allow for a variety of student approaches, particularly those that reflect possible learning trajectories toward more formal mathematical understandings.

Gravemeijer (1999) and Larsen (2018) argue that guided reinvention serves as a foundational heuristic—or even a mission statement—for any instructional design grounded in RME. Rather than offering prescriptive instructional strategies, guided reinvention highlights the importance of identifying contexts that are experientially real to learners. The goal is not merely to situate problems in everyday-life contexts, but to ensure that the contexts meaningfully resonate with students' lived experiences and support the emergence of mathematical reasoning.

The second design heuristic employed in this study is didactical phenomenology, a concept introduced by Freudenthal (1983). He defined it as the study of how mathematical concepts, structures, and ideas originate from the need to organize the phenomena of the physical, social, and mental world. Didactical phenomenology seeks to describe these concepts in relation to the phenomena for which they were created, extending this analysis to inform instructional design for learners. Freudenthal emphasized that mathematics instruction should reflect the ongoing process of learning that mirrors the historical development of mathematical ideas—not by reproducing history, but by identifying the living roots of mathematical thinking that learners can meaningfully access, while discarding unproductive historical detours.

Larsen (2018) interprets this notion by clarifying that the "thought-matter" in Freudenthal's work refers to the mathematical ideas that the teacher or curriculum designer aims for students to learn. The process of "organizing" refers to mathematizing—the act of structuring phenomena mathematically. Didactical phenomenology, then, requires instructional designers to identify contexts that can be effectively organized by the target mathematical ideas, thereby supporting learners in constructing those ideas themselves.

Gravemeijer and Terwel (2000) further emphasize that didactical phenomenology encourages the use of phenomenologically rich situations—those that are "begging to be organized"—as opposed to simply presenting abstract concepts through concrete materials. The instructional objective is to identify contexts in which students can construct the intended mathematical objects through engagement with meaningful situations. In this way, the "thought-matter" (noumenon) is constructed through the learner's effort to make sense of the "phenomenon."

The third design heuristic utilized in this study is the development and use of student-generated models, which play a critical bridging role between informal strategies and formal mathematical reasoning. Streefland (1985) articulated how models can evolve from being closely tied to a specific context (model of) to becoming generalized tools for mathematical reasoning (model for). Initially, students develop situation-specific models as they attempt to represent and solve real-world problems. Through generalization and abstraction, these models gradually gain autonomy and become applicable across a broader range of problems, enabling students to reason mathematically beyond the original context (Van den Heuvel-Panhuizen, 2003).

Gravemeijer (1994) stated that the aim of instruction is to create opportunities for learners to construct their own models, which can eventually evolve into more formal systems of reasoning. As students engage in problem-solving within real-world contexts and share their strategies through discussion, they begin to build increasingly sophisticated models. These models, initially grounded in specific situations, undergo a process of generalization and formalization and ultimately serve as frameworks for abstract reasoning (Yackel et al., 2003).

Gravemeijer (1994) uses the term model to refer to both situation models and mathematical models that are constructed by students. In the early stages of learning, a model functions as a model of a familiar situation (see Figure 1). As students gain experience, the model becomes a model for reasoning across multiple contexts, thereby facilitating a transition from informal to formal mathematical understanding (Zulkardi, 2010).

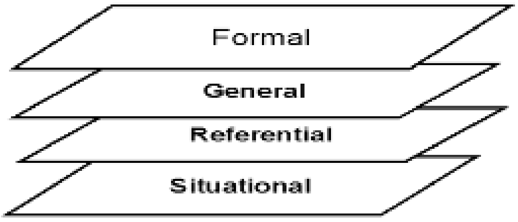


Figure 1. Levels of models in RME (Gravemeijer, 1994)

Da (2022) emphasizes the central role of modeling in bridging the gap between informal and formal mathematical understanding. For modeling to serve this function effectively, learners must transition from constructing a model of a specific contextual situation to developing a model for more generalized mathematical reasoning. This transition involves a progression across four levels, as articulated by Zulkardi (2010):

1. Situational level – learners apply domain-specific knowledge and strategies that are directly tied to the context of the problem;
2. Referential level (model of) – models and strategies refer to the original situation described in the problem, enabling initial structuring and interpretation;
3. General level (model for) – the model becomes a generalized mathematical tool, where strategies and reasoning begin to dominate over the contextual reference;
4. Formal mathematical level – learners operate with abstract mathematical procedures and conventional notations, independent of the initial context.

Zulkardi (2010) also highlights four key components that should be integrated into RME-based lesson planning: learning goals, instructional content, methodology, and assessment. These elements ensure that instructional design supports students' progression through the levels of modeling and fosters deep conceptual understanding. Furthermore, in the domain of teaching conditional probability, numerous studies have investigated effective pedagogical strategies at both secondary and tertiary levels. These studies can be grouped into three major thematic categories.

1. Studies on contextual teaching of conditional probability

   Vicent et al. (2012) employed a structural approach to identify, classify, and analyze ternary problems in Spanish mathematics textbooks, emphasizing the importance of context in understanding conditional probability. Shao (2015) warned that neglecting the cognitive processes involved in students' knowledge construction may result in superficial understanding, particularly regarding the relationship between theoretical probability and empirical frequency. To address this, the study advocated for the use of rich, practical examples. Similarly, Huerta (2009) proposed a context-based, problem-solving approach—especially with ternary problem structures—to improve students' understanding and probabilistic literacy.

2. Studies on visual tools and cognitive strategies
   Gigerenzer (2002) conducted a series of experimental studies to determine which representational formats are most effective for teaching conditional probability. His findings demonstrated that using tree diagrams with absolute frequencies (referred to as “natural frequencies”) significantly improved student performance and comprehension, outperforming traditional symbolic approaches such as the Bayesian formula. These visual tools also facilitated greater solution acceptance. Similarly, Díaz et al. (2010) noted that although independence and conditional probability can be defined intuitively, learners often struggle to apply these definitions in specific contexts, leading to cognitive biases and reasoning errors.
   Binder et al. (2020) examined the effectiveness of “net diagrams” as visual supports in a study involving 249 university students. Their results indicated that net diagrams could improve students’ reasoning about conditional probabilities to a degree comparable with the use of 2×2 tables and double-tree diagrams. Across all studies, frequency-based visualizations consistently outperformed probability-based representations. In particular, frequency nets were found to be as effective as frequency doubletrees, suggesting that visual strategies may help mitigate the cognitive load associated with abstract probabilistic reasoning.
3. Studies addressing curricular challenges and epistemological conflicts
   Numerous studies have identified the inherent challenges in teaching and learning conditional probability, particularly those arising from its conceptual complexity and epistemological ambiguity. Tomlinson and Quinn (1997) noted that students often struggle with the abstract nature of conditional probability and its formal representations, which frequently conflict with their intuitive, experience-based understandings. Carranza and Kuzniak (2008) conducted a curricular analysis and highlighted the tension between objectivist and subjectivist interpretations of probability. While many curricula emphasize probability as relative frequency (the objectivist perspective), they simultaneously incorporate examples requiring subjective reasoning, particularly those involving the Bayesian formula. This inconsistency often leads to confusion among students, especially when probabilistic reasoning is conflated with causal thinking.
   Díaz and Batanero (2009) further illustrated that even students with adequate mathematical knowledge tend to struggle with paradoxical situations involving conditional probability. Learners often fail to recognize recurring structures or transfer insights across problems, suggesting that instruction in this area must address each type of problem individually. One possible reason for these difficulties is the essential role of context in conditional probability problems, which increases their cognitive demands. Huerta (2009) criticized existing research for its narrow focus on a limited subset of problem types, thereby failing to represent the diversity of challenges that learners face. Borovcnik (2011; 2012) argued for a more central role of probability within stochastic curricula, stating that probability serves three essential functions: as a tool for understanding inferential statistics, as a model for representing reality, and as a form of thinking that enables reflection on uncertainty. According to Borovcnik (2012), conditional probability lies at the heart of probability theory. Given the conceptual difficulty of probability itself, conditional probability is often even more challenging to understand. While conditional probabilities can be intuitively inferred from context, they are more commonly introduced through formal calculation using the Bayesian formula. Borovcnik characterizes conditional probabilities and the Bayesian formula as theoretical "twins"—central not only to Bayesian interpretations but to probabilistic thinking more broadly.

Batanero et al. (2016) emphasized that while the concept of dependence in probability is bidirectional, conditional probability entails an inherently asymmetric relationship that is often misinterpreted as causal. For instance, a positive test result may indicate the presence of a disease, but a disease does not cause the test result. In many real-world cases, causal and probabilistic reasoning are intertwined. The challenge, often addressed through statistical methods, is to distinguish between causal influence and random variation.

Batanero and Álvarez-Arroyo (2024) synthesized recent research on the teaching and learning of probability, identifying key topics such as children's probabilistic thinking, the impact of visualizations on conditional probability reasoning, teacher education, and probability modeling. Their epistemological analysis cautions that while simulation-based approaches may help reduce abstraction, they also risk reducing probability to a purely frequentist perspective, neglecting its broader conceptual foundations.

Conditional probability was formally introduced into the Vietnamese national mathematics curriculum through the 2018 Vietnam Mathematics General Education Curriculum (VMGEC). However, it was not until the 2024–2025 academic year that this topic began to be taught in Grade 12. As a result, research on the teaching and learning of conditional probability in Vietnam is currently limited. A review of the available literature reveals only a few relevant studies.

Nguyen (2022) conducted two experiments focused on conditional probability. The first experiment involved the use of teaching situations designed to support the formation of conceptual understanding and enhance students' conditional probability problem-solving skills. The second experiment evaluated students' attainment of knowledge and skill standards as specified in the VMGEC. Nguyen and Quach (2023) proposed several pedagogical innovations for teaching statistics and probability in Grade 6, emphasizing the integration of information and communication technology. Meanwhile, Bui and Tran (2024) introduced a teaching model grounded in social constructivist theory. Their study presented a five-step instructional cycle for solving conditional probability problems and illustrated the model through ten designed teaching situations.

Notably, no studies—either within Vietnam or internationally—have explored the teaching of conditional probability through the lens of the Realistic Mathematics Education (RME) framework. While existing literature addresses various challenges in teaching this topic—including the role of conditional probability in deductive reasoning, the presence of epistemological tensions, and the design of instruction using ternary problems—several critical questions remain unanswered:

1. How do students develop a meaningful understanding of conditional probability?
2. How can students explore and construct formal models of conditional probability in real-life contexts?
3. How can instruction support the long-term retention and application of conditional probability concepts in solving real-world problems?

These questions form the basis for the present study. Specifically, we aim to design a teaching process for conditional probability grounded in RME principles and to conduct a teaching experiment to evaluate the feasibility and effectiveness of this instructional process in a high school context.

Given that conditional probability has only recently been added to the Vietnamese high school curriculum, there is growing interest among educators in how to design meaningful instructional tasks for this topic. However, teachers face challenges in creating real-world contexts that enable students to rediscover the concept and formula of conditional probability through guided reinvention. According to the VMGEC (Vietnam Ministry of Education and Training, 2018), the expected learning outcomes for Grade 12 students include the ability to recognize and define the concept of conditional probability, interpret the meaning of conditional probability in familiar practical situations, describe the total probability and Bayes' formulas using 2×2 statistical tables and tree diagrams, apply Bayes' formula to compute conditional probabilities in relevant real-world problems, and use tree diagrams to calculate conditional probabilities in statistical contexts.

In light of these curriculum goals and the existing research gap, this study is guided by the following objectives: (1) to design a conditional probability teaching process based on the RME framework, and (2) to evaluate the feasibility and effectiveness of this teaching process through an empirical classroom experiment. The study addresses the following research questions:

1. RQ1. How can a conditional probability teaching process be designed based on RME principles to enable students to reinvent the concept and computational rules of conditional probability under teacher guidance?
2. RQ2. How can the feasibility and effectiveness of the RME-based conditional probability teaching process be evaluated in the context of high school mathematics instruction?

METHODS

Design of the RME-based Lesson

Grounded in the theoretical foundations of Realistic Mathematics Education (RME) and drawing upon the lesson design framework proposed by Zulkardi (2010), we developed a seven-step process for designing mathematics lessons consistent with the RME approach:

1. Step 1: Select the mathematics topic to be taught in accordance with RME.
2. Step 2: Find situations in real contexts that contain the mathematical content to be taught.
3. Step 3: Choose a reasonable situation as the starting point of the learning process.
4. Step 4: Design learning activities to support students in creating specific models (models of).
5. Step 5: Design learning activities to support students in creating general models (models for).
6. Step 6: The teacher creates opportunities for students to state definitions or theorems from the general model results in step 5, and the teacher institutionalizes the mathematical results.
7. Step 7: The teacher chooses a similar real-life situation for students to apply the knowledge they have just learned.

This seven-step design process was implemented to construct a teaching sequence for conditional probability grounded in the RME framework (see Appendix A).

Data Collection Techniques

Two Grade 12 classes with comparable academic backgrounds at Can Thanh High School (Ho Chi Minh City, Vietnam) participated in the study. One class (12A2, n = 42) served as the experimental group and received instruction through the RME-based teaching process. The other class (12A3, n = 40) served as the control group and received instruction using traditional (non-RME) teaching methods.

To verify the comparability of the two groups prior to the intervention, we analyzed students' mid-term examination scores using descriptive statistics via SPSS software (see Appendix B). Both groups had prior instruction in basic probability concepts, including intersection and independence of events, in Grades 10 and 11. After the instructional intervention, both classes completed a conditional probability post-test designed to assess conceptual understanding and problem-solving ability. One month later, a delayed post-test was administered to the experimental group to evaluate students' retention of knowledge.

Data Analysis Techniques

Quantitative and qualitative methods were employed to analyze the data:

1. Pre-intervention comparison: Descriptive statistical analysis was conducted using SPSS to confirm the equivalence of academic performance between the experimental and control classes. Normal distribution was verified (see Appendix B).
2. Qualitative analysis of students' work: Throughout the instructional intervention, students' written work was analyzed using qualitative methods to trace the development of mathematical reasoning. Particular attention was paid to the emergence of model-of and model-for reasoning, as well as horizontal and vertical mathematization processes.
3. Post-test analysis: The immediate post-test results were analyzed quantitatively using SPSS to compare learning outcomes between the two groups. The results indicated improved performance in the experimental class relative to the control class (see Appendix B).
4. Retention analysis: One month after the intervention, we compared the delayed post-test scores of the experimental group to their immediate post-test results. The analysis showed no statistically significant difference, suggesting that students retained their understanding of conditional probability over time (see Appendix B).

Participants

The study involved students from two intact Grade 12 classes (12A2 and 12A3) at Can Thanh High School. The experimental class (12A2) comprised 42 students, while the control class (12A3) comprised 40 students. In the experimental class, students were organized into 14 groups of three. Group formation was based on academic performance categories commonly used in Vietnam: Good, Fair, and Satisfactory. As neither class included students in the "Unsatisfactory" category, each group was intentionally composed of students with varying academic levels to foster peer learning and collaborative problem solving.

Ethical Clearance

This research was designed, conducted, and reported in accordance with the ethical standards of the Committee on Publication Ethics (COPE). We ensured the following throughout the study, such as transparency in research procedures and analysis, originality of research and proper citation of prior work, informed participation and protection of student confidentiality, disclosure of any potential conflicts of interest, accountability for authorship and research integrity, and ethical approval for the study from the appropriate institutional review body. We affirm our full adherence to the COPE Code of Conduct and best practice guidelines and are committed to maintaining high standards of academic integrity throughout the research and publication process.

RESULTS AND DISCUSSION

Results of Solving 4 Problems

In Table 1 are the results of solving 4 problems in 4 situations of the teaching process including 5 steps as proposed.

Step 1. Problem setting: In this step, groups learn the content of problem 1.
Step 2. Experience: In this step, all 14/14 groups (100%) answered correctly the three questions 1a/, 1b/, 1c/ about calculating the probability of events M, C and $M \cap C$, because they had learned how to calculate the probability of a random event in grade 10, and the probability of the intersection of two events in grade 11.
For the requirement to calculate and explain how to calculate the two probabilities $P(M/C)$ and $P(C/N)$ in the two questions 1d/, 1e/, only 5/14 groups (35.7%) answered correctly, because this was the first time they approached the requirement to calculate the probability of an event when knowing that another related event had occurred before and they needed to think about how to choose the data of the events in the 2×2 two-way table reasonably. At the end of step 2, the teacher asks a representative of a group with a correct solution to present their group's work to the class. Finally, the teacher explains how to select data from the given statistical table to calculate a conditional probability.
Step 3. Rediscovering conditional probability: In this step, all 14/14 groups (100%) were successful in calculating the probabilities of events $P(A)$, $P(B)$, $P(A \cap B)$ in questions 2a/ and 2b/, because they used the same calculation method for events M, C and M∩C in problem 1. For the calculation of $P(A/B)$, $P(B/A)$ in questions 2c/ and 2d/, students do the same as the teacher's instructions in calculating the probabilities $P(M/C)$ and $P(C/N)$ in problem 1.
Particularly for the proof of the two formulas $P(A/B) = P(A \cap B)/P(B)$ and $P(B/A) = P(A \cap B)/P(A)$ in the last part of questions 2c/ and 2d/, all 14/14 groups of students proved that the expressions on both sides of the formula are equal to $a/(a+b)$. This successful proof is partly due to the students following the teacher's instructions on how to use the formula $P(M/C) = n(M \cap C)/n(C)$ in problem 1, and partly due to the Vietnamese students having good algebraic transformation skills when proving that $P(A \cap B)/P(B) = \frac{\frac{a}{a+b+c+d}}{\frac{a+b}{a+b+c+d}} = \frac{a}{a+b}$.
For the requirement to compare two probabilities $P(A/B)$ and $P(B/A)$ in question 2e/, all 14/14 groups were aware that the two values were different. However, only 7/14 (50%) groups explained in detail that these two probabilities are different because the order of occurrence of the events A and B is different, while the remaining 7/14 groups (50%) only explained that the two fractions $\frac{P(A \cap B)}{P(B)}$ and $\frac{P(A \cap B)}{P(A)}$ have different denominators.
Step 4. Forming conditional probability: In this step, all 14/14 groups (100%) were able to state the definition of conditional probability and the formula for conditional probability. The high success rate shows that this is a positive result of the experiential activity after students solved problems 1 and 2.
Step 5. Application: In this step, all 14/14 groups (100%) solved problem 3 correctly when calculating the conditional probability $P(F/W)$, after correctly calculating the component probabilities $P(W)$ and $P(W \cap F)$. This positive result also shows that students have understood and applied conditional probability to solve a simple practical situation.
For problem 4, all 14/14 groups (100%) answered question 4a/ correctly about determining the requirements of the problem.
For question 4b/, all groups correctly calculated the two probabilities $P(M)$ and $P(F)$ based on the assumption about the ratio between the number of male and female students. However, 3/14 groups (21.4%) incorrectly calculated the two probabilities $P(T/M)$ and $P(T/F)$. The reason is that these three groups understood the hypothesis “4% of male students are over 1.80 m tall” to mean “4% of students are male and over 1.80 m tall”, and “1% of female students are over 1.80 m tall” to mean “1% of students

are female and over 1.80 m tall”.

For question 4c/, the results of the groups were the same as for question 4b/, because the three groups incorrectly filled in the two probability values $P(T/M)$ and $P(T/F)$ in the probability tree diagram.

For question 4d/, all 14 groups correctly identified two routes from O to T, $O - M - T$ and $O - F - T$, and stated how to calculate the probability of event T using the multiplication rule and the probability addition rule. From the results of answering question 4d/, all groups proved the formula $P(T) = P(T/M).P(M) + P(T/F).P(F)$ to calculate the probability of event T in solving question 4e/.

For question 4f/, all 14 groups gave the formula for calculating the probability of event F given that event T has occurred as $P(F/T) = \frac{P(F \cap T)}{P(T)}$. This positive result was achieved thanks to the institutionalization of knowledge about conditional probability organized by the teacher in step 4.

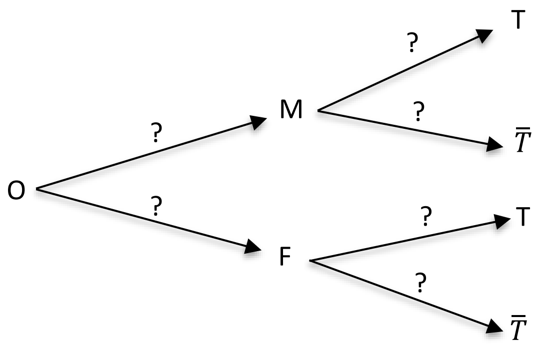
For question 4g/, all groups proved the formula $P(F/T) = \frac{P(T/F).P(F)}{P(T/F).P(F)+P(T/M).P(M)}$ using the answer to question 4e/, and used the conditional probability formula $P(T/F) = P(F \cap T)/P(F)$.

For question 4h/, 11/14 groups correctly calculated the probability $P(F/T)$ by substituting the values of the component probabilities $P(T/F)$, $P(F)$, $P(T/M)$, $P(M)$ calculated in the answer to question 4b/ into the formula $P(F/T) = \frac{P(T/F).P(F)}{P(T/F).P(F)+P(T/M).P(M)}$. The remaining three groups, which were the three groups that gave incorrect answers to question 4b/, calculated the probability value $P(F/T)$ incorrectly, because they used incorrect probability values $P(T/M)$ and $P(T/F)$ in their answers to question 4b/.

Table 1. Student response statistics table

Steps	Questions	Correct	False	No answer
1 and 2	1a/ Calculate the probability that the patient will be treated with drug M. 1b/ Calculate the probability that a patient has been cured. 1c/ Calculate the probability that a patient is cured and is treated with drug M.	42 100%	0 0%	0 0%
	1d/ Calculate the probability that the patient will be treated with drug M, knowing that the patient is cured. Explain how to calculate it. We denote the probability to be calculated as $\mathrm{P(M/C)}$. 1e/ Calculate the probability that the patient will be cured, given that the patient was treated with drug N. Explain how to calculate it. We denote this probability as $\mathrm{P(C/N)}$.	15 35.7%	27 64.3%	0 0%
3	2a/ Calculate $\mathrm{P(A)}$, $\mathrm{P(B)}$; 2b/ Calculate $\mathrm{P(A \cap B)}$;	42 100%	0 0%	0 0%
	2c/ Calculate the probability of event A, knowing that event B has occurred before, $\mathrm{P(A/B)}$, and prove that $\mathrm{P(A/B)} = \mathrm{P(A \cap B)/P(B)}$;	42 100%	0 0%	0 0%
	2d/ Calculate the probability of event B, knowing that event A has occurred before, $\mathrm{P(B/A)}$, and and prove that $\mathrm{P(B/A)} = \mathrm{P(A \cap B)/P(A)}$;	42 100%	0 0%	0 0%
	2e/ Compare $\mathrm{P(A/B)}$ and $\mathrm{P(B/A)}$. Explain why?	42	0	0

		100%	0%	0%
4	Students state the definition and formula of conditional probability, and then the teacher institutionalizes the knowledge of conditional probability.	42 100%	0 0%	0 0%
5	3/ Calculate the probability that the health insurance buyer works in a foreign company, knowing that the person is a woman.	42 100%	0 0%	0 0%
	4a/ What is the problem's requirement?	42 100%	0 0%	0 0%
	4b/ Calculate the probabilities $\mathrm{P(M)}$, $\mathrm{P(F)}$, $\mathrm{P(T/M)}$, $\mathrm{P(T/F)}$.	33 78.6%	9 21.4%	0 0%
	4c/ In the following tree diagram, called a probability tree, fill in all the probability values in the place of the question marks:	33 78.6%	9 21.4%	0 0%



	4d/ In the probability tree diagram completed in question c/, show the paths from O to T, and state how to calculate the probability that the selected student is taller than 1.80 m.	42 100%	0 0%	0 0%
	4e/ Show that: $P(T) = P(T/M).P(M) + P(T/F).P(F)$.	42 100%	0 0%	0 0%
	4f/ Write a formula to calculate the probability that the selected student is female, knowing that the student is taller than 1.80 m.	42 100%	0 0%	0 0%
	4g/ Show that: $P(F/T) = \frac{P(T/F).P(F)}{P(T/F).P(F)+P(T/M).P(M)}$.	42 100%	0 0%	0 0%
	4h/ Calculate $P(F/T)$.	33 78.6%	9 21.4%	0 0%

Teaching Process

Step 1. Problem setting

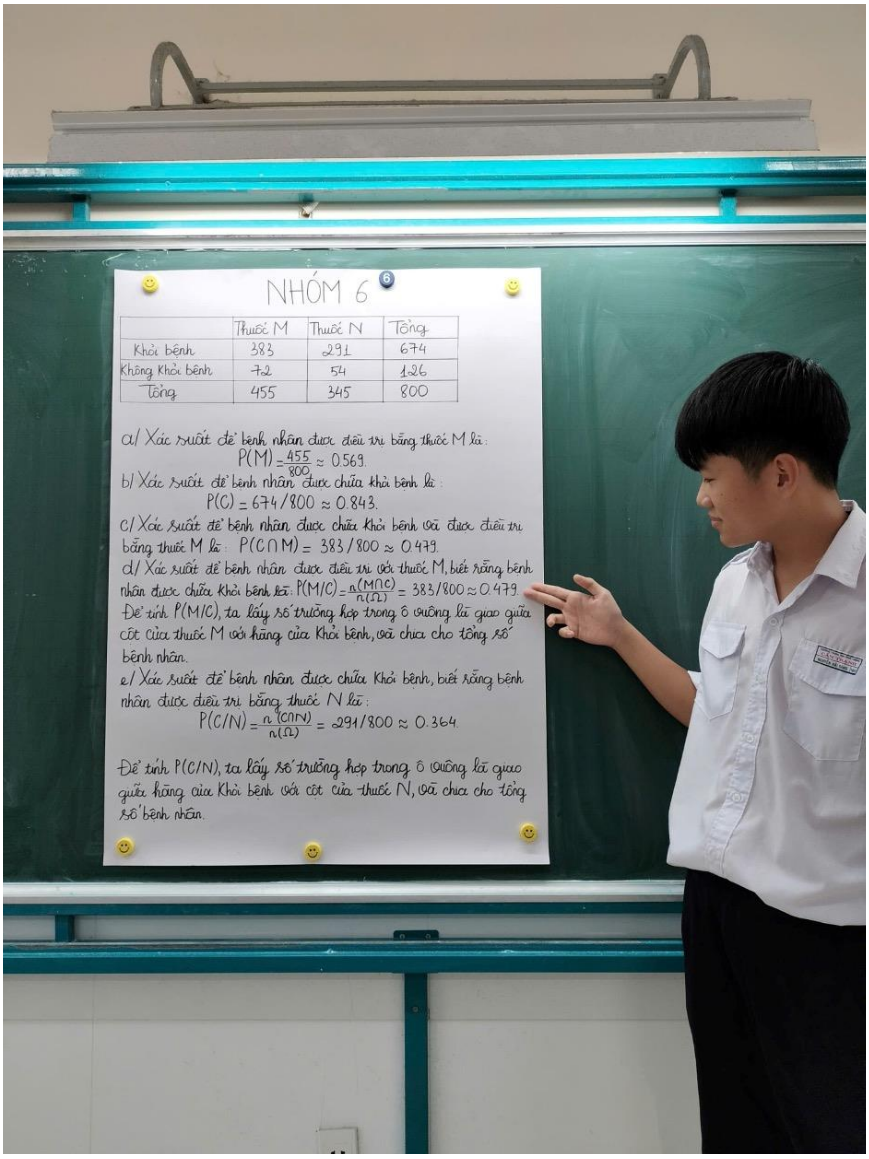
The teacher gives problem 1 and asks students to solve it.

Step 2. Experience

For problem 1, all 42 students successfully answered three questions 1a/, 1b/, 1c/ because in grade 10 students have learned and know how to calculate the probability of a random event, as well as in grade 11 the probability of intersection and union of two events.

For questions 1d/ and 1e/, 9 groups (27 students) did not calculate $P(M/C)$ and $P(C/N)$ correctly, because they could not distinguish between calculating the probability of event M knowing that C has occurred, and calculating the probability of the intersection of two events M and C. Thus, two "models of" the problem appeared in the students: $P(M/C) = \frac{n(M \cap C)}{n(C)}$ for the 5 groups (15 students) that correctly analyzed the situation and solved the problem correctly, and $P(M/C) = \frac{n(M \cap C)}{n(\Omega)}$ for the remaining 9 gourps (27 students).

The teacher asked group 6, the group with the incorrect solution, to present their work to the class. In Figure 2, a representative of group 6 is presenting their group's work in which, the group builds an incorrect specific model for $P(M/C)$ and $P(C/N)$.



Group 6

	Drug M	Drug N	Total
Cured	383	291	674
Not cured	72	54	126
Total	455	345	800

a/ The probability that the patient is treated with drug M is

$$\mathrm{P(M)} = \frac{455}{800} \approx 0.569.$$

b/ The probability that the patient is cured is

$$\mathrm{P(C)} = \frac{674}{800} \approx 0.843.$$

c/ The probability that the patient is cured and treated with drug M is

$$\mathrm{P(C \cap M)} = \frac{383}{800} \approx 0.479.$$

d/ The probability that the patient is treated with drug M, knowing that the patient is cured is:

$$\mathrm{P(M/C)} = \frac{\mathrm{n(M \cap C)}}{\mathrm{n(\Omega)}} = 383/800 \approx 0.479.$$

To calculate $\mathrm{P(M/C)}$, we take the number of cases in the square that is the intersection of the column of drug M and the row of Cured and divide it by the total number of patients.

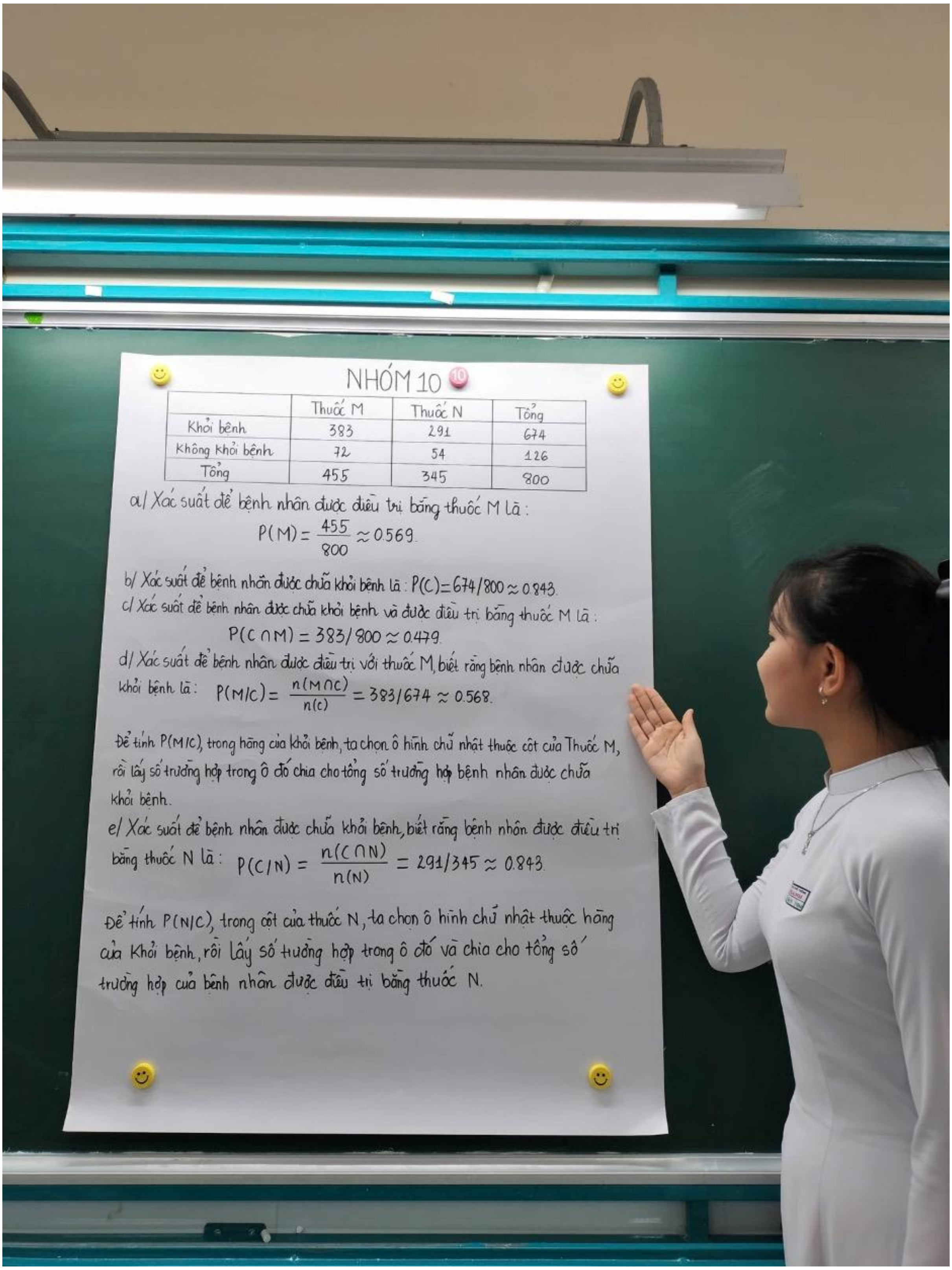
e/ The probability that the patient is cured, given that the patient is treated with drug N, is

$$\mathrm{P(C/N)} = \frac{\mathrm{n(C \cap N)}}{\mathrm{n(\Omega)}} = 291/800 \approx 0.364.$$

To calculate $\mathrm{P(C/N)}$, we take the number of cases in the square that intersects the row of Cured with the column of Drug N and divide it by the total number of patients.

Figure 2. Group 6 uses the same sample space in constructing the specific models P(M⌒C) and P(M/C)

Group 10 argued that Group 6 calculated the probabilities correctly in the first three questions a/, b/, c/, and incorrectly calculated the probabilities in questions d/ and e/. Group 10 argued that the probability $P(M/C)$ in question d/ could not be equal to the probability $P(C \cap M)$ in question c/, and group 6 used the same formula to calculate the probabilities in two different cases. The whole class agreed with Group 10's opinion. Next, the teacher asked group 10 to come to the board to present their group's work to the whole class. After listening to the representative of group 10 explain their group's solution, the whole class agreed with group 10 about the solution. In Figure 3, group 10 representatives are presenting their group's work, in which they have accurately constructed specific models for $P(M/C)$ and $P(C/N)$.


NHÓM 10
Thuốc M
Thuốc N
Tổng
Khỏi bệnh
383
291
674
Không khỏi bệnh
72
54
126
Tổng
455
345
800
a/ Xác suất để bệnh nhân được điều trị bằng thuốc M là:
P(M) = 455/800 ≈ 0.569
b/ Xác suất để bệnh nhân được chữa khỏi bệnh là: P(C) = 674/800 ≈ 0.843
c/ Xác suất để bệnh nhân được chữa khỏi bệnh và được điều trị bằng thuốc M là:
P(C ∩ M) = 383/800 ≈ 0.479
d/ Xác suất để bệnh nhân được điều trị với thuốc M, biết rằng bệnh nhân được chữa khỏi bệnh là: P(M/C) = n(M∩C)/n(C) = 383/674 ≈ 0.568
Để tính P(M/C), trong hàng của khỏi bệnh, ta chọn ô hình chữ nhật thuộc cột của Thuốc M, rồi lấy số trường hợp trong ô đó chia cho tổng số trường hợp bệnh nhân được chữa khỏi bệnh.
e/ Xác suất để bệnh nhân được chữa khỏi bệnh, biết rằng bệnh nhân được điều trị bằng thuốc N là: P(C/N) = n(C∩N)/n(N) = 291/345 ≈ 0.843
Để tính P(N/C), trong cột của thuốc N, ta chọn ô hình chữ nhật thuộc hàng của Khỏi bệnh, rồi lấy số trường hợp trong ô đó và chia cho tổng số trường hợp của bệnh nhân được điều trị bằng thuốc N.

Group 10

	Drug M	Drug N	Total
Cured	383	291	674
Not cured	72	54	126
Total	455	345	800

a/ The probability that the patient is treated with drug M is

$$\mathrm{P(M)} = 455/800 \approx 0.569.$$

b/ The probability that the patient is cured is

$$\mathrm{P(C)} = 674/800 \approx 0.843.$$

c/ The probability that the patient is cured and treated with drug M is

$$\mathrm{P(C \cap M)} = 383/800 \approx 0.479.$$

d/ The probability that the patient is treated with drug M, knowing that the patient is cured is

$$\mathrm{P(M/C)} = \frac{\mathrm{n(M{\cap}C)}}{\mathrm{n(C)}} = \frac{383}{674} \approx 0.568.$$

To calculate $\mathrm{P(M/C)}$, in the Cured row, we select the rectangular cell in the Drug M column, then take the number of cases in that cell and divide it by the total number of cases in which the patient is cured.

e/ The probability that a patient is cured, knowing that the patient is treated with drug N is

$$\mathrm{P(C/N)} = \frac{\mathrm{n(C{\cap}N)}}{\mathrm{n(N)}} = \frac{291}{345} \approx 0.843.$$

To calculate $\mathrm{P(C/N)}$, in the column of Drug N, we select the rectangular cell in the row of Cured, then take the number of cases in that cell and divide it by the total number of cases of patients treated with Drug N.

Figure 3. Group 10 constructs correct specific models for $\mathrm{P(M/C)}$ and $\mathrm{P(C/N)}$

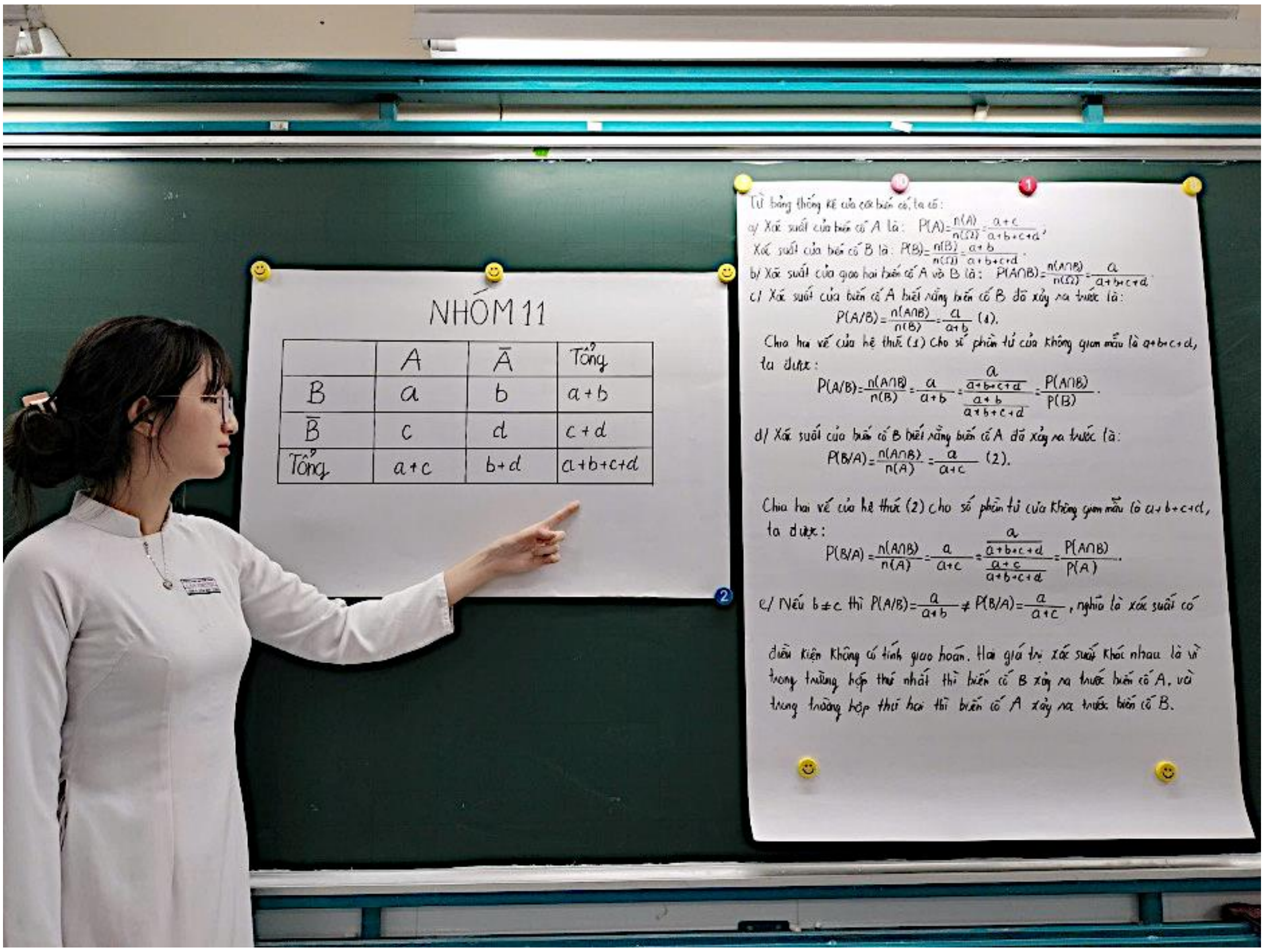
After commenting on the student's work, the teacher explains how to calculate the probability of event M knowing that event C has occurred, using the ratio $P(M/C) = n(M \cap C)/n(C)$, following these steps: first determining the number of patients who have recovered, then determining the number of patients treated with drug M among the recovered patients, and finally establishing the ratio between the two numbers determined in turn. Similarly, the teacher explains the sequence of steps for calculating the probability of event C knowing that event N has occurred by the ratio $P(C/N) = n(C \cap N)/n(N)$.
Then, the teacher introduces to the students that the two probabilities P(M/C) are called the conditional probability of event M given that event C has occurred before, and $P(C/N)$ is called the conditional probability of event C given that event N has occurred before.
Step 3. Rediscovering conditional probability
The teacher gives problem 2 and asks students to solve it.
For questions 2a/ and 2b/, all 42 students correctly calculate $P(A)$, $P(B)$, $P(A \cap B)$, by doing the same as calculating the probabilities $P(M)$, $P(C)$, $P(M \cap C)$ in problem 1.
For questions 2c/ and 2d/, all 42 students successfully calculated $P(A/B)$ and $P(B/A)$ because they did the same as the way to calculate the probability of the two events $P(M/C)$ and $P(C/N)$ in problem 1 that the teacher explained at the end of step 2.
To prove the formula $P(A/B) = \frac{P(A \cap B)}{P(B)}$, all 42 students proved that the expressions on both sides of

the formula are equal to $\frac{a}{a+b}$.The students argued that in the statistical table, in the row of event B, choose the number of elements in the cell belonging to the column of event A and divide by the total number of elements of event B to get $P(A/B) = \frac{a}{a+b}$; transform the expression $\frac{P(A\cap B)}{P(B)} = \frac{\frac{a}{a+b+c+d}}{\frac{a+b}{a+b+c+d}} = \frac{a}{a+b}$; from there deduce $P(A/B) = P\frac{A\cap B}{P(B)}$.

To prove the formula $P(B/A) = \frac{P(A\cap B)}{P(A)}$, all students argued similarly to the previous formula.

For question 2e/, all 42 students correctly answered that $P(A/B) \neq P(B/A)$ and explained that the two different probabilities were due to the different order of occurrence of the events. In particular, 18 students added that the conditional probability of an event is not commutative.

After commenting on the groups' work, the teacher asked group 11 to come to the board to present their group's work to the class. In Figure 4, group 11 representatives are presenting their group's work, in which they have proven two formulas $P(A/B) = \frac{P(A\cap B)}{P(B)}$ and $P(B/A) = \frac{P(A\cap B)}{P(A)}$.



Group 11

	A	$\overline{A}$	Total
B	a	b	a + b
$\overline{B}$	c	d	c + d
Total	a + c	b + d	a + b + c + d

From the statistical table of events, we have

a/ The probability of event A is $P(A) = \frac{n(A)}{n(\Omega)} = \frac{a+c}{a+b+c+d}$.

The probability of event B is $P(B) = \frac{n(B)}{n(\Omega)} = \frac{a+b}{a+b+c+d}$.

b/ The probability of the intersection of two events A and B is

$$P(A \cap B) = \frac{n(A \cap B)}{n(\Omega)} = \frac{a}{a+b+c+d}.$$

c/ The probability of event A knowing that event B has occurred before is

$$P(A/B) = \frac{n(A \cap B)}{n(B)} = \frac{a}{a+b} \qquad (1).$$

Dividing both sides of the formula (1) by the number of elements of the sample space $a + b + c + d$, we get

$$P(A/B) = \frac{n(A \cap B)}{n(B)} = \frac{a}{a+b} = \frac{\frac{a}{a+b+c+d}}{\frac{a+b}{a+b+c+d}} = \frac{P(A \cap B)}{P(B)}.$$

d/ The probability of event B knowing that event A has occurred before is

$$P(B/A) = \frac{n(A \cap B)}{n(A)} = \frac{a}{a+c} \qquad (2).$$

Dividing both sides of the equation (2) by the number of elements in the sample space $a + b + c + d$, we get

$$P(B/A) = \frac{n(A \cap B)}{n(A)} = \frac{a}{a+c} = \frac{\frac{a}{a+b+c+d}}{\frac{a+c}{a+b+c+d}} = \frac{P(A \cap B)}{P(A)}.$$

e/ If $b \neq c$ then $P(A/B) = \frac{a}{a+b} \neq P(B/A) = \frac{a}{a+c}$, which means that the conditional probability is not commutative. The two probability values are different because in the first case event B occurs before event A, and in the second case event A occurs before event B.

Figure 4. Group 11 proves two formulas correctly $P(A/B) = P(A \cap B)/P(B)$ and $P(B/A) = P(A \cap B)/P(A)$

Thus, in step 3, students have built "specific models" of problem 2 (model of) which is also the "general model" of problem 1 (model for) as follows:

$$P(A/B) = \frac{a}{a+b}, \frac{P(A \cap B)}{P(B)} = \frac{\frac{a}{a+b+c+d}}{\frac{a+b}{a+b+c+d}} = \frac{a}{a+b}, P(A/B) = \frac{P(A \cap B)}{P(B)}, P(B/A) = \frac{a}{a+c},$$

$$\frac{P(A \cap B)}{P(A)} = \frac{\frac{a}{a+b+c+d}}{\frac{a+c}{a+b+c+d}} = \frac{a}{a+c}, P(B/A) = \frac{P(A \cap B)}{P(A)}.$$

These specific models will be developed into general models for the problems of calculating the conditional probability of an event in step 4.

Step 4. Forming conditional probability

From the "model for" problem 1 identified in step 3, the teacher asks students to state the definition and formula for calculating the conditional probability of an event.

Then, the teacher corrects the students' statements and confirms the definition and formula (formal models) for calculating the conditional probability of an event.

Step 5. Application

The teacher gave problem 3 and asked the students to solve it. The results showed that all 14 groups calculated the probability values correctly $P(W) = 0.62$, $P(W \cap F) = 0.41$, and $P(F/W) = \frac{P(W \cap F)}{P(W)} = 0.41/0.62 \approx 0.661.$

This shows that all students initially identified the order of occurrence of events, understood the concept of conditional probability, and were able to use the formula for calculating conditional probability.

Next, the teacher gave problem 4 and asked students to solve it.

For question 4a/, all 14 groups gave the correct answer that it was necessary to calculate the probability that a randomly selected student was female, given that the student was taller than 1.80 m, or to calculate $P(F/T)$.

For question 4b/, 11/14 groups (33 students) correctly calculated the probabilities $P(M)$, $P(F)$, $P(T/M)$, $P(T/F)$; the remaining 3 groups (9 students) incorrectly calculated $P(T/M)$, $P(T/F)$ as follows:

$$P(T/M) = \frac{P(T \cap M)}{P(M)} = \frac{0.04}{\frac{2}{5}} = 0.1; P(T/F) = \frac{P(T \cap F)}{P(F)} = \frac{0.01}{\frac{3}{5}} \approx 0.017.$$

These three groups assumed that the hypothesis "4% of men are over 1.80 m tall" means "4% of the students are male and taller than 1.80 m", and "1% of women are over 1.80 m tall" means "1% of students are female and taller than 1.80 m".

For question 4c/, 11/14 groups (33 students) filled in the probabilities on the branches at the question marks completely and correctly; the remaining 3 groups (9 students), which were the three groups that incorrectly calculated $P(T/M)$ and $P(T/F)$ when answering question 4b/, filled in the probabilities on the branches originating from M and F of the probability tree diagram incorrectly.

For question 4d/, all 14 groups correctly identified two paths from O to T, namely $O - M - T$ and $O - F - T$. They also proposed a method to calculate the probability that the selected student is taller than 1.80 m, $P(T)$, by multiplying the probability values on the branches of the same route, then adding the two values together. In other words, the 14 groups were able to apply the multiplication rule and the addition rule of probabilities. Thus, the groups were able to construct the first "specific model" of problem 4 (model of).

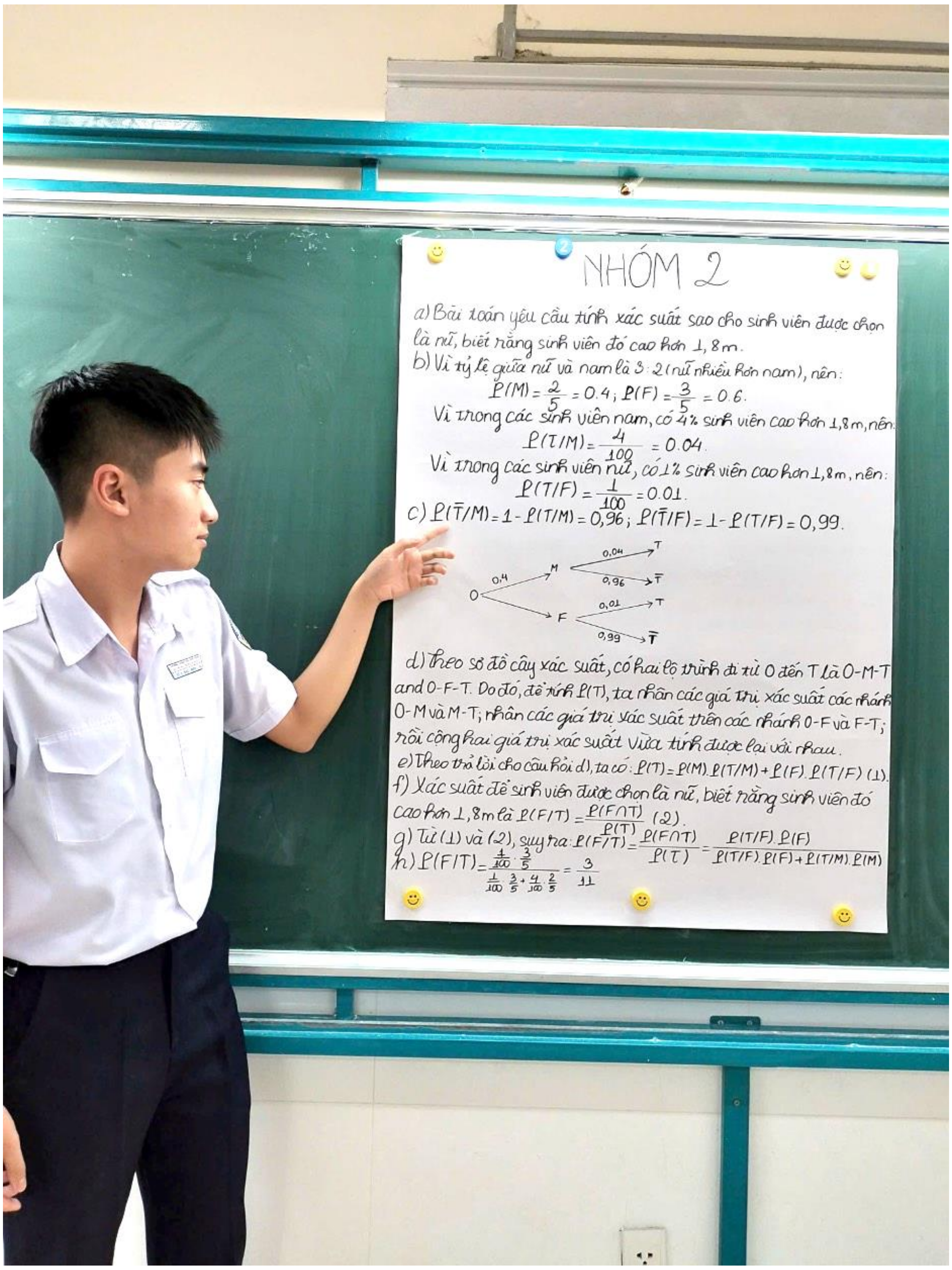
For question 4e/, all 14 groups interpreted from the calculation of $P(T)$ in question 4d/ to arrive at the formula $P(T) = P(T/M).P(M) + P(T/F).P(F)$. Thus, all groups were able to construct the first general model of problem 4 (model for).

For question 4f/, all 14 groups came up with the formula $P(T \cap F) = P(T/F).P(F)$, of which 9 groups arrived at the formula $P(F/T) = \frac{P(F \cap T)}{P(T/F).P(F)+P(T/M).P(M)}$. Thus, 9 groups were able to construct the second "specific model" of problem 4 (model of).

For question 4g/, all 14 groups proved the formula $P(F/T) = \frac{P(T/F).P(F)}{P(T/F).P(F)+P(T/M).P(M)}$ by using the formula for calculating the conditional probability of the event $P(F/T) = (P(F \cap T))/(P(T))$ and the result of question 4f/. This means that all groups were able to construct the second "general model" of problem 4 (model for).

For the calculation of probability $P(F/T)$ in question 4h/, 11/14 groups (33 students) calculated the value of $P(F/T)$ correctly; the remaining 3 groups calculated the probability value incorrectly. These 3 groups (9 students) were the 3 groups that calculated the values of $P(T/M)$ and $P(T/F)$ incorrectly when answering question 4b/, and they substituted those incorrect values into the formula $P(F/T) = \frac{P(T/F).P(F)}{P(T/F).P(F)+P(T/M).P(M)}$, thus leading to incorrect results.

After giving general comments on the groups' results, the teacher asked group 2, the group with the correct solution, to present their work to the class. The whole class agreed with group 2's solution. Next, the teacher asked the groups to edit their work to make it accurate and complete. In Figure 5, the representative of group 2 is explaining their group's work, in which they have correctly proven the total probability formula and Bayes' formula.



Group 2

a/ The problem requires calculating the probability that the selected student is female, knowing that the student is taller than 1.8 m.

b/ Since the ratio between female and male is 3:2 (females are more than males), then

$$P(M) = 2/5 = 0.4; P(F) = 3/5 = 0.6.$$

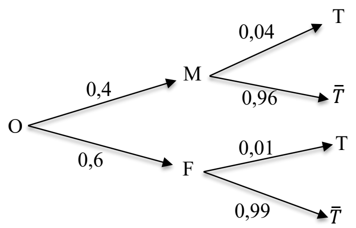
Since among the male students, 4% of the students are taller than 1.8 m, then

$$P(T/M) = 4/100 = 0.04.$$

Since among the female students, 1% of the students are taller than 1.8 m, then

$$P(T/F) = 1/100 = 0.01.$$

c/ $P(\overline{T}/M) = 1 - P(T/M) = 0{,}96$; $P(\overline{T}/F) = 1 - P(T/F) = 0{,}99$.



d/ According to the probability tree diagram, there are two paths from O to T, which are $O - M - T$ and $O - F - T$. Therefore, to calculate $P(T)$, we multiply the probability values on the branches $O - M$ and $M - T$; multiply the probability values on the branches $O - F$ and $F - T$; then add the two probability values just calculated together.

e/ From the answer to question d/, we have

$$P(T) = P(M).P(T/M) + P(F).P(T/F) \text{ (1).}$$

f/ The probability that the selected student is female, knowing that the student is taller than 1.8 m is

$$P(F/T) = P(F \cap T)/P(T) \text{ (2).}$$

g/ From (1) and (2), we deduce:

$$P(F/T) = \frac{P(F \cap T)}{P(T)} = \frac{P(T/F).P(F)}{P(T/F).P(F) + P(T/M).P(M)}.$$

h/ $P(F/T) = \dfrac{\frac{\frac{1}{100}.3}{5}}{\frac{\frac{1}{100}.3}{5} + \frac{\frac{4}{100}.2}{5}} = \dfrac{3}{11}.$

Figure 5. Group 2 correctly proves the total probability formula and Bayes' formula

Next, the teacher notes to the students that the formula $P(T) = P(M).P(T/M) + P(F).P(T/F)$ that the groups proved in question e/ is called the total probability formula, and the formula $P(F/T) = \frac{P(T/F).P(F)}{P(T/F).P(F)+P(T/M).P(M)}$ proved in question g/, is called Bayes' formula.

Then, the teacher declares the total probability theorem and Bayes' theorem as follows: "Given two events A and B with $0 < P(B) < 1$, then $P(A) = P(B).P(A/B) + P(\overline{B}).P(A/\overline{B})$ is called the total probability formula"; "Suppose A and B are two random events satisfying $P(A) > 0$ and

$0 < P(B) < 1$, then $P(B/A) = \frac{P(B).P(A/B)}{P(B).P(A/B)+P(\bar{B}).P(A/\bar{B})}$ is called Bayes' formula" (formal models). The Bayesian formulation process is consistent with Borovcnik (2012) view, Conditional probability is the core of probability. Conditional probabilities can be easily attributed to information from a given context, or they can be computed from the given data essentially using the Bayesian formulation. Therefore, conditional probability and Bayesian formulation are "twin brothers".

Discussion

The teaching experiment results show that the teaching process of conditional probability including 5 steps designed according to the principles of RME is feasible and effective, creating opportunities for students to start from the initial real-life problem to generalize specific models into general models, thereby developing into the concept and formula of conditional probability, and from the fourth practical problem to arrive at the total probability theorem and Bayes' theorem.

In the study, we do not use Frequency Net and Net Diagrams by Binder et al. (2020), because in Vietnam, the 2018 VMGEC for only mentioned two visual tools: 2x2 tables and tree diagrams. Gigerenzer (2002) mentioned the uncertainty of statistics in breast cancer diagnosis. However, in Vietnam, the examples in textbooks are idealized so as not to cause this uncertainty.

During the teaching experiments, students were assigned to solve four main problems: (1) an initial real-life problem that required analytical and comparative thinking to find a solution; (2) a generalized problem of the original problem that facilitated the rediscovery of the concept and formula of conditional probability; (3) a practical application problem in which students used the definition and formula of conditional probability to solve it; (4) a real-life problem that required the calculation of a number of conditional probabilities, the results of which, together with the use of probability trees, served as aids to the rediscovery of the total probability theorem and Bayes' theorem. The four problems were presented in four scenarios and were integrated into a five-step conditional probability teaching process under the teacher's focused guidance. The experimental teaching process was conducted in the context where students had learned about the probability of a random event in grade 10, the probability of the intersection of two events, and the probability of two independent events.

This teaching method does not rely on historical phenomenology due to the complexity and length of the development of conditional probability. Instead, a purely phenomenological framework is applied, starting with a contextual scenario: calculate the probability that a randomly selected patient is treated with drug M, given that the patient is cured; and calculate the probability that a randomly selected patient is cured, given that the patient is treated with drug N.

In step 1, the teacher presents the initial real-life problem and asks students to read, understand, and identify the requirements of the problem. A sample of 800 patients treated with drug M, or drug N, is either cured or not cured, and the data are presented in a 2 × 2 two-way statistical table.

In step 2, the teacher asks students to solve problem 1. In three questions 1a/, 1b/, 1c/, students are asked to calculate the probabilities of events M, C, and $M \cap C$. Calculating these probabilities is a familiar computational activity, since students have learned to calculate the probability of a random event, of the intersection of two events. However, in this case, students need to have the skills to read the numbers along the row of event C, along the column of event M, and read the numbers in the cell at the intersection of the row of C and the column of M. These activities represent a horizontal mathematization process, aiming to prepare students to build a "model of" the problem, which is how to calculate the two probabilities $P(M/C)$ and $P(C/N)$, required in two questions 1d/ and 1e/. The results show that there are two "models of" in students: $P(M/C) = \frac{n(M \cap C)}{n(C)}$ and $P(C/N) = n\frac{C \cap N}{n(N)}$ for 15 students who solved correctly; two "models of" $P(M/C) = \frac{P(M \cap C)}{n(\Omega)}$ and $P(C/N) = \frac{n(C \cap N)}{n(\Omega)}$ for 27 students who chose the wrong solution. At the end of step 1, the teacher explains how to correctly select the data in the statistical table to calculate the two probabilities $P(M/C)$ và $P(C/N)$, and asks the groups to correct their group's work. From there, the "models of" problem 1 are the two formulas $P(M/C) = \frac{n(M \cap C)}{n(C)}$ and $P(C/N) = \frac{n(C \cap N)}{n(N)}$. These are considered two new models that arise in the process of horizontal mathematicalization. The mistake of using the same sample space for group 6 shows a re-enactment of the confusion between independence and mutual exclusivity of two probabilities in the study of Carranza and Kuzniak (2008).

In step 3, the teacher presents problem 2, which is a general form of problem 1, and asks students to solve it. In questions 2a/ and 2b/, students are asked to calculate the probabilities of events M, C, and $M \cap C$, to prepare them to construct a "model for" problem 1. In questions 2c/ and 2d/, based on the "model of" problem 1, $P(M/C) = \frac{n(M \cap C)}{n(C)}$ and $P(C/N) = \frac{n(C \cap N)}{n(N)}$, students construct a "model of" problem 2, $P(A/B) = \frac{a}{a+b}$ and $P(B/A) = \frac{a}{a+c}$. Next, by algebraic transformations, students prove that $\frac{P(A \cap B)}{P(B)} = \frac{a}{a+b}$ và $\frac{P(A \cap B)}{P(A)} = \frac{a}{a+c}$, to arrive at the "model for" problem 1, which is $P(A/B) = \frac{P(A \cap B)}{P(B)}$ and $P(B/A) = \frac{P(A \cap B)}{P(A)}$. It is clear that the "model for" of problem 2 is also the "model for" of problem 1. In question 2e/, the student realizes that conditional probability is not commutative, by comparing the two denominators of the two fractions $\frac{P(A \cap B)}{P(B)}$ and $\frac{P(A \cap B)}{P(A)}$. The meaning of this comparison is to emphasize the order of occurrence of related events. Thus, the activities leading to the "model for" of problems 1 and 2 in this step are the manifestation of the vertical modeling process, in which mathematical rules have been used to develop the "model of" into the "model for" within mathematics.

In step 4, all 42 students built a formal knowledge of conditional probability, albeit not entirely mathematically precise, through the formulation of definitions and formulas. The activities in this step represent a vertical process of mathematization, developing the "model for" into a "formal model."

In step 5, the teacher gives problem 4 and asks students to solve it. Four questions 4a/, 4b/, 4c/, 4d/ require students to apply knowledge about the probability of an event and conditional probability to solve it. The results of these four questions aim to prepare the conditions for students to rediscover the total probability theorem through solving question 4e/. Similarly, questions 4e/ and 4f/ aim to prepare the conditions for students to rediscover Bayes' theorem through solving question 4g/. Finally, question 4h/ requires the application of Bayes' theorem to solve it. Thus, the activities in this step represent a vertical mathematization process, developing the "formal models" achieved in step 4, conditional probability definition and formula, into newly emerged "models for" which are special cases of the total probability theorem and Bayes' theorem. The formula $P(T) = P(T/M).P(M) + P(T/F).P(F)$ that students need to prove in question 4e/ is the first "model for", and the formula $P(F/T) = \frac{P(T/F).P(F)}{P(T/F).P(F)+P(T/M).P(M)}$ is the second "model for" problem 4. The two corresponding "formal models" problem 4, the total probability theorem and Bayes' theorem, are claimed by the teacher to be based on these two "models for". The process of deriving the total probability theorem and validating Bayes' theorem from a few individual cases in problem 4, although not mathematically sound, is pedagogically beneficial in this teaching context. Moreover, proving these formulas and theorems is not a goal to be aimed at in the current 2018 VMGEC. The probability tree diagram integrated into problem 4 serves as a means to support students' discovery of the total probability theorem, which in turn leads to the discovery of Bayes' theorem. Figure 6 shows the mathematization process with emerging models.

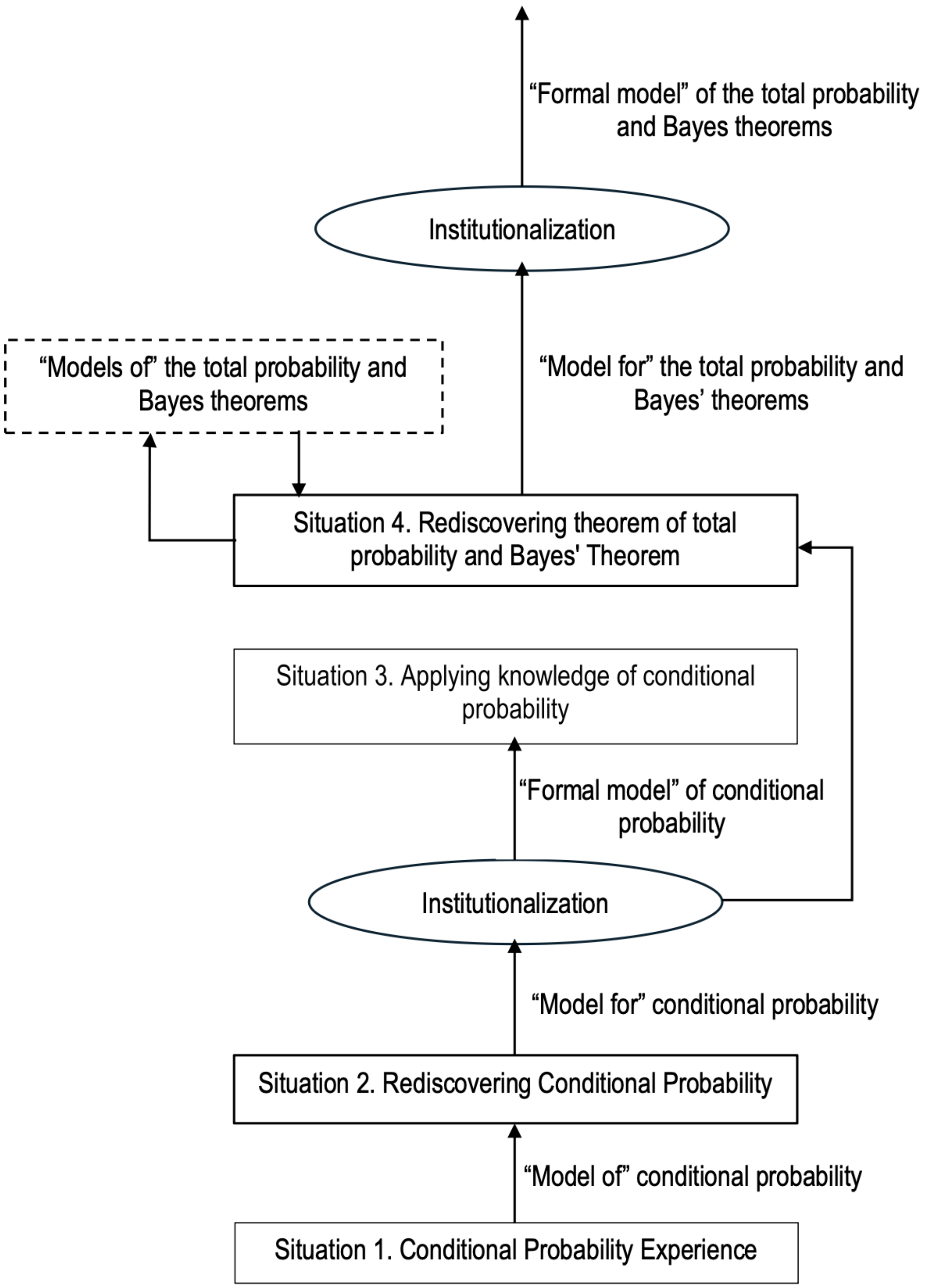


Figure 6. Mathematization process

On the other hand, in question 4b/, all students correctly calculated the two probabilities $P(M)$

and $P(F)$, but some students (9/42) incorrectly identified the two probabilities $P(T/M)$ and $P(T/F)$ when they did not correctly analyze the meaning of the two hypotheses “4% of men are over 1.80 m tall” and “1% of women are over 1.80 m tall”, and considered them as the intersection of the two events T and M, and of T and F. This mistake affected the answer to question 4c/ when they filled in the wrong values of the two probabilities $P(T/M)$ và $P(T/F)$ in the tree diagram, and thus also affected the answer to question 4h/. This phenomenon raises the question of how practical life experience affects conditional probability calculations? On the other hand, if we consider the way students fill in the probabilities on the branches of the probability tree, in which the order of events when following the paths from the root to the nodes must be respected, then the above two mistakes can be limited, and therefore question 4b/ is not needed in the problem. Then, students read each path on the tree diagram, determine the probabilities for each branch, and write the probability on that branch.

The research results show that there is a clear difference between teaching conditional probability according to RME in Vietnam and other studies in the world, such as the studies of Carranza and Kuzniak (2008), Binder et al. (2020), Gigerenzer (2002), and Borovcnik (2012). The new result of the article is to build a mathematicization process for teaching conditional probability for 12th grade students. The results of the teaching experiment show that teaching conditional probability according to RME helps students understand knowledge better and retain knowledge better.

CONCLUSION

This study examined the design and implementation of a teaching process for conditional probability grounded in the Realistic Mathematics Education (RME) framework, with a specific focus on the Grade 12 mathematics curriculum in Vietnam. The experimental findings demonstrate that students in the RME-based instructional setting actively engaged in horizontal and vertical mathematization processes. They were able to transform real-world contexts into mathematical representations and subsequently apply formal mathematical tools, including probability formulas and tree diagrams, to solve problems. The guided reinvention approach enabled students to develop a deeper conceptual understanding and to apply the conditional probability knowledge they had constructed to new, contextually meaningful situations. These results suggest that the RME-based teaching sequence was pedagogically effective, both in facilitating students' cognitive development and in promoting the transfer of knowledge to real-life contexts. The design’s adherence to core RME principles—including the use of contextual starting points, model development, and teacher scaffolding—played a critical role in achieving these outcomes.

Furthermore, a primary challenge observed during the teaching experiment was students’ difficulty in interpreting and navigating two-way tables when tasked with identifying relevant data for calculating conditional probabilities. Specifically, items 1d and 1e revealed that students struggled to conceptually grasp the conditioning event in the presence of tabular data. Additionally, in questions such as 4b, 4c, and 4h, errors were traced to misidentification of the event structure when tree diagrams were not initially provided. These findings raise the question of whether incorporating tree diagrams at earlier stages of instruction could better support students' conceptual clarity, particularly by emphasizing the sequential nature of conditional events. Future research should investigate the impact of explicit visual tools, such as probability trees, on students’ reasoning processes and accuracy when working with compound events.

Given the novelty of conditional probability within the Vietnamese high school curriculum and the limited number of empirical studies on its pedagogy, this research offers both timely and practical contributions. However, the experimental intervention was limited to a single class of 42 students from Can Thanh High School, and thus the generalizability of findings remains constrained. The high-stakes nature of Grade 12 education in Vietnam, particularly due to university entrance examinations, made it difficult to secure broader participation across multiple schools. Future studies should aim to replicate and extend this work across more diverse educational settings and student populations to enhance its external validity. Additionally, future research should explore the application of the RME framework to other advanced mathematical topics at the upper secondary level. Incorporating elements of historical epistemology into RME-based instructional design may further enrich students' understanding by situating mathematical concepts within their developmental and practical origins. Such interdisciplinary integration has the potential to deepen students' appreciation for mathematics as a human activity and to strengthen the link between formal mathematics and the realities of everyday life.

Declarations

Author Contribution	:	QAN: Conceptualization, Methodology, Writing - Original Draft, Editing and Visualization. NHD: Writing - Review & Editing, Formal analysis. TAH and NGN: Validation and Supervision.
Funding Statement	:	No funding.
Conflict of Interest	:	We declare no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.
Additional Information	:	Additional information is available for this paper.

REFERENCES

Batanero, C., & Álvarez-Arroyo, R. (2024). Teaching and learning of probability. *ZDM – Mathematics Education, 56*(1), 5–17. https://doi.org/10.1007/s11858-023-01511-5

Batanero, C., Chernoff, E. J., Engel, J., Lee, H. S., & Sánchez, E. (2016). Research on Teaching and Learning Probability (pp. 1–33). https://doi.org/10.1007/978-3-319-31625-3_1

Binder, K., Krauss, S., & Wiesner, P. (2020). A New Visualization for Probabilistic Situations Containing Two Binary Events: The Frequency Net. *Frontiers in Psychology, 11*. https://doi.org/10.3389/fpsyg.2020.00750

Borovcnik, M. (2011). Strengthening the Role of Probability Within Statistics Curricula (pp. 71–83). https://doi.org/10.1007/978-94-007-1131-0_11

Borovcnik, M. (2012). Conditional Probability-a Review of Mathematical, Philosophical, and Educational Perspectives. http://www.icme12.org/sub/tsg/tsgload.asp?tsgNo=11

Bui, A. K., & Tran, H. P. (2024). Using Social constructivist theory in the teaching of solving problems of conditional probability grade 12. *CTU Journal of Science, 60*(4), 187–196. https://doi.org/10.22144/ctujos.2024.390

Carranza, P., & Kuzniak, A. (2008, December 30). Duality of probability and statistics teaching in French education. Joint ICMI/IASE Study: Teaching Statistics in School Mathematics IASE Roundtable Conference. https://doi.org/10.52041/SRAP.08206

Da, N. T. (2022). Designing a teaching model based on the Realistic Mathematics Education (RME) approach and its application in teaching calculus. *Journal of Mathematics and Science Teacher, 2*(1), em006. https://doi.org/10.29333/mathsciteacher/11918

De Lange, J. (2006). Mathematical literacy for living from OECD-PISA perspective. Tsukuba Journal of Educational Study in Mathematics. Special Issue on the APEC- TSUKUBA International Conference "Innovative Teaching Mathematics through Lesson Study," 25, 13–35. http://www.criced.tsukuba.ac.jp/math/apec2006/

Díaz, B., Batanero, C., & Contreras, M. (2010). Teaching Independence and Conditional Probability. *BEIO Revista Oficial de La Sociedad de Estadística e Investigación Operativa, 26*(2), 149–162. www.seio.es/BEIO.

Díaz, C., & Batanero, C. (2009). University students' knowledge and biases in conditional probability reasoning. *International Electronic Journal of Mathematics Education, 4*(3), 131–162. https://doi.org/10.29333/iejme/234

Freudenthal, H. (1968). Why to teach mathematics so as to be useful. *Educational Studies in Mathematics, 1*(1–2), 3–8. https://doi.org/10.1007/BF00426224

Freudenthal, H. (1971). Geometry between the devil and the deep sea. *Educational Studies in Mathematics, 3*(3–4), 413–435. https://doi.org/10.1007/BF00302305

Freudenthal, H. (1983). Didactical Phenomenology of Mathematical Structures. Reidel Publishing Company.

Freudenthal, H. (2002). Revisiting Mathematics Education: China lectures (Vol. 9). Kluwer Academic Publishers. https://doi.org/10.1007/0-306-47202-3

Gigerenzer, G. (2002). Calculated risks: how to know when numbers deceive you. In Choice Reviews Online (illustrated, Issue 04). Simon & Schuster.

Gravemeijer. (1994). Developing realistic mathematics education. CD β Press. https://www.fisme.science.uu.nl/publicaties/literatuur/1994_gravemeijer_dissertation_0_222.pdf

Gravemeijer. (1999). How Emergent Models May Foster the Constitution of Formal Mathematics. *Mathematical Thinking and Learning, 1*(2), 155–177. https://doi.org/10.1207/s15327833mtl0102_4

Gravemeijer, K., & Terwel, J. (2000). Hans Freudenthal: A mathematician on didactics and curriculum theory. *Journal of Curriculum Studies, 32*(6), 777–796. https://doi.org/10.1080/00220270050167170

Huerta, M. P. (2009). On Conditional Probability Problem Solving Research – Structures and Contexts. *International Electronic Journal of Mathematics Education, 4*(3), 163–194. https://doi.org/10.29333/iejme/235

Inci, A. M., Peker, B., & Kucukgencay, N. (2023). Realistic mathematics education. In book: Current Studies in Educational Disciplines 2023 (pp.66-83) Edition 1, Chapter 5. Publisher: ISRES Publishing. ISRES Publishing. https://www.researchgate.net/publication/377209600_Realistic_Mathematics_Education

Larsen, S. (2018). Didactical phenomenology: the engine that drives realistic mathematics education. *For the Learning of Mathematics, 38*(3), 25–29. https://www.jstor.org/stable/26548508

Le, T. T., Pham, A. G., & Nguyen, T. T. (2021). Applying realistic mathematics education theory in teaching: Some challenges, principles and recommendations. *Vietnam Journal of Education, 494*(2), 37–43. https://tcgd.tapchigiaoduc.edu.vn/index.php/tapchi/article/view/22

Loc, N. P., Tran, N., & Tien, T. (2020). Approach to realistic mathematics education in teaching mathematics: A case of cosine theorem-geometry 10. *International Journal of Scientific & Technology Research, 8*, 1173–1178.

Nguyen, A. Q. (2022). Teaching conditional probability in grade 12 according to the general education curriculum in mathematics in 2018. *Vietnam Journal of Educational Sciences, 11*(3), 40–46. https://doi.org/10.15625/2615-8957/12211107

Nguyen, D. N. (2020). Some problems in realistic mathematics education. *Vietnam Journal of Education, 487*(1), 15–21. https://mysite.tnu.edu.vn/files/users/nguyendanhnam/487_Nguyn_Danh_Nam.pdf

Nguyen, T. T. T., & Quach, T. Sen. (2023). Some methods of teaching statistic and probability topics for 6th grade students with the support of information technology. *Vietnam Journal of Educational Sciences, 19*(7), 27–34. https://doi.org/10.15625/2615-8957/12310705

Nguyen, T. T., Trinh, T. P. T., Ngo, H. T. V., Hoang, N. A., Tran, T., Pham, H. H., & Bui, V. N. (2020). Realistic mathematics education in Vietnam: Recent policies and practices. *International Journal of Education and Practice, 8*(1), 57–71. https://doi.org/10.18488/journal.61.2020.81.57.71

Shao, X. (2015). An analysis of difficulties in learning Probability in high school. 7th ICMI-East Asia Regional Conference on Mathematics Education, 561–568. https://mathted.weebly.com/uploads/7/8/5/0/7850000/pp_561_shao_xiayan__final_paper_edited.pdf

Streefland, L. (1985). Wiskunde als activiteit en de realiteit als bron (Mathematics as an activity and reality as a source). *Nieuwe Wiskrant, 5*(1), 60–67. http://www.fisme.science.uu.nl/wiskrant/artikelen/artikelen00-10/051/0501september_streefland.pdf

Tomlinson, S., & Quinn, R. (1997). Understanding conditional probability. *Teaching Statistics, 19*(1), 2–7. https://doi.org/10.1111/j.1467-9639.1997.tb00309.x

Tran, T. T. H., & Tran, C. (2021). Potential of applying realistic mathematics education to the teaching of probability - statistics at the universities of medicine and pharmacy. *Journal of Science Educational Science, 66*(5), 209–216. https://doi.org/10.18173/2354-1075.2021-0249

Treffers, A. (1987). Three dimensions. A model of goal and theory description in mathematics instruction: The Wiskobas Project. Springer Dordrecht.

Treffers, A. (1991). Didactical background of a mathematics program for primary education. In: L. Streefland (ed.). Realistic mathematics education in primary school. In Realistic mathematics education in primary school (pp. 21–56). CDß press/Freudenthal Institute, Utrecht University, Utrecht, The Netherlands. https://www.fisme.science.uu.nl/publicaties/literatuur/1991_streefland_0-209.pdf

Van Den Heuvel-Panhuizen, M. (2000). Mathematics education in the Netherlands: A guided tour 1. Freudenthal Institute CD-rom for International Commission on Mathematics Education (ICME) 9.

https://p4mriunismuh.wordpress.com/wp-content/uploads/2010/08/mathematics-education-in-the-netherlands.pdf

Van Den Heuvel-Panhuizen, M. (2003). The didactical use of models in realistic mathematics education: An example from a longitudinal trajectory on percentage. *Educational Studies in Mathematics, 54*(1), 9–35. https://doi.org/10.1023/B:EDUC.0000005212.03219.dc

Van Den Heuvel-Panhuizen, M. (2005). The role of contexts in assessment problems in mathematics. *For the Learning of Mathematics, 25*(2), 2–23.

Van den Heuvel-Panhuizen, M., & Drijvers, P. (2014). Realistic Mathematics Education. In *Encyclopedia of Mathematics Education* (pp. 521–525). Springer Netherlands. https://doi.org/10.1007/978-94-007-4978-8_170

Van den Heuvel-Panhuizen, M., & Drijvers, P. (2020). Realistic Mathematics Education. In *Encyclopedia of Mathematics Education* (pp. 713–717). Springer International Publishing. https://doi.org/10.1007/978-3-030-15789-0_170

Vicent, M. L., Palau, M. P. H., & Fariña, M. C. (2012). Conditional probability problems in textbooks an example from Spain. *Revista latinoamericana de investigación en matemática educativa, 15*(3), 319-337. https://www.scielo.org.mx/pdf/relime/v15n3/v15n3a4.pdf

Vietnam Ministry of Education and Training. (2018). Mathematics General Education Curriculum of Vietnam (32/2018/TT-BGDĐT).

Yackel, E., Stephan, M., Rasmussen, C., & Underwood, D. (2003). Didactising: Continuing the work of Leen Streefland. *Educational Studies in Mathematics, 54*(1), 101–126. https://doi.org/10.1023/B:EDUC.0000005213.85018.34

Zulkardi. (2010). *How to Design Mathematics Lessons based on the Realistic Approach?* http://repository.unsri.ac.id/id/eprint/6362