



Probabilistic literacy and reasoning of prospective secondary school teachers when interpreting media news

Rocío Álvarez-Arroyo¹ · Carmen Batanero¹ · María M. Gea¹

Accepted: 30 April 2024
© The Author(s) 2024

Abstract

Educating students to apply probability literacy and reasoning in out-of-school contexts requires the previous preparation of teachers, which should include the assessment of their mathematical and didactic knowledge. Consequently, we investigated three research questions concerning the probabilistic reasoning and literacy of 66 prospective secondary school teachers when solving probability questions related to a report taken from the media news, their ability to identify the fundamental stochastic ideas needed to solve the task, and their capacity to predict their students' potential difficulties. The participants needed to compute the probability of a complementary event, several conditional probabilities, and perform critical reading of the information in the report. Despite the participants' good mathematical knowledge, a high percentage showed reasoning biases, such as the fallacy of the transposed conditional, confusion between conditioning and causation, and few identified the information missing to solve a question. There was a poor capacity to determine the stochastic ideas needed to solve the questions, beyond probability, sample space, conditional probability, and sampling. The participants quoted a few potential errors of their students, mainly conceptual mistakes with scarce recognition of interpretation errors. These results reveal the need to reinforce teachers' probabilistic literacy, reasoning, and related didactic knowledge.

Keywords Teacher education · Didactic-mathematical knowledge · Probabilistic reasoning and literacy · Secondary education · Media news

1 Introduction

There is increasing interest in teaching probability at different educational levels, given that recent research suggests that even young children develop ideas of probability with adequate instruction (Batanero & Álvarez-Arroyo, 2023). Probability is an area with the same relevance as calculus and algebra; however, it is the only branch that deals with uncertainty, leading to a new type of reasoning (Sharma, 2015). It applies many mathematical methods and has connections to proportionality, combinatorics, functions, and mathematical logic (Van Dooren, 2014). Another peculiarity

of probability is that some conclusions are not validated deductively, as they are sometimes based on the analysis of statistical data. The existence of propositions (such as the law of large numbers), given with a probability, is also specific (Batanero & Borovcnik, 2016).

Probabilistic knowledge is essential for students, because of the existence of randomness in many daily situations, such as reacting to media reports or making decisions when facing risk (Gigerenzer, 2015). In addition, a basic understanding of probability is also needed to apply statistical inference, which is useful in varied fields, and simulation makes its study now possible for all educational levels (Bargagliotti et al., 2020).

To respond to this need, probability is introduced in primary education in Spain and continues being studied in secondary and high schools with conditional probability, Bayes' theorem, and probability distributions (Ministerio de Educación y Formación Profesional, 2022). However, teaching is based almost exclusively on learning definitions and properties and solving textbook problems (Muñiz-Rodríguez et al., 2016). It is unusual to confront students with real problems

✉ Carmen Batanero
batanero@ugr.es

Rocío Álvarez-Arroyo
rocioaarroyo@ugr.es

María M. Gea
mmgea@ugr.es

¹ University of Granada, Granada, Spain

in which they must apply their probabilistic literacy and reasoning (Álvarez-Arroyo et al., 2022a).

However, developing students' adequate probabilistic literacy and reasoning requires previous education of teachers in this component of mathematical knowledge and related didactic knowledge. Although research on teachers' knowledge to teach probability is increasing, we have not found investigations that assess teachers' probabilistic literacy and reasoning in everyday situations or their related didactic–mathematical knowledge.

The main objective of this paper is to present the design and administration of a questionnaire to a group of prospective secondary school mathematics teachers (PTs), asking them to interpret a report on traffic accidents taken from the media. Specifically, three research questions guided the research: 1) Do prospective teachers use adequate probabilistic reasoning and literacy to solve probabilistic tasks related to the news report? 2) Are they capable of analysing the probabilistic content of the tasks? and 3) Can they predict the students' difficulties in solving the same task?

Below we present the previous research, theoretical framework, and results for these questions.

2 Previous research

We summarise research dealing with probabilistic reasoning and literacy, and with teachers' knowledge to teach probability directly related to this paper.

2.1 Probabilistic reasoning and literacy

Most previous research on probabilistic reasoning and literacy (see Section 3.2) has been linked to decision-making under uncertainty and was developed under the programme heuristics (intuitive strategies in problem solving) and biases (Kahneman et al., 1982), summarised in Gilovich et al. (2002). This research described different heuristics, such as *representativeness* (expecting that a small sample will have the main characteristics of the population) or *availability* (judging more probable events that are easy to remember). In the *fallacy of composition* (Chernoff et al., 2018), individuals incorrectly generalise to a whole (e.g. compound events in flipping several coins) a property of a part (equiprobability of events when flipping only a coin).

Other reasoning biases appear in conditional probabilities (Díaz & de la Fuente, 2007; Falk, 1986), which are involved in the task posed in this study. One of them is the *fallacy of the transposed conditional*, which consists of confusing a conditional probability $P(A|B)$ with its transposed $P(B|A)$ without understanding the asymmetry of conditional probabilities (Borovcnik, 2016). Another bias is *confusing conditioning and causation*. Díaz & de la Fuente (2007) investigated these biases

in a wide sample of university students who correctly solved probability problems. The pervasiveness of these biases has produced recent research that analysed different visualisations that might help students overcome these biases (described in Batanero & Álvarez-Arroyo, 2023).

The probabilistic reasoning of high school students in sampling tasks is supported by their understanding of variability, independence, and randomness (Sánchez & Valdez, 2017). Moreover, their reasoning about distributions is based on the ideas of sample space, combinatorial, classical and frequentist approaches to probability, random variables, distribution, and variability (Sánchez et al., 2018).

Álvarez-Arroyo et al. (2022a, b) analysed high school students' probabilistic reasoning when interpreting news reports. They reported that many students presented the transposed conditional bias and confused conditioning and causation when interpreting conditional probabilities (Álvarez-Arroyo et al., 2022b). Other students failed to compute compound probabilities and interpret small probabilities (Álvarez-Arroyo et al., 2022a).

2.2 Teachers' knowledge to teach probability

Previous research on teachers' probability knowledge has mainly focussed on probabilistic knowledge of specific concepts. For example, Chernoff et al. (2018) described the fallacy of composition in Canadian secondary school PTs' understanding of randomness. Gómez et al. (2016) and Vásquez and Alsina (2015) studied Spanish primary school PTs and Chilean in-service teachers' mathematical knowledge of the different meanings of probability, as did Ingram (2022) with secondary mathematics and science PTs. The general conclusion was the lack of capacity to relate the different meanings of probability. In Hourigan and Leavy (2020), Irish primary school PTs demonstrated an understanding of probability fairness, but some of them presented representativeness bias. Díaz et al. (2012) analysed the understanding of conditional probability of secondary school PTs and found a good capacity to compute probability, but a high incidence of transposed conditional bias, confusion between conditional and compound probability, and incorrect conceptions of independence. Similar biases were reported by Brückler and Milin Šipuš (2023) in PTs from Croatia, adapting items from Díaz and de la Fuente (2007) to the context of the COVID-19 pandemic.

We can classify research on teachers' didactic knowledge into different components (see Section 3). Concerning the knowledge of students and specialised knowledge, Gómez et al. (2013) analysed primary school PTs' knowledge to teach the frequentist meaning of probability. Although participants recognised the correct and incorrect answers of students and could explain the reasons for their errors, they identified few mathematical objects in the task when asked to do so.

Martin et al. (2022) investigated the attitudes and beliefs of in-service teachers and found agreement in the social usefulness and interest of probability in education and in the improvement of teachers' attitudes after teaching probability. Regarding knowledge of curriculum, teachers in Martin et al. (2022) agreed on the relevance of the theoretical approach to probability, the main probabilistic concepts to teach, and the use of traditional tasks.

Concerning the knowledge of teaching resources, in-service primary school teachers successfully adapted games of chance to be used in teaching (Malaspina & Malaspina, 2020) and constructed fair and unfair games for specific random generators (Hourigan & Leavy, 2020). However, some primary school PTs in Alonso-Castaño et al. (2021) posed problems that were inadequate for their students. In-service secondary school teachers used various tasks, manipulative resources, and technology to exploit the frequentist approach to probability, and there was agreement on the value of technological tools in teaching (Martin et al., 2022). Moreover, simulation activities educated teachers on mathematical and technological knowledge (Biehler et al., 2018).

In summary, research related to the current study has focussed on heuristics and biases in decision making under uncertainty, students' probabilistic reasoning in different situations, and isolated components of teachers' knowledge. Such research has not considered the probabilistic literacy and reasoning of teachers when interpreting media news or their related didactic knowledge.

3 Theoretical framework

We base our study on the ideas of probabilistic literacy and reasoning and the Teacher Didactic Mathematical Knowledge model described below.

3.1 Probabilistic literacy and reasoning

Gal (2005) defined *probability literacy* as “the knowledge and dispositions that students may need to develop to be considered literate regarding real-world probabilistic matters” (p. 40). According to the author, it implies appropriate use of probability language and competence to compute probabilities in varied situation and helps people understand probability statements in situations such as interpreting news reports. A probabilistic literate person appreciates the role of probability in different contexts and can pose critical questions when encountering information about random situations.

Probability literacy also involves a basic knowledge of main probability concepts, which includes the fundamental stochastic ideas (Heitele, 1975): 1) *Probability* in its different interpretations (classical, frequentist, subjective); 2) *Sample space* (set of observable events in the random

experiment); 3) *Equidistribution* (observing that in some experiments all the possible outcomes have the same probability); 4) *Addition rule*, to compute the probability of a compound event by the sum of probability of simple disjoint events that compose it; 5) *Independence and product rule*, that help solving probabilities in compound experiments; 6) *Conditional probability*, that formalises the change in our degree of belief in a random event with new information; 7) *Combinatorics* or techniques to compute the number of elements in the sample space; 8) *Random variable*, that introduces the study of probability distributions, such as the normal model; 9) *Law of large numbers*, which connects statistics and probability, and is the base for the frequentist meaning of probability; and 9) *Sampling*, which serves to generalize information from samples to populations.

Probability literacy also implies a critical stance towards probabilistic information; understanding and controlling one's own feelings regarding uncertainty (Lilleholt, 2019); overcoming mistaken beliefs, and positive attitudes towards probability.

In this study, we also assess *probabilistic reasoning*, which arises when solving probability problems and using arguments to prove a probabilistic assertion (Sánchez & Valdez, 2017). Following Ben-Zvi and Garfield's (2004, p. 7) definition of statistical reasoning, we can see probabilistic reasoning as how people reason with probability ideas and make sense of probability information. It requires judgments to establish the credibility of evidence (whether the information is sufficient and relevant to the problem) and its inferential strength (whether such information can be generalised) (Schum, 2001).

According to Borovcnik (2016), probabilistic reasoning includes competence in choosing the probabilistic model to apply in each situation, ability to discriminate between conditioning and causation, and correctly interpreting probabilistic statements. Other components of probabilistic reasoning (Batanero & Borovcnik, 2016) are:

1. *Becoming aware of the influence of prior probabilities on probability judgements.* Realising that many probabilities depend on other (prior) probabilities and must be considered.
2. *Recognising the asymmetry of conditional probabilities.* While independence is a symmetric relationship, conditional probability is not.
3. *Correctly interpreting small probabilities.* When an event has a negligible probability, we usually assume that it will not occur. However, when the experiment is repeated many times, the event tends to occur (Burns et al., 2010).

3.2 Teacher didactic-mathematical knowledge

Extensive research has focused on the education of mathematics teachers (e.g., Carrillo et al., 2018; Llinares, 2023), and frameworks have been suggested to describe the components of teachers' knowledge.

We used the Didactic-Mathematical Knowledge (DMK) model (Godino, 2009; Pino-Fan & Godino, 2015; Pino-Fan et al., 2015), which is divided into three dimensions: mathematical, didactic, and metadidactic. The mathematical dimension refers to mathematical knowledge “per se”, the didactic facet includes the knowledge needed to teach a given topic, and the metadidactic dimension allows the teacher to identify and manage the norms and contextual restrictions that condition the teaching of mathematics (Breda et al., 2017).

First, we focus on PTs’ mathematical knowledge related to probabilistic literacy and reasoning. According to the DMK model, teachers need *Common Content Knowledge* (CCK), which is shared by teachers and students and allows teachers to solve the problem-situations of the educational level at which they teach. To solve the tasks posed to them, the PTs in the sample should use their probabilistic literacy and reasoning skills, which are part of their probabilistic CCK.

Teachers also need *Advanced Content Knowledge* (ACK) of the mathematical concept being studied at a more advanced level; this knowledge provides the teacher with the mathematical foundations necessary to link the mathematical object to other mathematical or extra mathematical ideas, and to the same topic in more advanced courses. In the paper, we centre on the teachers’ ability to solve problems related to the probability of a complementary event and conditional probability, as well as the critical reading of probability information. Such knowledge appears in the Spanish secondary school curriculum guidelines and should be taught to high school students and therefore is part of teachers’ CCK.

The DMK model considers the following facets of didactic knowledge: *epistemic* (specialised mathematical knowledge or mathematical knowledge specific to teaching); *ecological* (knowing the relation of the topic with other themes in the curriculum and with society); *cognitive* (understanding the students’ learning, difficulties and reasoning); *affective* (managing students and own attitudes, beliefs and emotions); *meditational* (familiarity with teaching resources and technology) and *interactional* (managing the classroom discourse).

This study focuses on the epistemic and cognitive components. The *epistemic facet* includes specific mathematical knowledge that allows teachers to use different mathematical representations and procedures, various arguments and justifications, and identify the knowledge at play when solving a mathematical task (Pino-Fan et al., 2015). To evaluate part of the PTs’ knowledge in this facet, we asked them to identify the fundamental stochastic ideas in a task they had previously solved.

The *cognitive facet* consists of knowledge about the students’ reasoning and learning and helps teachers predict their possible solutions to mathematical tasks. It also serves to foresee students’ difficulties and misconceptions while solving mathematical problems. To evaluate part of this knowledge, we asked the PTs to predict their students’ difficulties when solving a task that they had previously solved.

4 Method

4.1 Sample and setting

In Spain, any person who wants a mathematics teacher position in a public or private secondary (students aged from 12 to 16 years) or high school (16 to 18 years) must complete a university degree (4 years of university studies). Afterward, they should study for a specific master’s programme oriented to provide didactic competencies, curricular knowledge, and teaching practise to graduates and also compulsory to obtain a position in private secondary and high schools.

The sample consisted of 66 PTs enrolled in the master’s programme, which included all participants in the academic year 2021–2022 at the University of Granada. Half of the participants had completed a university degree in mathematics, and the remainder had undertaken other scientific degrees (e.g., statistics, physics, architecture, or engineering). All of them had taken one or several university courses in advanced statistics, including probability and inference. Their probabilistic knowledge was homogeneous; therefore, we did not separate the results by the participants’ career. These participants had been taught fundamental statistical ideas in a previous lesson and were given a summary of Heitele (1975) for consultation. Furthermore, they had previously worked on identifying the mathematical content and students’ errors in other tasks.

4.2 Questionnaire

The PTs individually completed a questionnaire as part of a course unit on statistics education, which included the identification of students’ possible difficulties and the design of additional probabilistic tasks using other news reports. In this paper, we analyse the questions displayed in Fig. 1, which are based on a news report about traffic accidents published in Spanish media. The first six questions were adapted from Álvarez-Arroyo et al. (2022b), who selected the topic to foster students’ interest in statistics and make them aware of the possible effects of alcohol and drug consumption.

In this study, we proposed the same tasks on the participants to evaluate their probabilistic CCK and investigate our first research question. To evaluate the epistemic facet of their DMK and answer the second research question, we added question 7, in which participants were asked to analyse the mathematical content of the task. Finally, to evaluate the cognitive facet of their DMK, we added question 8 in which they should identify students’ potential difficulties in the task (research question 3). Once the written reports with the participants’ solutions to the different questions were collected, we performed a content analysis (Neuendorf, 2016). The coding was reviewed by the authors until a consensus was reached in case of disagreement.

Fig. 1 Questionnaire given to participants

On the TGDⁱ web page the following figures appear regarding the accident rate on Spanish roads:

Toxicological analyses carried out on 751 people died in 2018 in road accidents (535 drivers, 143 pedestrians and 73 accompanying persons) showed that the consumption of alcohol and other drugs still has a negative influence on road safety.

- More than 43% (232) of drivers killed tested positive for alcohol or drugs and this figure has not changed significantly from previous years.
- Another article published in a newspaper reported that 23% of those killed in cars in 2018 were not wearing a seatbelt.

Analyse the situation according to the information provided and answer the following questions:

1. What percentage of the drivers who died tested negative in alcohol and drugs?
2. What is the probability that a deceased person was wearing a seatbelt?
3. Can we know the percentage of drivers who died due to alcohol or drug consumption?

In view of your previous answers, please reply:

4. Is it more likely to die if you have consumed alcohol or drugs, or if you have not consumed any of these substances?
5. Which is more likely: to be killed if you were wearing a seatbelt or if you were driving without a seatbelt?
6. Do you think the way the information is provided is correct, and why? If not, what information would be missing to give a proper picture of the situation?
7. Identify the fundamental stochastic ideas needed to reply the above questions.
8. Describe the students' potential difficulties in solving the above questions.

ⁱTraffic General Direction, <https://revista.dgt.es/es/noticias/nacional/2019/07JULIO/0718-Informe-alcohol-drogas.shtml>

4.3 Categories of responses to the mathematical questions

To answer the first six questions, which evaluate their CCK (first research question), the PTs should read the report and understand the questions posed. Let us consider the events: D = “the person died”, TP = “testing positive for alcohol or drugs”, TN = “testing negative”, and S = “wearing a seatbelt”. Then, the problem data are: $P(TP|D) = 0.43$; $P(S|D) = 0.23$, where S (not wearing the seatbelt) is complementary to S . Participants should identify these data to answer questions 1 to 5. In Question 6 they should provide a reasoning that requires critical reading of the report.

We describe below the categories of responses in questions 1–6, which develop those used by Álvarez-Arroyo et al. (2022b). Table 6 with examples of responses to different categories is presented in the Appendix (participants have been coded as Px).

4.3.1 Probability of the complementary event

In Question 1, participants should compute the conditional probability $P(TN|D) = 0.57$, which is complementary to $P(TP|D)$. In question 2, we demand the probability $P(S|D) = 0.77$, complementary to $P(S|D)$. In questions 1 and 2, we classified the responses as follows: R1): Correct answers; R2): Confusing the probability requested and its complement; R3): Indicating that there was not enough data in the report to compute the probability; and R4): Other arithmetical errors or confusing the data.

4.3.2 Conditioning and causation

Questions 3 to 5 test the ability to distinguish a conditional probability and its transposed and to differentiate between conditioning and causation, two components of probabilistic reasoning included in the participants' CCK. The requested probabilities are $P(D|TP)$ in Question 3, $P(D|TP)$ and $P(D|TN)$ in Question 4, and $P(D|S)$ and $(D|S)$ in Question 5. These probabilities should be obtained using Bayes' theorem; however, this is not possible because the report omits some data, specifically, the probabilities of consuming alcohol or drugs, wearing a seatbelt, and dying in both cases. Thus, PTs should notice the missing data and justify that a conditional probability does not imply cause and effect.

We classified the responses in the following way: R1) Correct answer, where the PT distinguishes conditional probability and its transposed, as well as conditioning and causation; R2) Offering missing information but not clearly pointing to the difference between conditional probability and cause-effect; R3) Applying contextual knowledge without indicating which data were missing to solve the problem, not clearly demonstrating that the PT distinguished conditioning and causation; R4) Distinguishing a conditional probability and its transposed, with no reference to conditioning and causation; R5) Distinguishing conditioning and causation but confusing a conditional probability and its transposed; and R6) Confusing conditional probability with its transposed and mixing up causation and conditioning.

4.3.3 Identifying missing data

The correct response to Question 6 implies identifying the information missing in the report to answer the different questions. It involves posing critical questions, which is part of statistical literacy (Gal, 2005), and interpreting probabilistic statements, a component of probabilistic reasoning (Borovcnik, 2016). Both components are part of the teachers' CCK. We coded the answers as follows: R1) Correct responses, providing the missing data; R2) Correct responses not providing the missing data; R3) Offering irrelevant information; R4) Suggesting that the information is incorrect because we deal with a sample; R5) Pointing to the need of public awareness on the issue, without reference to the problem data; and R6) Not questioning the information or providing confusing information.

4.4 Categories of responses to didactical questions

4.4.1 Knowledge in the epistemic facet

In Question 7, related to the epistemic facet (research question 2), the PTs should analyse the mathematical content of questions 1–6 they had previously solved and identify the fundamental ideas (Heitele, 1975) involved in its solution. An a priori analysis of the task revealed the following fundamental ideas, which were used as categories for this question:

1. *Probability*. All the questions proposed involve computation or comparison of some probability. There is no equiprobability in the different events intervening (for example, wearing or not a seatbelt). Consequently, it is not possible to apply the Laplace rule, and the probabilities should be estimated from the frequency information about the different events.
2. *Sample space*. Participants should identify the different events in the random experiments intervening in each question: consumption or not of alcohol or drug; wearing or not a seatbelt; and dying or not.
3. *Equidistribution*, for each elementary event intervening (for example, each of the drivers who tested positive in drug or alcohol consumption). However, the compound events in the situation (e.g., testing positive, testing negative) are not equiprobable.
4. *The addition rule* does not appear explicitly but intervenes in questions 1 and 2 through the probability of the complementary event. Similarly, although no question is directly related to the *product rule*, we indirectly identify this rule when computing each conditional probability because we deal with compound experiments.
5. *Dependence and independence of events* should be analysed because the likelihood of dying is different depending on wearing a seatbelt and drug consumption.
6. *Conditional probabilities* are requested in the task, where the PTs should identify the event and the condition and not confuse a conditional probability with its transposed or causation.
7. *The law of large numbers* appears indirectly because we estimate the probability using the frequentist approach.
8. *Sampling*, because although the report refers to all traffic accidents, it can be seen as a sample of accidents in other years.

4.4.2 Knowledge in the cognitive facet

To assess the cognitive facet of the PTs' DMK, in Question 8, we asked the participants to identify the foreseeable errors of high school students when solving questions 1–6. We built the categories of analysis by considering the difficulties that high school students revealed in Álvarez-Arroyo et al. (2022b). These categories were classified into four groups: a) Formulating the problem; b) Conceptual understanding; c) Application of procedures; and d) Interpreting the results. Table 7 with examples of responses to different categories is presented in the Appendix.

FE: Errors in formulating the problem. These errors appear when the report is incorrectly read, there is incorrect identification of the data provided, or faulty translation of the questions into mathematical problems (mathematisation of the questions). This may happen when: FE1) The student might not understand the questions posed or not identify the intervening variables or events; FE2) The problem is not solved because not all the information needed was available; and FE3) There is difficulty in progressing from the verbal statement of the task to developing a mathematical problem.

CE: Conceptual errors. Solving the task requires identifying and applying different concepts, and students might confuse some of these concepts. These conceptual errors were classified as follows: CE1) Not recognising the conditional probability, since in the task this probability is defined verbally instead of symbolically; CE2) Confusing an event and its complementary in questions 1 and 2; CE3) Not distinguishing a conditional probability and its transposed or misapplying the Bayes' theorem in questions 3 to 5; CE4) Confusing dependence and independence of events "testing positive in alcohol or drugs" and "wearing the seatbelt" and "dying in the accident"; and CE5) Identifying conditioning and causation.

PE: Procedural errors. The procedures required to reply to the different questions are mostly reduced to computing the probabilities of the complementary event. Potential errors in these procedures were classified as follows: PE1) incorrect computation of probability, PE2) incorrect computation of proportion or percentages, and PE3) misapplication of the addition or product rules.

RE: Reasoning and interpretation of results. An objective of the task was to increase the critical attitude of PTs when facing incomplete probabilistic information or interpreting counterintuitive results. Potential errors in this interpretation are: RE1) Misinterpreting counterintuitive results; RE2) Probabilistic reasoning biases, such as confusing condition and causation; and RE3) Lacking a critical attitude to face missing or biased information.

5 Results

Below we present the results for each of the three research questions.

5.1 Teachers' probabilistic reasoning and literacy

To answer the research question 1 the responses involving teachers' CCK (Godino, 2009; Pino-Fan et al., 2015) are classified into dealing with computing the probability of a complementary event, discriminating causation and conditioning, and identifying missing data.

5.1.1 Probability of a complementary event

Table 1 presents the results obtained by computing the probability of a complementary event. Most responses to the first two questions were correct; thus, in general, there was a good understanding of the question statement and the participants correctly identified the data in the news report. In addition, the probability of the complementary event was correctly computed, and its application was recognized by the PTs.

A small group of participants repeated the data in the report by incorrectly interpreting the question posed and confusing a probability with its complementary (R2). They were unable to identify the type of probability intervening in the situation and thus missed a component of probability literacy (Gal, 2005). Another failure in statistical literacy was shown in response R3, which implies a-critical reading of the information, since

Table 1 Percentages of responses to questions 1 and 2

Response	Q1	Q2
R1. Correct response	90.9	84.9
R2. Confusing a probability and its complementary	1.5	6.1
R3. Suggesting that more data are needed	1.5	0.0
R4. Other errors	6.1	7.6
Does not reply	0.0	1.5

Table 2 Percentages of responses to questions 3, 4 and 5

Response		Q3	Q4	Q5
Correct	R1. Distinguishing conditional probability and its transposed, as well as conditioning and causation	10.6	10.6	13.6
Partly Correct	R2. Offering missing information	10.6	10.6	10.6
	R3. Applying contextual knowledge	1.5	3.0	3.0
	R4. Distinguishing conditional probability and its transposed with no reference to condition and causation	1.5	4.6	3.0
	R5. Distinguishing between conditioning and causation but confusing a conditional probability and its transposed	25.8	10.6	6.1
	Incorrect	R6. Confusing conditional probability and its transposed, as well as conditioning and causation	31.8	48.5
	R7. Other incorrect responses	18.2	12.1	13.7

both the number of drivers killed (232) and the percentage of those who tested positive (43%) were provided in the report. Other errors were related to improper calculation of the percentage by considering an incorrect total.

These results were similar to those of high school students in Álvarez-Arroyo et al. (2022b), although high school students produced a slightly higher number of errors in Question 2.

5.1.2 Conditioning and causation

Table 2 presents the results for questions 3 to 5. The percentage of correct responses (R1), where the participant rightly argued that the question could not be answered due to lack of information and pointed to some missing data required to solve the question, was small. However, the partly correct responses (categories R2 to R5) totalled 39.4%, 28.8 and 22.7%, respectively, in the three questions.

Globally, the results in these questions were poor because of the existence of biases affecting the participants' probabilistic reasoning (Borovcnik, 2016) that they might transmit to their future students. In all three questions, a high percentage in our study (more than a half) explicitly reasoned according to the *transposed conditional* bias (Falk, 1986) (R5 and R6). These participants assumed that the requested probability was given as data in the problem, because confused a probability $P(D|TP)$ with its transposed $P(TP|D)$. The report provides the probability that a person who died tested positive, and the PT interpreted this data as the probability that a person who had consumed drugs or alcohol died (similar reasoning regarding the use of seatbelt). Moreover, these PTs attributed a relationship between cause and effect to the conditional probability.

Only 19.7% to 36.4% of participants explicitly distinguished between conditioning and causation, and understood that there might be other causes of death apart from the information provided in the report (categories R1 and R5). Finally, we found PTs with both biases (R6) or that provided other incorrect responses. The biases in conditional probability reasoning confirm those found in other prospective teachers by Díaz et al. (2012), where 40% of prospective teachers reasoned according to the transposed conditional bias in a multiple-choice item and by Brückler and Milin

Šipuš (2023). Other incorrect responses included suggesting that the data were based on a sample and were not reliable, or confusing conditional and compound probability.

The percentage of PTs in our sample with biases in conditional probability reasoning was similar to that reported by Álvarez-Arroyo et al. (2022b) with high school students.

In other words, the results suggest the prevalence of biases regardless of the high probabilistic training of participants; this could be explained because formal teaching of probability at the university does not pay attention to these misconceptions.

5.1.3 Identifying missing data

Table 3 reproduces the results for question 6, which evaluates the critical interpretation of probabilistic statements by PTs (another component of probabilistic literacy, Gal, 2005). The results in this question were better. Even when only 45.5% of the PTs correctly identified the missing data to solve the questions posed (R1), another 21.2% answered correctly without specifying the data required to solve the problem. These PTs were aware that the information was incomplete but did not specify what information was needed.

Adding these two categories, 66.7% of the PTs provided correct responses, revealing a critical capacity to interpret probabilistic information in this question. The participants critically analysed the information in the report, correctly interpreted the probability data, and indicated some missing data, applying all of these characteristics of probability reasoning.

It is worth noting that 15.2% of PTs suggested information irrelevant to the solution, which would not help solve the problem. A small percentage alluded to the need for greater awareness of the risk of driving when there is substance consumption or the seatbelt is not used as criticism of the information provided by the media in the task; although this response is irrelevant from the point of view of probability, it is valuable in relation to the formation of PTs. Finally, 6% did not question the information or did not reply. The participants in category R4 suggested that we would not trust the information obtained from the sample. However, even if the data were obtained from a sample, the sample

Table 3 Percentages of responses to question 6

Response	Percentage
R1. Correct, providing the missing data	45.5
R2. Correct, not providing the missing data	21.2
R3. Offering irrelevant information	15.2
R4. Suggesting the information is incorrect as we deal with a sample	9.1
R5. Pointing to the need of public awareness	3.0
R6. Not questioning the information	4.5
Does not reply	1.5

size was large enough to assure a good estimation of the probabilities involved.

In Álvarez-Arroyo et al. (2022b), high school students proved worse probabilistic reasoning in this question, since only 23.7% and 17.1% of them replied correctly or identified the missing data. In addition, 30.3% of them suggested irrelevant information, 7.9% pointed to the need for public awareness, and 11.8% did not reply. Moreover, no participant in our study relied on the authority principle to interpret the data, as did 7.9% of students in Álvarez-Arroyo et al. (2022b). In summary, the PTs in the sample revealed a higher level of critical reading of the information than those students.

5.2 Knowledge in the epistemic facet

To reply research question 2, Table 4 presents the frequency of each of the fundamental stochastic ideas (Heitele, 1975) identified by the PTs as a solution to question 7. Because each participant generally provided several fundamental ideas, the sum of frequencies was higher than 66 (sample size). Thus, an average of 3.1 fundamental ideas per participant were quoted, which is a small number compared with our a priori analysis (Section. 4.4).

Ideas easily identified by the PTs were, in decreasing order of frequency: probability, sample space, conditional probability, and sampling; however, not all of the participants identified these ideas despite their relevance in the task. Less frequent was the understanding that dependence and independence or the addition rule were involved in solving the task. Even if participants implicitly applied these ideas to answer the questions, they could not recognise them explicitly when we asked them to.

Other ideas correctly quoted, but to a much lesser extent (frequencies from 15 to 1), were random variable, the product rule, equidistribution, law of large numbers, and

Table 4 Frequency and percentage of participants identifying different fundamental ideas (question 7)

	Frequency	Percentage
Probability	51	77.3
Sample space	35	53.0
Conditional probability	32	48.5
Sampling	31	47.0
Dependence–independence	15	22.7
Addition rule	14	21.2
Random variable	7	10.6
Product rule	5	7.6
Equidistribution	3	4.5
Law of large numbers	3	4.5
Combinatorics	1	1.5
Ideas incorrectly quoted	7	10.6

combinatorics. Thus, all the ideas of our a priori analysis (Section 4.4) were identified, but the number of PTs who quoted them was negligible. With regard to ideas incorrectly identified, because they do not intervene in the situation, five PTs suggested simulation, one correlation, and another variance.

In summary, we observed scarce knowledge in the epistemic facet of probability in the PTs in our sample. These results coincide with those of Gómez et al. (2016) and Vásquez and Alsina (2015), whose participants also showed limited competence in identifying the mathematical content in probability tasks.

5.3 Knowledge in the cognitive facet

To reply research question 3, Table 5 presents the frequency of students' potential errors identified by the PTs in the task solution (question 8). Because each participant provided several potential difficulties, the sum of frequencies was higher than 66 (sample size).

Most frequent was quoting conceptual errors, although only a third of participants argued that students might not recognise the conditional probabilities in the task. 15 PTs warned of not distinguishing a conditional probability and its transposed or misapplying the Bayes' theorem, and a few of them proposed that the students might not appreciate this dependence or argued the confusion between conditioning and causation.

Errors in interpreting the problem statement or in its formulation was the second most frequently cited category. Most often, these PTs indicated that the students might not

understand the statement or not identify the intervening variables or events; other PTs answered that the students might not reply by considering that not all the information needed was available. Finally, the difficulty in progressing from the verbal statement of the task to developing a mathematical problem was noticed by a few participants.

Regarding procedural errors, some PTs alluded to incorrect computation of probabilities, and a few reported that there might be incorrect computation of proportion or percentages or misapplication of sum or product rules.

The data provided in the report could suggest ideas contrary to students' knowledge and experience about traffic accidents. They required probabilistic reasoning (Borovcnik, 2016) to interpret these results correctly, and some of them may fail. However, only a few participants pointed to the difficulty of interpreting the results: three of them argued that it was difficult to work with counterintuitive information; therefore, these participants did not pay attention to the critical stance towards probabilistic information, which is a component of probabilistic literacy (Gal, 2005) and was needed to solve the task. A few participants pointed to reasoning biases related to conditional probability, and only four recognised that the lack of a critical attitude to face missing or biased information might be an obstacle for students. Probabilistic reasoning requires sound judgement to establish the credibility of the evidence (Schum, 2001), but only a few participants recognised that this might be an obstacle for the students.

Table 5 Frequency and percentage of potential errors identified by participants in the task solution (question 8)

Type of error	Frequency	%	Sum
<i>FE. Errors in formulating the problem</i>			49
FE1. Not understanding the statement or not identifying the events or variables involved	37	56.1	
FE2. Not solving the problem because lack of information	9	13.6	
FE3. Not formulating the problem mathematically	3	4.5	
<i>CE. Conceptual errors</i>			66
CE1. Not recognising the conditional probability	21	31.8	
CE2. Confusing an event and its complementary	15	22.7	
CE3. Not distinguishing a conditional probability and its transposed or misapplying the Bayes' theorem	15	22.7	
CE4. Confusing dependence and independence	10	15.2	
CE5. Confusing conditioning and causation	5	7.6	
<i>PE. Procedural errors</i>			22
PE1. Incorrect computation of probabilities	14	21.2	
PE2. Incorrect computation of proportion or percentages	5	7.6	
PE3. Misapplying the addition or product rules	3	4.5	
<i>RE. Reasoning and interpretation of results errors</i>			12
RE1. Misinterpreting counterintuitive results	3	4.5	
RE2. Probabilistic reasoning biases	5	7.6	
RE3. Lacking a critical attitude to face missing or biased information	4	6.1	
<i>Errors not applying in the task proposed</i>			30

Finally, 30 incorrect responses pointed to errors that did not apply in the tasks, such as confusing the sample and population or incorrectly solving equations.

In summary, we observed poor knowledge of PTs in the cognitive facet, given the scarce number of difficulties predicted (2.3 in average). Again, our results agree with those of Gómez et al. (2016) and Vásquez and Alsina (2015), whose participants also showed poor competence to recognise students' errors in probability tasks.

6 Discussion and conclusions

This study investigates the probabilistic reasoning and literacy of a sample of prospective secondary school teachers when interpreting news from the media. Their capacity to identify the fundamental stochastic ideas needed to solve the task, and their ability to predict their students' potential difficulties when solving the same task were also analysed.

Regarding the first research question, although it is not possible to generalize to other teachers, as we worked with a non-random sample, the results of our study indicate that most participants in the sample put into practice their probabilistic knowledge and reasoning, which are part of their probabilistic CCK. These participants showed their knowledge of probabilistic language and competence in identifying the problem data and computing probabilities, particularly, in the first two questions posed to them.

A substantial part of the sample correctly interpreted probabilistic statements and chose the probabilistic model to be applied in the situation in the other questions. In fact, and according to Sánchez and Valdez (2017), their probabilistic reasoning was evidenced when solving a non-routine probability problem and using arguments to prove the truth of a probabilistic assertion. Moreover, in some questions some participants made efficient use of contextual knowledge, which they used to compensate for the lack of data and as a support to discriminate between conditioning and causation, two additional components of probabilistic reasoning, according to Borovcnik (2016). All of these results were better than those of high school students in Álvarez-Arroyo et al. (2022b), and predictable, given the high formal knowledge in probability of participants in our study.

The study outcomes, however, point to important shortcomings in the participants' probability reasoning, with a notable proportion of them not recognising the difference between conditioning and causation (questions 3 to 5) and incorrectly interpreting probabilistic statements, two basic components of probabilistic reasoning suggested

by Borovcnik (2016) and Batanero and Borovcnik (2016). Moreover, noteworthy is the confusion of a conditional probability with its transposed, which has already been shown in previous research with students (e.g., Díaz & de la Fuente, 2007; Falk, 1986) and prospective teachers (Brückler & Milin Šipuš, 2023; Díaz et al., 2012). The percentage of PTs in our sample with biases in conditional probability reasoning was similar to that reported by Álvarez-Arroyo et al. (2022b) with high school students. This prevalence of biases regardless of the high probabilistic training of participants raises a concern about the formal teaching of probability at university.

Finally, some participants lacked the capacity to make judgments to establish the credibility of evidence in the report, another component of probabilistic reasoning (Schum, 2001). This capability is not taught formally to the students, but it is a knowledge needed by the teachers to adequately teach probability.

In response to the second and third research questions in this study, the didactic knowledge of these prospective teachers in both the epistemic and cognitive facets was poor, even after master's degree training. The responses to Question 7 proved the participants' limited competence to identify the fundamental stochastic ideas (Heitele, 1975) in the task they had previously solved and their failure to predict the potential errors of their students. Similar results were also found by Gómez et al. (2013) and Vásquez and Alsina (2015) in primary school prospective teachers.

Consequently, we should reflect that high mathematical preparation alone does neither assure the teachers' needed competence to adequately teach probability, nor it is enough to overcome their probabilistic reasoning biases. A limitation of the teaching experiment was that the final questions in the task did not help develop sufficiently the epistemic and cognitive facets of teachers' knowledge and advice more practice with performing epistemic and cognitive analyses are needed in the education of teachers.

These results open a line of research to develop new tasks such as the one analysed in this paper, which complement traditional problem solving in teaching and teacher education. On the one hand, these activities visualise the usefulness of probability in everyday life while reinforcing probabilistic reasoning; on the other hand, they make teachers aware of their own reasoning biases. These tasks should be complemented with questions that address the remaining didactic facets (affective, interactional, mediational, and ecological) and the meta didactic knowledge in teachers' education. Simulation of the situations analysed in the tasks might also help improve the probabilistic reasoning and didactic knowledge of prospective teachers (Biehler et al., 2018).

Appendix. Examples of prospective teachers' responses

Table 6 Examples of responses to mathematical questions

Questions	Category	Example of response
1–2. Probability of complementary event	R1. Correct response	P64: The probability that a deceased person wore the seatbelt is $1 - 0.23 = 0.77$ (question 2)
	R2. Confusing a probability and its complementary	P54: Less than 43% (question 1)
	R3. Suggesting there were not enough data	P21: We cannot answer this question. Although we know that more than 43% tested positive, the report does not inform, if the same person was positive and wore the seatbelt (question 1)
3–5. Conditioning and causation	R1. Correct response	P6: We know that 43% of the sample was drugged, and the possibility of an accident increases in this case. Interpreting that because 57% were not drugged it is more likely to die in that case is a typical error. We need P(Death Drug) and only know P (Drug Death) (question 4)
	R2. Offering some missing information; not pointing to the difference between conditional probability and cause-effect	P3: No, the data are taken from the analyses made to 535 drivers, but we do not know the total number of deaths (question 3)
	R3. Applying contextual knowledge	P12: It seems more likely to die without alcohol and drugs consumption, but it is because of the lack of clarity in the news report (question 4)
	R4. Distinguishing conditional probability and its transposed with no reference to conditioning and causation	P18: We are given a conditional probability P(Drug Death). To reply to the question, we need P(Death Drug) (question 4)
	R5. Distinguishing conditioning and causation but confusing a conditional probability and its transposed	P3: We do not know the real causes of the deaths and if the seatbelt could have saved these lives. Considering the percentages, more people died with the seatbelt (question 4)
	R6. Confusing conditional probability and its transposed, and conditioning and causation	P45: The probability is higher with the belt (question 5)
6. Critical interpretation	R1. Correct, providing missing information	P6: It is incorrect. This confusion is provoked by the fact that the data given are P(S D) and P(TP D); however, we need P(D S) and P(D TP)
	R2. Correct answer, not providing the missing data	P9: I do not think it is correct. Someone might think that it is better to drink
	R3. Offering irrelevant information	P14: I believe mixing all the data is an error, since it is not equally relevant to be pedestrian, companion, or driver
	R4. Suggesting that the information is incorrect as we deal with a sample	P4: It is incorrect since we have no description of the population from which the data have been collected
	R5. Need to raise public awareness,	P18. If the TGD want to raise public awareness, they should give the probability of dying when you do not wear a seatbelt

Table 7 Examples of responses to Question 8

Type of error	Category	Example of response
FE. Understanding the problem	FE1. Not understanding the statement	P14: The biggest difficulty is the interpretation of data in the report. Students can get confused by so many data and make a mess of it
	FE2. Not solving the problem because lack of information	P11: As I read some questions, I was unsure what to answer because I felt the information was incomplete and an answer was impossible
	FE3. Not formulating the problem mathematically	P11: Not considering the data as conditional probabilities
CE. Conceptual errors	CE1. Not recognising the conditional probability	P1: A potential difficulty in the task is the concept of conditional probability, although the formula is not needed in the resolution
	CE2. Confusing an event and its complementary	P53: Not identifying the complementary events that appear in the questions
	CE3. Not distinguishing a conditional probability and its transposed or misapplying the Bayes' theorem	P68: Difficulty to distinguish the conditional probabilities $P(A B)$ and $P(B A)$
	CE4. Confusing dependence and independence	P67: There might be incapacity to understand that the events are dependent
	CE5. Confusing conditioning and causation	P17: Not applying the difference between conditioning and causation
PE. Procedural errors	PE1. Incorrect computation of probabilities	P58: The student might confuse the computation of different types of probability
	PE2. Incorrect computation of proportion or percentages	P50: Error in computing percentages
	PE3. Misapplying the addition or product rules	P48: It is likely to be confused since they should know the product rule
RE. Reasoning and Interpretation	RE1. Misinterpreting counterintuitive results	P36: Interpretation of results. It must be shocking that more people die with a seatbelt
	RE2. Probabilistic reasoning biases	P53: Determine whether a specific event is the cause of death or simply a circumstance
	RE3. Lacking a critical attitude to face missing or biased information	P10: Some students might lack a critical sense and not reason using the data provided by the TGD that they might interpret incorrectly

Acknowledgements Grant PID2022-139748NB-I00 funded by MCIN/AEI/<https://doi.org/10.13039/501100011033> and Fondo Europeo de Desarrollo Regional (Una manera de hacer Europa).

Funding Funding for open access publishing: Universidad de Granada/CBUA.

Declarations

Disclosure No financial interest or benefits that create a conflict of interest are involved in this study.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will

need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Alonso-Castaño, M., Alonso, P., Mellone, M., & Rodríguez-Muñiz, L. J. (2021). What mathematical knowledge do prospective teachers reveal when creating and solving a probability problem? *Mathematics*, 9(24), 3300. <https://doi.org/10.3390/math9243300>
- Álvarez-Arroyo, R., Lavela, J. F., & Batanero, C. (2022a). Razonamiento probabilístico de estudiantes de bachillerato al interpretar datos de la COVID-19. *REDIMAT, Journal of Research in Mathematics Education*, 11(2), 117–139. <https://doi.org/10.17583/redimat.9741>
- Álvarez-Arroyo, R., Lavela, J. F., Gea, M. M., & Batanero, C. (2022b). High school probabilistic reasoning when interpreting media news. In S. A. Peters, L. Zapata-Cardona, F. Bonafini, & A. Fan (Eds.), *Proceedings of the 11th International Conference on Teaching Statistics (ICOTS11)*. International Association for Statistical Education. http://iase-web.org/Conference_Proceedings.php?p=ICOTS_11_2022. Accessed 1 Feb 2024.

- Bargagliotti, A., Franklin, C., Arnold, P., Gould, R., Johnson, S., Pérez, L., & Spangler, D. (2020). *Pre-K-12 guidelines for assessment and instruction in statistics education II (GAISE II)*. American Statistical Association & National Council of Teachers of Mathematics.
- Batanero, C., & Álvarez-Arroyo, R. (2023). Teaching and learning of probability. *ZDM Mathematics Education*. <https://doi.org/10.1007/s11858-023-01511-5>
- Batanero, C., & Borovcnik, M. (2016). Statistics and probability in high school. *Sense Publishers*. <https://doi.org/10.1007/978-94-6300-624-8>
- Ben-Zvi, D., & Garfield, J. (2004). Statistical literacy, reasoning, and thinking: Goals, definitions, and challenges. In D. Ben-Zvi, D., & Garfield, J. B. (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 3–15). Kluwer.
- Biehler, R., Frischmeier, D., & Podworny, S. (2018). Elementary pre-service teachers' reasoning about statistical modeling in a civic statistics context. *ZDM – Mathematics Education*, 50, 1237–1251. <https://doi.org/10.1007/s11858-018-1001-x>
- Borovcnik, M. (2016). Probabilistic thinking and probability literacy in the context of risk. *Educação Matemática Pesquisa*, 18(3), 1491–1516.
- Breda, A., Pino-Fan, L. R., & Font, V. (2017). Meta didactic-mathematical knowledge of teachers: Criteria for the reflection and assessment on teaching practice. *Eurasia Journal of Mathematics, Science and Technology Education*, 13(6), 1893–1918. <https://doi.org/10.12973/eurasia.2017.01207a>
- Brückler, F. M., & MilinŠipuš, Ž. (2023). Pre-service mathematics teachers' understanding of conditional probability in the context of the COVID-19 pandemic. *European Journal of Science and Mathematics Education*, 11(1), 89–104. <https://doi.org/10.30935/scimath/12436>
- Burns, Z., Chiu, A., & Wu, G. (2010). Overweighting of small probabilities. In J. Cochran, L. A. Cox, P. Keskinocak, J. Kharoufeh, & J. C. SmithWiley (Eds.), *Encyclopedia of Operations Research and Management Science*. John Wiley. <https://doi.org/10.1002/9780470400531.eorms0634>
- Carrillo, J., Climent, N., Montes, M., Contreras, L. C., Flores, E., Escudero, D., Vasco, D., Rojas, N., Flores, P., Aguilar, Á., et al. (2018). The mathematics teacher's specialised knowledge (MTSK) model. *Research in Mathematics Education*, 20, 236–253. <https://doi.org/10.1080/14794802.2018.1479981>
- Chernoff, E. J., Vashchysyn, I., & Neufeld, H. (2018). Comparing the relative probabilities of events. In C. Batanero & E. Chernoff (Eds.), *Teaching and learning stochastics. Advances in probability education research* (pp. 277–291). Springer. https://doi.org/10.1007/978-3-319-72871-1_16
- Díaz, C., & de la Fuente, I. (2007). Assessing students' difficulties with conditional probability and Bayesian reasoning. *International Electronic Journal of Mathematics Education*, 2(3), 128–148. <https://doi.org/10.29333/iejme/180>
- Díaz, C., Contreras, J. M., Batanero, C., & Roa, R. (2012). Evaluación de sesgos en el razonamiento sobre probabilidad condicional en futuros profesores de Educación Secundaria. *Bolema*, 26, 1207–1226. <https://doi.org/10.1590/S0103-636X2012000400006>
- Falk, R. (1986). Conditional probabilities: Insights and difficulties. In R. Davidson & J. Swift (Eds.), *Proceedings of the Second International Conference on Teaching Statistics* (pp. 292–297). International Statistical Institute.
- Gal, I. (2005). Towards "probability literacy" for all citizens: Building blocks and instructional dilemmas. In G. A. Jones (Ed.), *Exploring probability in school: challenges for teaching and learning* (pp. 39–63). Springer. https://doi.org/10.1007/0-387-24530-8_3
- Gigerenzer, G. (2015). *Calculated risks: How to know when numbers deceive you*. Simon and Schuster.
- Gilovich, T., Griffin, D., & Kahneman, D. (Eds.) (2002). *Heuristics and biases: The psychology of intuitive judgment*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511808098>
- Godino, J. (2009). Categorías de análisis de los conocimientos del profesor de matemáticas. *UNIÓN, Revista Iberoamericana de Educación Matemática*, 20, 13–31.
- Gómez, E., Batanero, C., & Contreras, J. M. (2013). Conocimiento matemático de futuros profesores para la enseñanza de la probabilidad desde el enfoque frecuencial. *Bolema*, 28(48), 209–229. <https://doi.org/10.1590/1980-4415v28n48a11>
- Gómez, E., Batanero, C., Díaz, C., & Contreras, J. M. (2016). Developing a questionnaire to assess the probability content knowledge of prospective primary school teachers. *Statistics Education Research Journal*, 15(2), 197–215. <https://doi.org/10.52041/serj.v15i2.248>
- Heitele, D. (1975). An epistemological view on fundamental stochastic ideas. *Educational Studies in Mathematics*, 6(2), 187–205.
- Hourigan, M., & Leavy, A. M. (2020). Pre-service teachers' understanding of probabilistic fairness: Analysis of decisions around task design. *International Journal of Mathematical Education in Science and Technology*, 51(7), 997–1019. <https://doi.org/10.1080/0020739X.2019.1648891>
- Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge University Press.
- Lilleholt, L. (2019). Cognitive ability and risk aversion: A systematic review and meta analysis. *Judgment and Decision Making*, 14(3), 234–279. <https://doi.org/10.1017/S1930297500004307>
- Llinares, S. (2023). Multidimensionality as a feature of the research in mathematics teacher education: Different targets to be noticed and different lenses to describe and explain. *Journal of Mathematics Teacher Education*, 26, 1–4. <https://doi.org/10.1007/s10857-023-09568-5>
- Malaspina, M., & Malaspina, U. (2020). Game invention as means to stimulate probabilistic thinking. *Statistics Education Research Journal*, 19(1), 57–72. <https://doi.org/10.52041/serj.v19i1.119>
- Martin, V., Thibault, M., & Homier, M. (2022). Convergence and divergence in probability teaching in primary and secondary school in Québec: A closer look at eight teachers' self-reported practices. *Canadian Journal of Science, Mathematics, and Technology Education*, 22, 659–678. <https://doi.org/10.1007/s42330-022-00241-2>
- Ministerio de Educación y Formación Profesional. (2022). *Real Decreto 217/2022, de 29 de marzo, por el que se establece la ordenación y las enseñanzas mínimas de la Educación Secundaria Obligatoria*. Madrid: Autor.
- Muñiz-Rodríguez, L., Velázquez, P. A., Rodríguez-Muñiz, L. J., & Valcke, M. (2016). Is there a gap in initial secondary mathematics teacher education in Spain compared to other countries? *Revista de Educación*, 372, 106–132. <https://doi.org/10.4438/1988-592X-RE-2015-372-317>
- Neuendorf, K. (2016). *The content analysis guidebook*. Sage. <https://doi.org/10.4135/9781071802878>
- Pino-Fan, L., & Godino, J. D. (2015). Perspectiva ampliada del conocimiento didáctico-matemático del profesor. *Paradigma*, 36(1), 87–109.
- Pino-Fan, L. R., Assis, A., & Castro, W. F. (2015). Towards a methodology for the characterization of teachers' didactic-mathematical knowledge. *EURASIA Journal of Mathematics, Science and Technology Education*, 11(6), 1429–1456. <https://doi.org/10.12973/eurasia.2015.1403a>
- Sánchez, E., & Valdez, J. C. (2017). Las ideas fundamentales de probabilidad en el razonamiento de estudiantes de bachillerato. *Avances de Investigación en Educación Matemática*, 11, 127–143. <https://doi.org/10.35763/aiem.v1i11.180>
- Sánchez, E., Carrasco, G., & Herrera, M. (2018). Fundamental ideas in the probabilistic reasoning of high-school students in binomial distribution activities. In M. A. Sorto, A. White, & L. Guyot (Eds.), *Proceedings of the Tenth International Conference on Teaching Statistics (ICOTS10)*. IASE. Available from http://iase-web.org/icots/10/proceedings/pdfs/ICOTS10_6C3.pdf?1531364279. Accessed 1 Feb 2024.
- Schum, D. A. (2001). *The evidential foundations of probabilistic reasoning*. Northwestern University Press.
- Sharma, S. (2015). Teaching probability: A socio-constructivist perspective. *Teaching Statistics*, 37(3), 78–84. <https://doi.org/10.1111/test.12075>

- Van Dooren, W. (2014). Probabilistic thinking: Analyses from a psychological perspective. In E. Chernoff, & B. Sriraman (Eds.), *Probabilistic thinking. Presenting multiple perspectives* (pp. 123–126). Springer. https://doi.org/10.1007/978-94-007-7155-0_7
- Vásquez, C., & Alsina, Á. (2015). El conocimiento del profesorado para enseñar probabilidad: Un análisis global desde el modelo del Conocimiento Didáctico-Matemático. *Avances de Investigación en*

Educación Matemática, 7, 27–48. <https://doi.org/10.35763/aiem.v1i7.104>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.