PROSPECTIVE TEACHERS' COMMON AND SPECIALIZED KNOWLEDGE IN A PROBABILITY TASK

J. Miguel Contreras[1], Carmen Batanero[1], Carmen Díaz[2], and José A. Fernandes[3]

[1]University of Granada, Spain; [2]University of Huelva, Spain,

[3]University of Minho, Portugal

The aim of this research was to assess the common and specialized knowledge of elementary probability in a sample of 183 prospective primary school teachers in Spain, using an open-ended task. Common knowledge of probability was assessed in the first part of the task, where teachers had to compute simple, compound and conditional probability from data presented in a two-way table. The specialized knowledge of probability was assessed in the second part of the task, were teachers were asked to identify and classify the mathematical content in the problem proposed. Results suggest participants' poor common and specialized knowledge of elementary probability in this task and point to the need of reinforcing the preparation of prospective teachers to teach probability.

INTRODUCTION

The reasons for including probability in schools have been repeatedly highlighted over the past years (e.g., Gal, 2005; Jones, 2005): usefulness of probability for daily life, its instrumental role in other disciplines, the need for a basic stochastic knowledge in many professions, and the important role of probability reasoning in decision making. Consequently, probability has recently been included in the primary school curriculum in many countries, where changes do not just concern the age of learning or the amount of material, but also the approach to teaching (Franklin et al., 2005). The success of these curricula will depend on the extent to which we can convince teachers that probability is an important topic for their students, as well as on the correct preparation of these teachers. Unfortunately, several authors (e.g., Franklin & Mewborn, 2006; Chick & Pierce, 2008) agree that many of the current programmes still do not train teachers adequately for their task to teach statistics and probability. The above reasons suggest to us the need to reinforce the specific and didactic preparation of primary school statistics teachers, and also the relevance of assessing the teachers' difficulties and errors in learning the topic.

Components in Teachers' Knowledge

An increasing number of authors have analysed the nature of knowledge needed by teachers to achieve truly effective teaching outcomes. Shulman (1987) described "pedagogical content knowledge" (PCK) as "that special amalgam of content and pedagogy that is uniquely the province of teachers, their own special form of professional understanding" (p. 6). Ball and her colleagues (Ball, Lubienski, & Mewborn, 2001; Hill, Ball, & Schilling, 2008) developed the notion of "mathematical knowledge for teaching" (MKT) in which they distinguished six main categories (see

Ball et al., 2001 for a comprehensive description). Our research was intended to assess two of these components in prospective primary school teachers in relation to elementary probability. More specifically we were interested in the following components of teachers' knowledge:

- *Common content knowledge* (CCK) or the mathematical knowledge teachers are responsible for developing in their students; that is, the mathematical knowledge that is typically known by competent adults (Hill et al., 2004). In this research we assess common knowledge of elementary probability with a task where teachers are asked to compute single, compound and conditional probability from a two-way table.
- *Specialized content knowledge* (SCK). In addition to common knowledge, teachers need to know the content they teach in ways that differ from what is typically taught and learned in mathematics courses. SCK is the mathematical knowledge that is used in teaching, but not directly taught to students (Hill et al., 2004). We include here the ability to recognise what probabilistic concepts or properties can be addressed in the teaching tasks and resources (that was considered by Chick & Pierce, 2008 as a part of PCK). To assess this knowledge, in this research participants are asked to identify and classify the mathematical content they used to solve the first part of the task.

Below we first summarise related previous research and then describe the method and results in this study.

PREVIOUS RESEARCH

Two-Way Tables and Conditional Probability

A two-way or contingency table serves to present in a summarised way the frequency distribution in a population or sample that was classified according to two statistical variables (an example is included in the task presented in Figure 1). Research on contingency tables, started with the pioneer study by Inhelder and Piaget (1955), and focused on students' strategies and conceptions when assessing association between the variables in rows and columns from the data presented in a two-way table (e.g., Batanero, Estepa, Godino, & Green, 1996). More recently research has focussed on students' performance when computing probabilities with data presented in a two-way tables (see Huerta, 2009, for an analysis of the structure of these problems).

Also relevant for this study is research related to conditional probability, such as that by Falk's (1986) who remarked that many students do not adequately discriminate between the two different conditional probability, that is, *P(A/B)* and *P(B/A).* Falk termed this confusion as *fallacy of transposed conditional.* Einhorn and Hogarth (1986) observed that some students confused joint and conditional probability because they misinterpreted the conjunction "and".

Teachers' Probabilistic Knowledge

The scarce research related to primary school teachers' understanding of probability indicates this understanding is weak. For example, Begg and Edward (1999) found that only about two-thirds of the in-service and pre-service primary school teachers in their sample understood equally likely events and very few understood the concept of independence. Batanero, Cañizares, and Godino's (2005) found three widespread probabilistic misconceptions in a sample of 132 pre-service teachers related to representativeness (Tversky & Kahneman, 1982), equiprobability (Lecoutre, 1992) and the outcome approach (Konold, 1991). Fernandes and Barros (2005) study with 37 pre-service teachers in Portugal suggested the teachers' difficulties to formulate events and to understand compound and certain events. In addition, these teachers frequently used additive reasoning to compare probabilities.

In relation to knowledge needed to teach probability, Stohl (2005) suggested that few teachers have prior experience with conducting probability experiments or simulations and many of them may have difficulty implementing an experimental approach to teaching probability. In Lee and Hollebrands's (2008) research, although the participant teachers engaged students in investigations based on probability experiments, their approaches to using empirical estimates of probability did not foster a frequentist conception of probability. Teachers almost exclusively chose small samples sizes and rarely pooled class data or used representations supportive of examining distributions and variability across collections of samples so they failed to address the heart of the issue.

Estrada and Díaz (2006) asked 65 prospective primary school teachers, who had followed a 60 hours long course in statistics education at the University of Lleida, in Spain, to compute simple, compound and conditional probability from data presented in a two-way table and analysed the solutions provided by these teachers. The authors found a large proportion of participants who were unable to provide any solution to the problems. There were a variety of errors, including confusion between compound and conditional probability, confusion between an event and its complementary, confusion between probabilities with possible cases (absolute frequencies), and assuming independence in the data. The aim of the present paper is to expand the research by Estrada and Díaz (2006) with a bigger sample of prospective teachers, who had not followed a specific course in statistics education. In addition, the second part of the task is intended to assess the SCK of probability that was not taken into account in Estrada and Díaz's research.

METHOD

The sample in the study consisted of 183 prospective teachers at the Faculty of Education, University of Granada, Spain. The task analysed in this paper was answered individually by each participant as a part of the final assessment in a course of Mathematics Education. In this course (60 teaching hours), the prospective teachers are introduced to the primary school mathematics curriculum, didactic resources, children's difficulties, and technological tools for teaching elementary mathematics. Most sessions are devoted to practical work, in which participants performed didactic analyses (including identification of mathematical content) of curricular guidelines, school textbooks, assessment items and children responses to these items, and teaching episodes. Several sessions of the course are devoted to probability and statistics education. The previous year all these prospective teachers took a Mathematics course (90 teaching hours) with about 10 hours of in-classroom work and 40 additional hours of extra-classroom work devoted to statistics and probability (data, distribution, graphs, averages, variation, randomness and probability, including some exercises of compound and conditional probability).

The task given to participants is presented in Figure 1 and is similar to another task used by Estrada and Díaz (2006), although the statement was simplified, in order to avoid the use of negative statements in the wording of the item and the use of inequalities in the definition of the events in the sample space. The three questions in the first part of the task, were aimed to assess the prospective teachers' CCK in relation to elementary probability. More specifically we were interested in the prospective teachers' ability to read the table and identify the data needed to compute a simple probability (question a), a compound probability (question b) and a conditional probability (question c). The second part was aimed to assess the participants' SCK knowledge of probability; more specifically we were interested in their ability to identify the mathematical problems, concepts, properties, language, procedures and language implicit or used to solve the task.

A survey in a small school provided the following results:

	Boys	Girls	Total
Liking tennis	400	200	600
Disliking tennis	50	50	100
Total	450	250	700

Part 1. Providing that we select one of the school students at random:

a. What is the probability that the student likes tennis?
b. What is the probability that the student is a girl and likes tennis?
c. The student selected is a girl. What is the probability that she does like tennis?

Part 2. Identify the mathematical content you used to solve the above tasks (specify the types of problems; concepts, procedures; properties, mathematical language and mathematical arguments you used to solve the task).

Figure 1. Task given to participants in the study

RESULTS

Common mathematical knowledge

The written reports produced by the participants in the study were analysed and the answers to each question were categorized, taking into account the correctness of the response, as well as the type of errors, in case of incorrect response, as follows:

Basically correct answers: We group in this category answers that showed students correctly read the two-way table, identified the probability required and provided a correct solution to the problem. We also include here responses that provided a correct numerical result, with incorrect symbolization of probabilities, such as for example: "*The probability of liking tennis is P(600/700)"* (Student 70). The percentage of basically correct responses is low, except for the first question (65,6%), in agreement with what was reported by Estrada and Díaz in their sample.

Confusing probabilities: Some participants confused simple, compound and/or conditional probabilities. The most frequent confusion (13,7%) was between conditional and compound probability: "*Probability of liking tennis assuming the student is a girl is 200/700"* (Student 73). This is an error described by Einhorn and Hogarth (1986) in university students and also found in 17% of prospective teachers in Estrada and Díaz's research. A few participants confused *P(A/B)* and *P(B/A),* an error that was described by Falk (1986): *"There is 33% probability that a girl likes tennis"* (Student 71). In the following example, instead of computing a simple probability, the student computed two conditional probabilities; we observe the student's inability to read the data in the two-way table as he did not reach the "reading between data" level (Curcio, 1989): *Probability of liking tennis is: 4/6=66,6% for boys and 2/6=33,3% for girls"* (Student 36). Other students confused simple probability with the probability of an elementary event: *"Probability of liking tennis if you select a student at random is 1/700, since there are 700 students"* (Student 82). The percentage of pre-service teachers confusing different probabilities was slightly lower than that reported by Estrada and Díaz, possibly because the task was simplified.

Confusing events: A few prospective teachers identified the probability but confused an event and its complement, an error described by Estrada and Díaz (2006); "*Probability of liking tennis is* $\frac{50}{250} = 20\%$ *"* (Student 102*)*, which again suggest the pre-service teachers' inability to read the two-way table. Additionally some prospective teachers *confused other different mathematical objects*; such as probability and frequency (or number of favourable cases) and for this reason obtained a probability higher than 1.

Confusing formulas: A small number of pre-service teachers identified correctly the probability to be computed and used correct symbols, but did not remember the formula, so that the final result was wrong. *Other errors* consisted in computing means of frequencies, or assuming independence in the data and applying directly the product's rule for independent events when computing compound probabilities.

Table 1. Frequency (and percentage) of responses to the three questions

Teacher's answer	P(A)	P(A∩B)	P(A/B)
Basically correct	120 (65,6)	75 (41,0)	80 (43,7)
Confuse probabilities	8 (4,4)	46 (25,1)	30 (16,4)
Confuse other objects	9 (4,9)	10 (5,5)	5 (2,7)
Confuse formulas	3 (1,6)	2 (1,1)	2 (1,1)
Confuse events	0 (0,0)	7 (3,8)	8 (4,4)
Other errors	1 (0,5)	4 (2,2)	11 (6,0)
Do not provide an answer	42 (23,0)	39 (21,3)	47 (25,7)
Total	183 (100)	183 (100)	183 (100)

The students' responses are presented in Table 1, where we use the following abbreviations: *A*="the student likes tennis"; *B*="the student is a girl". Although the majority of participants correctly computed simple probability, less than 45% of responses when computing compound and conditional probabilities were correct. Also, similarly to Estrada and Díaz's research, an important percent of participants in our study did not provide any solution. There were a variety of errors reported in previous research, in particular confusion between different probabilities, and at the same time we found other mistakes which have not been described in the literature, such as confusing a simple probability with the probability of an elementary event.

Specialized knowledge of content

In the second part of the task, we asked the participants to identify the probability content needed to solve the task. We included the following categories of objects:

- *Problems:* We expected the students to identify the three different specific problems in the task: A simple probability problem in part (a), a compound probability problem in part (b), and a conditional probability problem in part (c).
- *Language:* Verbal, numerical and tabular mathematical language appears in the task statement; depending on the solution, some students would also use symbolic and graphical language.
- *Concepts:* Implicit in the task we can identify the concepts of random experiment (selecting a school student at random); simple and compound events; sample spaces, favourable and unfavourable cases for each question; simple, compound and conditional probability, fraction, ratio and proportion, frequency and percentage, integer numbers, operations with integer numbers (division).
- *Properties* (or relationships between concepts). Some properties implicit in this task are: The probability axioms; the relation between the probability for an event and that of its contrary; the fact that sample space is restricted in computing conditional probability; equivalence of two fractions when dividing the two terms of the fraction by the same number; the Laplace rule; the relation between the total

sample size and the totals in rows or in columns; the relation between double, marginal and conditional frequencies.

- *Procedures* (or algorithms). Possible procedures that can be used in solving these tasks include doing numerical operations, such as division or addition, operating or simplifying fractions, reading a table; transforming a probability in percentage; applying the formulas for computing simple, compound and conditional probability, and computing percentages or proportions.
- *Arguments.* The main correct type of argument used to solve the task is *deductive* argument, which was identified by many students.

Many students were able to identify and correctly classify some of the above mathematical objects in the problem; although, in general, the number of objects identified was quite small, and an important proportion of students were unable to give examples in some categories. Other examples provided by the students were considered incorrect, due to some of the following reasons:

- Some responses were too imprecise, for example, answering that a mathematical problem was "*replying the questions that appear after the data table*" (Student 95) or that "*there are three different mathematical problems in the task*" (Student 125); these responses do not specify the type of problems (simple, conditional or compound probability problem).
- Some students confused the different types of mathematical objects; for example, some of them considered the procedures "*interpreting the table*" or "*performing a division*" to be concepts. Other students confused procedures with their solution or confused phenomenological elements with mathematical objects. For example, some students suggested *"girls"* or *"liking tennis*" instead of "*event*" as examples of mathematical concepts.
- Other students included in their responses some mathematical objects that were not needed to solve the task, such as, for example, "*median, mode, standard deviation*".

The number of correct and incorrect examples of mathematical objects provided by each participant in each category varied, ranging from not being able to identify a mathematical object in a given category to including several examples (in table 2 we present the mean and standard deviation). Results suggest that identifying the mathematical objects implicit in the task was not easy for the participants in the sample. On average, only half the students correctly identified a mathematical problem (even when three different problems were proposed in the task) and only a third identified correctly a property or the use of deductive argument. The easiest elements for the prospective teachers were concepts (2-3 concepts correctly identified per participant), procedures and language (1-2 correctly identified). Anyway, although some prospective teachers suggested incorrect mathematical objects in all the categories, the average number of correct responses was higher than the number of incorrect responses in all the categories and the differences were statistically significant, except for properties and arguments that were hardest to be identified in the task by participants.

Table 2. Mean and standard deviation for the number of correct and incorrect mathematical objects identified in the task

	Correct		Incorrect		p-value in the t-test of differences
Objects	Mean	Std. Dev.	Mean	Std. Dev.	
Problems	0,54	0,70	0,17	0,47	0,004*
Language	1,37	1,91	0,57	1,23	0,007*
Concepts	2,22	2,01	0,87	1,58	0,002*
Procedures	1,40	1,71	0,38	0,89	0,003*
Properties	0,32	0,77	0,26	0,62	0,076 N.S.
Arguments	0,37	0,62	0,26	0,62	0,066 N.S.

* Differences statistically significant at 0,05.

IMPLICATIONS FOR TRAINING THE TEACHERS

Our results suggest that computing simple, compound and conditional probabilities from a two-way table was not easy for participants in the sample who showed a weak common knowledge of probability to solve this task. Many teachers were unable to provide an answer to the problems, in agreement with Estrada and Díaz' (2006) research, or made errors reported in previous research, particularly by Einhorn and Hogarth (1986) and Falk (1986). We agree with Falk that the everyday language we use to state a conditional probability problem lacks precision and is therefore ambiguous. However, a future teacher should master both the concept and the language used in teaching, particularly the language which today is part of statistical literacy, which is important for their students, and which they should transmit them.

Participants also had difficulty in identifying and classifying mathematical objects in this task coinciding with Chick and Peirce's (2008) results, which suggest that the specialised knowledge of elementary probability was also poor. These results are cause for concern, since prospective teachers in our sample are likely to fail in future teaching of probability in some professional activities, such as "figuring out what students know; choosing and managing representations of mathematical ideas; selecting and modifying textbooks; deciding among alternative courses of action" (Ball, Lubienski, & Mewborn, 2001, p. 453). These activities involve mathematical reasoning and thinking, which were weak for these teachers when dealing with probability. To conclude these results suggest the need to reform and improve the probability education these future teachers are receiving during their training in the schools of education.

Acknowledgements: Research supported by the project EDU2010-14947 (MCIN), grant FPI BES-2008-003573 (MEC-FEDER) and group FQM126 (Junta de Andalucía).

REFERENCES

Ball, D. L., Lubienski, S. T., & Mewborn, D. S. (2001). Research on teaching mathematics: The unsolved problem of teachers' mathematical knowledge. In V. Richardson (Ed.), *Handbook of research on teaching* (pp. 433-456). Washington, DC: American Educational Research Association.

Batanero, C., Estepa, A., Godino, J. D., & Green, D. R. (1996). Intuitive strategies and preconceptions about association in contingency tables. *Journal for Research in Mathematics Education, 27*(2), 151-169.

Batanero, C., Godino, J. D., & Cañizares, M. J. (2005) Simulation as a tool to train Pre-service School Teachers. In J. Addler (Ed.), *Proceedings of ICMI First African Regional Conference* [CD]. Johannesburg: International Commission on Mathematical Instruction.

Begg, A., & Edwards, R. (1999, December). Teachers' ideas about teaching statistics. Paper presented at the *Annual Meeting of the Australian Association for Research in Education and the New Zealand Association for Research in Education.* Melbourne, Australia.

Chick, H. L., & Pierce, R. U. (2008). Teaching statistics at the primary school level: beliefs, affordances, and pedagogical content knowledge. In C. Batanero, G. Burrill, C. Reading, & A. Rossman (Eds.), *Joint ICMI/IASE Study: Teaching Statistics in School Mathematics. Challenges for Teaching and Teacher Education. Proceedings of the ICMI Study 18 and 2008 IASE Round Table Conference*. Monterrey: ICMI & IASE.

Curcio, F. R. (1989). *Developing graph comprehension.* Reston, VA: NCTM.

Einhorn, H. J., & Hogart, R. M. (1986). Judging probable cause. *Psychological Bulletin*, *99*, 3-19.

Estrada, A., & Díaz, C. (2006). Computing probabilities from two way tables: an exploratory study with future teachers. In A. Rossman & B. Chance (Eds.), *Proceedings of the Seventh International Conference on Teaching Statistics*. Salvador (Bahia, Brazil): International Association for Statistical Education.

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). Victoria, Canada: International Statistical Institute.

Fernandes, J. A., & Barros, P. M. (2005). Dificuldades de futuros professores do 1º e 2º ciclos em estocástica (Pre-service primary school teachers difficultis in stochastics). In *Actas do V Congresso Ibero-Americano de Educação Matemática* (CIBEM), Porto (Portugal): Faculdade de Ciências [CD].

Franklin, C., & Mewborn, D., 2006. The statistical education of PreK-12 teachers: A shared responsibility. In G. Burrill (Ed.), *Thinking and reasoning with data and chance − NCTM 2006 Yearbook* (pp. 335-344). Reston, VA: NCTM.

Gal, I. (2005). Towards "probability literacy" for all citizens: building blocks and instructional dilemas. In G. Jones (Ed.), *Exploring probability in schools: challenges for teaching and learning* (pp. 39-63). New York: Springer.

Hill, H. C., Ball, D. L., & Schilling, S. G. (2008). Unpacking pedagogical content knowledge: Conceptualizing and measuring teachers' topic-specific knowledge of students. *Journal for Research in Mathematics Education, 39*, 372-400.

Huerta, P. (2009). On conditional probability problem solving research – structures and contexts. *International Electronic Journal of Mathematics Education, 4*(3), 163-194.

Inhelder, B., & Piaget, J. (1955). *De la logique de l'enfant a la logique de l'adolescent* (From the logic of children to the logic of adolescent). Paris: PUF.

Jones, G. (2005). Introduction. In G. Jones (Ed.), *Exploring Probability in School: Challenges for Teaching and Learning* (pp. 1-12). New York: Springer.

Konold, C. (1991). Understanding students' beliefs about probability. In E. von Glasersfeld (Ed.), *Radical constructivism in mathematics education* (pp. 139-156). Dordrecht: Kluwer.

Lecoutre, M. P. (1992). Cognitive models and problem spaces in "purely random" situations. *Educational Studies in Mathematics*, 23, 557-568.

Shulman, L. S. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review, 57*, 1-22.

Stohl, H. (2005). Probability in teacher education and development. In G. Jones (Ed.), *Exploring probability in schools: Challenges for teaching and learning* (pp. 345-366). New York: Springer.

Tversky, A., & Kahneman, D. (1982). Judgement of and by representativeness. In D. Kahneman, P. Slovic y A. Tversky (Eds.), Judgement under uncertainty: Heuristics and biases (pp. 84 – 98). Cambridge, MA: Cambridge University Press.