THE NATURE OF PUPILS' PROBABILISTIC THINKING IN PRIMARY SCHOOLS IN CYPRUS

Iasonas Lamprianou & Thekla Afantiti Lamprianou

University of Manchester, U.K.

In this research work we explored the nature of 9-12 year old pupils' responses to probability problems. Analysis of pupils' arguments in 'Explain why' questions uncovered their thinking strategies, which we compared for pupils of different age and gender. The results revealed the existence of subjective elements and other errors in pupils' probabilistic thinking. The data were generated in year 2000 when the new mathematics books had just introduced probability extensively in the primary curriculum. Since the relevant literature in Cyprus is sparse, the results of the study form a general overview of the pupils' errors and build the basis for further in-depth and more focused research.

INTRODUCTION

Probabilistic thinking is a mode of reasoning attempting to quantify uncertainty, as a tool for decision making. A worldwide increased attention to probability and statistics is given and in Cyprus new mathematics books were written that introduced probabilities extensively in the primary curriculum. Nevertheless, pupils' probabilistic thinking is influenced by culture and is sensitive to cultural experience (Amir & Williams, 1994, 1999). There is, however, very little research in Cyprus focused on pupils' errors and probabilistic thinking (Gagatsis et al, 2001).

Research into pupils' capacity to compare two probabilities started with Piaget and Inhelder (1951). Following Piaget and Inhelder, other researchers have undertaken the study of pupils' abilities to compare probabilities (Green, 1983; Fischbein & Gazit, 1984; Canizares et al, 1997). Since comparing probabilities entails the comparison of two fractions, proportional reasoning is considered to be a basic tool of probabilistic reasoning. An important difference between comparing fractions and comparing probabilities is that the result of a proportional problem refers to a certain event, while the result of a probability problem implies a degree of uncertainty. Therefore, pupils' answers might be influenced by intuitive judgments. These intuitions are somehow some cognitive beliefs that sometimes may coincide with scientific accepted statements but some other times they may not or may contradict them. Furthermore, pupils consider subjective elements to assign probabilities.

In a previous study (Canizares et al, 1997) pupils' strategies, when comparing probabilities in tasks, were analysed and pupils' arguments were classified according to the following strategies: (a) Single variable strategies (comparing the number of possible cases; comparing the number of favourable cases and comparing the number of unfavourable cases), (b) two variables strategies (additive strategies, correspondence and multiplicative strategies) and (c) other types (equiprobability bias, outcome approach, taking decision depending on other irrelevant aspects in the task).

It has long been known that pupils' errors and misconceptions can be a starting point for effective diagnostically designed mathematics teaching (Williams & Ryan, 2000). If the teachers are aware of the most common errors and misconceptions that pupils have in probabilities, they will try to develop classroom strategies for helping students to confront them (Fischbein & Gazit, 1984; O'Connell, 1999).

This study aims to explore the probabilistic thinking of primary school pupils in Cyprus aged between 9-12 year old and to study the effect of age and gender.

METHOD AND INSTRUMENT

The research instrument was developed through several pilot steps. Informal unstructured interviews with pupils and teachers helped us to evaluate whether they interpreted the questions correctly. The pilot test was administered to pupils in Year 4 (90), Year 5 (51) and Year 6 (47).

The final version of the test consisted of nine questions clustered into three subtests. The first subtest consisted of Questions 1 and 2. Both tested the ability of the pupils to identify the most likely event from a single sample space (Single Variable questions). Question 1 was:

• Q1: There are 5 blue, 4 yellow and 3 green rubbers in a drawer. I pick a rubber randomly, without looking into the drawer. (a)What colour of rubber is more likely to pick? (b)Why?

The second subtest is cognitively more demanding than the first one. Questions 3, 4, 5 and 7 tested the ability of the pupils to define the probability of two independent events and to compare the two probabilities to decide which of the two events was the most likely.

• Q3: I have two bags with marbles. Bag A has 4 marbles, 2 blue and 2 green. Bag B has 6 marbles, 3 blue and 3 green. (a)From which bag do I have the largest probability of picking randomly, without looking in the bag, a blue marble? (b)Why?

• Q4: The yellow box contains 3 chocolate biscuits and 6 strawberry biscuits. The green box contains 1 chocolate biscuit and 2 strawberry biscuits. (a)From which box would you prefer picking randomly a biscuit in order to have a larger probability to get a chocolate biscuit? (b)Why?

Questions 6, 8 and 9 tested the reactions of the pupils to additional but irrelevant information gathered from the every day life. It was noticed from the pilot test that some pupils omitted the quantitative information they were given (numbers) and focused on irrelevant information like the size of an animal, the size of a marble and the number written on a card.

• Q6: Andrew wants to decorate his Christmas tree. On the carpet, there are 4 big and 6 small golden balls. Since he does not care if he starts the decoration with a small or with a big ball, he selects randomly a ball from the carpet. (a)What is it more likely to select firstly, a small or a big ball? (b)Why?

• Q8: In a zoo there are 2 elephants and 4 monkeys. Today the staff of the zoo wants to choose randomly an animal to wash. (a)What is it more likely to choose, an elephant or a monkey? (b)Why?

THE SAMPLE

The final instrument was administered to 426 pupils in four different district schools in Cyprus, 222 boys and 204 girls. The sample consisted of 169 pupils in Year 4, 132 pupils in Year 5 and 125 pupils in Year 6.

RESULTS & DISCUSSION

Single Variable Strategies: Pupils' correct responses can be based on intuition

In Question 1a, 381 pupils (90%) answered correctly but only 32 pupils (7.5%) explained their thinking strategy mentioning the concept of probability. Sub-questions Q1c, Q1d and Q1e asked pupils to identify the probability of drawing randomly a blue, a yellow and a green rubber. 44.8% of the sample gave a correct response to all three questions but 106 pupils (25%) failed to identify any of the probabilities correctly. 90 of them (21.2%) answered correctly that a blue rubber was more likely to be picked than a yellow or a green rubber. This result indicated that one to five pupils answered the question intuitively without being able to express in written any formal probabilistic thinking. Indeed, 87 of those 90 pupils explained that a blue rubber was more likely to be picked because "there are more blue rubbers in the drawer than rubbers of the other colours" but they failed to identify any of the probabilities of the three alternative events.

Single Variable Strategies: The effect of subjective elements

In Question 6, 69.7% of the sample answered correctly that it was more likely to pick randomly a small than a big ball. However, 72 pupils, approximately 17% of the sample, responded that it was more likely to select a big ball. Approximately 11% supported their response by saying that "it is easier to pick a large ball than a small one" or "because the small balls are harder to pick" and other similar statements.

In Question 8, 76.8% of the sample answered correctly that it was more likely for the staff to randomly select a monkey than an elephant. In total, 19% explained their thought mentioning physical and other characteristics of the two animals. For example, 15% of the pupils based their answer on the physical size of the elephant. Other wrong explanations included information about the elephants being quieter or dirtier than monkeys. Four pupils from different classes mentioned that one elephant equaled two monkeys (probably in size), therefore, two elephants equaled 4 monkeys and, consequently, the probability to select a monkey or an elephant was the same.

In Question 9 we had comparable results. The results showed that the probabilistic thinking of some pupils might be influenced by subjective elements. The percentages of those pupils according to the three questions above might lie in the area of 11% (Q6), 19% (Q8) and 10% (Q9). These percentages may be significantly underestimated because sometimes it was difficult to distinguish from the written explanation whether pupils' thought was influenced by subjective elements or not.

Two Variables Strategies: Comparisons between favourable or unfavourable events

In Question 3, only 21.1% of the pupils realized that the probability of getting a blue marble was the same in both bags. 257 (60%) of the pupils declared that they would have a larger probability of picking a blue marble from bag B. 246 of them (57.8%) explained their response in Question 3b by saying "bag B contains more blue marbles than bag A". From the responses of the pupils, it seems that they mainly compared the number of blue marbles in the two bags instead of the probability of getting a blue marble from each bag.

106 pupils (24.9%) gave correct responses to Q3c and Q3d that asked for the probabilities of getting a blue marble out of each of the two bags (they were ½ in both cases) but also claimed that it was more likely to get a blue marble out of bag B. Most of them explained that there were more blue marbles in bag B indicating that although these pupils knew how to find the relevant probabilities of the two events, they still used their own intuitive rules to take the decision of the most likely event.

In Question 4 we had similar results. Only 18.1% of the pupils realized that both boxes contained the same proportion of chocolate biscuits. 232

(54.5% of the sample) responded that the yellow box would give a larger probability to pick a chocolate biscuit. 218 of them (94%) explained their response by saying "there are more chocolate biscuits in the yellow rather than in the green box". Again, it seems that they focused on the number of chocolate biscuits in the two boxes instead of the proportion of chocolate biscuits in each box. 69 pupils (16.2% of the sample) gave correct responses to Q4c and Q4d that asked for the probabilities to get a chocolate biscuit out of each of the two boxes (1/3 in both cases) but used their own intuitive rule to answer incorrectly Question 4a. Questions 5 and 7 gave similar results with Questions 3 and 4.

Construction of an 'ability' measure

Eight of the nine questions of the test (excluding question 2) ask pupils to indicate the most probable between two events (e.g. to draw a blue marble from bag A or bag B) and to explain verbally their thinking strategy. A correct mark was awarded for a correct response to the question and one additional mark was awarded to the pupils whose explanation indicated a 'correct' probabilistic thinking strategy.

A Partial Credit Rasch model (Wright & Masters, 1982) was fitted on the data, using Quest. The mean difficulty for the questions was constrained to zero (SD=1.18 logits) and the reliability of question difficulty estimates was 0.96. The mean of the pupils' ability was -0.18 logits (SD=0.98 logits) and the reliability of pupil estimates was 0.66. All questions had a satisfactory fit. The test and the sample may be interpreted as falling into a hierarchy of three levels (Figure 1).

At level A (-2.5 to –1.5 logits) pupils can succeed on answering correctly questions that tested for the identification of the most likely outcome from a single sample space (Single Variable problem). At level B (-1.5 to 0 logits) pupils can succeed on answering correctly questions that tested for the identification of the most likely outcome from a single sample space (Single Variable problem) without being carried away from subjective elements (e.g. elephants are quieter than monkeys). At level C (0 to 2.5 logits) the pupils can succeed on answering correctly even harder questions, which ask them to compare the probabilities of two independent events and decide which of the two is the most likely to happen (Two Variables problems). Pupils at higher level can also answer correctly the questions of the previous levels.

According to Figure 1, a percentage of pupils was influenced by the 'Subjective Elements' effect (from 10% to 19%). In general, this type of error is more representative of pupils of low ability. The mean ability of the pupils who made the errors associated with the 'two variables' is near to the mean ability of the sample. The percentage of pupils making this error is relatively large (47% to 66%).

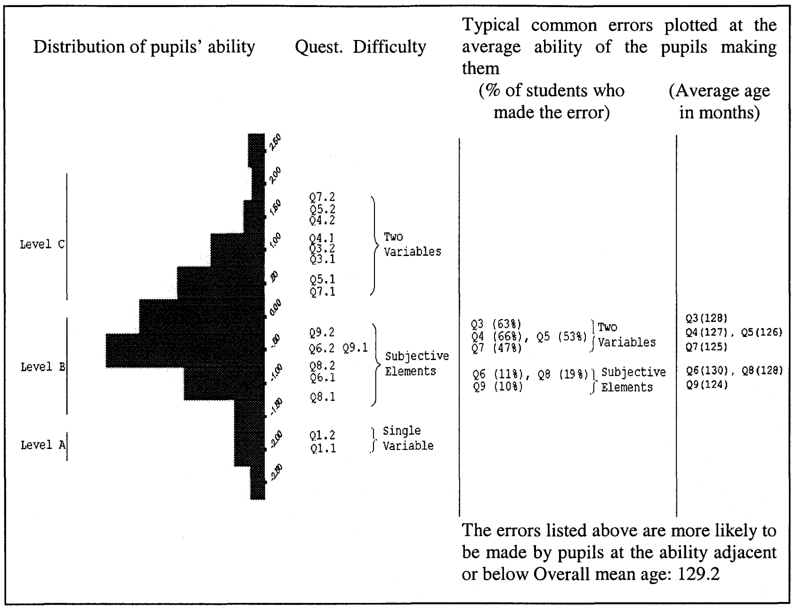

Figure 1: A 'probability performance' measure

The pupils in Year 4 are significantly less able than pupils in Year 5 and Year 6 ($F=17.35$, $df=2$, $p<0.01$). The difference in the average ability of the three year groups (average ability: Year 4= -0.52; Year 5= 0.03; Year 6= 0.12) was significant only between pupils in Year 4 and in Year 6.

Construction of the 'likelihood ability' measure

Some of the questions of the test were followed by sub-questions, which asked the pupils to identify the likelihood of certain events. A mark was awarded for each correct response and the simple Rasch model was fitted on the data, using Quest. The mean difficulty for the 13 questions was constrained to zero (SD=1.2) and the reliability of question estimates was 0.98. The mean ability for the pupils was –0.05 (SD=2.03) and the reliability of ability estimates was 0.83. The fit statistics were satisfactory except for one question, which was only marginally satisfactory.

The correlation between the 'likelihood ability' (ability to quantify the likelihood of events) and the 'ability' (performance to the test) of the pupils is 0.58, which is not as high as expected. Indeed, this showed that it was not necessary for the pupils to be able to quantify the probabilities of the events in order to give correct responses to the questions of the test.

It is possible that the pupils were able to answer the questions intuitively and the ability to numerically identify the likelihood of the events was not a prerequisite.

Predictors of pupils' 'ability'

The test included 8 questions, which asked the pupils to compare two fractions, similar to the ones used in the probability questions. This subtest was administered only to 223 pupils. A mark was awarded for each correct response and the distribution or pupils' scores had a mean of 4.7 and SD=2.5. The correlation between the 'ability' measure and the raw score in the fractions sub-test is 0.37; not as high as might be expected.

A multiple regression model was attempted using pupils' 'ability' as a depended variable and the 'fractions raw score', the 'likelihood ability', the gender and the age as predictors. The stepwise procedure accepted in the model only two terms: the 'fractions raw score' and the 'likelihood ability' (R2=0.38). When the 'fractions raw score' was not included in the predictors, the stepwise procedure kept only the predictor 'likelihood ability' in the model with almost no loss of information (R2=0.33). Gender and age were not good predictors of 'ability'.

CONCLUSION

We have managed to develop two scales describing pupils' responses to the instrument, which is revealing about their probabilistic thinking, especially as regards their inappropriate use of intuitions. We have further identified from pupils' responses that some pupils were influenced by other irrelevant aspects in the tasks (subjective elements). The knowledge that teachers would collect from these scales would enrich teachers' mental models of their learners and would help them improve their classroom practice.

Having collected these data, however, we should further continue interviewing pupils in order to shed more light into pupils' probabilistic thinking. Further research into pupils' probabilistic reasoning would be an essential step for selecting teaching and assessment situations. We will be studying this aspect in the next stage of the work.

REFERENCES

Amir, G. & Williams, J. (1994). *The influence of children's culture on their probabilistic thinking*. In J.P. Ponte & J.F. Matos (Eds.). *Proceedings of the 18th Conference of the International Group for the Psychology of Mathematics Education (PME)*. University of Lisbon.

Amir, G.S. & Williams, J.S. (1999). Cultural Influences on Children's Probabilistic Thinking. *Journal of Mathematical Behaviour,* 18(1), 85-107.

Canizares, M.J., Batareno, C., Serrano, L. & Ortiz, J.J. (1997). Subjective elements in children's comparison of probabilities. In E. Pehkonen (Ed.). *Proceedings of the 21st Conference of the International Group for the Psychology of Mathematics Education (PME).* University of Helsinki.

Fischbein, E., & Gazit, A. (1984). Does the teaching of probability improve probabilistic intuitions? *Educational Studies in Mathematics,* 15, 1-24.

Gagatsis, A., Kyriakides, L. & Panaoura, A. (2001). Construct validity of a developmental assessment on probabilities: A Rasch measurement model analysis. In M. van den Heuvel-Panhuizen (Ed.). *Proceedings of the 25th Conference of the International Group for the Psychology of Mathematics Education (PME).* Utrecht University.

Green, D.R. (1983). A survey of probability concepts in 3000 pupils aged 11-16 years. In D.R. Grey, P. Holmes, V. Barnett & G.M. Constable (Eds.). *Proceedings of the First International Conference on Teaching Statistics* (v.2, 766-783). University of Sheffield.

O'Connell, A.A. (1999). Understanding the nature of errors in probability problem-solving. *Educational Research and Evaluation,* 5(1), 1-21.

Williams, J., & Ryan, J. (2000). National Testing and the Improvement of Classroom Teaching: can they coexist? *British Educational Research Journal, 26*(1), 49-73.

Wright, B.D., & Masters, G.N. (1982). Rating Scale Analysis: Rasch Measurement. Chicago: MESA Press.