A Four Phase Model for Predicting the Probabilistic Situation of Compound Events

Irma Jan and Miriam Amit
Ben-Gurion University, Israel janirma@gmail.com

Abstract

This paper presents an innovative construction of a probabilistic model for predicting chance situations. It describes the construction of a four phase model, derived from an intense qualitative analysis of the written responses of 94 mathematically talented middle school students to the probabilistic compound event problem: "How many doubles are expected when rolling two dice fifty times?" We found that the students' comprehension process of compound event situations can be broken down into a four phase model: beliefs, subjective estimations, chance estimations and probabilistic calculations. The paper focuses on the development of the model over the course of the experiment, identifying the process the students underwent as they attempted to answer the question. We explain each phase as it was reflected in the students' rationalizations. All phases, including their definitions and students' citations, will be presented in the paper. While not every student necessarily goes through all four phases, an awareness and understanding of them all allows for efficient, effective intervention during the learning process. We found that guidance and learning intervention helped shorten the preliminary phases, leading to more relative time spent on probabilistic calculations.

Theoretical Background

An important distinction on which this study rests is the separation between *simple events* such as one-dimensional experiments, and the more complex *compound events*, which include two-dimensional experiments and events such as $A \cup B$, $A \cap B$ (Polaki, 2005). The transition from simple events to compound events has been found to be difficult for students (Watson, 2005). For example, to understand compound events, such as the sums of two dice, one must be able to generate complete sets of outcomes for the events and use sample space symmetry and composition to make probability predictions (Polaki, 2005). Understanding sample space is complex in itself, since it requires the coordination of several cognitive skills: (a) recognizing different possible ways of obtaining an outcome, (b) being able to systematically generate those possibilities, and (c) being able to "map the sample space onto the distribution of outcomes (Horvath & Lehrer, 1998). Fischbein, Nello & Marino (1991) claimed that some failures in constructing the sample space of rolling two dice are due to the fact that there seems to be no natural intuition regarding the order of two dice. This paper advances research on probabilistic thinking by examining students' probabilistic thinking about compound events in depth. We present here a four phase model derived from our test subjects' reasoning process throughout the research, regarding the prediction of compound events that involve the throwing of two dice.

Methodology

Setting

Six groups of gifted and talented students from grades 6 to 7 (a total of 94), members of the "Kidumatica"[1] math club (Amit, Fried & Abu-Naja, 2007), participated in an extensive study that aimed to investigate how a teaching intervention, in a dynamic semi-structured learning environment, contributes to students' development and understanding of key concepts in probability. The learning intervention, which was based on the constructivist approach (Carpenter & Lehrer, 1999), was conducted during twelve sessions (75 minutes each) with a chain of consecutive probability assignments in the form of tasks, tailor made to prompt understanding and discussion. This study is an extension of our preceding studies (Amit & Jan,

[1] Ben Gurion University's "Kidumatica" math club was established in 1998 by one of the authors of this paper. Its goal is to create an after-school program in which students from 5-10 grades could develop their interest in mathematics and their mathematical thinking. Students participate in several "mini-courses," held weekly at the university campus. The mini-courses include: "Logical Problems," "Real-Life Mathematics," "Mathematical Games," "Number Theory," "Number Sequences," "Fractals," etc .

2006/2007; Jan & Amit, 2006; Jan & Amit, 2008) and aims, among other things, to create an infrastructure to build upon in a future, more formal approach to probability.

Source of data - The compound event question

The students were given a pre test questionnaire, to determine their initial perceptions concerning chance and probability. A post test was then given at the end of the study, to illuminate the changes in the students' perceptions. Both tests included a question concerning the theoretical probability of a compound event (see below). In particular, students were asked to predict the outcome of a compound event situation. (Note: The question in the pre-test involves 50 throws, while that in the post-test involves only 30.)

When playing a game with two dice, a player gets an extra turn when a 'double' is rolled, i.e. when he gets the same number on the two dice (for example, 2 on one die and 2 on the other). Explain and justify: how many doubles would you expect in fifty rolls of two dice? (Jones, Thornton, Langrall & Tarr, 1999)

The question presents the student with a compound event. The correct solution requires building a sample space for the simple event (rolling one die) and transferring it in order to generate the complete sample space of the compound event (rolling two dice).

In addition to its mathematical complexity, the question raises several affective/psychological difficulties: (1) it is difficult to accept that random situations can be predicted mathematically ;(2) It is not obvious that outcomes of random situations can't be controlled, do not depend on subjective judgment, on individual's luck etc.

Findings

Students' written responses were analysed quantitatively and qualitatively. A qualitative analysis was executed according to four categories, which were identified as reflective of students' probabilistic reasoning: (a) types of strategies; (b) representation; (c) use of probabilistic language; (d) the nature of cognitive obstacle. In this paper we address the first category: *the development of students' strategies in their justifications.*

A Four Phase Model of Strategies

Intense analysis of the data provided a four phase model of strategies of justification: (a) beliefs as a source for justification; (b) subjective estimations of the compound event; (c) chance estimations; (d) theoretical calculations of the compound event.

Phase I-Belief Strategies

In the pre test 44.7% of the students used their everyday beliefs about chance and stated that it is impossible to predict the number of doubles. Examples are illustrated below:

> Tal: "*It is impossible to know the number of doubles since each time it will get a different number of doubles. We are betting, and in games of chance we do not know what would be".*
>
> Ben: *"Maybe the player had no luck and didn't get any doubles but may be he was lucky. It can be said that the chances for doubles are 50%-50%".*
>
> Nurit: "*Doubles aren't rare events but are difficult to get. Hence, it is impossible to know how many doubles to expect. It is possible to get 50 doubles and it is reasonable not to get any double. It can be assumed that in most of the throwing there would be no doubles".*

For these students it was impossible to predict a chance situation. Their explanations came from different perspectives: (1) Tal and Ben believed that chance situations relate to luck and luck isn't something that can be measured or controlled ;(2) From their experiences with board games, it is hard to get a double though it is still possible. This perspective is expressed in Nurit's justification: "*Doubles aren't rare events but are difficult to get*"; (3) Ben's justification suggests another strategy: chance situation either happen or not, and therefore they have a 50-50 chance. This strategy can be explained according to Konold's (1989) *outcome approach*: when children are required to make a prediction about a chance situation, they will respond that it is impossible to say "It's just a matter of chance". It can be expressed also as an incorrect use of language: the tendency to think in phrases like *50-50 chance* to describe unknown events (Amir & Williams, 1999). Though in the post test only a minor change occurred on the *belief strategy* (percentage of students decreased to 28.7%), still one significant change occurred. Students cease to use the word "*luck*" or phrases like: *"the chances for double are 50%-50%"* instead, they explain that since the chance of getting a double is low, it is impossible to predict the number of doubles in several trials of rolling dice. This approach indicates that a new comprehension was formed.

Phase II-Subjective Estimations

We assigned the term *subjective estimation* to all of the students' attempts to estimate the number of doubles using mathematical procedures without probabilistic consideration. The relative amount of subjective estimation used by the students remained constant (31.9% used this strategy in the pre test; 30.9% continued using it in the post test).

Examples are illustrated below:

> Omer: *"there are 50 dice rolls. If all are doubles then, there are 50* (doubles) *and if there are no doubles then there are 0* (doubles)*. I chose the number in the middle".*
>
> Ron: *"since there are six possibilities for getting a double and* $6 \times 8 = 48$*, this is the closest number to 50, therefore in 50 rolls of two dice there will be 8 doubles".*

These examples are evidence that students had no previous probabilistic knowledge connected to the question, and therefore searched for arithmetical and logical solutions. According to Omer, since the total number of throws is fifty, the number of doubles will be between 50 and 0 therefore the logical answer for the number of doubles was the number in the middle. Omer was trying to discern impossible totals from possible ones (cf. Fischbein et al., 1991), to accomplish that, he used a strategy that took into consideration the extremes of possible outcomes (Nilsson, 2007). Ron used the available numbers in the question (50 times rolling the dice and 6 sides to one die) and upon it built an equation. Taking available numbers and performing mathematical manipulations on them is a well known approach in the theoretical literature; having already been mentioned by Polya (1957).

> Liron (post test): "*there are 5 doubles to be expected because in a die there are 6 numbers and therefore by my estimation, every 6 rolls we will get one double".*
>
> Ariel (post test): *"in my opinion by rolling two dice 30 times, 5 doubles are to be expected. I calculated it in this way: "I divided the 30 rolls (which I have to roll) by 6 possibilities in each die and got 5. So, I can get 5 doubles in the game. Sometimes maybe more and sometimes maybe less. This is more or less the number according to calculations. Eventually everything is a matter of luck".*

The post test examples (see above, Liron's and Ariel's citations) show some change in students' thinking towards probabilistic thinking. Liron and Ariel are beginning to pay attention to all the possible ways to get double, but are still seeking for arithmetical solutions.

Phase III-Chance Estimations

11.7% of students in the pre test reasoned that to predict the number of doubles they had to connect the question to the probability of getting a double in one throw, i.e. to simplify the situation into a *simple event* and then transfer it to the *compound event*. They therefore used a strategy we referred to as *chance estimation*. These students built a method (though they used wrong estimations, the building process is important): (1) first they simplified the question and focused on one trial; (2) then they estimated subjectively the percentage of getting a double in one trial; (3) finally, to find the number of doubles in the compound situation (of several trials) they calculated the percentage out of the given trials. The percentage of students using the *chance estimations strategy* in the post test was very low (4.3%) since they progressed towards the formal theoretical probability predictions. Examples are illustrated below:

> Iaron: *In most of the trials we won't get a double therefore, 20% that we get a double and 80% that we get no double. 20% of 50 are 10 doubles".*
>
> Galit (post test): *"in each roll there is 10% of chance to get a double. 10% out of 30 are 3 therefore there will be at least 3 double in rolling two dice 30 times".*

Phase IV- Probabilistic Calculations

In the pre test, only two (out of 94) students calculated the probability of "*receive a double in rolling two dice*" and connected it to the number of doubles to be expected in 50 trails. In the post test 19.1% of the students made progress towards finding the number of doubles to be expected (theoretically) according to the following strategy: (1) they generated a complete sample space of rolling two dice; (2) quantified the probability of getting a double in one trial; (3) multiplied it by the number of trails and got the number of doubles (theoretically) in several trails. This strategy is illustrated in Sharon's solution to the question (Figure 1): First Sharon generates a complete set of outcomes in rolling two dice, finds the sample space and writes a total of 36 possible outcomes. Then, he finds how many doubles are possible and writes 6 possible doubles. In the left side Sharon writes his justification in Hebrew. He builds a fraction [$6/36$] that represents the probability of getting a double as a ratio between the 6 possibilities to get double and the sample space (36). He reduces the fraction and receives [$1/6$].Sharon accomplishes his strategy in following manner: since 6 out of 36 is equal to 5 out of 30 he concludes that in 30 rolling of two dice there will be expected 5 doubles.

"If we reduce the possibilities to 30 then the number of doubles will be 5 since we have to keep the same ratio between the numerator and the denominator as in the first calculation".	

Figure 1-Sharon's solution

Conclusions and Discussion

The four phase model that expresses students' typical strategies in understanding compound events is summarised in Table 1:

	The Four Phase Model for Predicting Compound Situations			
Definition:	A compound event is a two dimensional random experiment.			
Compound situation	When playing a game with two dice, a player gets an extra turn when a 'double' is rolled. How many doubles would you expect in fifty rolls of two dice?			
The four phases	**Belief strategy** ⇒	**Subjective estimations** ⇒	**Chance estimations** ⇒	**Probabilistic calculations**
Phase definition	Impossible to predict chance situations.	Estimating probability of compound situation using mathematical procedures without probabilistic consideration.	Simplifying compound situation and estimating the chance of its occurrence.	Using formal probability quantifications to calculate the probability of compound situations.
Students' citation	*"I don't know the number of doubles; it is a mater of luck and a gamble".*	*"There are 50 dice rolls. If all are doubles, then there are 50 (doubles) and if there are no doubles then there are 0. I chose the number in the middle".*	*"In each roll there is 10% chance for a double.10% out of 50 are 5. There will be at least 5 doubles in 50 rolls".*	*"In rolling two dice there are 36 possible outcomes, 6 of them are double, meaning, 6/36=1/6. 1/6 out of 30 are 5. this is the expected number of doubles".*

It is not necessary for students to cross all four phases, but we must be aware of their existence and address them in the learning intervention. We contend that extending the duration of the learning intervention will proportionally diminish the presence of *belief* and the *subjective estimation* strategies in favour of the *probabilistic calculation* strategy for compound situations, with the latter displacing the former more prominently as the time allotted for the intervention grows.

References

Amir, G. S., & Williams, J. S. (1999). Cultural influences on children's probabilistic thinking. *Journal of Mathematical Behavior, 18*(1), 85-107.

Amit, M., & Jan, I. (2006). Autodidactic learning of probabilistic concepts through games. *Proceedings of the 30th Conference of the International Group for the Psychology of Mathematics Education,* Vol. 2, pp. 49-56. Charles University, Prague: PME.

Amit, M., & Jan, I. (2007). A learning environment that promotes concept construction: The case of probability, annual conference of the American Educational Research Association, Chicago, IL .Chicago.

Amit, M., Fried, M., & Abu-Naja, M. (2007). The mathematics club for excellent students as common ground for Bedouin and other Israeli youth. In: B. Sriraman (Ed.), International perspectives on social justice in mathematics education, monograph 1, the montana mathematics enthusiast (pp. 75–90). The University of Montana Press: USA.

Carpenter, T. P. & Lehrer, R. (1999). Teaching and learning mathematics with understanding. In E. Fennema, & T. Romberg (Eds.), Mathematics classrooms that promote understanding (pp. 19–32). Mahwah, NJ: Lawrence Erlbaum Associates.

Fischbein, E., Nello, M. S., & Marino, M. S. (1991). Factors affecting probability judgments in children in adolescents. *Educational Studies in Mathematics, 22,* 523-549.

Horvath, J. K., & Lehrer, R. (1998). A model-based perspective on the development of children's understanding of chance and uncertainty. In S. P. Lajoie (Ed.), Reflections in statistics: Learning, teaching, and assessment in Grades K-12 (pp. 121 -148). Mahwah, NJ: Erlbaum.

Jones, G.A., Thornton, C.A., Langrall, C.W., & Tarr, J.E. (1999). Understanding students' probabilistic reasoning. In: L.V. Stiff, & F.R. Curcio (Eds.), *Developing Mathematical Reasoning in Grades K-12, National Council of Teachers of Mathematics' 1999 Yearbook* (pp. 146-155). Reston VA: The Council.

Jan, I., & Amit, M. (2006). Game-task activities as initiators of learning basic concepts in probability. CIEAEM 58 Congress: Changes in society: a challenge for mathematics education, pp. 205-208, CIEAEM, Srni, Czech Republic.

Jan, I., & Amit, M. (2008). Overcoming a misleading intuition: the relation between ratio and sample size. *Proceedings of the joint meeting of the 32nd Conference of the International Group for the Psychology of Mathematics Education and the XXX North American Chapter,* Vol. 3, pp. 441-448. Morelia, Michoacán, México: PME.

Konold, C. (1989). Informal conceptions of probability. *Cognition and Instruction, 6*(1), 59-98.

Nilsson, P. (2007). Different ways in which students handle chance encounters in the explorative settings of a dice game. *Educational Studies in Mathematics, 66*, 273-292.

Polaki, M. V. (2005). Dealing with compound events. In G. A. Jones (Ed.), *Exploring probability in school: Challenges for teaching and learning* (pp. 191-214). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Polya, George, How to Solve It. 2d ed. Princeton, N.J.: Princeton University Press, 1957.

Watson, J. M. (2005). The probabilistic reasoning of middle school students. In G. A. Jones (Ed.), *Exploring probability in school: Challenges for teaching and learning* (pp. 145-169). Dordrecht, The Netherlands: Kluwer Academic Publishers.