

LEARNERS' USE OF PROBABILITY MODELS IN ANSWERING PROBABILITY TASKS IN SOUTH AFRICA

Lydia Mutara and Judah Makonye

University of the Witwatersrand, South Africa

INTRODUCTION

After attaining democracy in 1994 South Africa has had a number of educational reforms particularly in the mathematics curriculum with the probability topic being introduced as compulsory for the first time from grade 10 to 12 in 2012. An immediate problem with this re-introduction of probability was the danger that some teachers had incoherent probability content knowledge (see Chernoff, 2012). It can be argued that in general, most mathematics teachers in South Africa encounter many challenges in teaching this topic. Further, there were not sufficient teaching and learning support materials for dealing with this new topic. Given this scenario, in an attempt to understand challenges embedded in the teaching and learning of Grade 10 probability, we explored learners' errors and misconceptions allied to their solutions of probability tasks with different representations. As Polya (2014) has envisaged, to solve a problem, the first thing is to understand the problem and the second stage is to devise a plan. One of Polya's plans is to draw a picture, that is a visual or make a model. Probability problems are often abstract to learners that is why making visual representations is one of the best ways for students to make sense of the problems and as way to a correct solution.

The assumption is that analysis of learner errors in their representations would help stoke teacher probability pedagogical knowledge. This view is in agreement with the National Council of Teachers of Mathematics (NCTM) (2014) call that eliciting learner thinking on certain mathematical ideas informs instruction and is good practice. Aichele & Oxford (1994) also urged for the analysis of learner cognitive difficulties as a way to grow professional knowledge in teaching probability. This research aimed to explore the errors and misconceptions associated with grade 10 learners' different representations in solving probability problems.

PROBABILITY AS A SCHOOL MATHEMATICS TOPIC IN SOUTH AFRICA

In the South African Curriculum Assessment Policy Statements (CAPS) (Department of Basic Education (DBE), 2011), the topic of probability involves learning how to calculate the chances of events happening. The topic is sixth in importance in the Further Education and Training (FET) mathematics curriculum. The content taught embraces theoretical and experimental probability, dependent and independent events, simple and compound events and generalisation of the fundamental counting principle.

The weighting for probability is about 18% in Grade 10, 11 and 12. The weighting of this mathematics topic shows its importance. The curriculum expects learners to be able to "identify and solve problems and make decisions using critical and creative thinking" (DBE, 2011; p. 5). They are also expected to communicate effectively using visual and symbolic skills in various modes. This involves representing the sample spaces by use of Venn diagrams, contingency tables, tree

diagrams, two-way tables and outcome listings. How learners evoke these representations and the errors that occur as they solve Grade 10 curriculum based probability tasks is the target of this research.

LITERATURE REVIEW

Batanero & Díaz (2012) and Batanero (2005) outline the evolving conception of probability since the 1600s; from Pascal's objective probability versus probability as intuitive belief. But what is important in this study is the Laplacian view of probability as the calculation of desirable events using equi-probability points in a sample space. Probability and statistics is harder to understand than ordinary mathematics because it needs much more interpretation and understanding of context (Batanero & Díaz, 2012, p.5). The teaching and learning of probability is particularly challenging because it is an abstract concept. For this reason, it is important to help learners understand and employ different representations to mediate learning probability (Batanero & Díaz, 2012). Such representations could be tree diagrams, Venn diagrams, contingency tables, two-way tables or outcome listings. Also Batanero and Díaz (2012) argue 'in arithmetic or geometry an elementary operation can be reversed and this reversibility can be represented with concrete materials' (p.3), results of random experiments cannot be concretely reversed to help children formulate abstract concepts about probability. Batanero and Díaz argue that only later would representations like tree diagrams help learners understand how probabilities are calculated abstractly. Thus the use of probability representations and models are imperative in teaching and learning calculation of probabilities to beginners, as they are the closest representations to the concrete mode.

The community of mathematics and statistics educators (see Freudenthal, 1973; Shaughnessy, 1992) concluded that modelling probabilistic situations is complex for learners and is often hindered by the learners' wrong intuitions, biases and primitive conceptions to mention a few.

The *representativeness misconception* refers to estimation of uncertain events based on sample reflection of population events (Hirsch & O'Donnell, 2001). For example, learners with this misconception will think that if a fair coin is tossed six times, the probability of getting an ordered sequence HHHHHH is less likely than that of getting HTHTHT. Fischbein (1999) gave the following problem to learners:

The likelihood of getting heads at least twice when tossing three coins is smaller than/equal to/greater than the likelihood of getting heads at least 200 times out of 300 times.

The learners were to choose the correct answer. Fischbein (1999) found that most of the learners said that the probabilities are equal which is wrong. Most learners used the equality of ratios to justify their answers.

Another misconception is the *recency effect* which refers to the effect that recent events have on later events which stems from improper distinction between independent and dependent events (Fischbein, 1999). Thus learners thought that; if a couple already have two sons and are considering a third child, they are more likely to have a son than a daughter next time. Most learners fail to

realise that events such as gender of a baby are independent events and consequently the outcome of one event does not influence the outcome of the next.

Several researches (for example Suh, 2007) have noted that learners make more meaningful connections between new information and previously acquired knowledge when representing a mathematical idea in multiple modes. On the other hand, Zahner and Corter (2010) argue that external visual representations can facilitate probability problem solving at the stage of finding a solution strategy if an appropriate representation is chosen.

THEORETICAL FRAMEWORK

In mathematics education, it is often helpful to establish the breadth and depth of learners' understanding of mathematical concepts through error analysis and diagnosis (EAD), so as to commence teaching at a learners' conceptual level. EAD aims to analyse, expose and interpret error patterns for the purpose of modifying instruction (Borasi, 1994) in reaction to them. By assessing and evaluating the learners' independent work to identify specific error types, error analysis informs priorities for teaching (Makonye, 2012). As learner errors and misconceptions are often consistent rather than random, this suggests launching of necessary instructional approaches that can begin to resolve them. One such approach is didactical engineering (DE). Didactic engineering (Artigue, 1989; Chevallard, 1992) is a research approach in mathematics education. This approach compares a mathematics teacher's work to that of an engineer, in that though an engineer does a project overly guided by standard scientific engineering knowledge; when he/she actually carries out the project he/she is compelled to solve complex problems that science is not yet able to solve or "cannot yet take over" (Artigue, 1989, p. 283; Chevallard, 1992). Thus while mathematics educators may have educational theoretical knowledge of teaching and learning, these are general guiding theories. In the end, when teaching mathematics in the classroom, the teacher must deal with practical problems in the classroom, problems that educational theory has no clear cut solutions. Didactic engineering is concerned with educational classroom interventions, usually a series of lessons; informed by trusted approaches (Artigue, 2011, p. 20). One such approach that comes to mind is Realistic Mathematics Education (Freudenthal, 1973). DE phases involve analysis of the prior situation and analysis of what intervention may be done to address an identified problem. The next phases are implementation of intervention, and a posteriori analysis and evaluation. This is done to check if the praxis, the fusion of theory and practice has helped to alleviate the teaching and learning problem. In this study learners were asked to solve probability tasks. Their solutions were analysed in relation to the representations they used to solve these tasks with an eye to understand the errors and misconceptions in their thinking. These two stages are in effect the introductory phases of DE; in readiness of a teaching intervention. If the interventions were to be implemented they could be influenced by Borasi (1994) who view learners' errors as springboards to effective teaching and learning on related concepts. The errors and misconceptions that learners are pre-disposed to in a certain topic, and how they can be handled to help learners' understand becomes part of a teachers' pedagogical content knowledge (Shulman, 1986) for that topic.

METHODOLOGY

This study was carried out on a class of sixteen grade 10 learners in a Johannesburg, Soweto school, South Africa. First, grade 10 level probability tasks emanating from the CAPS curriculum incorporating aspects of representations were validated by other three grade 10 mathematics teachers and then piloted on five grade 10 learners at a different school in Soweto. When the tasks were deemed appropriate, they were given to participants emphasising the need to for them to show any diagrams, models and representations that would help them to answer each questions. Learners’ scripts were scrutinized for errors and misconceptions with respect to representations used to help answer the questions. Second, six learners were selected for interviews based on the type of errors noted in their scripts. This was done to obtain an understanding of the learners’ thought processes in representations as they engaged with the tasks and elicit an insight into their thinking. All research ethical protocols were adhered to. For example, pseudonyms are used in reporting to maintain anonymity of participants.

DATA ANALYSIS

The data was mainly analysed qualitatively and also quantitatively. The quantitative analysis looked at the frequency of the type of representations used and the number of correct and incorrect representations, and answers. Qualitative analysis particularly dealt with what thinking the learners had in interpreting the questions, what representations were constructed to make sense of the questions and how they interpreted back the representations to answer the questions. It is important to note that at this level the frequentist Laplacian view of probability is dealt with (Bernabeu, 2005).

The first task analysed was; *What is the probability of getting two 3s if a die is tossed two times? Show all your working in a diagram.* All learners encountered difficulties in constructing a diagram for this item. No one drew a two-way table for this task. Most learners failed to realise that the tree diagram ought to have six branches. Two learners who got the correct answer for item drew incorrect tree diagrams. Figure 1 shows Prisca’s tree diagram which was a typical example of the tree diagrams drawn by other learners. Prisca’s diagram was so tiny that the authors decided to re-draw it electronically as shown in Figure 1.

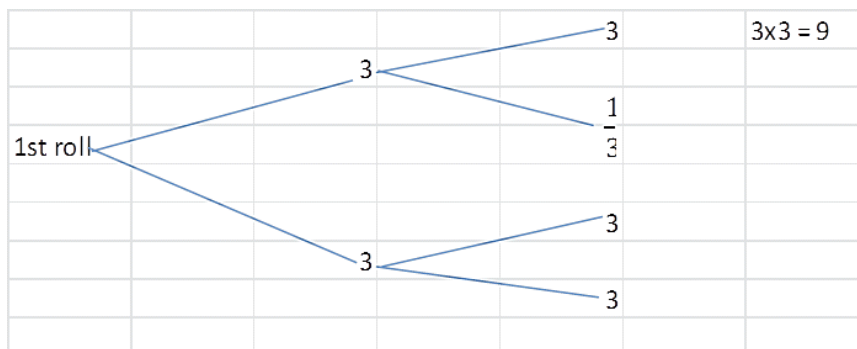


Figure 1: Prisca’s tree diagram in finding the probability of getting two 3s when a dice is tossed twice

At least five errors are evident in Prisca's representation. First, the tree diagram has two branches at each level (toss) instead of six. This signifies a fixation to a familiar type of representation, of tossing a coin that has two outcomes, a head and a tail so two branches. Brodie and Berger (2010) regarded this type of error as an error of familiarity, of generalising the known into the unknown. Second, 3 is assigned as an outcome in each case, the other five possible numbers of the die ignored. This is a case of selective processing of information and ignoring other attributes of the problem (Donaldson, 1963). Donaldson referred to this kind of error as an arbitrary error. Thirdly in the learner determining the probability of getting two 3s, 3 is appearing on each branch in the tree diagram which influenced her to write $3 \times 3 = 9$. If she had noted that the probability of getting each 3 in each toss is $\frac{1}{6}$ and multiplied those she would have been fine. Multiplying the probabilities is correct as they are independent events. Fourthly, the fact that calculation resulted to a probability greater than 1 is an error in itself (Prisca abandoned use of tree diagrams in subsequent tasks - she must have felt that this method of tree diagrams is unhelpful to her). The fifth error was the probability of $\frac{1}{3}$ shown on Fig. 1. It looked like the number 3 was so eminent in Prisca's thinking and overshadowed other possibilities 1, 2, 4, 5, and 6. The idea of considering and representing the sample space first before calculation of the required probability seemed to be inconsequential to her. The tree diagram was *one* of the correct representations, though not the most economic for this problem. We realize that often a representation can become a pitfall if it is not correctly constructed and interpreted as in Prisca's case.

The second task analysed in the following was on *a new family that decided to have 3 boy and 1 girl children*. The first part required students to *list all possibilities (the sample space) of having four children and state the number of ways of having 4 children*. The second part required them to *list the favourable outcomes for the family from all the possibilities*. Only 4 learners out of 16 successfully listed the possible outcomes for 3 boys and 1 girl. Eight learners listed less than 3 outcomes and 4 learners listed 3 outcomes. This error could have been due to inappropriate analysis in that out of four children, a girl could be born first, second, third or fourth.

Tendai's work was different and we became interested in knowing how she came up with her responses. Her response to this part item was in all there were 17 possible outcomes of having four children. In making the list she did it well but also listed GGG, which was the source of her error. When asked number of ways there could be 3 boys and 1 girl, she said there were two ways GBBB and BBBG conveniently ignoring other options BGBB and BBGB she had listed earlier. We regard this as a careless error because when we asked her about it in an interview she quickly corrected herself to say there were four favourable outcomes out of sixteen. Hence her errors were due to inappropriate visual scanning of the good representation she had on the task. Prisca was another learner who listed more than 16 possible outcomes, but in an interview she corrected herself to say 16 after some probing.

The third item required them to answer to: *Find the probability of finding 1 head, 2 heads, and 3 heads when a coin is tossed 3 times*. Thulani wrote:

- a) H b) HH c) HHH.

This answer was partially correct in that it showed the required outcomes in correct sequence; they were missing the Ts however. Further, the student did not pass through the threshold to determine the numerical probability which were required.

In dealing with the fourth task; *What is the probability of getting a sum of 3 if two dice are tossed?* Mandi provided a representation below (Fig. 2) to help answer the question.

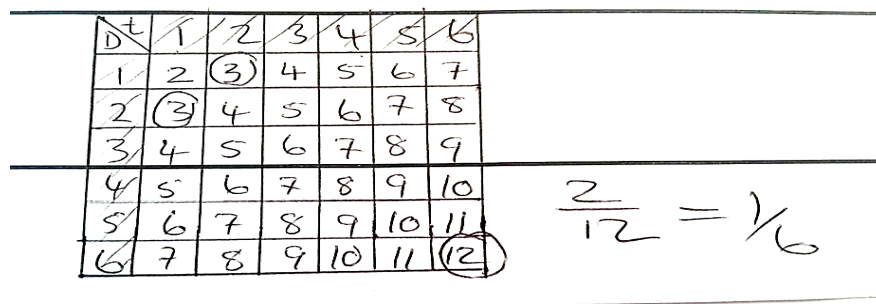


Figure 2: Mandi’s representation of getting sum of 3 when two dice are tossed

While Mandi had a helpful two-way representation, she could not interpret it correctly; that there were 36 sample points instead of 12. This again seemed to be a careless error.

In other cases, most learners presented outcome listings suggesting that they were conversant with that type of representation. About 85% of the learners were able to list possible outcomes from a tree diagram. However some lists were not exhaustive or had repetitions, indicating that the method was still not mastered well as shown by Prisca. For those with partially correct representations for example Thulani, most could not pass through the threshold to progress to the next step to compute and virtually utter the numerical probabilities required. Some had good representations for example Mandi, but wrongly decoded the solutions.

CONCLUSION

The research revealed learners’ behavioural patterns related to the different types of representations used in solving probability tasks. We conclude the following:

Tree diagrams: In order to determine the probability of a sequential experiment, say the Bernoulli type, 85% of the participating learners preferred to use the tree diagram over other types of representations. However, none of the tree diagrams were in essence a correct reflection of the problem. Learners had confidence in the efficiency of the tree diagram despite the fact that they experienced difficulties in constructing correct ones. About 90% of correct answers resulted from tree diagrams that were partially completed. Therefore, we conclude that learners though learners liked the tree diagrams, they often experienced difficulties in constructing them from scratch; but they were quite competent in completing partially constructed tree diagrams and proceed to calculate the required probabilities. This finding implies that learners need scaffolding in modelling probabilistic problems using tree diagrams as this is in their zone of proximal development (Vygotsky, 1986). Starting with partially completed representations would boost their success rate of solving the probability problems with this representation. Once completing partially completed

representations is mastered it will boost their confidence in learning more complex representation models.

Contingency tables: About 50% of the learners successfully solved the problems using the contingency tables. Learners showed a good grasp in the construction and interpretation of the contingency table representation.

Venn diagrams: Only five Venn diagrams were drawn in solving probability problems involving compound events. Two of the five Venn diagrams were drawn by one learner. It can then be said that learners do not have confidence in handling Venn diagrams representations.

Two-way tables: The two-way table was not a popular method of representation for the participants in this study. Only one two-way table was clearly and correctly drawn on the dice task. The other solution seemed to be a hybrid of a two-way table and a contingency table. The unpopularity could suggest that some of the learners may have not been familiar of this type of representation.

Outcome listings: This was fairly popular. About 85% of the learners were able to list possible outcomes by head or through a tree diagram. Most lists were not exhaustive or had repetitions, indicating that the representation is still in learners' zones of proximal development.

The authors argue that real competency in solving probability tasks is achieved if students can make multiple representations of probability problems they are working on. In particular they should be able to assess which representation is best for the problem as often more than one type of representation can be helpful. Yet in others, only one visual representation may do. The ultimate goal is for students to do away with visual representations and be able to work abstractly but this often cannot be achievable if students have not passed through the probability representation phase particularly in Laplacian classical probability dealt at this level. Through visual models, learners realise that probability problems make sense and are relevant in our interpretation of everyday events.

This study shows that participants made multiple errors in representations of the sample spaces as well as interpreting their representations when solving probability tasks. This study acts as a prior phase of didactical engineering. It helps to build teacher probability pedagogical content knowledge needed to plan and implement interventions lessons that resource learner modelling errors as a way to boost learning in this important curricula area.

RECOMMENDATIONS

It is a fact that probability is a new teaching topic in South African schools, where both teachers and learners are facing problems. So, it is recommended that the errors and misconceptions found in this study, particularly on representations in learners' zones of proximal development, inform didactical engineering teaching interventions that use learner errors and misconceptions as teaching resources, to build teacher knowledge on the probability topic.

References

Aichele, D. B. & Coxford, A. F. (1994). *Professional development for teachers of mathematics*. Reston, VA: National Council of Teachers of Mathematics.

- Artigue M. (2009). Didactical design in mathematics education. In, C. Winslow (ed.), *Nordic Research in Mathematics Education. Proceedings from NORMA08*, Copenhagen, April 21-April 25, 2008. pp. 7-16. Sense Publishers.
- Batanero, C., & Díaz, C. (2012). Training school teachers to teach probability: reflections and challenges. *Chilean Journal of Statistics*, 3(1), 3-13.
- Batanero, C. (2005). Significados de la probabilidad en la educación secundaria. *RELIME. Revista latinoamericana de investigación en matemática educativa*, 8(3), 247-264.
- Borasi, L. (1994). Capitalizing on Errors as "Springboards for Inquiry": A Teaching Experiment. *Journal for Research in Mathematics Education*, 25(2), 166-208.
- Brodie, K., & Berger, M. (2010). Toward a discursive framework for learner errors in mathematics. In V. Mudaly (Ed.) *Proceedings of the 18th meeting of SAARMSTE. University of KwaZulu-Natal*, 196-181.
- Chernoff, E. J. (2012). Recognizing revisitation of the representativeness heuristic: An analysis of answer key attributes. *ZDM*, 44(7), 941-952.
- Chevallard, Y. (1992). Concepts fondamentaux de la didactique: Perspectives apportées par une perspective anthropologique. *Recherches en Didactique des Mathématiques*, 12(1), 73-112.
- Department of Basic Education. (2011). *Curriculum Assessment Policy Statements: Grades R-12*. Pretoria: DBE.
- Donaldson, M. (1963). *A study of children's thinking*. London: Tavistock Publications.
- Fischbein, E. (1999). Intuitions and schemata in mathematical reasoning. *Educational Studies in Mathematics*, 38, 11-50.
- Freudenthal, H. (1973). *Mathematics as an Educational Task*. Reidel: Dordrecht.
- Godino, J. D., Batanero, C., Contreras, A., Estepa, A., Lacasta, E., & Wilhelmi, M. (2013). Didactic engineering as design-based research in mathematics education. In *Proceedings of the Eight Congress of the European Society for Research in Mathematics Education* (pp. 2810-2819).
- Hirsch, L. S., & O'Donnell, A. M. (2001). Representativeness in Statistical Reasoning: Identifying and assessing misconceptions. *Journal of Statistics Education*, 9(2).
- Makonye, J. P. (2012). *Learner mathematical errors in introductory differential calculus tasks: a study of misconceptions in the senior school certificate examinations* (Doctoral dissertation). University of Johannesburg.
- National Council of Teachers of Mathematics. (2014). *Principles to actions: Ensuring mathematical success for all*. Reston, VA: Author.
- Polya, G. (2014). *How to solve it: A new aspect of mathematical method*. NY: Princeton university press.
- Shaughnessy, J. M. (1992). 'Research in probability and statistics: Reflections and directions'. In D. A. Grouws (Ed.). *Handbook of Research on Mathematics Thinking and Learning* (pp. 465-494). New York: Macmillan.
- Shulman, L. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, 15(2), 4-14.
- Suh, J. M. (2007). Tying it all together: Building mathematics proficiency for all students. *Teaching Children Mathematics*, 14(3), 163-169.
- Vygotsky, L. S. (1978). *Mind in society*. MA: Harvard University Press.
- Zahner, D., & Corter, J. E. (2010). The process of probability problem solving: Use of External Visual Representations. *Mathematical Thinking and Learning*, 12(2), 177 - 204.