Cognitive Levels and Misconceptions of Grade 11 Lebanese Students in Probability

Shehayeb SANAA
Lebanese University

Abstract: The concept of probability is fundamental to students especially at the secondary school level. Learning probability is always linked to logic and reasoning. This paper aims at unveiling the cognitive levels in probability and exploring the probability misconceptions of a sample of 41 grade 11 Lebanese students at the end of their academic year during which they encountered probability as a subject in school for the first time. The approach to data collection is quantitative and qualitative. The 25 item questionnaire in Paul and Hlanganipai (2014) was used to determine the students' cognitive levels in probability based on SOLO taxonomy and using the rubrics used in Watson and Collis (1994). The questionnaire is divided into 5 categories: probability terms and definitions, theoretical probability, Venn diagrams, union and intersection and dependent and independent events. After that 10 students were randomly selected from the sample and interviewed to explore two probability misconceptions: representativeness and equiprobability bias. Results showed that grade 11 students attained level 3 in two of the categories and were able to reach level 2 in two other categories while they remained at level 1 in the fifth categories. As for the students’ misconceptions, representativeness misconception was rarely found while equiprobability bias was more prevalent.

Keywords: Mathematics, Probability, Secondary level, Misconceptions, Solo taxonomy

Introduction

Statistical reasoning is essential to many professions such as psychology through making judgments and decisions, medicine through interpretation of risks of medical outcomes, politics through analysis and interpretation of polls and elections, and journalism through explanation and criticism of statistical information (Garfield, 2002). However, all statistical reasoning is probabilistic (Schum, 2001). That is why several researches recommended teaching probability as early as elementary school (Fischbein, 1991).

The study of probability in the 1970s, became a subject in high schools, while later on, it was also introduced in the fifth and sixth grade of the primary school. However, since the 1990s, probability seem to hold an integrated part in the mathematics curriculum from early grades in primary schools internationally. (Andrew, 2009).

The Lebanese curriculum introduces probability concepts in grade 11 and continues in grade 12. Students in grade 11 get acquainted with the notion and vocabulary of probability, calculate probabilities and use the some properties of probability like incompatible and opposite events while in grade 12, they learn conditional probability and total probabilities with random variable studies. (CERD, 1997).

Problem Statement

Research findings showed that students face many difficulties when learning probability (Anggara et.al., 2018). The national council of teachers of mathematics recommended introducing probability as early as kindergarten (NCTM, 2000). Lebanese students start to develop probability concepts in grade 11. The notion of probability is not introduced before that class. (CERD, 1997). It was noticed that no published research concerning students’ hierarchichal probability levels nor misconceptions in probability was implemented in Lebanon. This research
will fill a gap in literature and shed a light on the importance of teaching probability and the way its probability concepts are developed. This study targets curriculum designers and math textbook authors who will take into consideration students’ levels and misconceptions in probability. Teachers and coordinators also may benefit from this paper in preparing their lesson plans.

Purpose

The purpose of this research is to unveil the cognitive levels of Lebanese grade 11 students in probability based on the SOLO taxonomy and using the rubrics used in Watson and Collis (1994). It also aims at exploring the students’ misconceptions, mainly the equiprobability bias and the representativeness misconception.

Research Questions

The research aimed at answering two research questions:

1. What are the cognitive levels of Lebanese grade 11 students in probability according to levels suggested by Watson and Collis (1993)?
2. What are the Lebanese grade 11 students’ misconceptions in the areas of equiprobability and representativeness?

Student Probability Levels

Piaget and Inhelder analyzed children’s thinking about probability into the usual stages (pre-operational, concrete-operational, formal operational). They have concluded that during the intuitive period (before the age of 6-7), the child is not able to distinguish clearly between chance and necessary phenomena and state that a child in a concrete-operational period is neither able to differentiate between certain and random predictions nor formulate predictions, taking into account his experiences from previous similar situations. The concept of probability, as a formal construct, develops only during the formal operational stage and it represents a synthesis between necessary and the possible, they also suggested that children in the primary grades were able to identify all possible outcomes in a one-stage experiment (Piaget & Inhelder, 1951). A child encounters the concept of probability at the level of his concrete operations, at which time he starts to differentiate between a certain and a possible event (Goldberg, 1966). He also noted that the systematic understanding of probability starts not earlier than between the ages of 9 and 12 years and even during that period children solve problems intuitively, and not on the basis of formal reasoning.

In the same opinion, Jones concluded that significant numbers of grades one through three children were not able to list the outcomes of one stage experiment (Jones, 1974). Consistent with Jones’s finding, Green observed that more than 62 percent of 11-year-olds students were not able to solve one stage sample space items (Green, 1989). According to him, most English pupils finish secondary school without achieving the level of formal operations.

Fischbein have suggested that the concept of “possible” may develop before the concept of “certainty”. According to results of his study, some children develop mathematically mature language for certain and impossible events before they can use it for possible events (Fischbein et.al.,1991). He also showed that some intuitions in young children’s thinking are important in helping their pre-formal probabilistic thinking. These intuitions are a product of personal experience (Fischbein et. al., 1997).

Polaki (2002) also researched probability and suggested four levels of probabilistic thinking, the first being subjective, at which pupils predict the most/least likely event based on subjective judgement, e.g. pupils predict the extracted color to be red, because this is their most favorite color. Transitional probabilistic thinking represents the second level, for which it is significant that students are able to predict the most and the least likely event based on quantitative judgement; which is often invalid, and besides, they may revert to subjective judgements. For the third, informal quantitative probabilistic thinking level it is significant, that pupils correctly predict the most and least likely events, based on quantitative judgements and use numbers informally to compare probabilities. At the fourth level of probabilistic thinking, which is a numerical one, pupils assign a numerical probability and make a valid comparison. (Polaki, 2002).
Another perspective that focused on the hierarchical features was defining the reasoning levels according to different psychological theories. Biggs and Collis (1989) proposed a SOLO model (Structure of the Observed Learning Outcome) to describe cognition levels of probability, which divided students’ probability cognition into five levels, namely pre-structural, uni-structural, multi-structural, relational and extended abstract levels.

Li & Pereira-Mendoza (2002) administered a questionnaire to 567 Chinese students from grades 6, 8 and 12 in order to determine their hierarchical levels in probability in reference to SOLO taxonomy, then 64 students were interviewed to determine their misconceptions. Results showed that most of the students were not able to reach the highest two levels of SOLO but grade 12 students performed better than grades 6 and 8 with no significant difference between grades 6 and 8.

Students’ Misconceptions in Probability

Lai Huat Ang et al. (2014) examined 177 Year 10 and 11 students from two schools who participated in the research study. The two instruments used for this study were ‘Misconception on Probability’ two-tier multiple choice questionnaire and interviews. Carelessness and Incorrect method were grouped as error-typed whereas, Representativeness, Equiprobability bias, Beliefs and Human control were the four identified specific misconceptions on Probability.

Method

Design of the Research

The design of this study can be classified as descriptive quantitative followed by qualitative as it analyzes data collected from a questionnaire with the aim of classifying students’ misconceptions and quantifies variations using Statistical Package for the Social Sciences (SPSS) software-version 23, where participants were generally measured using a questionnaire after which some of them were measured through an interview in order to explore their misconceptions.

The aim of a quantitative research study is to classify features, count them, and construct statistical patterns, paradigms and models in an attempt to explain and demonstrate what has been observed (Babbie, 2010). Qualitative research enriches the research as its main concern is the development of the observed phenomenon, like students misconceptions, through interviews with students (Morse & Field, 1996).
Research Instrument

A set of 25 questions were selected and administered to the students in order to examine the students’ SOLO taxonomy levels in probability. The analysis of the responses identified four levels according to level of sophistication in a similar manner to that of Watson and Collis (1994) as cited in Paul & Hlanganipai (2014). The levels were classified according to the following criteria.

Level 1. In interpreting probability situations no analysis or evidence of use of probability principles is demonstrated. Features may include: the use of irrelevant information, subjective judgements, disregarding quantitative information, guessing at random, belief in control of probability and absence of any reason. Responses that use recent experiences to predict or estimate probabilities, availability, are included in this level.

Level 2. Some evidence of the use of probability principles and appropriate quantitative information is evident, but they may be incomplete or are incorrectly used. Probabilistic reasoning based on the assumption of equal likelihood when none exists and the use of the representativeness heuristic is considered to be illustrative of this level.

Level 3. Probability principles are applied correctly used and an awareness of the role of quantification is evident. However, such quantification is precise or numerical.

Level 4. Probability principles are used correctly and relationships are explained quantitatively. The qualitative part was administered through an interview targeting two misconceptions, the equiprobability bias and the representativeness. Six multiple choice questions extracted from the Probability Assessment Test (PAT) prepared by Anyway & Bennet (2004).

The equiprobability bias is the tendency of students to view several outcomes of an experiment as equally likely. For instance, students who have an equiprobability bias think that when two dice are rolled, all the sums possible are equally likely. They do not realize that the sum of 6 for the two dice is more probable than the sum of 2. (Anyway & Bennet, 2004). People estimate the likelihood of a sample based on how similar it is to the population. A series of coin tosses that has an equal number of heads and tails is judged to be more likely than a series with a dominant occurrence of heads/tails. (Anyway & Bennet, 2003).

Reliability and Validity

The questionnaire reliability was tested by calculating Cronbach alpha by SPSS. The results showed a cronbach alpha value of 0.847 indicating a high consistency in measurement and good correlation between the items, so the questionnaire can be considered as reliable. Content validity was assured by asking three experts to check the questionnaire items and certify their alignment with the purpose of the research.

Data Analysis

Each of the 25 items of the questionnaire was corrected by the researchers and two math education experts and each answer was marked according to the rubric and students’ responses were recorded and coded according to Watson & Collis (1994)’s levels. A statistical computer package, SPSS version 21, was used to process the data. Descriptive statistics was used during data analysis of the questionnaire items. Misconceptions were revealed through analysis of students’ interpretations of the multiple choice questionnaire.

Results

Results of the questionnaire were analyzed for each category and for the overall questionnaire. Interviews were analyzed in terms of the targeted misconceptions

Results of the Questionnaire

The respondents’ cognitive level frequencies for each item were analyzed and are shown in table 1. The results showed that the overall mean (M) cognitive level was (M) =2.3, with a standard deviation (SD) of 0.51 which
indicates that Lebanese grade 11 students have some evidence of use of probability principles and appropriate quantitative information, but they may be incomplete or are incorrectly used. The modal (Mo) cognitive level for each item was 2.0, showing again that students the probability concept is mainly uni-structural according to SOLO taxonomy which indicates that students may have limited knowledge in probability or know just a few isolated facts. Students are able to answer isolated questions but are not able to see connections between ideas.

Table 1. Classification of Students’ Levels

<table>
<thead>
<tr>
<th>Variables</th>
<th>Mean</th>
<th>Mode</th>
<th>Standard deviation</th>
<th>Skewness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probability terms and definitions</td>
<td>2.57</td>
<td>2</td>
<td>0.74</td>
<td>0.92</td>
</tr>
<tr>
<td>Theoretical probability</td>
<td>2.88</td>
<td>3</td>
<td>0.73</td>
<td>-0.77</td>
</tr>
<tr>
<td>Venn diagrams and proportions</td>
<td>2.55</td>
<td>3</td>
<td>0.85</td>
<td>-0.33</td>
</tr>
<tr>
<td>Union and intersection</td>
<td>1.73</td>
<td>2</td>
<td>0.56</td>
<td>0.59</td>
</tr>
<tr>
<td>Dependent and independent events</td>
<td>1.77</td>
<td>1</td>
<td>0.56</td>
<td>-0.065</td>
</tr>
</tbody>
</table>

Table 1 shows that students performed best when answering items related to theoretical probability. An excerpt (figure 2) from a student’s answer to one of the items belonging to this category shows the correct use of probability terminology and relationships. Similar items were coded as level 4 according to the rubric.

On the other hand, students who showed no analysis or evidence of use of probability principles were coded on level 1. Figure 3 shows an excerpt of students’ work which indicates that the student is using a counting principle to answer the question.

Similar analysis was used to code the other categories. Results showed that students were up to level 3 in two categories only which are theoretical probability and Venn diagrams. Students were classified at level 2 in two categories probability terms and definitions and intersection. Excerpt of a student work in figure 4 showing a wrong answer for items 4 and 5. The student was asked to write the sample space of an experiment. The student answered a completely different question which is: in how many ways can we choose a letter of the word probability? This indicates that the student does not show understanding of the vocabulary related to probability unlike the student’s answer in figure 5 which shows a correct writing of the sample space.
As for the third category which is Venn diagram and proportions, the mean was 2.55 and the mode was 3 indicating that the majority of students are able to apply probability principles in this category as shown in figure 6. However, some students had a problem in identifying the prime numbers as in figure 7 and found no intersection between the two sets. Moreover, they failed to represent the Venn diagrams correctly.

Most of the students were not able to reach level 2 in the last category which is related to dependent and independent events. Students faced many difficulties caused by their inability to identify the sample space and the intended event space. The major reason is that students in Lebanon do not study this concept in grade 11 but in grade 12. Whereas the majority were able to attain level 2 in the category related to union and intersection of events though the mean was 1.73 indicating a difficulty in understanding the concept of sets and numbers.

Results of the Interviews

Ten students were chosen randomly from those who have completed the questionnaire and were interviewed in order to determine their equiprobability and representativeness misconceptions. Items related to these misconceptions were adapted from the PAT test prepared by Anyway (2003). Students were asked to answer each of the multiple choice questions and then justify their choice. Results showed that students have more equiprobability misconceptions than representativeness misconceptions. The equiprobability misconception is evident when students do not recognize that the probabilistic outcomes that they are faced with are not perfectly random. (Gauvrit & Morsanyi, 2014). Selecting choice b for question 2 in the probability misconceptions test which indicates that all the three results are equally likely when three dice are simultaneously thrown. The justification by most of interviewed students was that the probability for choosing a two 5’s and a 3 or three 5’s or 5, a 3 and a 6 in any order is $P= \frac{1}{6} \times \frac{1}{6} \times \frac{1}{6}=1/216$. The same justification was brought up by 5 out of the 10 interviewed students showing that students are confused by choosing the sample space and the event space. As for the representation misconception, it was represented by questions 5 and 6. All the interviewed students answered question 5 correctly insisting that heads and tails are equally likely to occur when a fair coin is tossed 5 times. On the other hand, three of the ten students failed to answer question 6 correctly. They insisted that alternatives (a) 3 5 1 6 and (b) 4 2 6 1 5 are equally likely to occur and have a greater chance to occur than that of (c) 5 2 2 2 2.
Discussion

Results from this study indicated that students’ cognitive level was 2.3 which conforms with Paul and Hlanganipai (2014) where the overall mean was 2.14. This signifies that most students operate at level 2 so they have incomplete information of the probability principles and its applications. Students performed least on questions related to dependent and independent events. The main reason is that students get introduced to this concept in grade 12. These results agree with Anggara et. al. (2018) that the most difficult of the probability concepts is the notion of independence. The misconceptions targeted in this study were equiprobability and representativeness which were considered as common misconceptions (Anyway & Benett, 2018).

Conclusion

The purpose of this research was to reveal students’ difficulties in learning probability. This study has many limitations in the sense that it targets grade 11 students only. Interviews were made with ten students only. However, the results were aligned with other studies.

Recommendations

Previous research asserts that transition from one level is related to instruction and not to age (Paul & Hlanganipai, 2014). Curriculum designers should take the cognitive levels into consideration while designing school curricula. Teachers also need to get acquainted with these levels in order to prepare their lesson in a way that promote students to the next level. Teachers should be aware of students’ misconceptions to address them in their assessment. Future research should address these issues on a larger scale in Lebanon especially the ability of students to grasp probability concepts at an early age.

References


---

**Author Information**

Sanaa Shehayeb  
Lebanese University  
Beirut- Lebanon  
Contact E-mail: sanaashehayeb@gmail.com
Appendix

Questionnaire

1. Define the following terms: probability of an event, An experiment, An outcome, Sample space, An event.
   ........................................................................................................................................
   ........................................................................................................................................
2. Write down the sample space for rolling a single die numbered 1 to 6?
   ........................................................................................................................................
   ........................................................................................................................................
3. Suppose that the probability of snow is 0.67, what is the probability that it will NOT snow?
   ........................................................................................................................................
   ........................................................................................................................................
4. What is the sample space for choosing a letter from the word probability?
   ........................................................................................................................................
   ........................................................................................................................................
5. For any event A, P (A) + P (\(\bar{A}\)) = _______, that is P (\(\bar{A}\)) = _____ - P (A).
   ........................................................................................................................................
   ........................................................................................................................................
6. A bag contains 6 red, 3 blue, 2 green, and 1 white balls. A ball is picked at random. Determine the probability is a blue.
   ........................................................................................................................................
   ........................................................................................................................................
7. What is the probability of getting an even number when rolling a single 6- sided die?
   ........................................................................................................................................
   ........................................................................................................................................
8. What is the probability of landing on an odd number after spinning a spinner with 9 equal sectors numbered 1 through 9?
   ........................................................................................................................................
   ........................................................................................................................................
9. What is the probability of getting a 0 after rolling a single die numbered 1 to 6?
   ........................................................................................................................................
   ........................................................................................................................................
10. A bag has 20 raffle tickets in it, numbered from 1 to 20. What is the probability of picking out an even number?
    ........................................................................................................................................
    ........................................................................................................................................
11. Let S denote the set of whole numbers from 1 to 16, X denote the set of even numbers from 1 to 16 and Y denote the set of prime numbers from 1 to 16. Draw a Venn diagram depicting S, X and Y.
    ........................................................................................................................................
    ........................................................................................................................................
12. a) Pieces of paper labelled with the numbers 1 to 12 are placed in a box and the box is shaken. One piece of paper is taken out and then replaced. What is the sample space, S? Find n(s).
    ........................................................................................................................................
    ........................................................................................................................................
b) Write down the set A, representing the event of taking a piece of paper labelled with a divisor of 12? Find n (A).
    ........................................................................................................................................
    ........................................................................................................................................
c) Write down the set B, representing the event of taking a piece of paper labelled with a prime number. Find n (B).


d) Represent A, B and S by means of a Venn diagram.


13. Let E and F be events such that Pr(E) = 0.6, Pr(F') = 0.3, and Pr (E∪F) = 0.8. Find Pr (E∩F).


14. A jar has purple, blue and black sweets in it. The probability that a sweet chosen at random will be purple is 0.2 and the probability that it will be black is 0.6. If I choose a sweet at random, what is the probability that it will be purple or blue.


15. If dice are the same color, what is the probability of getting 2 or 3 on at least one of the dice?


16. Suppose our experiment is flipping a coin three times in a row. Let B be the event that we do not get three heads in a row. Find P (B).


17. Two fair dice are rolled. What is the probability that the sum of the values is a prime number?


18. A school decided that its uniform needed upgrading. The colors on offer were beige or blue or beige and blue. 40% of the school wanted beige, 55% wanted blue and 15% said a combination would be fine. Are the two events independent?


19. A jar contains 4 white marbles, 5 red marbles, and 6 black marbles. If a marble were selected at random, what is the probability that it is white or black?


20. If D and F are mutually exclusive events, with P (D) = 0.3 and P(D or F) = 0.94, find P(F).


21. Given Pr (E) = 0.5, Pr (F) = 0.3, and Pr (E∩F) = 0.1. Determine if E and F are independent events?


22. A cloth bag has four coins, one R1 coin, three R2 coins and two R5 coin. What is the probability of first selecting a R1 coin and then selecting a R2 coin?


**Misconceptions Questionnaire**

1. Five faces of a fair die are painted black, and one face is painted white. The die is rolled six times. Which of the following results is more likely?
   a. Black side up on five of the rolls; white side up on the other roll
   b. Black side up on all six rolls
   c. a and b are equally likely

2. When two dice are simultaneously thrown it is possible that one of the following two results occurs: Result 1: a 5 and a 6 are obtained in any order. Result 2: a 5 is
obtained twice. Select the response that you agree with the most.

a. The probability of obtaining each of these results is equal.
b. There is a higher probability of obtaining result 1.
c. There is a higher probability of obtaining result 2.
d. It is impossible to give an answer.

3. When three fair dice are simultaneously thrown, which of the following results is MOST LIKELY to be obtained?
   a. Result 3: Two 5's and a 3
   b. All three results are equally likely.
   c. Result 1: A 5, a 3 and a 6 in any order
   d. Result 2: Three 5's

4. When three dice are simultaneously thrown, which of these three results is LEAST LIKELY to be obtained?
   a. Result 1: A 5, a 3 and a 6 in any order
   b. Result 2: Three 5's
   c. Result 3: Two fives and a 3 in any order
   d. All three results are equally unlikely.

5. Which of the sequences is least likely to result from flipping a fair coin 5 times?
   a. H H H T T
   b. T H H T H
   c. T H T T T
   d. H T H T H
   e. All four sequences are equally likely

6. If a fair die is rolled five times, which of the following ordered sequence of results, if any, is MOST LIKELY to occur?
   a. 3 5 1 6 2
   b. 4 2 6 1 5
   c. 5 2 2 2 2
   d. Sequences (a) and (b) are equally likely.
   e. All of the above sequences are equally likely.