MEASURING STUDENTS’ UNDERSTANDING IN COUNTING RULES AND ITS
PROBABILITY VIA E-LEARNING MODE: A RASCH MEASUREMENT
APPROACH

N. Baharun1*, N. F. M. Razi1, R. Z. Abidin1, N. A. C. Musa2 and Z. Mahmud2

1Faculty Computer and Mathematical Sciences, UniversitiTeknologi MARA, Perak, Malaysia
2Faculty Computer and Mathematical Sciences, UniversitiTeknologi MARA, Shah Alam, Selangor, Malaysia

Published online: 10 November 2017

ABSTRACT
Probability is a study of the rules that offers the foundational theory for the development of statistics. This sets out the investigation where students’ understanding of counting rules and its probability were explored using the Rasch measurement approach. A test instrument with 20 items was developed and administered to 74 students taking the STA150 Probability and Statistics course. Data were captured through an interactive e-learning platform that is Edmodo.com and analyzed using Winsteps 3.81.0. The results from the Wright map showed that 83.8% of the students have the ability that matched well with the difficulty of the items while 16.2% of the students need to be given more attention on the topic. The study was also able to show that the items can be replicated in other samples of similar characteristics.

Keywords: students’ understanding; counting rules; probability; Rasch measurement model; Wright map.

Author Correspondence, e-mail: norha603@perak.uitm.edu.my
doi: http://dx.doi.org/10.4314/jfas.v9i6s.33
1. INTRODUCTION

Probability courses have become a requisite in a wide field of study. It is designed to introduce the basic concepts and as a decision-making tool in many disciplines. Statistical concepts are the basis of learning statistics and should be given extra attention by every educational institution as statistics can be applied to a variety of areas including medicine, public health, psychology, biology, business, economics, engineering, education and sports among others.

Recently, research in learning probability and statistics concept has been well documented. The study was examined extent of which anxiety and attitudes towards learning statistics affects the academic performance for those students enrolled in statistics course [9]. For the above reason, more attention has been given for measuring students’ belief and feelings about statistical concepts. In [14] has focused on investigating the reasoning factors of belief and feelings about statistics and the extent of its relationship with statistics achievements. As a result of this study, some guidelines to improve the teaching and learning of statistics in education are offered.

Mathematical concepts of probability is introduced where permutations, combinations, probability and n factorials can be obtained. Due to this, many students facing problem to understand the probability concepts because of inadequacies in prerequisite mathematics skills problem [6]. Many middle and high school students have difficulty in understanding probability and statistical concepts. These difficulties may be due to less or no curriculum instruction for probability given at the elementary school level. As a result, it is challenging to relate student and engage them in learning experiences in which they construct their own understanding of probability concepts [5]. There are several difficulties in learning and teaching statistics and some of these might be stated as follows [3].

- Students have no motivation in learning statistics since statistical ideas and methods are difficult and complex and counter-intuitive students. Therefore, it is difficult for a teacher to motivate students of this subject.
- Many students have problems in statistically related mathematics subjects (fractions, decimal numbers, proportional judgement, algebraic formulas, etc.).
• Students consider that statistical learning is to memorize formulas and feel the results can be achieved only by operational knowledge. Do not taking into account the statistical conceptual of learning in this subject.
• Students consider mathematics and statistics is the same. Therefore, the students feel they need to emphasize just numbers, calculations, formulas and the one single correct answer.
• Many similar words are indeed different or have a statistically special significance.
• In most statistical learning problem, the context often leads students through a mistake where students rely on experience rather than choose the right way with evidence-based data. Thus, the context will mislead the students.
• Students feel anxious and nervous when receiving the scattered, chance and randomness data. This allows students have a variety of different interpretations in different things and it requires intensive skills of writing, reading, communication and cooperation.

Probability is the first place among subjects that students have difficulty in understanding and teachers handling. Many researchers have revealed the difficulties high school students encounter about likelihood events [14]. Therefore, focus of this study tends to be on probability and chances of uncertain outcomes. Most of students have develop perception about probability concepts. They have difficulties in facing learning probability concepts and the reasons why these concepts cannot be understand well are investigated. Regarding this problem, a cross-hatching model was used. As a result, there are six reasons why probability concepts cannot be learned and taught has been discovered. These categories were age, the insufficiency of advanced information, the deficiency of argumentation ability, teacher, error in concept and students’ negative attitudes [12].

Recently, development in nanotechnology bring to our everyday life advanced in engineering devices that can be explained by principle quantum mechanics where it is introduces the ideas of probability. A few research has investigated student difficulties in understanding and interpreting probability and its relevant technical terms as it relates to quantum measurement. From this research, it is evident that students have difficulties in understanding these terms and often fail to differentiate among similar but different concepts [16].

Besides probability understanding, students also have difficulties in viewing and interpreting graphs. They have trouble relating between data and graph in particular identifying variation
and distribution in a histogram. There are many different graphs that can look quite similar and students have problems identifying what is going [2]. The study was presented in three levels of students’ comprehension namely reading the data, reading between the data and reading beyond the data [4]. Furthermore, the analysis show that conceptual understanding of basic statistical concepts is more difficult to attain than statistical competence [11]. Therefore, the aim of this study is to explore students’ understanding of counting rule and its probability concepts and examine which concepts were found most difficult by the students to understand.

2. METHODOLOGY

This study involved the participant of 74 students from the Probability and Statistics (STA150) course in the semester of June-Nov 2016. Based on the previous examination results, many students did not perform well in the Counting Rules topic. Therefore, in this study, a set of structured test comprises of 20 items on Counting Rules topic was developed. Based on a new paradigm shift in the teaching and learning of statistics, this study had decided to use Edmodo.com, an open source e-learning platform as a medium of interactive learning for the students. However, the process of normal teaching method will continue to be held in the class. Test items were administered and data was collected after the students have covered the Counting Rules topic.

The data were analyzed using the Bond and Fox Winsteps version. The results are then presented in several sections namely data exploration, validation and calibration of items and responses. These analyses includes item and person fit statistics in the form of mean square values (MNSQ) and standardized z-score for infit and outfit. The analysis was carried out in order to calibrate between person ability and item difficulty towards the probability concept. Rasch Dichotomous Model was embedded in Winsteps to measure students’ ability in the particular probability topic. This instrument provides items with dichotomous answers which are correct and incorrect based on the student’s ability to answer the questions. The probability of person and item is defined mathematically as follows
\begin{equation*}
P_{ni}(x = 1|\beta_n, \delta_i) = \frac{\exp(\beta_n - \delta_i)}{1 + \exp(\beta_n - \delta_i)}
\end{equation*}

where \(P_{ni}(x = 1|\beta_n, \delta_i)\) is the probability of person \(n\) on item \(i\) scoring as correct \((x=1)\) response rather than an incorrect \((x=0)\) response, given person ability \(\beta_n\) and item difficulty \(\delta_i\).

3. RESULTS AND DISCUSSION
The analysis and result for this study are discussed.

3.1. Summary Statistics for Items Measure

Table 1. Summary statistics for items measure

| TOTAL SCORE | COUNT | MEASURE | MODEL S.E. | INFIT S.E. | INFIT ZSTD | OUTFIT S.E. | OUTFIT ZSTD |
|-------------|-------|---------|------------|------------|------------|-------------|-------------|
| MEAN        | 43.1  | .00     | .28        | 1.00       | 1.00       | 1.00        | 1.00        |
| P.SD        | 14.5  | .10     | .05        | .09        | 1.00       | 1.00        | 1.00        |
| S.SD        | 14.8  | .12     | .05        | .09        | 1.00       | 1.00        | 1.11        |
| MAX.        | 67.0  | 74.0    | 2.64       | .41        | 1.16       | 1.9         | 1.26        |
| MIN.        | 9.0   | 74.0    | -2.09      | .25        | .82        | -2.1        | .77         |

RELATION RMSE .29 TRUE SD 1.05 SEPARATION 3.62 ITEM RELIABILITY .93
MODEL RMSE .29 TRUE SD 1.05 SEPARATION 3.69 ITEM RELIABILITY .93

Item RAW SCORE-TO-MEASURE CORRELATION = -1.00
Global statistics: please see Table 44.
U MEAN = .0000 USCALE = 1.0000

Reliability for items measure can be interpreted on a 0 to 1 scale [15]. Table 1 shows that there is a high item reliability index at 0.93 suggesting that the items used in this study are ready to be replicated to other samples for whom it is suitable [15, 13]. The item separation index refers to the ability of the test to define a distinct hierarchy of items along the measured variables. Based on Table 1, the separation index of 3.62 indicates that the items can be separated into four difficulty levels.

Table 2 shows the summary for the person reliability measure. A reliability measure at 0.54 indicates a considerably low reliability coefficient. The reasons for this could be attributed to student’s making careless mistakes or guessing for answers.
Table 2. Summary statistics for Person measure

| TOTAL SCORE | COUNT | MEASURE | MODEL S.E. | INFIT MNSQ | ZSTD MNSQ | OUTFIT MNSQ | ZSTD MNSQ |
|-------------|-------|---------|------------|------------|-----------|-------------|-----------|
| MEAN        | 11.6  | .42     | .53        | 1.00       | .0        | 1.00        | .0        |
| P. SD       | 3.0   | .0      | .82        | .05        | .22       | .9          | .37       |
| S. SD       | 3.0   | .0      | .83        | .05        | .22       | .9          | .37       |
| MAX.        | 18.0  | 20.0    | 2.65       | .81        | 1.70      | 2.7         | 2.48      |
| MIN.        | 5.0   | 20.0    | -1.34      | .49        | .51       | -1.5        | .29       |

REAL RMSE: 0.56 TRUE SD: 60 SEPARATION: 1.08 Person RELIABILITY: 0.54
MODEL RMSE: 0.53 TRUE SD: 63 SEPARATION: 1.17 Person RELIABILITY: 0.58

S.E. OF PERSON MEAN = .10

Person raw score-to-measure correlation = .99
CRONBACH ALPHA (KR-20) Person raw score "test" reliability = .57 SEM = 1.96

3.2. Wright Map

Fig.1. Wright map for the distribution of persons and items
Fig. 1 shows the Wright map for the distribution of persons and items. Persons are distributed on the left side of the logit ruler (center vertical line) and items are distributed on the right side. “M” refers to person and item mean, “S” refers to 1 standard deviation away from the mean, and “T” is 2 standard deviation away from the mean. The logit mean for item measure is adopted by default as the 0 point where item Q16 is located. Item Q12 was found to be the hardest question to answer while item Q4 was the easiest question. Through this map, it is able to identify the most able and least able students in endorsing the items as correct. The most able student in endorsing items as correct is a female student from CS143 program while the least able student is a male student from the CS112 program. Overall, it shows that 83.8% of the students’ have the ability that matched well with the difficulty of the items. This indicates that majority of the students were able to understand about counting rules and its probability following their exposure to the interactive medium of instruction.

### 3.3. Item Fit

| ENTRY NUMBER | TOTAL SCORE | TOTAL COUNT | MEASURE | MODEL S.E. | INFIT MNSQ | ZSTD | OUTFIT MNSQ | ZSTD | CORR. EXP. | PTMEASUR-AL | EXACT MATCH | OBS% | EXPR% | Item |
|--------------|-------------|-------------|---------|------------|------------|------|------------|------|------------|-------------|-------------|------|-------|------|
| 2            | 52          | 74          | -0.56   | 0.27       | 1.05       | 0.5  | 1.26       | 1.4  | A.24       | 33          | 70.3        | 72.3 | Q3    |
| 3            | 38          | 74          | 0.35    | 0.25       | 1.15       | 1.9  | 1.20       | 1.9  | B.17       | 36          | 59.5        | 65.3 | Q4    |
| 18           | 39          | 74          | -0.23   | 0.25       | 1.12       | 1.4  | 1.20       | 1.8  | C.21       | 36          | 60.8        | 65.4 | Q18   |
| 12           | 9           | 74          | 2.64    | 0.37       | 0.96       | -1.1 | 1.18       | -0.6 | D.25       | 26          | 89.2        | 87.8 | Q12   |
| 14           | 38          | 74          | 0.35    | 0.25       | 1.11       | 1.4  | 1.14       | 1.3  | E.23       | 36          | 59.5        | 65.3 | Q14   |
| 6            | 64          | 74          | -1.66   | 0.35       | 1.10       | 0.5  | 1.06       | 0.3  | F.14       | 25          | 86.5        | 86.5 | Q5    |
| 10           | 45          | 74          | -0.09   | 0.25       | 1.08       | 0.3  | 1.02       | 0.2  | G.28       | 35          | 54.1        | 67.6 | Q10   |
| 19           | 46          | 74          | -0.15   | 0.28       | 1.07       | -0.8 | 1.04       | -0.4 | H.27       | 35          | 63.5        | 68.1 | Q19   |
| 15           | 24          | 74          | 1.26    | 0.28       | 1.04       | 0.3  | 1.03       | 0.2  | I.30       | 34          | 65.2        | 70.7 | Q15   |
| 17           | 61          | 74          | -1.33   | 0.32       | 1.02       | 0.3  | 1.03       | 0.2  | J.25       | 26          | 81.1        | 82.4 | Q17   |
| 8            | 47          | 74          | -0.22   | 0.27       | 0.99       | -1.9 | 0.94       | -4  | K.37       | 35          | 64.9        | 68.6 | Q8    |
| 9            | 51          | 74          | -0.43   | 0.27       | 0.96       | -3.9 | 0.99       | 0   | L.37       | 33          | 75.7        | 71.4 | Q9    |
| 1            | 56          | 74          | 0.87    | 0.29       | 0.93       | -1.8 | 1.07       | -6  | M.36       | 31          | 78.4        | 76.3 | Q1    |
| 20           | 14          | 74          | 2.07    | 0.31       | 1.97       | -1.9 | 0.92       | -2  | N.34       | 30          | 83.8        | 81.5 | Q20   |
| 13           | 36          | 74          | 0.35    | 0.25       | 0.95       | -6.9 | 0.94       | -6  | O.42       | 36          | 64.9        | 65.3 | Q13   |
| 4            | 67          | 74          | -2.09   | 0.41       | 0.93       | -1.7 | 1.07       | -4  | P.32       | 22          | 90.5        | 90.5 | Q4    |
| 7            | 40          | 74          | 0.23    | 0.25       | 0.92       | -1.0 | 0.87       | -1.2 | R.47       | 36          | 67.6        | 65.6 | Q7    |
| 11           | 40          | 74          | 0.23    | 0.25       | 0.90       | -1.2 | 0.86       | -1.3 | S.48       | 36          | 64.9        | 65.6 | Q11   |
| 5            | 48          | 74          | -0.29   | 0.28       | 0.85       | -1.5 | 0.79       | 0.5  | T.52       | 34          | 77.0        | 69.1 | Q5    |
| 16           | 44          | 74          | -0.02   | 0.25       | 0.82       | -2.1 | 1.17       | 1.8  | U.36       | 28          | 77.7        | 67.1 | Q16   |

**Table 3.** Items fit statistics prior to misfit order
Fit statistics can be used diagnostically to investigate the existence and potential impact of guessing. Item fit is expressed as indexes that indicate the mean square (MNSq) and a standardized value (MNZ). Mean squares (MNSq) indicate the size of misfit that is revealed in the actual data. An item may become misfit because it may too complex, confusing or it may actually measure a different construct.

The misfit response can be classified into two condition which are overfit where the mean square value (MNSq) is less than 1.0 and a negative standardized fit (MNZ) and underfit (noise) is defined when the mean square value (MNSq) is more than 1.2 and the Z standardized fit greater than 2.0. Table 3 shows the item fit statistics prior to misfit order. It was found that the items in the bracket (Q12, Q8, Q9, Q1, Q20, Q13, Q4, Q7, Q11, Q5 and Q16) were found to be overfit since the mean square value less than 1.0. This indicates that there could be too little variation in the item response pattern or perhaps indicating the presence of redundant items. There was no item found to be underfit since no MNSq is above 2.0. To overcome the overfit problem, it is suggested to either drop or restructure the items in order to increase its reliability index.
Table 4. Person fit statistics prior to misfit order

| ENTRY NUMBER | TOTAL SCORE | MEASURE | MODEL S.E. | INFIT ZSTD | OUTFIT ZSTD | PTMEASUR-AL CORR. EXP. | EXACT OBSERV | MATCH EXP | Person |
|--------------|-------------|---------|------------|------------|-------------|--------------------------|--------------|-----------|--------|
| 54           | 18          | 2.65    | .61 .146   | .9 .1 .91   | .1 .04 .35 .35         | 90 .90       | F0753C51L343C |
| 33           | 18          | 2.10    | .69 .192   | .89 .12 .25 | .25 .39 .80 .853       | F6253C51L343C |
| 69           | 17          | 2.10    | .69 .352   | .11 .45 .8 | .03 .39 .80 .853       | F6732C51L343C |
| 55           | 16          | 2.47    | .62 .455   | .62 .49 .71 | .45 .9 .80 .853       | F7142C51L343C |
| 60           | 16          | 2.17    | .62 .75    | .58 .61 .61 | .41 .9 .80 .853       | F3842C51L343C |
| 5            | 15          | 1.32    | .57 .99    | .42 .2 .46 | .42 .75 .75 .75       | F9471C51L343C |
| 12           | 15          | 1.32    | .57 .94    | .41 .2 .42 | .42 .80 .75 .75       | F5866C51L343C |
| 31           | 15          | 1.32    | .57 .99    | .41 .2 .42 | .42 .80 .75 .75       | F5338C51L343C |
| 59           | 12          | 1.32    | .57 .99    | .41 .2 .42 | .42 .80 .75 .75       | F6271C51L343C |
| 10           | 14          | 1.01    | .54 .80    | .4 .7 .70 | .6 .6 .43 .85 .75 .75 | F8424C51L343C |
| 24           | 14          | 1.01    | .54 .80    | .4 .7 .70 | .6 .6 .43 .85 .75 .75 | F8938C51L343C |
| 30           | 14          | 1.01    | .54 .80    | .4 .7 .70 | .6 .6 .43 .85 .75 .75 | F9142C51L343C |
| 36           | 14          | 1.01    | .54 .80    | .4 .7 .70 | .6 .6 .43 .85 .75 .75 | F9471C51L343C |
| 13           | 14          | 1.01    | .54 .80    | .4 .7 .70 | .6 .6 .43 .85 .75 .75 | F5866C51L343C |
| 14           | 14          | 1.01    | .54 .80    | .4 .7 .70 | .6 .6 .43 .85 .75 .75 | F5338C51L343C |

| MEAN         | 11.80       | 20.00   | .42 .531.00 | .01 .00 .00 | .00 | 7.18 | 72.6 |
| P.5D         | 3.00        | 0.00    | .82 .05 .22 | .9 .37 .9 | 10.8 | 5.7 |
On the other hand, person fit describes how well the persons have responded to the items in a consistent manner. Responses may fail to be consistent when people are bored and inattentive to the task when they are confused or when an item evokes an unusually salient response from an individual. Table 4 shows that there are two students found to have misfit responses (Female, CS143, 6332) and (Female, CS143, 8934) since the infit and outfit values for mean square (MNSq) is greater than 1.6 and standardized z greater than 2.0 respectively.

3.4. Item Characteristics Curve (ICC)

Fig. 2. Actual performance versus theoretical ICC for item Q2 with poor fit

Further analysis look at the endorsement of the items and responses based on the Item Characteristics Curve (ICC). Selected items that show the poorest fit have been selected in order to give a clearer picture of the existence of the misfit problem. Fig. 2 and Fig. 3 show that some of expected responses does not lie along the sigmoid curve. Fig. 3 explain Q2 “Find the number of ways that can be formed from all the letters of word PROPOSITION” able to detect the presence of misfit data or unusual responses since it shows some scores that are scattered away from the range of the confidence interval. This situation may be due to guessing of the answer or unexplained situation.

Fig. 3 however shows the misfit situation for item Q18 where some of the scores were plotted below the expected location. It indicates that item Q18 “It is required to seat 5 men and 4 women in a row so that the women occupy the even places. How many such arrangements are possible?” is unable to be answered by most students.
4. CONCLUSION
This study has shown that Rasch measurement model is a powerful tool to measure students’ understanding of counting rules and its probability. It also provides a good measurement between student ability and item difficulties. In this study, e-learning has been implemented to complement the conventional method in delivering understanding for this topic.

The results from the Wright map showed that 83.8% of the students have the ability that matched well with the difficulty of the items. This indicates that majority of the students were able to understand about counting rules and its probability following their exposure to the interactive medium of instruction while 16.2% of the students need to be given more attention on the topic. The study was also able to show that the items can be replicated in other samples of similar characteristics due to its high item reliability index at 0.93. Rasch measurement is also able to detect misfit in the responses and items. Therefore, it is recommended that the investigation on student’s learning be expanded to the next level.

5. ACKNOWLEDGEMENTS
The authors would like to thank UniversitiTeknologi MARA for funding this research through the ARAS Grant (600-RMI/DANA/5/3/ARAS (1/2015).
6. REFERENCES

[1] Arthur B. Reasoning about shape as a pattern in variability. Statistics Education Research Journal, 2004, 3(2):64-83

[2] Bakker A, Gravemeijer K. Learning to reason about distribution. In D. Ben-Zvi, & J. Garfield (Eds.), The challenge of developing statistical literacy, reasoning, and thinking. Amsterdam: Springer, 2004, pp. 147-168

[3] Ben-Zvi D., Garfield J. The challenge of developing statistical literacy, reasoning, and thinking. Dordrecht: Kluwer Academic Publishers, 2004

[4] Curcio F. Comprehension of mathematical relationships expressed in graphs. Journal for Research in Mathematics Education, 1987, 18(5):382–393

[5] Taylor F M. Why teach probability in the elementary classroom. Louisiana Association of Teachers Mathematics Journal, 2011, 2(1):1-10

[6] Garfield J, Ahlgren. Assessing students’ conceptual understanding after a first course in statistics; difficulties in learning basic concepts in probability and statistics. Journal for Research in Mathematics Education, 1988, 19(1):44-63

[7] Ivan B. Modelling in Teaching probability and statistics at the faculties of technology in the Republic of Croatia. The Holistic Approach to Environment, 2016, 6(4):147-157

[8] Lalonde R N, Gardner R C. Statistics as a second language? A model for predicting performance in psychology students. Canadian Journal of Behavioural Science, 1993, 251(1):108-125

[9] Nasser F. Structural model of the effects of cognitive and affective factors on the achievement of Arabic-speaking pre-service teachers in introductory statistics. Journal of Statistics Education, 2004, 12(1):1-28

[10] Nor A C M, Zamalia M, Nor A M G. Application of Rasch measurement model in validation and analysis of measurement instruments in statistical education research. In Recent Research in Education Conference, 2011, pp. 123-129

[11] Kassim N L, Ismail N Z, Mahmud Z, Zainol M S. measuring students' understanding of statistical concepts using Rasch measurement. International Journal of Innovation, Management and Technology, 2010, 1(1):13-19
[12] Memnun D S. Difficulties of learning probability concepts, the reasons why these concepts cannot be learned and suggestions for solution. Inonu University Journal of the Faculty of Education, 2008, 15(9):89

[13] Zamalia M, Anne L P. Using Rasch analysis to explore what students learn about probability concepts. Indonesian Mathematical Society Journal on Mathematics Education, 2015, 6(1):1-11

[14] Zieffler A, Garfield J, Alt S, Dupuis D, Holleque K, Chang B. What does research suggest about the teaching and learning of introductory statistics at the college level? A review of the literature. Journal of Statistics Education, 2008, 16(2):1-23

[15] Bond T., Fox C. M. Applying the Rasch model: Fundamental measurement in the human sciences. Abingdon: Routledge, 2015

How to cite this article:
Baharun N, Razi N F M, Abidin R Z, Musa N A C, Mahmud Z. Measuring students’ understanding in counting rules and its probability via e-learning mode: a rasch measurement approach. J. Fundam. Appl. Sci., 2017, 9(6S), 429-441.