APPLICATION OF RASCH MEASUREMENT MODEL FOR RELIABILITY MEASUREMENT INSTRUMENT IN VECTOR CALCULUS COURSE

H. OTHMAN*, N. A. ISMAIL, I. ASSHAARI, F. M. HAMZAH, Z. M. NOPIAH

Centre for Engineering Education Research, Unit of Fundamental Engineering Studies, Faculty of Engineering and Built Environment, Universiti Kebangsaan Malaysia, 43650 UKM Bangi, Selangor Darul Ehsan, Malaysia

*Corresponding Author: haliza@eng.ukm.my

Abstract

Competence in mathematical knowledge is of prime importance to most if not all tertiary course, especially for engineering related courses. Mastery in mathematical thinking and problem solving skills is a must for engineers and engineering students, in which mathematics is the key element to engineering and research studies. Most of undergraduate courses in higher institutions in Malaysia use final exam as measurement instrument to measure students’ academic achievement. Good constructed final exam questions/items would be able to measure both students’ academic achievement and their generic skills. This study used Rasch Measurement Model to evaluate the reliability final exam questions for KKKQ1124 Mathematics Engineering I (Vector Calculus) course in Faculty of Engineering and Built Environment (FKAB), Universiti Kebangsaan Malaysia (UKM). The items in the examination paper were studied and items that do not measure up to expectations were identified. Analysis on questions/items gave a good indication how a well-constructed questions/items will provide useful information about a student's ability. The results of this study found no exam questions/items which were identified as unsuitable (misfit), but the revision of the exam questions KQ1124 must be made on the level of difficulty of the questions/items. In addition, the Rasch Measurement Model shows that the performance of the students of Year One, is lower than expected based on the lower score of Individuals Min Items.

Keywords: Rasch measurement model, Mathematics engineering, Student’s ability.
1. Introduction

Fundamentally, mathematical knowledge is a crucial element to all courses in engineering and engineering research. Teaching mathematics to engineering students is to find a balance between the practical applications of mathematics and a deep understanding of real life [1]. Whereas, Cardella [2] states that the effect of teaching mathematical thinking skills in an engineer would enable them to use mathematics in their work practices. The findings made by Zainuri et al. [3] and Othman et al. [4] on the performance of engineering students in the Faculty of Engineering and the Built Environment (FKAB), Universiti Kebangsaan Malaysia (UKM) in the Pre-Test Mathematics is parallel and agrees with the findings of the Lawson [5], that there is a significant decline in many math skills among students in higher education institutions who follow courses that require knowledge of mathematics.

This paper applies the Rasch Measurement Model in analyzing the reliability of examination questions for Engineering Mathematics I: Vector Calculus (KKKQ1124). Fuad et al. [6] used this model to identify items difficulties in Pre-Test based on students pre-university background, which will guide lecturers in providing remedial actions for the respective students. Abdul Aziz et al. [7] stated that Persons Item Distribution Map (PIDM) can give an accurate indication of student achievement on a linear scale of measurement. While Rozeha et al. [8] and Saifuddin et al. [9] using Rasch measurement in the assessment of learning outcomes to study Electrical Engineering at UKM and research in Engineering Education, respectively. Osman et al. [10] used Rasch Measurement model to measure students’ performance in Civil Engineering Design II Course in FKAB, UKM. Meanwhile, Ashaari et al. [11] studied and analyzed the separation of Bloom’s cognitive level in the final exam questions (items) for Mathematics Engineering III (Differential Equation) using this measurement model. The paper revealed that Rasch Measurement Model accurately classified the questions/items according to their Blooms’ Taxonomy level and also accurately classified the students according to their observed achievements.

This paper focuses on the reliability of exam questions/items which have been constructed and formulated by lecturers. The exam questions/items reliability must maintain their relative difficulty on the same interval of logit scale, regardless of the student’s ability who answered and challenged the exam questions/items. In addition, Rasch Measurement model was also used to identify the questions/items that do not fit or suit (mismatch).

2. Methodology

Supporting data was obtained from the final exam questions KKKQ1124, Vector Calculus course for Year One students of FKAB, UKM. Data from 214 students from four different engineering departments (Department of Civil and Structure Engineering, Department of Electrical, Electronic and System Engineering, Department of Mechanical and Materials Engineering, and Department of Chemical and Process Engineering) have been collected. There are three sections in the exam at the end of Part A, Part B and Part C. However, this study only focused on Part A and Part B because these questions are the ones that must be answered and the two parts consist of 11 questions including sub-questions. It is
considered as an appropriate instrument to measure student’s learning capabilities and abilities as it covers most of the topics taught in courses like Calculus Vector KKKQ1124, such as partial derivatives, vector functions, integration and others. Program Outcomes (PO) and Course Outcomes (CO) are used as a guide in the construction of the final exam questions/items. However, only PO1 (The ability to acquire and apply knowledge of mathematics, science and engineering) and PO2 (ability to perform identification, formulation and solution of engineering problems) and partial CO taken into account in this study. These questions are labeled as entry CO as shown in Table 1. The rating grades are compiled in Excel *prn format. Numerical coding is required for the assessment of student achievement by using Rasch software, that is Winstep, and output obtained from Winstep was analysed.

<table>
<thead>
<tr>
<th>Section</th>
<th>Question</th>
<th>CO</th>
<th>PO</th>
<th>Blooms’ Taxonomy</th>
<th>Entry No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1a</td>
<td>1</td>
<td>1</td>
<td>Knowledge</td>
<td>A1a_CO1PO1_K</td>
</tr>
<tr>
<td></td>
<td>1b</td>
<td>2</td>
<td>1,2</td>
<td>Knowledge</td>
<td>A1b_CO2PO1PO2_K</td>
</tr>
<tr>
<td></td>
<td>1c</td>
<td>3</td>
<td>1</td>
<td>Understand</td>
<td>A1c_CO3PO1_C</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>Understand</td>
<td>A2_CO3PO1_C</td>
</tr>
<tr>
<td></td>
<td>3a</td>
<td>3</td>
<td>1</td>
<td>Knowledge</td>
<td>A3a_CO3PO1_K</td>
</tr>
<tr>
<td></td>
<td>3b</td>
<td>3</td>
<td>1</td>
<td>Understand</td>
<td>A3b_CO3PO1_C</td>
</tr>
<tr>
<td></td>
<td>4a</td>
<td>5</td>
<td>1</td>
<td>Knowledge</td>
<td>A4a_CO5PO1_K</td>
</tr>
<tr>
<td></td>
<td>4b</td>
<td>5</td>
<td>1</td>
<td>Understand</td>
<td>A4b_CO5PO1_C</td>
</tr>
<tr>
<td></td>
<td>4c</td>
<td>5</td>
<td>1</td>
<td>Knowledge</td>
<td>A4c_CO5PO1_K</td>
</tr>
<tr>
<td>B</td>
<td>a</td>
<td>3</td>
<td>1</td>
<td>Application</td>
<td>Ba_CO3PO1_A</td>
</tr>
<tr>
<td></td>
<td>b</td>
<td>4</td>
<td>1,2</td>
<td>Application</td>
<td>Bb_CO4PO1PO2_A</td>
</tr>
</tbody>
</table>

3. Results and Discussion

_Winstep_ (Rasch analysis software) is used to carry out the necessary analysis of data. Figures 1 and 2 show a summary of the statistical results for the measurement of individual (person) and measurement of items involved in this study. In the Rasch Measurement Model, individuals (persons), represents the students who take KKKQ1124 course and item represents the questions that were asked in the final examination of the KKKQ1124 Vector Calculus course. Basically, statistics summary provide information of the mean, standard deviation, and maximum and minimum standard values for individual (persons) and items. Persons-Item Distribution Map (PIDM) which shows the distribution of individuals (persons) and items, which dispersion maximum and minimum position of individuals (persons) and items demonstrated by the standard deviation (SD). Distribution of individuals (persons) and items is measured using a measuring standard known as the "logit". PIDM exhibits a clear illustration of the relationship between the distribution of individual (persons) and the distribution of items. In addition, the PIDM also may provide a clearer picture about the ability of individuals (persons) and difficulties of the related items.

Figure 1 shows the statistics summary for the individual (persons) of which the Cronbach $\alpha$, acceptable reliability, is 0.67 and the Individuals Mean $\mu_{person}$ are -0.21logit. In other words, the logit shows that individual (student) performance is above the expected performance. Whereas, as depicted in Fig. 2,
the Item Mean had lower rate than the threshold value $\mu_{item} = 0.0$, and it also provides a good statistics summary for the item with the Item Separation of 8.71 logit. Item Separation is the difference between Maximum Measurement Item and Minimum Measurement Item, which is, $1.22 - (-1.27) = 2.49$ logit. However, individuals (students) cannot be separated into several groups because the Individuals (students) Separation is 1.44 logit.

**Fig. 1.** Summary statistics for individual (person).

**Fig. 2.** Summary statistics for items (questions)

There is a three-step comparisons procedure that should be made to determine the misfit items, refer to Fig. 3, Item Measure for Fit Statistics. First, it refers to the Point-Measure Correlation, followed by the Outfit Mean Square (MNSQ) and Outfit z-Standard (ZSTD). These criteria are consecutively compared with the specific range of acceptable region, that are $0.4 < \rho < 0.8, 0.5 < MNSQ < 1.5, -2 < z < 2$, respectively. Analysis obtained from Fig. 3 shows that there are no questions/items which do not fit or not appropriately constructed. Therefore, no further action, such as checking or eliminating any questions, should be taken.
Figure 4 refers to the Person-Item Distribution Map (PIDM), where the discussion focused on the performance of the item and all of the 11 questions spread on logit scale. Item difficulty is shown by the height of the item rank, that is, the higher the rank the more difficult the item and vice versa. According to Fig. 4, only 41.6% (N=89) students were measured above Item Mean, $\mu_{item}$, while 58.4% (n=125) students are under $\mu_{item}$. PIDM shows that the exam questions/items can be divided into five categories, namely 'very difficult', 'difficult', 'moderate', 'easy' and 'very easy'. Question/item labeled by A4c_C05P01_K was most difficult questions/items to be answered by the students and the question/item A2_C03P01_C is the easiest. When correlations and comparisons between individuals (students) and item, we found that no student can answer the question/item which was identified as the most difficult question/item (A4c_C05P01_K) and, meanwhile, eight students failed to answer the easiest questions. There is a large gap between the two questions/items, A1a_C01P01_K and A3a_C03P01_K shown by the vertical arrows in Fig. 4, which are the questions/items in the category of 'hard' and 'medium'. Large gap means that students have problems answering the questions/items given in the final exam.

4. Conclusions
In this study, the final exam questions/items for KKKQ1124 Mathematics Engineering I (Vector Calculus) course was designed and constructed well and reliable because no question/item was found as misfit. However, the performance of students in Year One was lower than expected. It is proven by Person Mean (students) which is lower than the Items Mean (questions), indicating that the students could not answer all questions in the given and expected scope. Although there is no misfit question/item is found, the final exam questions/items level of difficulty should be revised and improved, so that the gap of level of difficulty between questions/items are reduced; which is to reduce the gap between A1a_C01P01_K and A3a_C03P01_K level of difficulty. In conclusion, the Rasch Measurement Model has provided a very useful tool to verify the reliability of the final exam questions for KKKQ1124 course and others Mathematics Engineering courses.
Fig. 4. Person-item distribution map.

Acknowledgement
This is a collaboration research between Centre for Engineering Education Research (www.ukm.my/p3k), and Unit of Fundamental Engineering Studies. The authors wish to express gratitude towards Universiti Kebangsaan Malaysia (UKM) under Action Research Grant (PTS-2014-036) for supporting the research.

References


