Measuring the Students’ Ability among Engineering Technology Student: A Preliminary Investigation

1Abdul Aziz, A., Phd, 2Jusoh, M.S., Phd, 2Omar, A.R., Phd, 2Amlus, M.H., Phd, 3Awang Salleh, T.S

1Infrastructure University Kuala Lumpur (IUKL), Faculty of Information and Computing Technology, 43000 Kajang, Selangor, Malaysia
2University Malaysia Perlis (UniMAP), School of Technopreneurship and Business Innovation, 01000 Kangar, Perlis, Malaysia
3University Kuala Lumpur Malaysian French Institute, Technical Foundation, 43650 Bandar Baru Bangi, Selangor, Malaysia

ARTICLE INFO

Article history:
Received 25 June 2014
Received in revised form
8 July 2014
Accepted 10 August May 2014
Available online 30 August 2014

Keywords:
Students’ assessment, Rasch measurement model, Engineering mathematics, Quality items.

ABSTRACT

Assessing the performance of students by measuring their level of performance is crucial in order to monitor their learning progress. Measurement should be done correctly for it will determine the rightful action towards the placement of the students in the grading system and will determine the enhancement process of the program outcome. Conventional grading are done by summation of scores from tests and or quizzes, without putting consideration on the difficulty of the tasks given to the student. This will influence the performance; hence provide reliable indication of the students’ performance. This paper explain on how Rasch measurement model enable convenient and easy assessment is done which takes into account students’ ability and difficulty of task on the logit measurement rules. The valuable and relevant assessment can only be done with correct measurement tool with reliable and quality items.

© 2014 AENSI Publisher All rights reserved.


INTRODUCTION

Assessment is usually associated with scoring the tests or quizzes. High scores will be regarded as achievers, without considering the given tasks to each student. Education programs are carefully tailored in accordance to the expected learning outcome by the academic institutions including the demand from industry. Combination of both world, added the challenges to academician on their delivery methodology and on the students’ outcome at the end of the program. Accreditation bodies are being set-up in ensuring that the programs carried out by academic institutions are of quality and meet the standard of the accreditation requirements. This is especially in the case of engineering program and specifically those countries which are in the Washington Accord signatories (Richard et al., 2000; Shuman et al., 2005; WA, 2013; EAC, 2010). Malaysia is one of the signatories, and has made an impact in ensuring that the requirement and standards are met on their academic institutions on their engineering programs.

Monitoring the progress of the learning process and making sure that students met the expected outcome prescribed has to be done continuously (Garfield, 1994). Monitoring the achievement by measuring the expected outcome is usually done by summatting the scores for tests, quizzes and even completed projects by the students. However, it is often too common that measurement are not the measurement that were done as what normally be done on scientific measurement. Measurement has to qualify certain properties before it can be considered as a measurement (Rasch, 1960; Wright, 1982; Wright & Mok, 2004; Saidfudin et al., 2010; Saidfudin et al., 2010). Correct measurement is results from good raw materials and worthy data (Wright, 1982). Instrument or test were developed according to the prescribed expected learning outcome thus enable correct assessment be conducted. The study applied Rasch measurement model to the analysis since Rasch model comply the measurement properties which enable objective comparisons, useful predictions and the construction of firm bases of the physiological development (Wright, 1982; Andrich, 1988; Wright & Mok, 2004). This paper explains the process of ensuring the high reliability of the instrument with quality items before proceeding with the assessment on the students.

Corresponding Author: Abdul Aziz, A., Infrastructure University Kuala Lumpur (IUKL), Faculty of Information and Computing Technology, 43000, Kajang, Selangor.
Tel: 603-89266993/Ext: 683; E-mail: azrilah@gmail.com,
**Background:**

The subject is a common mathematics for Diploma in Engineering Technology in all campuses of the university. The topics are functions including trigonometry functions, limit, differentiation and integral. Students enrolled in engineering technology have various mathematics backgrounds, some have experienced learning calculus in secondary school additional mathematics, but some have never learnt the topic before. The mode of teaching in the university is lecture and tutorial, so it is a challenging job to distribute the knowledge by using the same approach as the students’ background are different. Furthermore they prefer hands-on type of learning since their technical subjects are taught by using that nature.

**Instrument:**

The test comprised of 3 sections, that are section A, section B and section C. Section A are all the multiple choice questions total of 15 questions altogether, with options from A to D. Section B is the subjective questions that has 7 questions which the students has to answer all questions. Each question is further divided into subsections, which will be treated as an item by itself. Therefore, section B has 11 items. In section C, there are 3 questions which students were asked to answer two out of three questions. Similarly, each of the questions is further divided into subsections and will be treated as individual items for convenient of measurement and analysis. Altogether, there will be 34 items in the test which will be referred to as the instrument.

**Sample and Data:**

125 students from the program Diploma in Engineering Technologies took the test. The students were coded according to their respective student matrix. The multiple choice answers were keyed-in exactly as to what the students have chosen, into the Excel formatted file. This method of keying in is essential in analyzing the students’ ability on common mathematics and their probability of choosing other answers. It will provide an indication on how the students react towards the construct of the questions; whether the questions are behaving as it is supposed to behave or otherwise.

The short essay marks are keyed-in accordingly to the exact marks the students obtained in the test. However, in order to avoid bias treatment to each of the marks, all were converted to similar base; all the marks are converted to have a full mark of 10 points, since all subsection items are to be treated as individual items.

**Measurement model:**

This paper is applying the Rasch Measurement Model as the model of measuring the students’ performance on the common mathematics subject and will be correlated to the difficulty of the items prescribed in the test. Rasch measurement model, subsequently will be referred to as Rasch model, is used in measuring on ordinal type of data and enable measurement be done as per ordinary scientific type of measurement. The data will be analyzed using Winsteps, one of Rasch model analysis software

**Findings:**

Preliminary findings indicate that the data has high acceptable Cronbach alpha of 0.79 with person reliability at 0.76 and item reliability at 0.95. It indicates that the raw scores consistency is acceptable and the range of person ability and item difficulty is of acceptable range on the measurement ruler (Linacre, 2006; Fisher Jr., 2007). Reference is made to the Table 1 for summary statistics of person and item.

The mean person is -0.21, a negative mean measure indicating that the test is found to be a bit difficult for the students. Their ability measure in general is lower than the mean item measure at +0.00 logit. Rasch theorised that a person have a chance of 50:50 in succeeding a given task a difficulty 0.00 logit. As the item gets harder, the odd of success is reduced; getting less chance to success. Therefore in the case of this test, the items are more difficult and the chance of success for the student is low (Wright & Stone, 1979). The summary statistics also revealed the maximum person measure located at +0.12 logit indicating the location of the most competent student in this group of students. On the other hand, the minimum person measure located at -0.49 logit which represents the location of the least competent student in the group. The spread of the person ability on the measurement ruler is only at +0.37 logit apart, revealing a small gap for the person ability, however with a small standard error at +0.01 logit.

The maximum item is located at +0.90 logit whereas the minimum item is located at -0.49 logit, yielding a gap of +0.41 logit of item difficulty. The standard error for item difficulty is provided at 0.06 logit measure. The good item reliability at +0.95 provides person separation of 1.80 able to roughly segregate the person into two (2) ability level. Referring to the ability spread of the person, the students may be grouped into those with mediocre performance and those with non-performance. Refer to the Wright map in Figure 1, which shows the distribution of person and items according to the ability and difficulty measures on the measurement ruler.

The left most numbers represent the logit measurement on the ruler. The dashed line in the middle separates between the distribution of the students and the item. Due to limitation of space in viewing the whole map, the
Wright map in Figure 1 displays only the distribution of the items and hides the demography details of the students. The hash sign ‘#’ and the dot ‘.’, signifies the approximate location of the students.

Table 1: Summary statistics of person and item.

<table>
<thead>
<tr>
<th>Person ID</th>
<th>Score</th>
<th>Count</th>
<th>Error</th>
<th>MRep</th>
<th>Item</th>
<th>MRT</th>
<th>Item ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.5</td>
<td>122.2</td>
<td>0.6</td>
<td>0.3</td>
<td>0.8</td>
<td>0.1</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>2.4</td>
<td>84.0</td>
<td>0.6</td>
<td>0.3</td>
<td>0.8</td>
<td>0.1</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>3.5</td>
<td>110.5</td>
<td>0.6</td>
<td>0.3</td>
<td>0.8</td>
<td>0.1</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>4.2</td>
<td>130.0</td>
<td>0.6</td>
<td>0.3</td>
<td>0.8</td>
<td>0.1</td>
<td></td>
</tr>
</tbody>
</table>

Fitting Items:

Based on the distribution of the person, in this case the students, and the items, those students located above the mean person of -0.21 logit measures can be grouped as mediocre performer students. The mean person is indicated by the point -0.21 and the ‘M’ sign on the left side of the Wright map. These groups of student are located on the upper half portion of the logit ruler. There are 53 students in this group, about 0.42% students having their ability logit measures above -0.21 logit. The average performer students finds 24%, eight (8) from the total of 34 items, are easy to be accomplished, since those items are located below their ability logit measures.

However, looking back at the detail logit measures for each of the items in item measure table as in Table 2, shows that there are few items appears to be very difficult for this group of students. The items are those located at the upper end portion of the Wright map from Figure 1.

Fig. 1: Wright map of Person and Items Distribution.

Rasch manages its fit item by looking at the 3 fit parameters values that is the point measure correlation, the mean square and the z-standard values (Smith, 1996; Linacre, 2006). Those items depicted in square in Table 2, are the difficult items, with negative or low point measure correlation values. Being difficult items with negative and low point measure correlation values, indicates that those items behaving oppositely than expected. However, for outfit mean square (MNSQ) and z-standard (Z-Std), the values are still within acceptable range that is between 0.4 to 0.8 for MNSQ and within +2.0 to -0.2 for Z-Std (Linacre, 2006; Bond & Fox, 2007).
Comparing those items with those items marked with the horizontal box, this are items with negative point measure correlations values and their MNSQ and Z-Std values are relatively high. Therefore, misfit items 11, 8 and 14 will be dropped (Bond & Fox, 2007), in order to gain better insights of the students’ performance. After deleting the 3 items, the summary statistics yield better reliability compared to initial data. It yields 0.84 Cronbach alpha value compared to previous value at 0.79 with higher person reliability at 0.83. It yielded a bigger person separation at 2.20 compare to previously at 1.80 and still small standard error of 0.02. The deleting of the items are done from the Winsteps control file with insertion of the ‘idelete’ command and the items entry number indicating which items to be deleted. It is a much convenient way in deleting items without having to physically deleting the items from the raw data (Linacre, 2006; Bond & Fox, 2007).

The high difficulty items that are items 4, 10, 5, 7 and 2 with negative point measure correlations seems to distract the reliability of the measures Thus effecting the person ability spread. Therefore, those items will be deleted too, leaving with only 26 items.

Table 2: Item Measure Table.

<table>
<thead>
<tr>
<th>Item</th>
<th>Measure Correlations</th>
<th>MNSQ</th>
<th>Z-Std</th>
</tr>
</thead>
<tbody>
<tr>
<td>Item 1</td>
<td>-0.20</td>
<td>0.95</td>
<td>1.00</td>
</tr>
<tr>
<td>Item 2</td>
<td>-0.15</td>
<td>0.85</td>
<td>0.95</td>
</tr>
<tr>
<td>Item 3</td>
<td>-0.25</td>
<td>0.80</td>
<td>0.90</td>
</tr>
<tr>
<td>Item 4</td>
<td>-0.10</td>
<td>0.75</td>
<td>0.85</td>
</tr>
<tr>
<td>Item 5</td>
<td>-0.12</td>
<td>0.70</td>
<td>0.80</td>
</tr>
</tbody>
</table>

The summary statistics revealed a high Cronbach alpha value of 0.85 with small 0.02 standard error. The person reliability is a 0.83 and person separation increases to 2.24, yielding a better person mean at -0.10 logit measure. Noticeably, the item reliability increase to 0.96 from initial 0.95 with better item separation of 4.93 and a smaller standard error at 0.04 compare to 0.06 earlier. The summary statistic is as shown in Table 3.

Table 3: New Summary Statistics (after deleting 8 misfit items).

<table>
<thead>
<tr>
<th>Item</th>
<th>Measure Correlations</th>
<th>MNSQ</th>
<th>Z-Std</th>
</tr>
</thead>
<tbody>
<tr>
<td>Item 1</td>
<td>-0.20</td>
<td>0.95</td>
<td>1.00</td>
</tr>
<tr>
<td>Item 2</td>
<td>-0.15</td>
<td>0.85</td>
<td>0.95</td>
</tr>
<tr>
<td>Item 3</td>
<td>-0.25</td>
<td>0.80</td>
<td>0.90</td>
</tr>
<tr>
<td>Item 4</td>
<td>-0.10</td>
<td>0.75</td>
<td>0.85</td>
</tr>
<tr>
<td>Item 5</td>
<td>-0.12</td>
<td>0.70</td>
<td>0.80</td>
</tr>
</tbody>
</table>

Profiling:

Upon getting a better fit items and better reliability, it is now possible to get better reading of students’ performance. The new Wright map which displays all students involved in this study and 26 items is as in Figure 2, with the person being represented by the hash sign and dots.

The 0.96 item reliability able to separate the person into roughly 2 groups indicated by the person separation of 2.24. The benchmark is set at the mean item value of 0.00 logit measure. Those students located above mean item 0.00 logit measures are those with average performance. There are 32 students (26%) in this group compared to the previous findings before deletion of misfit items. There are 11 items in which the average performer students’ found easy to accomplished, compared to the other 15 items in which they find more challenging to accomplished. Those 11 items are located below the logit measure of 0.00, lesser difficulty than the students’ ability measures (Lunz, 2010). Reading the Wright map is like looking at the person ability to jump on the high jump bar. The higher the person and the lower the jump bar, the possibility of the person to jump is high; easy jump bar. However, when the high jump bar is near the height of the person, the chance of the person to be able to jump over is either successful or not; a 50:50 chance of successful to failure.
Subsequently, if the high jump bar is higher than the person, the possibility of the person to be able to jump over is lesser.

The second group of students are those located below mean item of 0.00 logit measure. There are 93 students at 74% are categorized as non-performer students. They have difficulties in accomplishing 15 out of all 26 items at 56%, because those items are located above the ability measures of the students.

Conclusion:
Tests are conducted as one of the many ways in measuring the understanding of the students on a particular subject. Tests usually developed based on the program objective and expected learning outcomes, prepared earlier. The results will be used to plan for worthy actions to better enhance the process of learning for the students. However, careful development of the test itself is crucial in making sure that the items are relevant and reliable as to measure the students’ performance. It would be an added advantage for lecturers to be able to use convenient and easy measurement tool in ascertaining the reliability of the test thus achieving the objectives of the measurement. It is very critical in making conclusions on the performance of the students, based on their ability and with consideration of the difficulty of tasks given. This would reflect the true ability of the students, compared to analysing raw score achievement (Wright & Mok, 2004; Granger, 2008).

Tests reliability are checked in making sure that the items fit according to the Rasch fit parameters. Upon finding that there are 8 items with negative point measure correlation values, with MNSQ and Z-Std value out of the acceptable range, those items were deleted. Item deletions are conveniently done within the control file of Winstep software. The reliability of person and item increased and yield a good person separation index.

The students can be grouped into 2, students who are located above the item mean 0.00 logit measures are those students with mediocre or average performance on the subject of basic mathematics. 26% of the students are in this group and conveniently able to accomplished only 11 items out of the 26. While those students below 0.00 logits are the non-performer, where they find it difficult to accomplished 56% of the items prescribed in the test.

This method of measuring the ability of the students able to provide information on good items that are used in the test. Academician not only able to measure the students’ performance but also counter checking the quality of the items. Both provide a valuable information for better enhancement of the learning process for a quality education.

REFERENCES


