Development of Objective Standard Setting Using Rasch Measurement Model in Malaysian Institution of Higher Learning

Nuraini Khatimin[1], Azrilah Abdul Aziz[1], Azami Zaharim[1] & Siti Hanani Mat Yasin[1]

[1] Centre for Engineering Education Research, Faculty of Engineering and Built Environment, Universiti Kebangsaan Malaysia, Selangor, Malaysia

Correspondence: Nuraini Khatimin, Centre for Engineering Education Research, Faculty of Engineering and Built Environment, Universiti Kebangsaan Malaysia, 43600 UKM Bangi, Selangor, Malaysia. Tel: 60-389-216-030. E-mail: nuraini305@gmail.com

Abstract

Measurement and evaluation of students' achievement are an important aspect to make sure that student really understand the course content and monitor students' achievement level. Performance is not only reflected from the numbers of high achievers of the students, but also on quality of the grade obtained; does the grade 'A' truly reflective of a high achiever student. As part of quality improvement, standard setting for students' examination scoring should be set up for courses offered in the institutions. Setting the cut scores of the performance standard is important in making sure the performance is of expected standard. Commonly standard settings done through experts' judgement based on their knowledge and experience on the subject's matter and estimation made often too difficult and confusing for many. This paper discusses implementation of Objective Standard Setting (OSS) using Rasch measurement model in a Malaysian Institution of Higher Learning. This method gives an advantage for academic administrators to scientifically establish the standards or cut scores focusing on validity of the test being used and the Rasch measurement properties of the resultant scale.

Keywords: objective standard settings, measurement and evaluation, Rasch measurement model, systematic

1. Introduction

1.1 Introduce

Engineering programs in Malaysia have been progressively going through changes in making sure higher learning institutions produced quality graduates fit for the working environment. Graduates not only obtained good grades but are being prepared for the industry through trainings as well. It is in line with the aims of the ministry who wants to produce graduates who are competent and able to obtained employment in the relevant field. In fact, the Ministry of Higher Education (MOHE) has formulated a strategic plan for nurturing a knowledgeable, innovative and first class human capital in 2020. The plan emphasizes on the improving quality of teaching and learning including enhancing the system of student assessment by paying attention to the students' achievement other than the writing test (http://www.mohe.gov.my/psptn/).

Besides MOHE, the Malaysian Qualifications Agency (MQA) provided the standards and framework for engineering programs in making sure that the institutions of higher learning are able to produce quality graduates (MQA, 2011). MQA also ensures quality of the courses offered either in local (IPTA) or private institutions of higher learning (IPTS). Code of Practice for Programme Accreditation (COPPA) and Code of Practice for Institutional Audit (COPIA) were developed to assess the courses offered in higher education institutions. Some key aspects which are guaranteed by MQA are a continuous quality improvement and student assessment (http://www.mqa.gov.my). Besides ensuring the goal and learning outcomes of the courses are met, students' achievement also are being monitored to produce competitive and quality students in the face of global competition in line with the aspirations and the National Education Philosophy. Measurement, evaluation and assessment were done with continuous effort in meeting the expected outcomes, without foregoing quality.

At the same time, Faculty of Engineering & Built Environment (FKAB) performs a variety of teaching and learning methods to meet the needs of the industry hence accordance with accreditation requirements. Subsequently, FKAB make improvements from time to time to make changes in the methods of teaching and

learning in stages. Measurement of quality of students always been discussed especially in determining the achievement lines. Not only in IPTA and IPTS, Ministry of Education also agrees with the importance of achievement lines. In year 2012, only 3,000 among 12,000 polytechnics graduates offered to further their studies in local institutions of higher learning. It is because the scoring standard, Cumulative Grade Point Average is lower than the Matriculation and Malaysian Higher School Certificate (STPM) colleges. Because of that issue, the Ministry of Education has called all polytechnics' management to revise the examination standard and grading system to ensure their graduates managed to further their study in local universities (Utusan, 2012).

Given that this issue arose, it clearly shows that setting the benchmark standard for the students to pass the exam is also important to improve quality of teaching and learning, thus improving quality of students in an educational institution. Quality assessment is expected to produce credible information that facilitates critical decision making (Zamaliah et al., 2011; Stone, 1995). It depends on what are expected and calibrated instruments that provide the test-takers various levels of measures for decision makers to judge what is best and less achievement (Wright, 2000). It has always been an issue when deciding where the lines of achievement should be, taking consideration of students' measures and quality that been set.

2. Standard Setting Review

The process of teaching and learning is expanding and the standard setting procedure is one of the important aspects to improve teaching and learning. In preserving and ensure the test results will be useful and defensible, standard setting ideally involves with policy makers, developers test and measurement expert, not just a methodological process (Bejar, 2008).

Standard setting can be defineas a procedure of establishing one or single cut score (such as pass/fail and allow/deny) or more (multiple) cut score (e.g. level of achievement such as weak, moderate, excellent) in a single test depends on what test requirement. This cut score function as separation of two or more categories required by the test (Cizek & Bunch, 2007). Meanwhile, other definitions of standard setting besides determines the level of achievement or mastery is a method applied to obtain the corresponding cut score which can classify students who are under cut score to a higher level (Bejar, 2008).

There are many procedures for standard setting such as Angoff, Nedelsky, Ebel, Bookmark and Objective Standard Setting. Abu Kassim (2007) argued that educational standard is lack of objectivity due to the unsystematic judgement of the experts when constructing the cut scores. Angoff had its way by averaging the subject-matter experts predicted difficulty of the items (James et al., 1998) in which described by Sick (2009) as "too difficult and confusing". Objective (OSS) model and Bookmark represent clearly departures from Angoff model, Ebeland Nedelsky (Stone, 2011). Compared with both model, Angoff failed to define a legitimate and stable construct and avoids the item by item judgement may be tedious and difficult to judges (Stone, 2011; MacCann & Stanley, 2006).

Objective Standard Setting through Rasch measurement model has shown efficacy in psychometric of the stable formation of more reasonable standard (Stone, 1995). It can lead to the acceptable benchmark of the test taker. This method asks the experts who act as judges to determine the areas of knowledge required directly from the items or questions in the test. These experts will also need to set standard set for the course or subjects, at appropriate benchmark on the measuring line between essentialand higher items.

This paper explores the objective standard setting with constructed-response items with the use of a research-based approach in the quantification of the judges' ratings.

3. Method

Objective Standard Setting method is based on the study conducted by Wright and Grosse (Wright & Grosse, 1993; Wright & Grosse, 1987). This method combines between expert judgement on the test content and consideration of the ability and the difficulty of students-item respectively.

3.1 Objective Standard Setting Methodology

This method involves evaluative decisions and translation in the quantification of the standard. The evaluative decisions steps are as below;

a) Expect judgement on essential and optional items within the test.

b) The level of mastery required and the quantification of the standards.

a) Calculation of the mean item difficulty for all the essential items within respective grade separation; μ difficulty.

b) Translating the mastery level decision onto the logit measurement.

c) Calculating the standard error (SE) of each grade separation; for the allowance of grade demarcation changes.

The steps to fix the cut score is outlined are;

a) Judges to determine the essentiality of items in the prescribed test; E item, thus establishing total of E items; ΣN_E

b) Considering the number of grade separation, and to divide the number of total essential items with the number of grade separation; $\Sigma N_E/N_S$ to give the number of items per grade.

c) Calculate the mean difficulty of essential items within each grade; $\mu_{item.}$ Include also calculation for mean SE; $\mu_{SE.}$

d) Decision by the expert to set the 'mastery' level, which is normally set at 60%. Therefore, 60% out of the total number of essential items would give the number of items actually needed to achieve the mastery level taking consideration from the easiest essential items, in ascending order from item measure table, thus having the logit measurement.

Upon having the logit measure on accepTable 30% essential items, then map it on to the person measure table; to consider the number of accepted students in mastery level.

3.2 Participant

217 of the first year students of engineering faculty at a Malaysian institution of higher learning involved in this study. It comprises of four (4) disciplines of the engineering program; civil structural, mechanical, chemistry process, electric electronic. The students distributions based on department and gender are shown in Table 1.

Table 1. Distribution based on department and gender

		N	%
Department	Civil & structural	46	21.2
	Electric & electronic	56	25.8
	Chemical & process	47	21.7
	Mechanics & materials	68	31.3
Gender	Male	134	61.8
	Female	83	38.2

3.3 Instruments

The students' raw scores gathered from final examination. 19 items of linear algebra (KKKM1224) final examination paper were use in this study. Linear Algebra is one of the engineering mathematics subjects in UKM which discussed on complex number, hyperbolic functions, power series expansion, matrix and determinant and important applications of linear algebra in engineering. This course will highlight to the students on importance of linear algebra in engineering (Undergraduate Handbook, 2011). All items were given to judges to review their essentiality. Essential items refer to the major item which significant to achieve learning outcome and meet the objective of the course or subject. The details of test items are summarise in Table 2.

Table 2. Topics for each examination question

Part	Question. No		Marks
A	A1	Determinant matrix	10
	A2	Positive integer	10
	A3	Eigenvectors	12
	A4	Diagonalizable	8
B	B5	Transition matrix	9
	B6	Vector transition	4
	B7	Linear combination vector	4
	B8	Graphic linear	3
C	C9	Subspace 2x2	9
	C10	Subset of subspace	6
	C11	Basis of subspace	4
	C12	Not subspace	1
	C13	Orthogonal product	2
	C14	Evaluate orthogonal	4
	C15	Gram-Schmidt orthogonal	14
	C16	Elastic deformation	8
	C17	Deformation boundary	2
	C18	Conic quadratic	8
	C19	Quadratic graph	2

3. Results

The raw scores were then run using Winsteps version 3.68.2, a Rasch analysis software. To ensure that there are no misfit items which can affect the result for standard setting, the test items need verification procedure before proceed with OSS procedure. In Rasch analysis, there are three-step comparison procedures to determine the test items misfit. The procedure start with point measure correlation (x), followed by outfit MNSQ and then with outfit ZSTD (z) value. Each control have accepted region given as follows; $0.4 < x < 0.8$, $0.5 < MNSQ < 1.5$ and $-2.0 < z < 2.0$. Item misfit if all controls did not meet the accepted region respectively. Figure 1 shows the Item-Measure Order.

Item STATISTICS: MEASURE ORDER

```
-------------------------------------------------------------------------------------------------------
|ENTRY   TOTAL                    MODEL|   INFIT  |  OUTFIT  |PT-MEASURE |EXACT MATCH|                      |
|NUMBER  SCORE  COUNT  MEASURE    S.E. |MNSQ  ZSTD|MNSQ  ZSTD|CORR.  EXP.| OBS%  EXP%| Item                G |
|-------------------------------------+----------+----------+-----------+-----------+----------------------|
|   13     91    194    1.27      .08| .71  -1.3| .86   -.5| .22   .30| 58.8  61.0| C13 Orthogonal product   1 |
|   11     14     46    1.08      .18|1.20    .6|1.04    .3| .31   .38| 80.4  78.6| C11 Basis of subspace    0 |
|    8    238    217     .40      .06| .95   -.8| .94   -.7| .40   .34| 15.2  17.6| B8 Graphic linear        F |
|   19    152    192     .35      .08| .89  -2.1| .87  -2.3| .46   .26| 22.9  18.8| C19 Quadratic graph      0 |
|   17    195    192     .07      .08| .90  -2.3| .89  -2.4| .44   .26| 17.7  17.4| C17 Deformation boundary 0 |
|    7    461    217     .02      .04| .92  -1.3| .91  -1.2| .49   .41| 16.1  16.9| B7 Linear combination vector 0 |
|    6    468    217     .01      .04| .87  -2.3| .83  -1.9| .54   .44| 10.6  11.0| B6 Vector transition     7 |
|    3   1603    217    -.03      .03|1.03    .3|1.02    .3| .60   .60| 14.7  13.9| A3 Eigenvectors          B |
|    5   1156    217    -.04      .02|1.08   1.0|1.01    .1| .60   .61| 16.1  16.8| B5 Transition matrix     9 |
|    4   1135    217    -.08      .03|1.15   1.6|1.38   3.0| .43   .50| 22.1  20.0| A4 Diagonalizable        1 |
|    9    258     46    -.12      .05|1.09    .5|1.40    .9| .55   .58| 26.1  24.6| C9 Subspace 2x2          0 |
|   16   1102    192    -.13      .04| .99    .0| .94   -.3| .51   .49| 25.0  23.5| C16 Elastic deformation  0 |
|   18   1179    192    -.19      .03| .99   -.1| .90   -.4| .51   .48| 34.4  26.9| C18 Conic quadratic      0 |
|   15   1940    194    -.22      .02|1.08    .8|1.22   1.7| .59   .63|  8.2  11.1| C15 Gram-Schmidt orthogonal J |
|    1   1500    217    -.35      .03|1.31   3.4|1.43   3.3| .39   .56| 12.9  15.1| A1 Determinant matrix    2 |
|   10    223     46    -.35      .07| .91   -.3| .56    .1| .47   .42| 37.0  34.3| C10 Subset of subspace   K |
|   12     28     46    -.40      .31| .95   -.7| .94   -.6| .31   .17| 67.4  61.5| C12 Not subspace         H |
|   14    661    194    -.55      .06| .91   -.7| .76  -1.1| .43   .33| 52.1  49.4| C14 Evaluate orthogonal  0 |
|    2   1980    217    -.74      .04|1.00    .1| .90   -.4| .41   .40| 35.9  36.1| A2 Positive integer      0 |
|-------------------------------------+----------+----------+-----------+-----------+----------------------|
| MEAN  757.1  172.1     .00      .07|1.00   -.2|          |           | 30.2  29.2|                      |
| S.D.  647.1   66.0     .48      .07| .13   1.4|          |           | 20.1  19.0|                      |
-------------------------------------------------------------------------------------------------------
                                                 2          3           1
```

Figure 1. Item-Measure order

From Item-Measure order table, there are no misfit items in this examination question paper. Then, we can proceed to the OSS procedure with all 19 items.

Raw scores always mislead our perception on students' ability. It does not indicate the marks earned are from difficult items to confirm their valid ability. Figure 2 shows that student 197 with score 93, student 65 with score 91, student 67 with score 86 are better ability students as compared to their corresponding higher scores students namely student 29 with score 96, student 138 with score 92, student 16 with score of 87. The latter students obtained higher scores from easy items, where else the former, despite having low scores but obtained the marks from more difficult items. Rasch position them of better ability according to their ascending ability order. Raw scores is only indication of achievement but not a measure of ability. OSS gives a better psychometry of the students' ability.

```
-----------------------------------------------
|ENTRY   TOTAL  TOTAL             MODEL|   |       |
|NUMBER  SCORE  COUNT  MEASURE    S.E. |   | Person|
|-------------------------------------+-- +-------|
|  197     93     16    1.03      .32|   | 1974MC|
|   29     96     15    1.02      .27|   | 0291FC|
|  113     95     15     .95      .27|   | 1132MM|
|  160     94     15     .94      .25|   | 1603FC|
|   65     91     16     .85      .29|   | 0652MC|
|  138     92     15     .81      .25|   | 1383MC|
|  100     92     15     .75      .25|   | 1002FC|
|   58     89     16     .70      .26|   | 0582MC|
|  190     91     15     .69      .24|   | 1904MC|
|   61     90     15     .63      .23|   | 0612MC|
|  178     90     15     .63      .23|   | 1784MC|
|  105     89     15     .58      .22|   | 1052MM|
|  174     88     15     .58      .22|   | 1744MC|
|    4     88     15     .53      .21|   | 0041FM|
|  182     88     15     .53      .21|   | 1824MM|
|   67     86     16     .53      .22|   | 0672MC|
|   16     87     15     .49      .20|   | 0161FC|
|   51     87     15     .49      .20|   | 0512MM|
```

Figure 2. Snippet of the person measure table

Setting the standards starts by identifying the essential items among the items prescribed in the test. Experts in the area of linear algebra are to identify essential items within the 19 items in the test. This is crucial in

determining the topics which deemed important and for the students to achieve in mastering linear algebra. For the purpose of this paper, an expert has identified 16 items as essential items. Those items indicated with 'E' in Figure 3.

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	ZSTD	OUTFIT MNSQ	ZSTD	PT-MEASURE CORR.	EXP.	EXACT MATCH OBS%	EXP%	Item	
13	91	194	1.27	.08	.71	-1.3	.86	-.5	.2	Group 8		61.0	C13 Orthogonal product	E
11	14	46	1.08	.18	1.20	.6	1.04	.3	.3			78.6	C11 Basis of subspace	E
8	238	217	.40	.06	.95	-.8	.94	-.7	.4	Group 7		17.6	B8 Graphic linear	E
19	152	192	.35	.08	.89	-2.1	.87	-2.3	.4			18.8	C19 Quadratic graph	
17			.07	.08	.90	-2.3	.89	-2.4	.4			17.4	C17 Deformation boundary	E
			.02	.04	.92	-1.3	.91	-1.2	.4	Group 6		16.9	B7 Linear combination vector	E
			.01	.04	.87	-2.3	.83	-1.9	.5			11.0	B6 Vector transition	E
			.03	.03	1.03	.3	1.02	.3	.6	Group 5		13.9	A3 Eigenvectors	E
5	1156	217	-.04	.02	1.08	1.0	1.01	.1	.6			16.8	B5 Transition matrix	E
4	1135	217	-.08	.03	1.15	1.6	1.38	3.0	.4	Group 4		20.0	A4 Diagonalizable	E
9	258	46	-.12	.05	1.09	.5	1.40	.9	.5			24.6	C9 Subspace 2x2	E
16	1102	192	-.13	.04	.99		.94	-.3	.5	Group 3		23.5	C16 Elastic deformation	E
18	1179	192	-.19	.03	.99	-.1	.90	-.4	.5			26.9	C18 Conic quadratic	
15	1940	194	-.22	.02	1.08	.8	1.22	1.7	.5			11.1	C15 Gram-Schmidt orthogonal	E
1	1500	217	-.35	.03	1.31	3.4	1.43	3.3	.3	Group 2		15.1	A1 Determinant matrix	E
10	223	46	-.35	.07	.91	-.3	.56	.1	.4			34.3	C10 Subset of subspace	E
12	28	46	-.40	.31	.95	-.7	.94	-.6	.3	Group 1		61.5	C12 Not subspace	E
14	661	194	-.55	.06	.91	-.7	.76	-1.1	.4			49.4	C14 Evaluate orthogonal	E
2	1980	217	-.74	.04	1.00	.1	.90	-.4	.41	.40	35.9	36.1	A2 Positive integer	
MEAN	757.1	172.1	.00	.07	1.00	-.2	.99	-.1			30.2	29.2		
S.D.	647.1	66.0	.48	.07	.13	1.4	.22	1.5			20.1	19.0		

Cut score 60% at -0.03

Figure 3. Item measure table

Usually, the grading system has 5 separations; A, B, C, D, F. However, due to leniency of grading, this academic institution has 11 separations; A, A-, B+, B, B-, C+, C, C-, D+, D, F; then dividing all the items in the test with the grades separations yielding approximately about 2 items for each grade. The separation should starts from the lowest item in measure table as in Figure 3 which sorted according to their respective difficulty logit measure with the lowest is the easiest items.

Distribution of essential items according to the 11 grading system causing the items being segregated into 8 groups altogether. Table 3 list out the items according to their respective groups and difficulty measures, from the lowest;

Table 3. Groups and respective logit measures

	Items	Logit measure
Group 1	C14	-0.55
	C12	-0.40
Group 2	C10	-0.35
	A1	-0.35
Group 3	C15	-0.22
	C16	-0.13
Group 4	C9	-0.12
	A4	-0.08
Group 5	B5	-0.04
	A3	-0.03
Group 6	B6	+0.01
	B7	+0.02
Group 7	C17	+0.07
	B8	+0.40
Group 8	C11	+1.08
	C13	+1.27

The logit measures for all the items within each groups are then summed up and divided by the number of items for each group giving the mean value for each group. The mean value will act as the cut score value for that

particular group. Similarly, mean of each standard error has calculated for each group. Table 4 shows the list of groups with mean and mean error. The mean standard error is allowance of range systematically provided in situations where decision need be made on reducing or increasing the cut score for each group grades.

Table 4. Groups with mean and mean error

	Mean logit; μ	μS.E
Group 1 F	-0.48	0.19
Group 2 D	-0.35	0.05
Group 3 D+	-0.18	0.03
Group 4 C-	-0.10	0.03
Group 5 C	-0.04	0.03
Group 6 C+	0.02	0.04
Group 7 B-	0.24	0.07
Group 8 B	1.18	0.13

However, refer to Figure 3 again, take note that some of the items have similar logit measures and some have narrow measures between the items. For example items in Group 2, both items; A1_Determinant matrix and C10_Subset of subspace, both having same measures at -0.35 logit. Rasch analysis provides empirical evidence that the items are redundant. Lucky enough that both items are from different topics or dimension.

Items in Group 5 and 6, which the measures are so narrow, are vague in their grade separation. Considering the measurement for item B6_Vector transition at +0.01, SE 0.04; on the high side the item could be located at +0.05 logit and the lower end at -0.03 logit. At the worst scenario, the item has crossed over the Group 6 separation instead. This is another consideration to note before embarking to OSS where items need calibrated properly. However, for this exercise in view of insufficient items, so the process has skipped.

The students expected to meet 60% 'mastery' level of among the essential items in order for them being considered as meeting the objective of linear algebra subject. This gives the demarcation point of roughly meeting or achieving 10 out of the essential 16 items; marked at -0.03 logitmeasure or at A3_Eigenvectors item. Mapping the mean logit measures on to the person measure table where the person sorted in ascending order according to their respective ability measures.

Figure 4. Snippet of person measure table

```
. ENTRY  . TOTAL . TOTAL .          . MODEL .
! NUMBER ! SCORE ! COUNT ! MEASURE ! S.E  !     NAME !
.   197 .   93 .   16 .   1.03 .  0.32 .  1974MC .
!   201 !   79 !   15 !   0.25 !  0.15 !  2014MM !    B-
.     2 .   78 .   15 .   0.23 .  0.15 .  0021MM .

!   152 !   66 !   15 !   0.02 !  0.12 !  1523FI !    C+
    36 !   65 !   15 !   0.01 !  0.12 !  0361FM !    C
    62 !   64 !   15 !   0.01 !  0.12 !  0622MC !
    94 !   65 !   15 !   0.01 !  0.12 !  0942FM !
   142 !   65 !   15 !   0.01 !  0.12 !  1423FM !
   154 !   65 !   15 !   0.01 !  0.12 !  1543FM !
   177 !   64 !   15 !   0.01 !  0.12 !  1774MC !
   161 !   63 !   16 !      0 !  0.12 !  1613MC !
   168 !   63 !   16 !      0 !  0.12 !  1684MM !
    37 !   64 !   15 !  -0.01 !  0.12 !  0371MM !
    43 !   64 !   15 !  -0.01 !  0.12 !  0431FM !
    72 !   62 !   15 !  -0.01 !  0.11 !  0722MM !
   156 !   64 !   15 !  -0.01 !  0.12 !  1563FM !
   205 !   64 !   15 !  -0.01 !  0.12 !  2054MM !
     9 !   63 !   15 !  -0.02 !  0.12 !  0091FM !
    26 !   63 !   15 !  -0.02 !  0.12 !  0261MM !
 ┌─────────────────────────┐  -0.02 !  0.12 !  0972FM !
 │ Cut score 60% at -0.03  │  -0.02 !  0.12 !  1233FC !
 └─────────────────────────┘  -0.02 !  0.12 !  1303FM !
   130 !   63 !   15 !  -0.02 !  0.12 !  1303FM !
   170 !   63 !   15 !  -0.02 !  0.12 !  1704FM !
     5 !   62 !   15 !  -0.04 !  0.12 !  0051FM !
    23 !   62 !   15 !  -0.04 !  0.12 !  0231FM !
    27 !   62 !   15 !  -0.04 !  0.12 !  0271MM !
   109 !   49 !   12 !  -0.04 !  0.13 !  1092MC !
   171 !   60 !   15 !  -0.04 !  0.11 !  1714FM !
    74 !   61 !   15 !  -0.05 !  0.12 !  0742MM !    C-
    89 !   61 !   15 !  -0.05 !  0.12 !  0892FM !
   153 !   61 !   15 !  -0.05 !  0.12 !  1533FM !
   167 !   61 !   15 !  -0.05 !  0.12 !  1674FM !
    95 !   60 !   15 !  -0.06 !  0.12 !  0952FM !
   115 !   60 !   15 !  -0.06 !  0.12 !  1153MM !
   119 !   59 !   16 !  -0.06 !  0.11 !  1193MM !
 ┌─────────────────────────┐  -0.08 .  0.12 .  0301MC .
 │ Cut score 60% at -0.06  │
 └─────────────────────────┘
    35 !   57 !   15 !   -0.1 !  0.12 !  0351MM !
   122 !   54 !   15 !  -0.11 !  0.11 !  1223MC !    D+
   202 !   52 !   15 !  -0.17 !  0.12 !  2024FM !
    66 !   51 !   15 !  -0.18 !  0.12 !  0662MM !    E
   207 !   41 !   15 !  -0.32 !  0.12 !  2074MM !
    75 !   38 !   15 !  -0.36 !  0.12 !  0752MM !    F
```

Figure 4 of the person measure reveals that the 'mastery' point is within 'C' group, however, five (5) of the students did not reach the -0.03 logit measures even though there are within 'C' group in the conventional fixed-grade setting. This reveals, even though that the five (5) students are within grade 'C' but they do not reached 'mastery' acceptance of linear algebra; they may only managed to answer correctly on easy items and have some difficulties answering difficult items.

Having considering passing 'mastery' level of linear algebra at 60% which marked at -0.03 logit, revealing that 139 out of 217 students managed to reached the mastery level. This is equal to 64% passes. However, if the academic administrators decide to have high number of passes, the SE value could be use. Rasch applies the SE of measures as the limit for variance. Referring to Table 5, the 60% mastery level measure is at -0.03 logit, SE 0.03, will have a lower limit of -0.06 logit. This allows 12 students for possible inclusion as deemed mastery. Now the number of passes has increased to 151 (from 139) or 70% (from 60%). The academic administrators can report such achievement at high 70% of mastery skill achieved and yet does not breach the ethics and values of morality. This is contrary to many practices to achieve higher percentages of passes by arbitrarily increased the marks of all students.

However, the exercise in increasing the number of passes allowed the grade C- students to pass the test. Therefore, here, the academic administrators have to decide whether to allow the increase of passes but on other hand jeopardizing quality of the 'passes'.

4. Conclusions

The Ministry of Higher Education has formulated a strategic plan to improve quality of teaching and learning in effort producing quality and competitive students. In the meantime, education standards have always been an issue of learning institutions where it involved experts' judgments that is still chaotic in some ways. Judges perception on essentiality items may vary depending on their scope requirement, knowledge and expectation of their expert nature. Objective Standard Setting (OSS) using Rasch measurement model facilitates the various perceptions of experts in deciding systematically on the cut-off line of mastery and performance levels of the test-takers without jeopardizing quality of the performance. Rasch logit can measure students' ability and an approach where the decision is being more objectively on measure scale (Zamaliah et al., 2011). This method could be used by academics and the grade coordinator as an alternative method for each course offered at education institutions. This method is one of the efforts to overcome factors affecting examination quality.

References

Abu Kassim, & Noor Lide. (2007). Using TheRasch Measurement Model for Standard Setting of the English Language Placement Test at the IIUM. *PhD thesis, Universiti Sains Malaysia*. Retrieved September 27, 2012, from http://eprints.usm.my/8328/

Bejar, I. (2008). Standard Setting: What Is It? Why Is It Important? *R & D Connections, 7,* 1-6. Retrieved January 25, 2013, from http://www.ets.org/Media/Research/pdf/RD_Connections7.pdf

Cizek, G. J., & Bunch, M. B. (2007). What is Standard Setting. In *Standard Setting: A Guide to Establishing and Evaluating Performance Standards on Tests* (Chapter 2, pp. 13). *SAGE Publications.*

Faculty of Engineering and Built Environment, UKM. (2011). 2011-2012 Undergraduate Handbook for Faculty of Engineering and Built Environment. Malaysia, Bangi: UniversitiKebangsaan Malaysia

Grosse M. E., & Wright B. D. (1987). Criterion Item Standard Setting. *Research Papers and Memoranda*. Retrieved August 3, 2012, from http://www.rasch.org/memo84.pdf

James, C. I., & Barbara S. P. (1998). Teacher's Ability to Estimate Item Difficulty: A Test of the Assumptions in the Angoff's Standard Setting Method. *Journal of Educational Measurement, 35*(1), 69-81.

MacCann R. G., & Stanley Gordan. (2006). The Use of Rasch Modeling to Improve Standard Setting. *Practical Assessment Research & Evaluation, 11*(2), 1-17. Retrieved September 20, 2012, from http://pareonline.net/pdf/v11n2.pdf

Malaysian Qualification Agency. (2011). *Programme Standards: Engineering and Engineering Technology*. Malaysia: Selangor. Retrieved from http://www.mqa.gov.my/portal2012/garispanduan/standard%20kejuruteraan_en.pdf

Malaysian Qualifications Agency. (2012). Quality Assurance. Retrived November 20, 2012, from http://www.mqa.gov.my/

Ministry of Higher Education. (2012). Improvement on quality teaching and learning. 2011. Strategic Plan, National Higher Education. Retrieved November 6, 2012, from http://www.mohe.gov.my/psptn/

Sick, J. (2009). Rasch Measurement in Kanguage Education Part 3: The family of Rasch Models. *Shiken: JALT Testing & Evaluation SIG Newsletter, 13*(1), 4-10.

Stone, G. E. (1995). Standard Setting Methods. *Rasch Measurement Transactions, 9*(3), 452.

Stone, G. E., Kristin, L. K. Koskey, & Toni A. Sondergeld. (2011). Comparing Construct Definition in the Angoff and Objective Standard Setting Models: Playing in a House of Cards without a Full Deck. *Educational and PsychologicalMeasurement, 71*(6), 942-962. http://dx.doi.org/10.1177/0013164410394338

Utusan Malaysia. (August 5, 2012). *Scoring Standards in Polytechnic must be review*. Retrieved August 8, 2012, from

http://www.utusan.com.my/utusan/Dalam_Negeri/20120805/dn_08/Piawaian-peperiksaan-politeknik-perlu-dikaji---Khaled

Wright, B. D. (2000). How to Set Standard. *Rasch Measurement Transactions, 14*(1), 740.

Wright, B. D., & Grosse M. (1993). How to Set Standards. *Rasch Measurement Transactions, 7*(3), 315-316.

Zamaliah, M., Saidfudin, M., Azrilah, A. A., & Zaharim. (2011). Standard Setting for Engineering Education Achievement Test in an Institution of Higher Learning Malaysia. *Proceeding 10th WSEAS International Conference on Education and Educational Technology (EDU'11)*, Penang, Malaysia, October 3-5, 2011.