SOCIAL and MANAGEMENT RESEARCH JOURNAL

Institute of Research Management Innovation (IRMI)

Social and Management Research Journal is a journal by Institute of Research Management & Innovation (IRMI),Universiti Teknologi MARA, Bangunan Wawasan, Level 3, 40450 Shah Alam, Selangor Darul Ehsan,Malaysia. E-mail: irmiuitm@salam.uitm.edu.my

The views, opinions and technical recommendations expressed by the contributors and authors are entirely their own and do not necessarily reflect the views of the editors, the publisher and the university.

SOCIAL and MANAGEMENT RESEARCH JOURNAL

Institute of Research Management Innovation (IRMI)

Simplified Reliable Online Essay Test Marking for Massive Open Online Course (MOOC) using Rasch Model Analysis

Mohd Nor Mamat[1], Zawawi Temyati[2], Siti Fatahiyah Mahamood[1], Hanifah Musa[1]

[1]Academy of Contemporary Islamic Studies
Universiti Teknologi MARA Malaysia, Shah Alam 40450 Selangor

[2]Institute of Neo Education
Universiti Teknologi MARA Malaysia, Shah Alam 40450 Selangor

E-mail: mohdnoor@salam.uitm.edu.my

ABSTRACT

Manual practice in formal examination does not assess accurate measure of a student's ability, as it merely counts the score of every question to be considered for the student's grade. There are many educators who have used raw score as a form of measurement for a student's ability, but it never truly measures the right measurement. The raw score should be converted into the right linear metrics for ability measurement. This procedure contains measuring score of accurate student's ability in LOGIT unit, providing of student's result profile, and measuring reliability of the test set and the student's answers. The procedure is designed for massive open online learning and paperless essay-based test which is more difficult to be analysed. This procedure converts the student's answer into rubrical ratio-based scale to be more accurately measured. It is definitely better than the common practice of merely analysis on raw marks for each question. It would show true student's performance of cognitive performance (test) which represents the true student's ability (in LOGIT unit), in order to accurately measure the right outcome. This new paradigm of assessment is fit to be applied for massive numbers on online students. It uses Rasch model which offers reliable solution in producing accurate ability marks for students, together with scientific reliability score for student's answer.

Keywords: *student's performance, accurate measure, MOOC, rasch model*

INTRODUCTION

Evaluation is an important component of education. The right instrument and right method make assessment easier and more accurate. Outcome-based learning calls for the evaluation of the subject learning outcomes, as specified in every course description. For years, the evaluation method has been largely dependent on student's performance in tests, quizzes or submission of assignments with major assessment in the final examination. Traditionally, marking on papers and accumulating the raw marks are common practice. However, measurement of student's achievement was merely counted from the marks allocated for every question. The observed outcomes remain vague. Very rarely in our practice do we provide the construct validity score or the questionnaire reliability score for the set of questions. It would be more difficult in huge class or massive open online course. This is a conceptual paper which describes a simple procedure for assessing student's online performance using Rasch Analysis that can be used to measure any course's learning outcomes.

PROBLEM STATEMENT

In educational assessment, there are still many educators who use raw scores as a form of measurement for student's ability, but raw scores have never been a true measure (Aziz, A. A., 2012; 2008; Bond, 2007; Wright, 1989). Not all educators are proper researchers. Therefore, there is a crucial need to provide a simplified model as a bridge between accurate scientific analysis in research and raw educational assessment, as well as to upgrade the quality of student's ability assessment to be more accurate and reliable. In addition, it is difficult to analyse student's response especially in essay-based test. In addition to that, online marking and paperless essay test for a huge open number of students like MOOC would be more complicated, as it should be printed and marked one by one. This simplified procedure proposes the method of rubrical marks for student's responses without any need of printed answer script.

RESEARCH OBJECTIVES

(a) To implement a rubrical analysis or assessment for the massive open online essay-based test using the Rasch Measurement Model.

(b) To produce a reliable and accurate measurement procedure for students' essay-based performance.

DISCUSSION 1: MOOC ONLINE MARKING AND ANALYSIS PROCEDURE USING RASCH MODEL

MOOC is the massive open online learning which is open to all. It was purposely designed for huge numbers of students and require a procedure that caters to the needs of massive numbers of exercises and tests. We need a new paradigm of test evaluation online marking with simple but reliable steps. The proposed simplified procedure needs only five simple steps to be applied. The answers will be marked online according to designated rubrics. This pilot study was conducted based on the Environmental Ethics (IPK661) June 2016 Final Examination paper of a class of 15 students from the Environmental Technology programme (EVT229), Faculty of Applied Science, Universiti Teknologi MARA Malaysia. The set has one question with four sub questions, and three questions with three sub questions. Each sub question carried one, two or three marks. The answers have been submitted via online drawer and the rubrical marks would be keyed into EXCEL format and saved as prn format. With one click, it opens the .prn file via Bond& Fox Steps. With a few more clicks, it displays the reliability score for the question set and student's answer. The next step is to relook or review the question item to determine if the reliability score is low or not acceptable. With one more click, we have the result of students (in rank order) within log odd unit (logit) unit or probabilistic ratio-based measure. Any kind of human ability measure should be assessed using this kind of measurement, not merely calculating the marks given to each question. This method provides a more comprehensive analysis of each student's unique problems. The last step was to provide percentile marks converted from the logit unit which is based on probabilistic calculation in accordance to norm-based assessment and difficulty level of question set (refer Table 1).

A comprehensive but simplified procedure of measurement has been done and it could be simplified as follows:

Table 1: Five Simplified Steps

STEPS	ACTIVITIES	TOOLS
1st step	Students' answers scripts will be read online (paperless)	Student's Answer Online
2nd step	Marks will be given in according to rubrics	Rubrics Table
3rd step	Reliability Analysis (using analysis by Rasch model)	Excel and Bond& Fox Steps
4th step	Students' profiling, reliability score will be produced	Bond&Fox Steps
5th step	Probability Analysis for Percentile Educational Grade	M2M Template

Table 2: Raw Marks vs Rubrical Score

Type	Rubrics Code			
Marks/Answer	1	2	3	4
1 mark	0	0.1 - 0.33	0.34 - 0.67	0.68 - 1
2 marks	0	0.1 - 0.66	0.68 - 1.33	1.34 - 2
3 marks	0	0.1 - 1	1.1 - 2	2.1 - 3

The rubrical value would be keyed in to EXCEL format (refer Table 2) and saved as .prn format. The .prn file would be opened using Bond&Fox Steps. A few clicks resulted in the display of reliability score for the question set and students' answers which were acceptably reliable (ir=0.56) and (pr=0.63) with Cronbach alpha value at 0.65 (refer Figure 1). In future practice, this result will help teachers and educators to evaluate the quality of test questions before distributing to the students.

```
TABLE 3.1 FINAL EXAM RESULT                          ZOU713W5.TXT Aug 18 12:15 2016
INPUT: 15 Persons  44 Items  MEASURED: 15 Persons  44 Items  4 CATS        1.0.(

      SUMMARY OF 15 MEASURED Persons
+------------------------------------------------------------------------------+
|           RAW                       MODEL       INFIT          OUTFIT         |
|          SCORE     COUNT   MEASURE  ERROR    MNSQ   ZSTD    MNSQ   ZSTD       |
|------------------------------------------------------------------------------|
| MEAN     130.9     44.0      .45     .16     1.02     .0    1.00    -.1       |
| S.D.      10.6      .0       .29     .02      .24    1.2     .25    1.1       |
| MAX.     151.0     44.0     1.06     .20     1.49    2.3    1.52    2.1       |
| MIN.     114.0     44.0      .03     .15      .63   -2.3     .61   -2.1       |
|------------------------------------------------------------------------------|
| REAL RMSE   .17  ADJ.SD   .23  SEPARATION 1.32  Person RELIABILITY  .63      |
| MODEL RMSE  .16  ADJ.SD   .24  SEPARATION 1.44  Person RELIABILITY  .67      |
| S.E. OF Person MEAN = .08                                                    |
+------------------------------------------------------------------------------+
Person RAW SCORE-TO-MEASURE CORRELATION = 1.00
CRONBACH ALPHA (KR-20) Person RAW SCORE RELIABILITY = .65

      SUMMARY OF 44 MEASURED Items
+------------------------------------------------------------------------------+
|           RAW                       MODEL       INFIT          OUTFIT         |
|          SCORE     COUNT   MEASURE  ERROR    MNSQ   ZSTD    MNSQ   ZSTD       |
|------------------------------------------------------------------------------|
| MEAN      44.6     15.0      .00     .29      .98    -.1    1.00     .0       |
| S.D.       5.7      .0       .47     .06      .44    1.5     .48    1.5       |
| MAX.      57.0     15.0      .98     .59     1.99    2.6    2.17    2.8       |
| MIN.      30.0     15.0    -1.62     .24      .28   -3.5     .28   -3.2       |
|------------------------------------------------------------------------------|
| REAL RMSE   .31  ADJ.SD   .35  SEPARATION 1.12  Item   RELIABILITY  .56      |
| MODEL RMSE  .29  ADJ.SD   .37  SEPARATION 1.25  Item   RELIABILITY  .61      |
| S.E. OF Item MEAN = .07                                                      |
+------------------------------------------------------------------------------+
UMEAN=.000 USCALE=1.000
Item RAW SCORE-TO-MEASURE CORRELATION = -.97
660 DATA POINTS. APPROXIMATE LOG-LIKIHOOD CHI-SQUARE: 1536.50
```

Figure 1: Summary of Reliability and Cronbach Alpha Values

The last procedure in this study was to analyse the rubrical values of student's answers to be in a linear metrics, which is in logit unit.

Out of the 15 students, 13 passed the examination with scores between 51-80 marks, while two of them failed with 41 and 46.5 marks (refer Figure 3). After all these marks were converted into LOGIT unit, the result showed that all students have positive result, but only six of them exceeded the mean value (0.43) which means above 50%. Student's ability score was between 0.03 to 1.06 logit. It could be concluded that the use of raw marks for every question does not show the real values of student's ability. From the table below (refer Figure 2), we could conclude student with ID S6454 is excellent and able to answer all questions, and all students could answer the mean level of questions.

```
TABLE 17.1 FINAL EXAM RESULT                    ZOU713WS.TXT Aug 18 12:15 2016
INPUT: 15 Persons  44 Items  MEASURED: 15 Persons  44 Items  4 CATS      1.0.0
-------------------------------------------------------------------------------
Person: REAL SEP.: 1.32  REL.: .63 ... Item: REAL SEP.: 1.12  REL.: .56

         Person STATISTICS:  MEASURE ORDER

+------------------------------------------------------------------------------+
|ENTRY    RAW                  MODEL|  INFIT  |  OUTFIT |PTMEA|EXACT MATCH|      |
|NUMBER  SCORE  COUNT MEASURE  S.E. |MNSQ ZSTD|MNSQ ZSTD|CORR.| OBS%  EXP%|Person|
|------------------------------------------------------------------------------|
|   11    151    44    1.06     .20|1.49  1.7|1.52  1.6| .09| 38.6  51.0| S6454 |
|   12    147    44     .90     .19|1.11   .5| .99   .1| .50| 59.1  48.4| S2876 |
|    2    143    44     .77     .18| .83  -.7| .78  -.8| .43| 52.3  44.5| S4204 |
|    8    140    44     .67     .17|1.16   .8|1.19   .8| .17| 43.2  42.4| S3678 |
|    3    135    44     .53     .16|1.20  1.0|1.14   .7| .30| 40.9  37.5| S8264 |
|    4    135    44     .53     .16| .85  -.7| .82  -.8| .56| 38.6  37.5| S2457 |
|   10    130    44     .40     .16| .63 -2.3| .61 -2.1| .44| 40.9  37.0| S1634 |
|    1    128    44     .35     .16|1.11   .6|1.01   .1| .59| 27.3  36.9| S5522 |
|    6    128    44     .35     .16|1.45  2.3|1.47  2.1| .01| 31.8  36.9| S4537 |
|   14    128    44     .35     .16| .93  -.3| .92  -.4| .48| 36.4  36.9| S6712 |
|    7    127    44     .33     .16|1.02   .2|1.09   .5| .24| 43.2  36.2| S8145 |
|    5    123    44     .23     .15| .73 -1.7| .66 -2.0| .48| 47.7  34.9| S3886 |
|   15    118    44     .12     .15| .86  -.8| .85  -.8| .51| 25.0  33.0| S3498 |
|    9    117    44     .09     .15|1.07   .5|1.11   .7| .08| 34.1  31.4| S5116 |
|   13    114    44     .03     .15| .84 -1.0| .86  -.8| .56| 20.5  31.4| S7568 |
|------------------------------------------------------------------------------|
| MEAN  130.9   44.0    .45     .16|1.02   .0|1.00  -.1|    | 38.6  38.4|      |
| S.D.   10.6    .0     .29     .02| .24  1.2| .25  1.1|    |  9.9   5.6|      |
+------------------------------------------------------------------------------+
```

Figure 2: Person Fit Order

From the above analysis, all students fit into the model, which means those students are normal and their answers are reliable. The rank of student's ability was shown clearly, in which the student with ID S6454 scores 1.06 logit to be the most capable students in answering questions, while the student with ID S7566 scores the lowest value, which is 0.03 logit (refer Table 3).

Table 3: Raw Marks vs Ability Metrics Score

Students	Raw Marks (%)	Ability Metrics (LOGIT)
S2876	80	0.9
S7568	74	0.03
S6454	70	1.06
S4204	66.5	0.77
S8264	62.5	0.53
S3678	62.5	0.67
S2457	61	0.53
S1634	60.5	0.4
S6712	60.5	0.35
S4537	58.5	0.35
S5522	57.5	0.35

S3886	57	0.23
S5116	51	0.09
S8145	46.5	0.33
S3498	41	0.12

Using this procedure, students would be assessed according to their ability, not merely on counting raw marks provided for each question. After the analysis, it was found that the best student with raw marks (80) is not the best able student in logit (0.9). Similar to that finding, the weakest student with raw marks (41) is not the least able student in logit (0.12). The most able student (1.06) is the student who obtained 70 marks, who was the third in rank, while the second best student with raw marks (74) was actually the least able student in logit. For further study, it would be interesting to find the correlation between student's marks in the final examination and student's ability performance using logit unit.

Face validity essentially looks at whether the scale appears to be a good measure of the construct 'on its face'. While construct validity is referring to the analysis or outcome of the theories and ideas on the study being carried out. The actual instrument construct that was developed should reflect the theories initiated or chapters taught, in case of assessing educational course outcomes. For the face validity, the committee endorsed the question set as the valid and reliable instrument for the course assessment. Reliability wise, it was agreed that the use of such instrument would lead the way to understand the course accordingly. In educational practice, students are respondents and their answers could be analysed for the reliability of test questions or instrument's construct (refer Figure 3).

```
TABLE 3.1 FINAL EXAM RESULT                       ZOU713WS.TXT Aug 18 12:15 2016
INPUT: 15 Persons  44 Items  MEASURED: 15 Persons  44 Items  4 CATS        1.0.0
-------------------------------------------------------------------------------
       SUMMARY OF 15 MEASURED Persons
+------------------------------------------------------------------------------+
|          RAW                         MODEL        INFIT          OUTFIT       |
|          SCORE    COUNT   MEASURE    ERROR    MNSQ   ZSTD    MNSQ    ZSTD     |
|------------------------------------------------------------------------------|
| MEAN     130.9    44.0      .45       .16     1.02    .0     1.00    -.1      |
| S.D.      10.6     .0       .29       .02      .24   1.2      .25    1.1      |
| MAX.     151.0    44.0     1.06       .20     1.49   2.3     1.52    2.1      |
| MIN.     114.0    44.0      .03       .15      .63  -2.3      .61   -2.1      |
|------------------------------------------------------------------------------|
| REAL RMSE   .17 ADJ.SD   .23 SEPARATION 1.32  Person RELIABILITY   .63       |
| MODEL RMSE  .16 ADJ.SD   .24 SEPARATION 1.44  Person RELIABILITY   .67       |
| S.E. OF Person MEAN = .08                                                    |
+------------------------------------------------------------------------------+
Person RAW SCORE-TO-MEASURE CORRELATION = 1.00
CRONBACH ALPHA (KR-20) Person RAW SCORE RELIABILITY = .65
       SUMMARY OF 44 MEASURED Items
+------------------------------------------------------------------------------+
|          RAW                         MODEL        INFIT          OUTFIT       |
|          SCORE    COUNT   MEASURE    ERROR    MNSQ   ZSTD    MNSQ    ZSTD     |
|------------------------------------------------------------------------------|
| MEAN      44.6    15.0      .00       .29      .98   -.1     1.00     .0      |
| S.D.       5.7     .0       .47       .06      .44   1.5      .48    1.5      |
| MAX.      57.0    15.0      .98       .59     1.99   2.6     2.17    2.8      |
| MIN.      30.0    15.0    -1.62       .24      .28  -3.5      .28   -3.2      |
|------------------------------------------------------------------------------|
| REAL RMSE   .31 ADJ.SD   .35 SEPARATION 1.12  Item   RELIABILITY   .56       |
| MODEL RMSE  .29 ADJ.SD   .37 SEPARATION 1.25  Item   RELIABILITY   .61       |
| S.E. OF Item MEAN = .07                                                      |
+------------------------------------------------------------------------------+
UMEAN=.000 USCALE=1.000
Item RAW SCORE-TO-MEASURE CORRELATION = -.97
660 DATA POINTS. APPROXIMATE LOG-LIKELIHOOD CHI-SQUARE: 1536.50
```

Figure 3: Reliability of Questions and Students

In this study, the analysis showed that the question set has low reliability which is 0.56 value, student's answer also has low reliability value at 0.63 score. The Cronbach Alpha value also showed that the test reliability is low at 0.65 score. This means, although the examination questions prepared by the university has low reliability and the student's answer is also of low reliability, still it was accepted. In our current practice, the reliability of question paper and student's answers are seldom checked or analysed. Our examination therefore remains untested and possibly unfair for the students.

```
TABLE 17.1 FINAL EXAM RESULT                    ZOU713WS.TXT Aug 18 12:15 2016
INPUT: 15 Persons  44 Items  MEASURED: 15 Persons  44 Items  4 CATS      1.0.0
-------------------------------------------------------------------------------
Person: REAL SEP.: 1.32  REL.: .63 ... Item: REAL SEP.: 1.12  REL.: .56

           Person STATISTICS:  MEASURE ORDER

+------------------------------------------------------------------------------+
|ENTRY    RAW                  MODEL|  INFIT  | OUTFIT |PTMEA|EXACT MATCH|      |
|NUMBER  SCORE  COUNT  MEASURE  S.E. |MNSQ  ZSTD|MNSQ  ZSTD|CORR.| OBS%  EXP%| Person|
|------------------------------------------------------------------------------|
|   11    151    44    1.06    .20|1.49  1.7|1.52  1.6| .09| 38.6  51.0| S6454|
|   12    147    44     .90    .19|1.11   .5| .99   .1| .50| 59.1  48.4| S2876|
|    2    143    44     .77    .18| .83  -.7| .78  -.8| .43| 52.3  44.5| S4204|
|    8    140    44     .67    .17|1.16   .8|1.19   .8| .17| 43.2  42.4| S3678|
|    3    135    44     .53    .16|1.20  1.0|1.14   .7| .30| 40.9  37.5| S8264|
|    4    135    44     .53    .16| .85  -.7| .82  -.8| .56| 38.6  37.5| S2457|
|   10    130    44     .40    .16| .63 -2.3| .61 -2.1| .44| 40.9  37.0| S1634|
|    1    128    44     .35    .16|1.11  -.6|1.01   .1| .59| 27.3  36.9| S5522|
|    6    128    44     .35    .16|1.45  2.3|1.47  2.1| .01| 31.8  36.9| S4537|
|   14    128    44     .35    .16| .93  -.3| .92  -.4| .48| 36.4  36.9| S6712|
|    7    127    44     .33    .16|1.02   .2|1.09   .5| .24| 43.2  36.9| S8145|
|    5    123    44     .23    .15| .73 -1.7| .66 -2.0| .48| 47.7  34.9| S3886|
|   15    118    44     .12    .15| .86  -.8| .85  -.8| .51| 25.0  33.0| S3498|
|    9    117    44     .09    .15|1.07   .5|1.11   .7| .08| 34.1  31.4| S5116|
|   13    114    44     .03    .15| .84 -1.0| .86  -.8| .56| 20.5  31.4| S7568|
|------------------------------------------------------------------------------|
| MEAN  130.9   44.0    .45    .16|1.02   .0|1.00  -.1|    | 38.6  38.4|      |
| S.D.   10.6    .0     .29    .02| .24  1.2| .25  1.1|    |  9.9   5.6|      |
+------------------------------------------------------------------------------+
```
Figure 4: Student's Fit Order

Not only can this procedure verify validity of the examination paper, it can also recognise true normal good students or students with problems, or abnormal (unique) students, individually (refer Figure 4). This is very important to be observed and wisely taken into consideration, as the assessment is done onto human ability. From the above analysis, all students fit with the model, which means they are normal and their answers are reliable. If necessary, this accurate logit unit could be simply converted into percentile marks, as our usual practice in examination grading, using the template below (refer Figure 5).
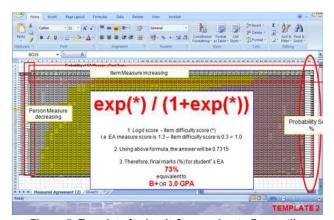

Figure 5: Template for Logit Conversion to Percentile

All these procedures have been done online and paperless, using Excel and Bond & Fox Steps software to measure the accurate student's performance via essay-based examination, as well as for instrument and responses' reliability, and it is simplified.

DISCUSSION 2: WHY DO WE NEED RASCH FOR STUDENT'S PERFORMANCE ASSESSMENT?

Measurement is fundamental in education. A comprehensive and effective education contains of good content, wise objective, positive outcomes, effective instructions and right assessment with fair and accurate measurement (Fisher, 2007; Linacre, 2000). Rasch measurement model (RMM) has been accepted worldwide as a better method to verify the validity of measurement construct (test question construct) and accurate measure of student's ability in logit (log odd unit) scale. Recent years have seen increasing interest among researchers to apply RMM in social science studies, especially on psychometric measurement, but not many teachers, lecturers or educators use RMM in educational assessment. This procedure will help educators to precisely and accurately measure the ability of students within simplified steps. This model is in parallel with the aspiration of the national educational policy which stresses much on outcome-based education. This model has produced comprehensive templates and procedures of analysis for lecturers/ teachers/ educators to use as basic reliable assessment. With Bond & Fox Steps software, educators can transform ordinal data (answer's score/ marks) into equal interval scale (probability of student's ability) which is more precise and accurate in psychometrical result, rather than normal result with calculation of raw score (Linacre, 2000). This should be used for accurate student's performance as a better alternative compared to the current practice which counts merely the raw score marks. This procedure provides linear metrics of ratio-based measurement in LOGIT as practiced in most developed countries. The common practices in our current examination use essay based approach and this rubrical analysis is the most compatible solution. This simplified procedure is easy and friendly for all lecturers without the need to understand the details of Rasch statistical analysis or Bond & Fox Steps software in depth.

The use of Rasch Measurement Model (RMM) is based on the several distinguished reasons, as follows:

1. Rasch offers a new paradigm in education longitudinal research.
2. Rasch is a probabilistic model that offers a better method of measurement construct, hence a scale.
3. Rasch gives the maximum likelihood estimate (MLE) of an event outcome.
4. Rasch reads the pattern of an event thus predictive in nature which ability resolves the problem of missing data. Hence, it is more accurate.
5. Rasch transforms ordinal data into equal interval scale, which is more appropriate for humanities and social science research.
6. Rasch measures item and task difficulties, separately and accurately (Mohd Nor Mamat, 2012; Wu, 2007).

CONCLUSION

A well-designed scientifically reliable set of examination questions will give us valid data of student's ability and performance in any course for meaningful accurate assessment. Rasch offers a new paradigm of accurate assessment on human's ability and performances and provides a means for verifying question items, and subsequently the validity of the question set as well as reliability of student's answers. The instrument that measures outcomes of teaching and learning found in student's understanding, comprehending, applying, analysing, synthesizing and evaluating what has been learned is now scientifically vetted, tested and validated, and more importantly is reliable for the purpose. Logit unit, the probabilistic ratio-based measure scale found in this method is more accurate for human performance assessment. In addition, this essay outlines a simplified procedure that may provide the optimal solution for students' assessment the digital and online era, especially for massive numbers of students, online based course, and paperless test through MOOC.

REFERENCES

Aziz, A. A. (2008). Developing an Instrument Construct Made Simple. Winstep and Rasch Model Workshop. Shah Alam: Universiti Teknologi MARA 30-31 December 2008.

Aziz, A. A., Masodi, M. S., Zaharim, A. (2012). *Asas Model Rasch: Pembentukan Skala dan Struktur Pengukuran*. Bangi: Universiti Kebangsaan Malaysia Press.

Bond, T. G. and Fox, C. M. (2007). *Applying the Rasch Model*, 2nd Edition. New Jersey: Lawrence Erlbaum Associates Publishers.

Fisher, W. P. Jr. (2007). Living capital metrics. *Rasch Measurement Transactions, 21*(1), 1087-1096.

Linacre, J. M & Wright, B. D. (2000). *Winsteps: A Rasch Computer Program.* Chicago: MESA Press.

Mohd Nor Mamat (2012). Development of Hadhari Environmental Ethics among Environmental Ethics Student at Tertiary Education. Unpublished PhD Dissertation, Universiti Teknologi MARA.

Wright B. & Linacre J. (1992). Combining and splitting categories. *Rasch Measurement Transactions 6*(3), 233-235.

Wu, M. & Adams, R. (2007). A*pplying the Rasch Model to Psyho-Social Measurement*. Melbourne: Educational Measurement Solutions.