Future Academy®
Offering excellence in Social and Behavioural Sciences

The European Proceedings of
Social & Behavioural Sciences
EpSBS

eISSN: 2357-1330

ICEEPSY 2016 : 7th International Conference on Education and Educational Psychology

Measurement of Rubrical Essay-based Test Using Rasch Model

Mohd Nor Mamat[a]*, & Siti Fatahiyah Mahamood[a]

* Corresponding author: Mohd Nor Mamat, mohdnoor@salam.uitm.edu.my

[a]*Universiti Teknologi MARA, 40450 Shah Alam, Malaysia, mohdnoor@salam.uitm.edu.my, +60 19 281 9003/ +60 12 391 7175*

Abstract

There are many educators use raw score as measurement for student's ability, but it never truly measures the right measurement. The raw score should be converted into a right linear metrics for ability measurement. This study is to implement rubrical analysis or assessment for essay-based test using Rasch Model (RM). It is to produce a reliable and accurate measurement procedure for student's performance in essay-based test. The initial study has identified the concept of RM to be applied in educational test or examination. The design and development stages has produced a comprehensive but simplified procedure template of scientific analysis within Bond&Bond Steps software. The proposed procedures contains of measuring score of accurate student's ability in LOGIT unit, providing of student's result profiling, and measuring reliability of the test set and the student's answers,. The procedure is designed for essay-based test which is more difficult to be analysed, compared to the multiple choice type. This procedure converts the student's answer into rubrical ratio-based scale to be most accurately measured and definitely better than the common practice of merely analysis on raw marks for each question. It will show true student's performance of cognitive performance (test) which represents the true student's ability, in order to accurately measure the right outcome.

Keywords: Student's performance, learning assessment, rubrical analysis, rasch measurement model.

1. Introduction

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. Rasch measurement model (RMM) was accepted worldwide as 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. Lately, many researchers are aware of the reliability of RMM to be applied in social sciences study, especially on psychometric measurement, but not many

teachers, lecturers or educators use RMM in educational assessment. This procedure will help much educators, to precisely and accurately measure the ability of students within simplified steps. This model is in parallel with aspiration of national educational policy which stresses much on outcome-based education. This has produced a comprehensive templates and procedures of analysis for lecturers/ teachers/ educators to use as basic reliable assessment. The procedure will engage with Bond&Fox Steps software to 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.

This should be used for accurate student's performance as better alternative for the current practice which counts on merely the raw score of question marks. This procedure provides linear metrics of ratio-based measurement in LOGIT, as practiced in most developed countries. Most 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 any necessary of understanding Rasch statistical analysis or Bond&Fox Steps software in depth.

1.1 Problem Statement and Objective of the Study

In educational assessment, there are still many educators uses raw score as measurement for student's ability, but raw scores never been true measure (Azrilah, 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 the student's response especially in essay-based test. This simplified procedure proposes the method of rubrical analysis as student's responses.

The research questions are:

a) How to analyse student's responses in essay-based test using Rasch Measurement Model?
b) How to produce reliable and accurate student's performance in essay-based assessment?

The research objectives are:

a) To implement rubrical analysis or assessment for essay-based test using Rasch Model.
b) To produce a reliable and accurate measurement procedure for essay-based student's performance.

1.2 Methodology

As common practice in education, test is administered among the students, to sit normal examination process. Student's answers then will be keyed in to Microsoft Excel format, in which lecturers are very familiar. The next step is to convert the student's answers (data) into the Bond&Fox Steps software for reliability analysis. The data in Microsoft Excel's format needs to be resaved in .prn format, to allow them to be opened using Bond & Fox Steps for more friendly use in educational purposes. This study has determined a class of 15 students from Environmental Technology programme (EVT229), from the Faculty of Applied Science, Universiti Teknologi Mara Malaysia.

http://dx.doi.org/10.15405/epsbs.2016.11.80
eISSN: 2357-1330 / Corresponding Author: Mohd Nor Mamat
Selection and peer-review under responsibility of the Organizing Committee of the conference

The rest is only to extract person measure and item measure by only one click within Steps software. This can automatically identify and sort good students with the highest score to weak students with the lowest score. These simple steps may be used as tools for accurate assessment of reliable student's ability or performance in such course in a linear ruler with LOGIT measure. This procedure may allow us to easily understand and utilize in any courses or programs which aim for learning outcome of the course as reliable student's performance.

The study used the Final Examination set of Environmental Ethics course (IPK661) for June 2016 examination. The set has one question with four sub questions, and three questions with three sub questions. Each sub question carries one, or two or three marks provided for each answer. Using analysis function in the software, the result of item and student's reliability is also can be shown. This also can determine the reliability of the overall student's answers and detect any problems appeared in detailed result of individual analysis. The observed variance and the error variance values will be taken into calculation to find reliability level, as shown below:

(Observed Variance – Error Variance) / Observed Variance = Reliability Score

This study uses common statistical measurement where it accepted the range of reliability score between 0.5 to 1.0 as reliable and acceptable. It is similar to Cronbach Alpha value which accepts the value between 0.5-1.0. However, in the real practice as formal assessment towards students, it is recommended to accept only excellent reliability of student's answer to be more valid and highly accepted which is between 0.8 and 1.0 only. This is more equivalent to common measurement scheme in education which is between 80% - 100% to be considered A. Otherwise, the students will be required to retake the test, or it would be recommended to investigate the real factors behind of the inconsistency of the responses.

2. Discussion and Findings

2.1. Why Rasch Model?

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).

2.2 Procedure of Student's Ability Measurement

A comprehensive procedure of measurement could be simplified as follows:

1.2.1 Students' answers scripts will be marked as usual

1.2.2 Students' raw score will be recoded to rubrical score for analysis

1.2.3 Data will be analysed using Bond&Fox Steps software

1.2.4 Results of students' accurate performance score (LOGIT) will be delivered

1.2.5 Students' profiling, reliability score and test's reliability score will be produced

This simplified procedure needs only few simple steps to be applied. For the first steps, after the answer's script has been marked (as usual practice), it need to be recoded into rubrics scale, like Likert scale. As for IPK661 question set has different values for each answer, it is necessary to recode accordingly.

Raw Marks vs Coded Rubrics

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

Fig. 1. Raw marks vs coded rubrics.

The rubrical value would be keyed into EXCEL format and saved as .prn format. The .prn file would be opened using Bond&Fox Steps and few clicks, resulted the display of reliability score for the question set and student's answer which is acceptably reliable (i*r*=0.56) and (p*r*=0.63) with Cronbach alpha value at 0.65. In the future practice, this result will help much for teachers and educators to evaluate the quality of test questions before distributing to the students.

```
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
```

Fig. 2. Summary of reliability and Cronbach alpha values.

http://dx.doi.org/10.15405/epsbs.2016.11.80
eISSN: 2357-1330 / Corresponding Author: Mohd Nor Mamat
Selection and peer-review under responsibility of the Organizing Committee of the conference

The last procedure in this study is to analyse the rubrical values of student's answers to be in a linear metrics, LOGIT unit. With one more click, we have the result of students (in rank order) within log odd unit (LOGIT) 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. In this case, the more comprehensive analysis may be made to identify unique problems of every single student. This procedure would identify the right measure of student's ability and put them in rank order.

2.3 Result of Student's Ability

Among 15 students involved, 13 passed the examination with score of between 51-80 marks, while 2 of them failed with 41 and 46.5 marks. After all these marks were converted into LOGIT unit, the result showed that all students have positive result, but only six of them have passed over mean value (0.43) which means above 50%. Student's ability score are 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 map below, 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 NUMBER	RAW SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PTMEA CORR.	EXACT MATCH OBS%	EXACT MATCH 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	

Fig. 3. Person fit order.

From the above analysis, all students are fit with the model, which means those students are normal and their answers are reliable. The rank of student's ability is 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.

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

Fig. 4. Raw marks vs ability metrics score.

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 got 70 marks, who is the third at the rank, while the second best student with raw marks (74) is actually the least able students in logit. For further study, it is an interesting study to find correlation between student's marks in the final examination and student's ability performance using logit unit.

2.4 Reliability of Test and Student's Answer

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 the analysis or outcome of the theories and ideas on the study being carried out. The actual instrument construct that is 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.

http://dx.doi.org/10.15405/epsbs.2016.11.80
eISSN: 2357-1330 / Corresponding Author: Mohd Nor Mamat
Selection and peer-review under responsibility of the Organizing Committee of the conference

```
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
```

Fig. 5. Reliability of questions and studens.

In this study, the analysis showed that the question set has low reliability which is 0.56 value, while student's answer 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, the examination question, prepared by the university is low reliable and the student's answer is also low reliable, but still accepted. In our current practice, this reliability of question paper and student's answers are seldom checked or analysed, made our examination remains untested and possibly unfair for the students.

```
TABLE 6.1 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:  MISFIT 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|A .09| 38.6  51.0| S6454 |
|     6    128     44      .35    .16|1.45   2.3|1.47   2.1|B .01| 31.8  36.9| S4537 |
|     3    135     44      .53    .16|1.20   1.0|1.14    .7|C .30| 40.9  37.5| S8264 |
|     8    140     44      .67    .17|1.16    .8|1.19    .8|D .17| 43.2  42.4| S3678 |
|    12    147     44      .90    .19|1.11    .5| .99    .1|E .50| 59.1  48.4| S2876 |
|     9    117     44      .09    .15|1.07    .5|1.11    .7|F .08| 34.1  31.4| S5116 |
|     1    128     44      .35    .16|1.11    .6|1.01    .1|G .59| 27.3  36.9| S5522 |
|     7    127     44      .33    .16|1.02    .2|1.09    .5|H .24| 43.2  36.2| S8145 |
|    14    128     44      .35    .16| .93   -.3| .92   -.4|g .48| 36.4  36.9| S6712 |
|    15    118     44      .12    .15| .86   -.8| .85   -.8|f .51| 25.0  33.0| S3498 |
|    13    114     44      .03    .15| .84  -1.0| .86   -.8|e .56| 20.5  31.4| S7568 |
|     4    135     44      .53    .16| .85   -.7| .82   -.8|d .56| 38.6  37.5| S2457 |
|     2    143     44      .77    .18| .83   -.7| .78   -.8|c .43| 52.3  44.5| S4204 |
|     5    123     44      .23    .15| .73  -1.7| .66  -2.0|b .48| 47.7  34.9| S3886 |
|    10    130     44      .40    .16| .63  -2.3| .61  -2.1|a .44| 40.9  37.0| S1634 |
|-------------------------------------+----------+----------+-----+-----------+-------|
| 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|       |
+-------------------------------------------------------------------------------------------+
```

Fig. 6. Student's fit order.

Instead of verifying validity of the examination paper, this procedure with Rasch measurement model can also recognize true normal good students or students with problems, or abnormal (unique) students, individually. 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 are fit with the model, which means those students are normal and their answers are reliable.

3. Conclusion

A well scientifically reliable-proven question set of examination will give us valid data of student's ability and performance in any course for the meaningful accurate assessment. Rasch offers a mean of verifying question item, subsequently the validity of the question set as well as student's answers reliability. The certainty of an instrument to measure outcomes of the teaching and learning, which is student's ability of answering the questions of what has been learned is now scientifically vetted, tested and validated, and more important is reliable for the purpose. Logit unit, the probabilistic ratio-based measure scale is more accurate for human performance assessment. In additional, this esay and simplified procedure would be the best solution to be implemented by all educators without knowing statistical elements in detail.

Acknowledgment

Acknowledgment to Universiti Teknologi Mara Malaysia for funding this research and publication.

References

Azrilah Abdul Aziz et al (2012). Asas Model Rasch: Pembentukan Skala dan Struktur Pengukuran. Bangi: Universiti Kebangsaan Malaysia Press.

Azrilah Abd Aziz (2008). "Developing an Instrument Construct Made Simple": *Winstep and Rasch Model Workshop*. Shah Alam: Universiti Teknologi Mara 30-31 December 2008.

Bond, Trevor G. & Fox, Christine M. (2007). *Applying the Rasch Model: Fundamental Measurement in The Human Sciences*. New York: Routledge.

Bond, Trevor G. and Fox, Christine M. (2007). *Applying the Rasch Model* (Second Edition). New Jersey: Lawrence Erlbaum Associates Publishers.

Fisher, William P. Jr (2007). *Rasch Measurement Transactions* (http://www.rasch.org/rmt).

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. Shah Alam: PhD Thesis (unpublished)

Wright B. & Linacre J. (1992). "*Combining and splitting categories*": Rasch Measurement transactions 6: 233-235.