How good was the test set up? From Rasch Analysis Perspective

Zulkifli Mohd Nopiah [a], **Mohd Haniff Osman** [a], **Noorhelyna Razali** [a], **Fadiah Hirza Mohammad Ariff** [a], **Izamarlina Asshaari** [a]

[a] Fundamental of Engineering Unit,
Faculty of Engineering and Built Environment,
Universiti Kebangsaan Malaysia,
43600 UKM Bangi, Malaysia

Abstract

This research analysed the validity and reliability of given questions in the Linear Algebra course at Universiti Kebangsaan Malaysia (UKM). Through Rasch analysis, the difficulty of the questions with respect to students' ability level was assessed. The 20 questions covering such topics as hyperbolic function, eigenvalue and eigenvector, vectors, etc were administrated to 240 first-year faculty of engineering and built environment students. The dichotomous Rasch model using 0 (if student guessed wrong) and 1 (if student guessed right) scoring of item response data was applied. The analysis distinguished item hierarchy, items gaps, unidimensionality, fit-statistic, person reliability and separation statistics. The report provides interesting information that may be used to improve the validity and reliability of items.

Keywords: Rasch Analysis; multiple-choice questions; fit-statistics

1. Introduction

In order to measure student's understanding of course material, tests or examinations need to be conducted. Yet, there are two general categories of test items; (a) objective items which require students to select the correct response from several alternatives or to supply a word or short phrase to answer a question or complete a statement; and (2) subjective or essay items which permit the student to organize and present an original answer (Cory, 1979). A standard multiple-choice test item consists of a *stem* and a list of alternatives. The *stem* refers to a question or problem and the list of alternatives contain one correct alternative (answer) and a number of incorrect alternatives called distractor (Kehoe, 1995).

Multiple-choice format is the most common form of teacher-constructed tests. This type of assessment is an efficient and effective way to assess a wide range of knowledge, skills, attitudes and abilities (Haladyna, 1999). It is a vital issue to construct a well designed test which is relevant and meets the needs of the syllabus. The taste of test designers to make sure that the test provides the channel to justify that the student has achieved the required level of competence.

Analysis of latent trait, which is also known as construct in Rasch Analysis terminology such as intelligence, attitudes, knowledge, quality of tests cannot be performed directly. For this reason, Rasch analysis was used to review how well the multiple-choice items were constructed in Linear Algebra course. Rasch analysis is a statistical technique used in education and psychology to measure abstract constructs (Obermeier, 2009).

2. Experimental details

2.1. The measurement tool

The Rasch dichotomous model was applied to analyze the data provided from students' score. Rasch first converts an instrument's ordinal data into interval data, thus meeting the minimum requirement of true measurement (Haley, McHorney & Ware, 1994; Merbitz et al., 1989; Wright & Linacre, 1989). The value of 1 is assigned to students for guessing correct item response and 0 for getting item response wrong. Then by converting item scores to logarithm, and then calculating the odds of each student to answer each item correctly, the scores become meaningful calibrations of difficulty and ability (Obermeier, 2009).

Rasch analysis offers validation on unidimensionality, separation statistics, item-person distribution map and fit statistics. Unidimensionalty refers to how well each item measures of "fits" a construct (Bond & Fox, 2001). Rasch analyses also provide information item difficulty (also known as item hierarchy). Item difficulty is described on a measurement continuum from less difficult to more difficult (Draugalis & Jackson, 2004). The person separation statistic reveals how efficiently the instrument assigned respondents into several distinct level of ability. While item separation indicates distance between items of varying difficulty which provide additional evidence to support construct validity. In addition, Rasch analysis provides person reliability which refers to how well the instrument consistently reproduces a participant's score (Bond & Fox, 2001). An item-person distribution maps place all the items according to level of difficulty determined by students' performance and charts student's ability based on how many items they were able to answer correctly. Finally, fit statistics provide the indices for fit of the data to the model and usefulness of the measure (Green, 2002). Fit statistics include the mean square (MNSQ) infit and outfit of student and items.

2.2. Software

Rasch Analysis was performed using Bond & Fox Steps (Bond & Fox, 2006) which is a customized WINSTEPS. The WINSTEPS program provides detailed statistics for each item as well as for the overall instrument.

2.3. Participants

The subjects used in this study are the first year students of Faculty of Engineering and Built Environment, Universiti Kebangsaan Malaysia (UKM), class of 2008/2009. The sample consisted of 97 women (40.4%) and 143 men (59.6%).

3. Results and discussions

3.1. Student and Item map

The heart of Rasch analysis is provided in Fig. 1, the map of students and items simulatenously. The vertical dashed line represents the ideal less-to-best continuum of quality. Items and students share the same linear measurement units known as logits. Logits is a unit derived from transforming ordinal data into an interval scale (Bond & Fox, 2001). On the right hand side of the dashed line, the items are aligned from too easy to too hard, starting from the bottom. The distribution of student positions is on the left side of the vertical dashed line in increasing order of ability. Letter "M" denotes the student and item mean, "S" is one standard deviation away from the mean and "T" marks two standard deviations away from the mean. Each "#" symbol represents three students, "*"equals to two student and "x" for one student in this figure.

The right side of the Fig. 1 depicts the test items which are represented by letter "I" and followed by the number of the question. For instance, I020 represents the 20th question. Items at the top of the scale are more challenging whereas items at the bottom of the map are easy to resolve. Item I6 is the most difficult since it's charted towards the top of the map.

In Fig. 1, 240 students in this study are mapped on the left side of the map under the heading "STUDENTS". The most able students are number 57 and 109, followed by a group of 21 best students (1, 16, 71, 75,…,195) and two students which numbers 41 and 209 are categorized as least able. The ability of student number 52 matches the difficulty level of item I17. That means this student had a 50% chance of getting item I17 correct. Those students who are mapped far above the position of I17 all probably answered item I17 correctly. On average, this batch of students agreed that the items were easy since the mean of students' ability is 1.52 higher than the mean of the items; which is calibrated as zero.

According to Item fit statistics shown in Table 1, I6 is the most challenging item. It has the highest measure of 2.74 logits. The transcript of the item I6 follows in Fig. 2. Item 16 is an example of a statement-based question and no calculation is required to it. The purpose of establishing this type of question is to challenge students on concepts, theorem, and properties of certain topics or subtopics. As it happens, students were asked to choose any true statements on the topic of system of linear equations. Students need to have a good comprehension on properties of row operations.

Distracter "ii" responds to the statement about row equivalent. Now, suppose that we have two matrices of the same size C and E. These matrices are row equivalent if matrix E can be formed by applying a finite number of row operations to C. Consequently, the given statement is absolutely false. Choice 'iii' is clearly more straightforward and it is incorrect because if the system has no solution then we call the system inconsistent and if there is at least one solution to the system, we call it consistent. By neglecting any combinations consists choice "ii" and/or "iii", the correct answer is "c".

Looking at 'raw score' column in the Table 1, only 64 students managed to answer I6 correctly. It reflects that near to three-quarters (~75%) of students failed to answer this type of question. By ignoring the chosen topic or subtopic, we hardly believed that the students will counter problem while dealing with the statement-based question. The item I14 supports the earlier findings. Item I14 which is another statement-based question was found to be the second hardest question in this test. The transcript of Item I14 follows in Fig. 3. Item I14 is about properties of eigenvalues and eigenvectors. The correct answer is "b" which statement "i" and "iii" are incorrect. There is no such theorem or properties that stated any similar matrices have the same eigenvectors. Instead, there are two questions (understanding) viz.; I8 and I13 which about dealing with eigenvalues and eigenvectors which are located towards lower order of difficulty where students were found easy to solve. It can be deduced that students do not go for reading and understanding the fundamental concepts. Their preference is towards mechanical application and algorithm solving questions.

The Rasch analysis revealed that the majority of the students were unable to define or even to determine the null space of matrix. In the last section of the test, students needed to determine the basis of column space, row space and null space of the given matrix. These three definitions were taught in class simultaneously and if the students get any question correct, then they should get the other two correct. The three items namely I17 (basis of column space), I18 (basis of row space) and I19 (basis of null space). Surprisingly, students felt item I19 (1.06 logits) more difficult than I17 (-0.52 logits) and I18 (-1.13 logits). Fig. 1 illustrates the big gap between I19 and I18, approximately 2.30 logits. In addition, Item I15 and I16 were tested on dimension of null space and dimension of column space respectively. In Fig. 1, it was noted that there is a huge gap between I15 and I16 denoted by ($\longleftrightarrow$) indicating the extent of difficulty the students encountered in attempting the questions. Again, it shows that students found the topic of null space difficult to grasp. Therefore, in future class, lecturer should give priority to this topic.

Investigating the ordering of the test items proposed that item 5 was the easiest item and should be placed in the beginning of future test. Then followed by item I001 and I010. By reordering the items, it will help to relieve student's anxiety and also boost student's confidence. It is worthwhile to setup some easy items in the beginning of the test as to help student to bring their full ability without being distracted by confusing items when they start the test.

Fig.1 demonstrates a major gap between item I5 and I1. This gap suggests that the quality of question denoted by I5 is questionable. It is highly probably that this item did not impart a sufficient challenge for the ability level of students who took the test. To improve assessment's reliability, the test designer is encouraged to rephrase or substitute I5 in order to improve the assessment's reliability.

3.2. Fit Statistics

Summary fit statistics of students and items measures were next captured. The result will explain whether the data fit a construct (Bond & Fox, 2001). Rasch experts examine item fit by looking at two types of fit values known as infit and outfit (Pomeranz et. al, 2008). Item fit is an index of how well items function in reflection of the trait (Green, 2002). Rasch analysts typically examine infit scores, which indicate how well the observations fit the Rasch modeled expectations (Bond & Fox, 2001) whereas "outfit" is less threatening to measurement and easier to manage (Linacre, 2002).

Table 2 shows comprehensive information about whether the data showed acceptable fit to the model. The mean square (MNSQ) infit and outfit for students and items are expected to be 1.0. Evaluation of MNSQ infit and outfit were between 0.8 and 1.2 (Bade & Wright, 1999). Thus this suggests that the provided data exhibited fit the model and supported the unidimensionality. Note that the mean of items is 0.0. The measurement shown in Table 3 reveals that the students mean, $\mu_{student} = 1.52$ logits which suggests these items, on average, were easy to answer.

Table 1 presents items in order of difficulty. Entry number is the item's location in this scale of 20 items. Raw score is the total number of students who obtained correct for the corresponding item. Count tells us that all 240 students responded to the items. Measure is the logit position of the item, the more positive the value indicates more difficult the item. For this data, mean square (MNSQ) infit of all items less than 1.20 (20% variance), which was recommended by Wright & Linacre (1994). Therefore we can consider the items fit well

The next overall statistics investigated is separation. Item separation is the distance in logits between items of varying difficulty (Draugalis, 2004). The greater values of separation represent spread of items along the continuum and lower the values indicate redundancy in the items (Green, 2003). The item summary gives a good summary with Separation, G = 6.81. This value indicates that the items were sufficiently well separated in difficulty. For students, separation is 1.23. This small value indicates that there is not enough differentiation among students to separate them into distinct performance level or strata. Strata can be calculated using the formula: Strata = (4 X student separation +1)/3. Thus, a student separation of 1.23 was placed into the strata formula and equaled 2.96, almost three distinct groups (excellent, moderate, poor).

Table 3 showed that student number 118 who appeared to have score 13 over 20 is found to be a misfit (too unpredictable) where mean square (MNSQ) outfit is 3.58, exceed the recommendation. This is further confirmed by the Point Measure Correlation which is negative in nature; -0.21. This mean the student is responding in the reverse direction where he can answer difficult question when others can't and *vice-versa*. The pattern of the student's answers in Fig. 5 (S118's answer scheme) is an attestation to this phenomena, the student answered the I6 and I14 correctly, the two most difficult items while got wrong for the two easiest item I01 and I10. These outcomes did not meet Rasch model expected outcomes. This major finding raised some conclusions, for example, the student underestimated the easiest items and miscalculated the matrix operations. Conversely, for the difficulty items, suspects probably have special interest or knowledge on the topic and/or comfort answering statement-based question. On the hand, it makes sense that the student may simply guess the questions. Rasch has this particular predictive properties embedded in the model to make it very reliable.

4. Conclusion

Through Rasch analysis, there are several recommendations that can be proposed in order to increase the quality and reliability of test construct such as; 1) reduce item gaps by rephrasing or replacing the current items 2) items need to be rearranged in order of item difficulty to allowing students to spend more time to answer tougher items. This approach may offer researchers an opportunity to understand his/her participant's ability. However, the recommendations apply for the first part of the instrument. In a nutshell, this study concluded positively that the given questions were suitable for assessing student knowledge and comprehension on Linear Algebra.

Acknowledgements

The authors would like to acknowledge the financial support received from National University of Malaysia, UKM as Action Research grant and also the valuable input of Mohd Saidfudin Masodi, Program Coordinator for the Pro-Cert in Quality Management System at SPACE, UTM. His expertise in Rasch measurement theory was especially helpful in producing this article.

References

1. A. Obermeier, Improving English test questions through Rasch Analysis, Journal of Educational Research and Development (2009) 107-113
2. B.D. Wright, G. N. Masters, Observations are always ordinal: Measure, however must be interval, Archives of Physical Medicine and Rehabilitation 70 (1989) 857-860.
3. B.D. Wright, J.M. Linacre, Reasonable mean-square fit values, Rasch Measurement Transactions 8(3) (1994) 370.
4. C. Merbitz, J. Morris, J.C. Grip, Ordinal scales and foundations of misinference, Archives of Phsyical Medicine and Rehabilitation 70 (1989) 303-313.
5. J.C. Cory, Improving Your Test Questions, USA, 1979.
6. J. Kehoe, 1995, Writing multiple-choice test items, Practical Assessment, Research & Evaluation 4(9), Retrieved February 12, 2008.
7. J.L. Pomeranz, K.L. Byers, M.D. Moorhouse, C.A. Velozo, R.J. Spitzngel, Rasch Analysis as a technique to examine the psychometric properties of a career ability placement survey subtest, Rehabilitation Counseling Bulletin 51(4) (2008) 251-259.
8. J.M. Linacre, What do infit and outfit, mean square and standardized mean?, Rasch Measurement Transactions 7(4) (2002) 878.
9. J.T. Draugalis, T.R. Jackson, Objective curricular evaluation: Applying the Rasch model to a cumulative examination, American Journal of Pharmaceutical 68(2) Article 35 (2004) 1-12.
10. K.E. Green, Survey development and validation with the Rasch model, Conference paper of International Conference on Questionnaire Development, Evaluation and Testing, South Carolina, USA, 2002.
11. R.K. Bode, B.D. Wight, Rasch measurement in higher education, In J.C. Smart & W.G. Tierney Higher Education: Handbook of theory and research Vol.(XIV), Agathon Press, New York, 1999.
12. S.M. Haley, C.A. McHorney, J.E. Ware Jr., Evaluation of the MOS SF-36 physical functioning scale (PF-10): Unidimensionality and reproducibility of the Rasch item scale, Journal of Clinical Epidemiology 47 (1994) 671-684.
13. T.G. Bond, C.M. Fox, Applying the Rasch model: Fundamental measurement in the human science, Lawrence Erlbaum Associates, Mahwah, New Jersey, 2001.
14. T.G. Bond, C.M. Fox, Bonds & Fox Steps (customised version of WINSTEPS), [Computer Software], 2006.
15. T.M. Haladyna, Developing and validating multiple-choice test items, 2nd ed., Lawrence Erlbaum Associates, Mahwah, New Jersey, 1999.

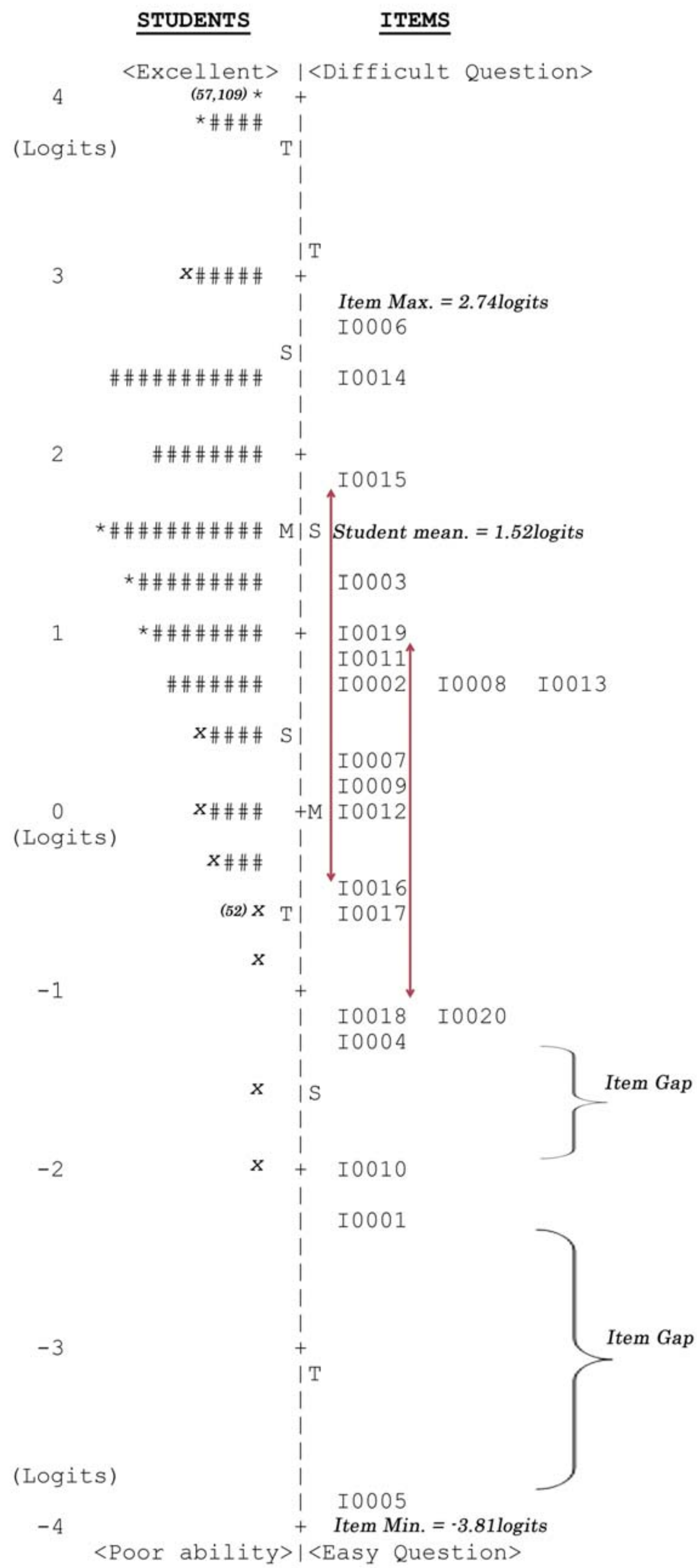
STUDENTS
ITEMS
<Excellent> |<Difficult Question>
4
(57,109) *
*####
(Logits)
3
X#####
Item Max. = 2.74logits
I0006
###########
I0014
2
########
I0015
*########### M|S Student mean. = 1.52logits
*#########
I0003
1
*########
I0019
I0011
#######
I0002 I0008 I0013
X####
I0007
I0009
0
X####
I0012
(Logits)
X###
I0016
(52) X
I0017
X
-1
I0018 I0020
I0004
X
Item Gap
-2
X
I0010
I0001
Item Gap
-3
(Logits)
I0005
-4
Item Min. = -3.81logits
<Poor ability>|<Easy Question>

Fig. 1. Student-item distribution Map

Item I6

Which of the following statements are **TRUE**?

i. Elementary row operations in an augmented matrix never change the solution set at the associated linear system.
ii. Two matrices are row equivalent if they have the same number of rows.
iii. An inconsistent system has more than one solution.
iv. Two linear systems are equivalent if they have the same solution set.

A. i and ii
B. i and iii
C. i and iv
D. ii and iii
E. ii and iv

Fig. 2. Transcript of the most difficult item

Item I14

Which of the following statements are **TRUE**?

i. If 0 is the only eigenvalue of a square matrix *A*, then *A* is the zero matrix.
ii. If an *n* x *n* matrix has *n* distinct eigenvalues, then it must be diagonalizable.
iii. Similar matrices have the same eigenvectors
iv. If 0 is an eigenvalues of an *n* x *n* matrix *A*, then *A* is singular.
v. Two diagonalizable matrices *A* and *B* with the same eigenvalues and eigenvectors must be the same matrix.

A. i and ii
B. ii, iv and v
C. i, iii and v
D. ii and v
E. iii, iv and v

Fig. 3. Transcript of item I14

Item I05

Balance the given chemical equation.

$$Fe_3O_4 + C \rightarrow Fe + CO$$

A. $2Fe_3O_4 + 2C \rightarrow 4Fe + CO$
B. $Fe_3O_4 + 3C \rightarrow 4Fe + 2CO$
C. $Fe_3O_4 + 4C \rightarrow 3Fe + 4CO$
D. $Fe_3O_4 + C \rightarrow 2Fe + 2CO$
E. $3Fe_3O_4 + C \rightarrow 3Fe + CO$

Fig. 4. Transcript of the easiest item

MOST DIFFICULT

Item	06	14	15	03	19	11	13	02	08	07
Logits	2.74	2.44	1.84	1.30	1.06	0.91	0.78	0.69	0.69	0.23
Score	✓	✓	✗	✓	✗	✓	✓	✗	✓	✗
Item	09	12	16	17	18	20	04	10	01	05
Logits	0.13	-0.04	-0.49	-0.52	-1.13	-1.19	-1.30	-2.00	-2.35	-3.81
Score	✗	✗	✓	✗	✗	✗	✓	✗	✗	✓

LEAST DIFFICULT

Fig. 5. Student's (S118) answer scheme

TABLE 13.1 KKKQ1114(08/09) ZOU304WS.TXT Apr 27 23:49 2010
INPUT: 240 Persons 20 Items MEASURED: 240 Persons 20 Items 2 CATS 1.0.0

Person: REAL SEP.: 1.23 REL.: .60 ... Item: REAL SEP.: 6.18 REL.: .97

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%	Item
6	64	240	2.74	.16	1.12	1.5	1.27	1.8	.32	74.4	76.9	I0006
14	76	240	2.44	.15	1.13	1.7	1.25	2.0	.33	72.7	73.5	I0014
15	103	240	1.84	.15	.94	-1.1	.90	-1.1	.50	72.7	69.6	I0015
3	129	240	1.30	.14	1.16	2.8	1.27	3.0	.29	62.2	68.5	I0003
19	140	240	1.06	.15	.95	-.8	.90	-1.1	.47	70.2	69.3	I0019
11	147	240	.91	.15	.98	-.3	.95	-.5	.44	68.9	70.0	I0011
13	153	240	.78	.15	1.12	1.9	1.16	1.5	.31	65.5	71.1	I0013
2	157	240	.69	.15	1.06	1.0	1.13	1.2	.35	69.7	71.7	I0002
8	157	240	.69	.15	.97	-.4	.91	-.8	.44	71.4	71.7	I0008
7	176	240	.23	.16	.96	-.5	.88	-.8	.42	76.9	76.4	I0007
9	180	240	.13	.16	.96	-.5	.81	-1.3	.43	78.6	77.5	I0009
12	186	240	-.04	.17	1.01	.2	.92	-.4	.36	78.2	79.4	I0012
16	200	240	-.49	.19	.99	-.1	.89	-.4	.34	84.0	84.6	I0016
17	201	240	-.52	.19	.88	-1.0	.63	-1.8	.45	84.5	85.0	I0017
18	215	240	-1.13	.23	.99	.0	.71	-.9	.31	90.3	90.6	I0018
20	216	240	-1.19	.24	.89	-.6	.55	-1.5	.39	90.8	91.0	I0020
4	218	240	-1.30	.25	1.05	.3	1.02	.2	.22	91.6	91.8	I0004
10	227	240	-2.00	.32	.88	-.4	.64	-.7	.29	95.4	95.4	I0010
1	230	240	-2.35	.37	.83	-.4	.41	-1.3	.32	96.6	96.6	I0001
5	236	240	-3.81	.72	1.01	.3	.61	-.3	.11	99.2	99.2	I0005
MEAN	170.5	240.0	.00	.22	.99	.2	.89	-.2		79.7	80.5	
S.D.	49.1	.0	1.60	.13	.09	1.0	.24	1.3		10.8	10.1	

Table 1. Item fit statistics

Summary of 240 Measured Students

	RAW SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD
MEAN	14.3	20	1.52	0.67	1.00	0.1	0.89	0.1
S.D.	2.9	0	1.08	0.14	0.28	0.9	0.63	0.6
MAX.	19	20	3.84	1.08	0.86	0.1	1.25	0.0
MIN.	4	20	-1.97	0.67	1.20	0.6	1.14	0.5
RMSE 0.68	SD.0.84		SEPARATION 1.23			RELIABILITY 0.60		

Summary of 20 Measured Items

	RAW SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD
MEAN	170.5	240	0.0	0.22	0.99	0.2	0.89	0.1
S.D.	49.1	0	1.6	0.13	0.09	1.0	0.24	1.3
MAX.	64	240	2.74	0.16	1.12	1.5	1.27	1.8
MIN.	236	240	-3.81	0.72	1.01	0.3	0.61	-0.3
RMSE 0.26	SD.1.58		SEPARATION 6.18			RELIABILITY 0.97		

Table 2. Overall model fit information, separation and mean logit

TABLE 17.1 KKKQ1114(08/09) ZOU840WS.TXT Apr 28 1:25 2010
INPUT: 240 Persons 20 Items MEASURED: 240 Persons 20 Items 2 CATS 1.0.0

Person: REAL SEP.: 1.23 REL.: .60 ... Item: REAL SEP.: 6.18 REL.: .97

ENTRY	RAW			MODEL	INFIT		OUTFIT		PTMEA	EXACT MATCH		
NUMBER	SCORE	COUNT	MEASURE	S.E.	MNSQ	ZSTD	MNSQ	ZSTD	CORR.	OBS%	EXP%	Person
222	10	20	.07	.54	.98	.0	.81	-.2	.57	75.0	74.4	222
231	10	20	.07	.54	1.27	1.1	1.59	1.2	.38	65.0	74.4	231
24	9	20	-.22	.54	1.17	.7	1.00	.2	.49	65.0	75.4	024
29	9	20	-.22	.54	1.57	2.0	2.04	1.8	.25	55.0	75.4	029
49	9	20	-.22	.54	1.08	.4	.94	.0	.53	75.0	75.4	049
50	9	20	-.22	.54	.75	-1.0	.58	-.9	.69	75.0	75.4	050
79	9	20	-.22	.54	1.09	.4	.95	.1	.52	75.0	75.4	079
118	9	20	-.22	.54	2.34	4.0	3.58	3.5	-.21	45.0	75.4	118
193	9	20	-.22	.54	1.22	.9	1.54	1.1	.42	65.0	75.4	193
216	9	20	-.22	.54	.91	-.3	.84	-.2	.60	75.0	75.4	216
220	9	20	-.22	.54	1.02	.2	.80	-.3	.57	65.0	75.4	220
228	9	20	-.22	.54	1.16	.7	.98	.1	.50	65.0	75.4	228
52	8	20	-.53	.55	1.57	1.9	3.21	3.0	.16	60.0	76.3	052
240	7	20	-.84	.57	1.46	1.5	1.48	1.0	.33	65.0	78.0	240
41	5	20	-1.55	.63	1.69	1.8	2.27	1.5	.16	70.0	82.1	041
209	4	20	-1.97	.67	1.20	.6	1.14	.5	.43	75.0	84.6	209
MEAN	14.4	20.0	1.55	.65	1.00	.1	.89	.1		79.7	80.5	
S.D.	2.9	.0	1.12	.17	.28	.9	.63	.6		10.8	6.1	

Table 3. Student fit statistics