EVALUATING QUALITY AND RELIABILITY OF FINAL EXAMINATION QUESTIONS FOR MAT0144 USING RASCH MODEL

T018

Suhaila Bahrom[1,*], Siti Nubailah Yaacob[2], A'fifah Happas[3], Nurhafizah Saidin[4] and Mohd Nazim Mohd Nawi[5]

[1,2,3,4,5]*Department of Mathematics, Centre for Foundation Studies, International Islamic University Malaysia, 26300 Gambang, Pahang, Malaysia*

INTRODUCTION

Final examination is one of the assessment tools to measure academic achievement among the students in Malaysia. In evaluating the quality of these questions, a discussion of reliability is essential. Reliability is the degree to which an instrument consistently measures the ability of an individual or group. In determining the quality and reliability of examination question paper, the best method used by most researchers is analyzing items. Item analysis has been discovered to be the most effective method for determining the quality of constructed test items. In this study, Rasch model was used to analyze each question in examination paper.

OBJECTIVE

This study used aimed to evaluate the quality and reliability of the final examination questions for Statistics (MAT0144) for Biological Module in Centre for Foundation Studies, IIUM.

METHODOLOGY

RESULTS

Person	344 Input			INFIT		OUTFIT	
	Score	Count	Measure	MNSQ	ZSTD	MNSQ	ZSTD
Mean	25.9	21	-1.99	1.09	0.1	1.0	0.1
S.D	14.6	0.0	0.95	0.57	1.1	0.61	0.8
Separation : 2.62 (Good)							
Person Reliability : 0.87 (Good)							
Items	21 Input			INFIT		OUTFIT	
	Score	Count	Measure	MNSQ	ZSTD	MNSQ	ZSTD
Mean	424.6	344.0	0.00	1.05	0.0	1.00	0.0
S.D	366.5	0.0	1.23	0.33	3.3	0.45	3.0
Separation : 9.28 (Excellent)							
Item Reliability : 0.99 (Excellent)							

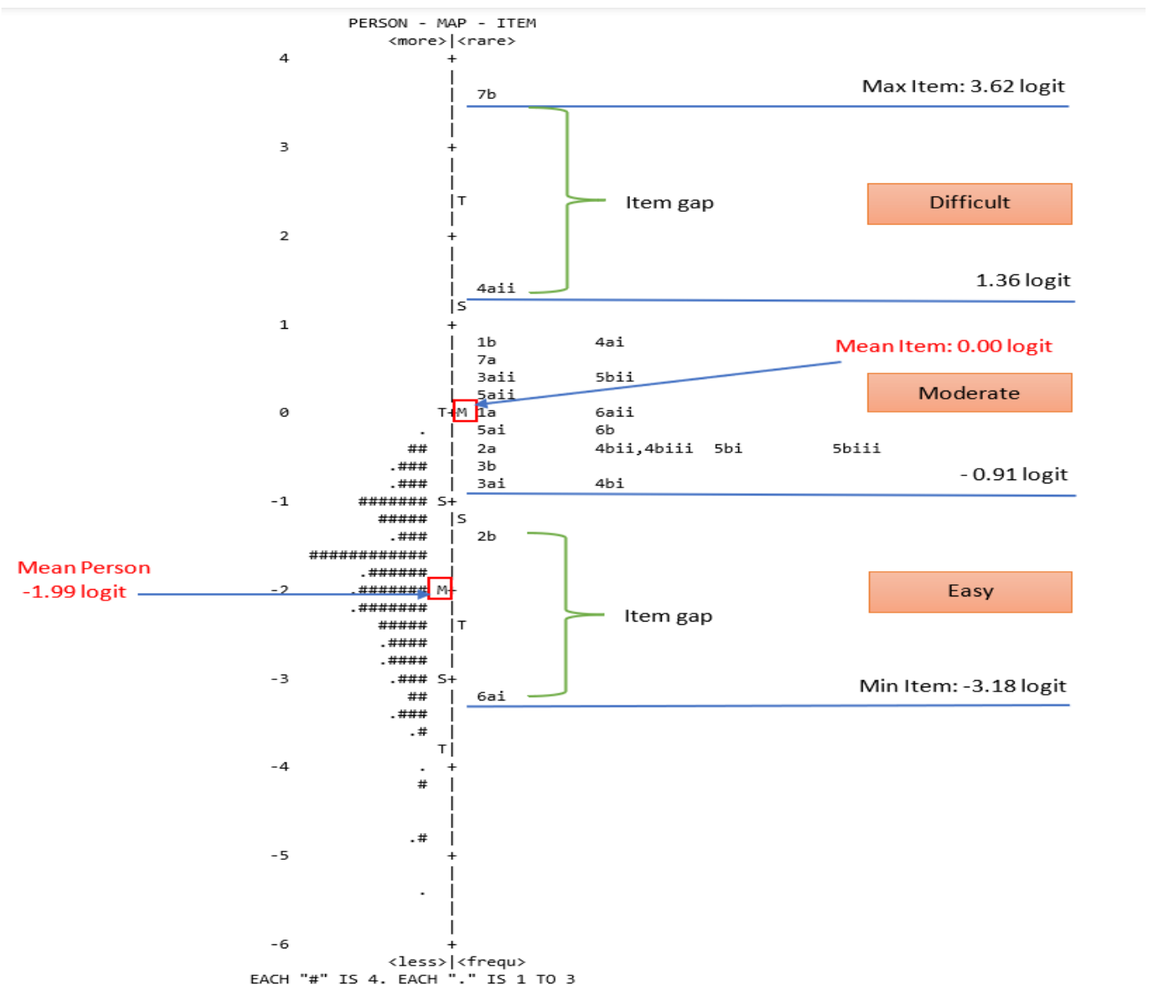
Person Raw Score-to-Measure Correlation=0.94
Cronbach Alpha (KR-20) Person Raw Score Reliability=0.90

ITEM STATISTICS: MISFIT ORDER

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PT-MEASURE CORR.	PT-MEASURE EXP.	EXACT MATCH OBS%	EXACT MATCH EXP%	ITEM
12	378	344	-.14	.06	1.13	1.6	2.46	9.1	A .35	.57	33.4	42.9	5ai
21	6	344	3.62	.40	1.77	1.5	.41	-1.2	B .17	.09	99.1	98.3	7b
13	267	344	.26	.06	1.71	6.9	1.64	3.9	C .47	.52	42.7	50.0	5aii

Statistics	Fit Indices
Outfit mean square values (MNSQ)	0.50 - 1.50
Outfit z-standardized values (ZSTD)	-2.00 - 2.00
Point Measure Correlation (PTMEA-CORR)	0.40 - 0.85

Items 5(a)(i) and 7(b) which was placed at the top was suggested to be misfit. These two items fail to fulfill all the three criteria suggested by Boone *et al.* (2014).



There is a large gap detected between item 7(b) and 4(a)(ii) which indicates that the examination paper is underrepresented in measuring students' ability accurately. This pattern also can be seen between item 2(b) and 6(a)(ii). Furthermore, there are 4 items tested on the same level of difficulties which are item 2(a), 4(b)(ii, iii), 5(b)(i) and 5(b)(iii) which these items supposedly fill in the gap on the difficult and easy level.

DISCUSSION/CONCLUSION

The item dimensionality shows that the examination question paper MAT0144 is unidimensional. It has a very high Cronbach's alpha (KR-20), and item and person reliability based on the analysis from the Rasch Model. The high index item separation value 9.28 indicates that the exam questions paper contains a wider range of items. However, this analysis shows that the mean of the items which is 0.00 index is higher than the mean of the students' performance –1.99 index, which indicates that the overall examination question paper was seen to be tough by most of the students. Thus, the questions' level of difficulty needs to be revised because there is a gap between the items in the two categories, difficult and easy. This finding provides valuable information for further item modification and future references for examination setter for MAT0144 subject. It is also encouraging other researchers to perform items analysis to ensure the quality and reliability of constructed examination questions.

ACKNOWLEDGEMENT

The authors would like to convey their thanks and gratefulness to Centre for Foundation Studies, IIUM and especially to Department of Mathematics for providing them the opportunity to run this study.

REFERENCES

1. Sumintono, B., & Widhiarso, W. (2015). *Aplikasi pemodelan rasch pada assessment pendidikan*. Trim komunikata.
2. Ridwan, I. I., Ali, R., Adam, Z., & Salim, K. R. (2017, March). Rasch Model Validation Of An Instrument To Measure Student Readiness For Embedded Systems Design Course. In *International Higher Education Conference (IHEC 2017)* (pp. 17-18).