Measurement of Rasch Analysis in Quality Model for a Web-based Integrated Student Assessment Application: Academician Perspective

Nur Razia Mohd Suradi[1], Saliyah Kahar[2], Nor Azliana Akmal Jamaludin[3]

[1,2]Faculty Communication, Visual Art and Computing, Universiti Selangor, 45600 Bestari Jaya, Selangor

[3]Faculty of Defense Science and Technology, Universiti Pertahanan Nasional Malaysia, 57000 Kuala Lumpur

[1]razia@unisel.edu.my

***Abstract*– Software quality is a complex issue in any applications. The purpose of an education web-based application in higher learning institution is to provide online facilities or services to academicians. The study attempts to identify quality attributes of web-based integrated Student Assessment Application (iCGPA) among academicians. Online survey questionnaires were distributed among academicians from Malaysian public institutions for pilot study. Rasch Measurement Model is used to identify the items needed for a web-based iCGPA system. Rasch Unidimensional Measurement Model software known as WiNSTEPS is used to assess and analyse respondent responses. The items were analysed for person-item map and misfit data. Each analysis is discussed in depth based on item and person. Results shows that the quality items proposed is agreed based on the responses retrieved. Finally, inputs from pilot study will be used to improve actual study.**

***Keywords*– Software Quality Model, Quality Model, Web-based Quality Model, Rasch Measurement Model, Rasch Analysis, Software and Educational Application**

I. INTRODUCTION

The purpose of implementing iCGPA (Integrated Cumulative Grade Performance Average) in HLI among five pilot public universities is to assess a more comprehensive performance of students by incorporating academic and non-academic elements in their academic transcript. A web-based integrated system is required to link between the Program Learning Outcome (PLO), Program Educational Objective (PEO) and Course Learning Outcome (CLO) for each particular program served by university also student assessment. Nonetheless, the quality aspect of the system, such as users' satisfaction should also be considered. The research conducted is to propose a quality model for the iCGPA application.

According to [1], five Public Institutions (PIs) have been selected to be pilot for this initiative. The five PIs are UKM, UMP, UiTM, UMK and UMT. Each of the PIs has independently developed an in-house integrated assessment system. According to the Ministry of Higher Education (MOHE), twenty (20) public institutions implement their own assessment techniques.

There are 5 research universities, namely UKM, UM, UTM, UPM and USM, but only one university implements iCGPA at present. For Comprehensive University (CU) category, the list includes UiTM, UIA, UNIMAS and UMS. UiTM is chosen to implement iCGPA. Lastly for FU, there are nine universities belongs to this category and three are selected to implement iCGPA which are UMP, UMK and UMT.

This paper emphasizes the quality model for web-based application addressing academic application needs.

II. BACKGROUND

[2] described various quality models suitable for web-based application in education domain. There is very limited study aim to propose a quality model for iCGPA application in HLI. A preliminary study has been conducted to identify quality attributes that are suitable for web-based iCGPA. Based on the study done, a set of questionnaires have been constructed to serve a quality model mentioned. The process in selecting suitable quality attributes are being discussed in detailed by [3].

Georg Rasch developed an analytical model of the item response theory (IRT) in 1960s and then later popularized by Benjamin Wright in United State [4]. As discussed by [5], Rasch is able to analysed item-question and ability of respondent towards the items.

This work was supported by Universiti Selangor as a part of PhD research.

Nur Razia Mohd Suradi. is with the Software Engineering Department, Universiti Selangor (Email: razia@unisel.edu.my)

Saliyah Kahar was with Universiti Selangor. She is now with the Department of Computer Science, Selangor Universiti, Malaysia, (Email: saliyah@unisel.edu.my)

Nor Azliana Akmal Jamaludin is with the Computer Science Department, National Defence University of Malaysia (UPNM), Kuala Lumpur, (Email: azliana@upnm.edu.my)

III. METHODOLOGY

The survey was conducted using online questionnaire among academicians from public institution since they are among the university implement using web-based iCGPA. For this research, the population is taken from PI's in Malaysia and are chosen by their university category. The sample is from Faculty which implements iCGPA. Cluster sampling is used to select respondents for the pilot study. [6] described that cluster sampling refers to a sampling technique that selected groups with similar characteristics and were chosen randomly.

The main objective of conducting the pilot study is to determine the accuracy of the questionnaire. According to [7], a pilot study is a trial run before having a real study to the instrument. Nevertheless, this does not guarantee success for actual study.

Each of the questionnaire item analysed using a measurement model which is capable to perform the following tasks: a) produce linear measures, b) overcome missing data, c) give estimates of precision, d) tools to detect misfit. All the operations mentioned are performed by Rasch.

IV. INSTRUMENTATION

The survey questionnaire for this research used Likert Scale (4-3-2-1) for descriptive response category for ("1-2-3-4") (Strongly Disagree-Disagree-Agree-Strongly Agree). The word Content validity refers to how to examine the aspects of the construct. One of the practical ways to assess content validity is to consult an expert's opinion [8]. Expert opinion or subject matter experts (SME) have to be formed to answer the related question or test.

According to [9] who developed method of measuring content validity, the content validity ratio (CVR) is as formulated below in eq. (1):

$$CVR = \frac{n_e - N_2}{N_2} \quad (1)$$

in which n_e is the number of panellists indicating "essential" and N' is the total number of panelists. Based on the formula above, Lawshe stated that a minimum of five SMEs is sufficient to perform content validity. This research used five subject matters expert to check content validity where the initial questionnaire is then distributed to them.

V. RASCH MEASUREMENT MODEL

Rasch Model is used to analyse data. Application of the Rasch model through software such as Winstep [4] and other Rasch software be responsible for estimates of person and threshold locations on the latent variable scale. The software also yields indices of item and person fit to show that the requirement of unidimensionality is met. Rasch answer on by what method to have the right measurement with valid instrument. Instrument is extremely crucial if involve human life.

VI. RESULT AND DISCUSSION

A) Summary Statistic

From Table II, it can be concluded that N=30 respondents is a valid response. A total of 1860 data points arising from 30 respondents on 62 items were examined. The value of Chi-Square X^2 is 2870.94.

Table I: Summary of Measured 62 Items

SUMMARY OF 62 MEASURED Items

	RAW SCORE	COUNT	MEASURE	MODEL ERROR	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD
MEAN	83.8	30.0	.00	.34	.99	-.1	.99	.0
S.D.	7.3	.0	.81	.02	.33	1.2	.47	1.2
MAX.	97.0	30.0	1.92	.38	2.02	2.9	3.49	4.5
MIN.	65.0	30.0	-1.68	.30	.44	-2.8	.39	-2.8

REAL RMSE .36 ADJ.SD .73 SEPARATION 2.01 Item RELIABILITY .80
MODEL RMSE .34 ADJ.SD .74 SEPARATION 2.16 Item RELIABILITY .82
S.E. OF Item MEAN = .10

UMEAN=.000 USCALE=1.000
Item RAW SCORE-TO-MEASURE CORRELATION = -1.00
1860 DATA POINTS. APPROXIMATE LOG-LIKELIHOOD CHI-SQUARE: 2870.94

The purpose of Cronbach's alpha is to measure internal consistency, indicating how closely correlated a set of items are as a group. The Cronbach-α value is 0.97 means that the high raw score for the instrument in identifying quality characteristics for iCGPA application. The value of Cronbach's-α is above 0.7 (reliability > 0.7). From Table III, the person reliability is 0.97 which showed high valid responses of 99.8%.

Table II: Summary of Measured 30 Persons

INPUT: 30 Persons 62 Items MEASURED: 30 Persons 62 Items 4 CATS 1.0.0

SUMMARY OF 30 MEASURED Persons

	RAW SCORE	COUNT	MEASURE	MODEL ERROR	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD
MEAN	173.1	62.0	.81	.25	1.02	-.5	.99	-.6
S.D.	26.0	.0	1.74	.07	.55	2.9	.54	2.8
MAX.	245.0	62.0	6.58	.60	2.17	4.5	2.14	4.2
MIN.	128.0	62.0	-1.50	.21	.17	-6.3	.15	-6.5

REAL RMSE .30 ADJ.SD 1.71 SEPARATION 5.66 Person RELIABILITY .97
MODEL RMSE .26 ADJ.SD 1.72 SEPARATION 6.58 Person RELIABILITY .98
S.E. OF Person MEAN = .32

Person RAW SCORE-TO-MEASURE CORRELATION = .99
CRONBACH ALPHA (KR-20) Person RAW SCORE RELIABILITY = .97

MeanItem from Table I is 0.00 and *MeanPerson* give a result as +0.81 logit.

B) Person Fit

The sum of Mean and Standard Deviation (SD) based on Point Man Correlation (PTMEA), Mean Square (MNSQ) and z-standard (ZSTD) determine either the item fit to be in the

model. Item whose MNSQ near to 1 and z-std nearer to 0 means that the item is fit. In Table III, it shows that person 24, 25, 23,17 and 19 are misfit with MNSQ >1.5 and z-std >±2.

Table III: Person Fit

Person STATISTICS: MISFIT 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
24	172	62	.57	.24	2.17	4.6	2.14	4.2	A .46	45.2	68.7	1
25	192	62	1.79	.25	2.09	4.0	2.04	4.0	B .54	40.3	70.5	1
23	177	62	.86	.24	1.90	3.6	1.88	3.4	C .69	51.6	70.5	1
17	139	62	-1.03	.21	1.82	4.0	1.82	4.1	D .53	35.5	55.7	1
19	241	62	5.61	.41	1.76	2.1	1.75	1.5	E .14	91.9	88.7	1
7	245	62	6.58	.60	1.68	1.9	1.04	.3	F .19	96.8	95.1	3
27	155	62	-.31	.22	1.50	2.6	1.60	2.9	G .63	43.5	59.2	1
9	166	62	.24	.23	1.59	2.7	1.53	2.4	H .01	50.0	65.4	1
28	160	62	-.07	.22	1.28	1.5	1.27	1.4	I .54	48.4	61.5	1
10	202	62	2.43	.25	1.24	1.4	1.22	1.3	J .33	54.8	66.2	3
30	156	62	-.26	.22	1.20	1.1	1.16	.9	K .47	67.7	59.4	3
21	160	62	-.07	.22	1.04	.2	1.01	.1	L .40	61.3	61.5	1
18	128	62	-1.50	.21	1.02	.2	1.00	.1	M .44	61.3	56.0	1
12	170	62	.46	.24	1.02	.1	1.01	.1	N .22	64.5	67.4	2
16	160	62	-.07	.22	.82	-1.0	.77	-1.3	O .66	64.5	61.5	1
8	213	62	3.12	.25	.81	-1.0	.80	-1.4	o .54	66.1	62.7	6
15	148	62	-.63	.21	.72	-1.6	.71	-1.8	n .26	64.5	56.5	1
2	162	62	.03	.23	.71	-1.7	.70	-1.7	m .44	67.7	62.8	2
3	162	62	.03	.23	.71	-1.7	.70	-1.7	l .44	67.7	62.8	2
4	162	62	.03	.23	.71	-1.7	.70	-1.7	k .44	67.7	62.8	2
14	150	62	-.54	.21	.66	-2.0	.70	-1.9	j .42	74.2	57.1	1
20	163	62	.08	.23	.70	-1.6	.68	-1.8	i .57	72.6	63.3	1
5	167	62	.29	.23	.67	-1.9	.67	-1.8	h .43	64.5	65.9	1
29	149	62	-.59	.21	.64	-2.0	.64	-2.4	g .51	59.7	56.5	1
6	193	62	1.86	.25	.58	-2.6	.58	-2.4	f .19	83.9	70.1	6
26	164	62	.14	.23	.57	-2.7	.56	-2.6	e .64	82.3	64.2	1
13	178	62	.92	.25	.55	-2.6	.51	-2.7	d .12	80.6	70.7	1
1	187	62	1.48	.25	.22	-5.6	.21	-5.7	c .05	96.8	71.5	7
11	186	62	1.41	.25	.17	-6.0	.15	-6.5	b .00	96.8	71.5	2
22	186	62	1.41	.25	.17	-6.0	.15	-6.5	a .00	96.8	71.5	1
MEAN	173.1	62.0	.81	.25	1.02	-.6	.99	-.6		67.3	65.9	
S.D.	26.0	.0	1.74	.07	.55	2.6	.54	2.8		17.2	8.6	

TABLE 6.4 C:\Users\Nur Razia\Documents\PHD\Softwa ZOU005WS.TXT Nov 22 23:00 2018
INPUT: 30 Persons 62 Items MEASURED: 30 Persons 62 Items 4 CATS 1.0.0

C) Item Fit

For item to fit in the model, the same thing occurs as for person fit. For a second time, the recommendation is based on Point Mean Correlation (PTMEA), Mean Square (MNSQ) and z-standard (ZSTD). Item whose MNSQ is nearer to 1 and z-std nearer to 0 is a better fit. The item will remain if the-0.5 <MNSQ < 1.5 and ZSTD ±2. Refer Table IV.

Table IV: Item Fit

Item STATISTICS: MISFIT 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%	Item
30	87	30	-.33	.35	2.02	2.9	3.49	4.5	A .18	53.3	68.6	Sb
15	86	30	-.21	.35	1.28	1.9	2.14	2.6	B .50	53.3	66.9	Rc
28	92	30	-.98	.37	1.75	2.1	1.61	1.3	C .27	63.3	72.1	Ff
52	70	30	1.46	.31	1.17	.8	1.74	2.5	D .48	76.7	60.2	Sef
40	65	30	1.92	.30	1.63	2.4	1.57	2.2	E .62	76.7	57.9	Sl
51	88	30	-.45	.35	1.61	1.9	1.45	1.2	F .47	53.3	69.5	See
45	81	30	.36	.33	1.47	1.7	1.46	1.4	G .45	60.0	62.2	Ae
21	81	30	.36	.33	1.46	1.6	1.44	1.4	H .65	60.0	62.2	Eb
16	85	30	-.09	.34	1.45	1.6	1.45	1.3	I .57	60.0	66.1	Rd
29	76	30	.88	.32	1.12	.6	1.41	1.4	J .49	70.0	60.9	Sa
39	66	30	1.83	.30	1.39	1.5	1.34	1.4	K .66	60.0	58.5	Sk
56	93	30	-1.11	.37	1.14	.6	1.10	.4	S .49	76.7	72.3	Id
11	69	30	1.55	.31	1.12	.6	1.11	.5	T .58	66.7	60.4	Uk
18	95	30	-1.40	.38	1.10	.4	1.01	.2	U .50	76.7	72.7	Rf
10	85	30	-.09	.34	1.07	.3	.97	.0	X .47	60.0	66.1	Uj
60	91	30	-.84	.37	.62	-1.8	.59	-1.0	g .68	83.3	71.7	Ih
6	85	30	-.09	.34	.62	-1.5	.55	-1.5	f .72	60.0	66.1	Uf
19	93	30	-1.11	.37	.60	-1.4	.54	-1.1	e .74	60.0	72.3	Rg
47	97	30	-1.68	.38	.59	-1.5	.55	-.8	d .57	83.3	72.4	Sea
26	90	30	-.71	.36	.56	-1.7	.53	-1.3	c .73	60.0	71.2	Fd
1	82	30	.25	.33	.50	-2.2	.44	-2.2	b .75	63.3	63.5	Ua
5	78	30	.68	.32	.44	-2.8	.39	-2.8	a .80	63.3	60.6	Ue
MEAN	83.8	30.0	.00	.34	.99	-.1	.99	.0		67.3	65.9	
S.D.	7.3	.0	.81	.02	.33	1.2	.47	1.2		11.3	4.5	

D) Unidimensionalty

To guarantee the measurement is measuring the specific objective, therefore, unidimensionality is essential. Rasch Analysis applies the Principal Component Analysis (PCA) of the residuals to know on how much variance of the instrument measuring that supposedly to measure.

Fig. 5 depict that the raw variance explained by measures is 74.8% closely match the expected 74.4%. Nevertheless, the analysis shows that only 20% if unidimensionality requirement minimum. Rasch cut-low point of 40% not achieved as reported by [10] .

	Empirical
Modelled	
Total variance in observations = 245.7 100.0% 100.0%	
Variance explained by measure = 183.7 74.8% 74.4%	
Unexplained variance (total) = 62.0 25.2% 100.0% 25.6%	
Unexplned variance in 1st contrast = 8.5 3.5% 13.7%	

Fig. 1: Standardized Residual variance (in Eigenvalue units)

E) Person-Item-Map

The Person-Item-Map is also known as Wright Map. Wright Map or Variable Map is a tool in Rasch model measurement that analyse the comprehensive outlook of the data. This map, also called as construct map, exemplifies person abilities and item difficulties using the same logit ruler that provides information about the result of a test [11]

Thirty respondents take took part in answering the questionnaire with different background and different experiences in using web-based application. The most excellent academicians are at highest ability 6.58 logit as seen at Fig. 2. The poor academicians are at -1.5 logit. A few items between Item*Max* and Person*Mean* can be classified as the most difficult items.

VII. CONCLUSION

The observations clarified above indicated that an instrument needs to be testing among the respondents before conducting the actual survey. The Rasch model has the ability to assess the quality attributes that belong to iCGPA application. The Rasch model enables researcher to identify missing data, Person and item misfit is discussed. Content validity is done based on Lawshe experience. Here, five subject expert matters (SME) has been introduced as suggested. Item which is supposed to deleted will be removed in current questionnaire as preparation for actual study.

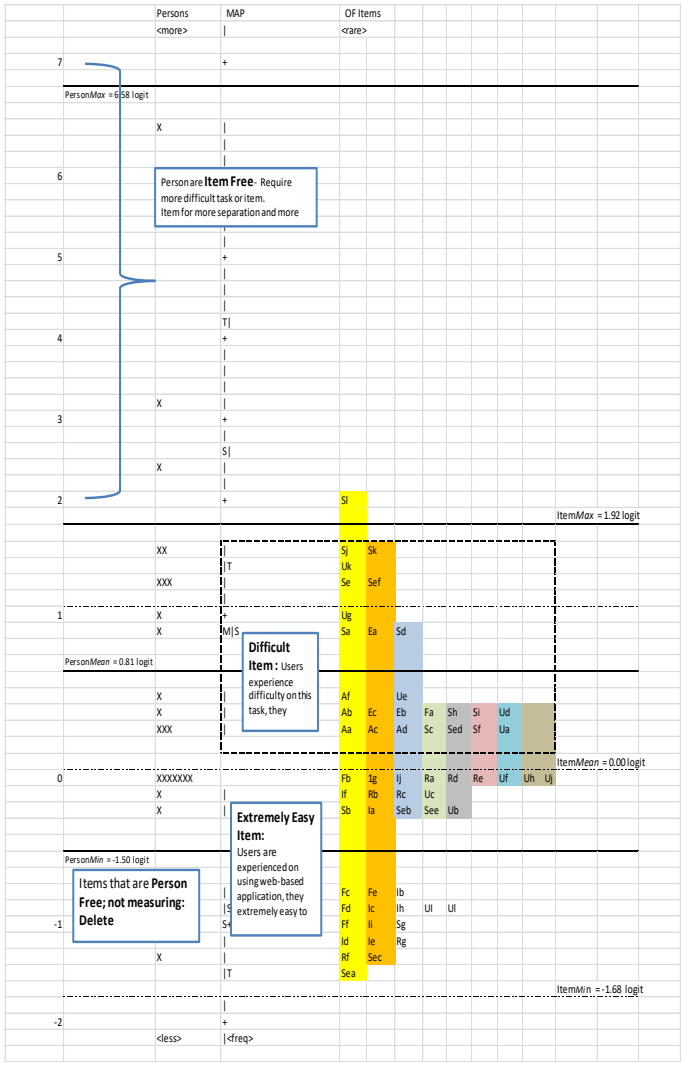

Fig. 2: Person-Item Map Distribution Map – Wright Map (PIDM)

ACKNOWLEDGEMENT

The author would like to express sincere gratitude to the supervisory team, Dr. Saliyah Kahar and Dr. Nor Azliana Akamal Jamaludin for continuous support in the journey of PhD research and Universiti Selangor for supporting the research.

REFERENCES

[1] BorneoPost, "Five Public universities to pilot iCGPA," BorneoPost online, 2015. [Online]. Available: http://www.theborneopost.com/2015/08/11/five-public-universities-to-pilot-icgpa-programme/.

[2] Nur Razia Mohd Suradi, Saliyah Kahar, and N. A. A. Jamaludin, "A Review on Software Quality Attributes for Web-Based Applications," in International Conference on Engineering and Applied Science Purwokerto, Indonesia, 2016, Vol. 1, pp. 181–191.

[3] Nur Razia Mohd Suradi, Saliyah Kahar, and Nor Azliana Akmal, "Identification of software quality characteristics on academic application in Higher Education Institution (HEI)," J. Telecommun. Electron. Comput. Eng., Vol. 10, No. 2–7, pp. 133–136, 2018.

[4] J. Linacre, "Winsteps Rasch Measurement Version 3.71. Winstep.com," 2011.

[5] B. Sumintono, "Rasch Model Measurement as Tools in Assessment for Learning Rasch Model Measurement as Tools in Assessment for Learning," in International Conference on Education, Universitas Negeri Surabaya at: Wyndham Hotel, Surabaya, Indonesia, 2018, October 2017.

[6] P. Singh, C. Y. Fook, and G. K. Sidhu, A Comprehensive Guide in Writing A Research Proposal (2015th ed.), Kuala Lumpur: Venton Publishing Sdn Bhd (M), 2015.

[7] Simon M.K., "Conducting Pilot Studies," Diss. Sch. Res. Recipes success, pp. 5–10, 2011.

[8] D. F. Polit and C. T. Beck, "The Content Validity Index: Are You Sure You Know What's Being Reported ? Critique and Recommendations," Res. Nurs. Health, Oct 29(5), pp. 489–497, 2006.

[9] C. H. Lawshe, "A Quantitative Approach to Content Validity," in Content Validity II, 1975, Vol. 28, No. 1, pp. 563–575.

[10] K. M. Conrad, M. L. Dennis, B. B. Riley, and R. Funk, "Validation of the Crime and Violence Scale (CVS) to the Rasch Measurement Model , GAIN Methods Report 1.1", July 2016, 2009.

[11] Mark Wilson, Constructing Measures an Item Response Modeling Approach. 2005.