Development of M-Learning Implementation Instrument Based on Competency-Based Education: Rasch Analysis

Siti Azwanee Abdullah, Muhammad Sukri Saud, Mohd Hizwan Mohd Hisham

Development of M-Learning Implementation Instrument Based on Competency-Based Education: Rasch Analysis

Siti Azwanee Abdullah, Muhammad Sukri Saud, Mohd Hizwan Mohd Hisham

Faculty of Social Sciences and Humanities, Universiti Teknologi Malaysia, Malaysia

Abstract

This study was undertaken to produce empirical evidence of validity and reliability of the m-learning implementation instrument constructed based on competency-based education using Rasch Model Analysis. It aims to validate the instrument by using tests in Rasch analysis such as item polarity, fit measure, principal component analysis and reliability. The questionnaire was distributed to 55 teachers teaching Upper Secondary Vocational Programme (PVMA) in the states of Johor and Selangor. Analysis of the Rasch Measurement Model was performed using Winsteps software version 3.69.1.11. The final instrument recorded 47 items that can be used to measure the eight constructs of the study. This study shows that the Rasch Measurement Model can help researchers build a good instrument as the items constructed offset psychometric standards.

Keywords: Mobile Learning, Competency-Based Education, Validity, Reliability, Rasch Measurement Model Approach

Introduction

The study of the validity and reliability of the instrument is very important to maintain the accuracy and obtain quality questionnaires. The higher the value and the higher the validity and reliability of the questionnaire, the more accurate the data will be. In the Rasch Measurement Model, the validity of an instrument can be identified by reference to key analyses such as item polarity, item matching, unidimensionality, and rating scale (Bond & Fox, 2007; Rasch, 1980). Accordingly, this study was conducted to generate empirical evidence on the validity and reliability of the m-learning implementation questionnaire based on Competency-Based Education using the Rasch Measurement Model.

Rasch Model Analysis is a mathematical formula that looks at the probability of an individual answering an item correctly or supporting an item depending on the individual's ability and capability and the change in difficulty of the item (Bond & Fox, 2015a). The Rasch measurement model can be defined as an idea, principles, guidelines or techniques that require measurement to be made of latent properties (Aziz, Masodi, & Zaharim, 2013). Furthermore, among the advantages of the Rasch model is that it provides linear

measurement, which is capable of detecting missing data, able to provide accuracy, and able to detect misfits (pattern matching) and clothing (outlier matching) (Aziz et al., 2013). The tests for the Rasch model are as follows:

i.   Item Polarity

For the polarity analysis of items, it is carried out to see the parallelism of the items on the instrument moving in the same direction. Concerning the Point Measure Correlation value (PTMEA CORR), this value must be positive to indicate that the item can distinguish the ability between respondents. If the PTMEA CORR value is positive, then it suggests that each item can achieve its purpose of measuring the construct to be measured. In contrast, if the PTMEA CORR value obtained is negative, the item is inconsistent in measuring the construct of a study (Bond & Fox, 2015a).

ii.   Item Fit Measure

The second analysis is Item Fit (Fit Measures), which detects any problematic items in the data or outlier items or misfits in measuring a construct. The recommended MNSQ range is between 0.5 and 1.5 (Aziz et al., 2013; Boone et al., 2014). Bond and Fox (2015a) argue that an MNSQ value should be between 0.5 and 1.5 for the Dichotomous Data. The value of the MNSQ Outfit index needs to be considered first to determine the fit items that measure a construct (Asbulah et al., 2018). If the MNSQ value exceeds the range, it means that the item confuse the respondent, while if the MNSQ value decreases from the range, it indicates that the item is too easy to expect (Harun & Ghani, 2016). Apart from the MNSQ outfit value, the Zstd value should also be noted. The value of Zstd should be in the range of -2 to +2 (Bond & Fox, 2015a). MNSQ outfit values that are not in the set range typically show Zstd values that exceed the range -2.0 <Zstd <+0.2 (Boone et al., 2014; Linacre, 2012).

iii.   Principle Component Analysis

Next, Principal Component Analysis (PCA), or Principal Component Analysis, is conducted to ensure the uniformity of the instrument's dimensions is in a unidimension and a common direction (Adzhar, Karim, & Sahrin, 2017; Aziz et al., 2013). Unidimensionality is defined as a single latent property on a latent variable to form a quality item for a research instrument (Brentari & Golia, 2007; Wu & Adams, 2007). The concept of unidimensionality is often defined as a single latent trait that can explain performance on the items that make up a questionnaire.

Rasch analysis also can determined unidimensionality. Unidimensionality was examined with Principal Component analysis (PCA) of the residuals that involve analysis to determine the value of variance explained by the measure, the level of item interference in the first contrast and the Eigen value. In unidimensional measures, it is expected that the observed variance explained by the measures roughly matches the expected variance in the model. The value of raw variance explained by the measurement (Raw Variance Explained by Measure) should exceed the minimum accepted value of 40% (Aziz et al., 2013) and the level of item interference based on the value of variance that is not explained in the first contrast (Unexplained variance in the 1st contrast) should be below the value of 15% (Fisher, 2007). Furthermore, an Eigen value of less than five indicates that the second dimension does not clearly exist (Linacre, 2005).

Objective

This study aims to test the validity and reliability in the implementation of mobile learning based on Competency-Based Education instruments for the teacher at the secondary school using Rasch analysis. The objectives of this study are to (1) identify the polarity item that measures the constructs, (2) examine the suitability of item (item fit) of the instrument, (3) detect unidimensionality of construct, and (4) measure the reliability of the construct.

Methodology

This study was conducted using a survey method on 55 people in the study sample involving teachers who teach Upper Secondary Vocational Programme (PVMA). The number of samples for this pilot study is sufficient based on the recommendation of Linacre (1994), who suggested that the minimum number for one sample is 30 people. The respondents were selected using a convenient sampling technique. The study sample is PVMA teachers from the states of Johor and Selangor. The selection of this sample is based on respondents' involvement in the teaching of skills programs according to NOSS issued by the Department of Skills Development (DSD) to produce students who have a Malaysian Skills Certificate (MSC). The analysis for this pilot study was using the Rasch Measurement Model analysis.

The research instrument used was a questionnaire. This questionnaire is divided into two parts. Section A contains the demographic information items of the respondents, while Section B is the item that measures the M-Learning implementation construct. The constructs are students, teachers, technology, learning environment, content, assessment, learning strategies and learning activities. Details of the questionnaire instrument used are as in the Table 1 below.

Table 1:   Details of the questionnaire instrument

Part	Item	No. of item
A	Respondent demographics	2
B Construct	Students	6
	Teachers	7
	Technology	6
	Learning environment	5
	Content	6
	Assessment	7
	Learning strategies	5
	Learning activities	5

The process of building this m-learning implementation instrument involves three main stages, namely exploring the constructs and items for the constructs studied through interviews, providing measurement guidelines such as Likert scales, and expert review. The details of each stage are as follows:

Stage 1: Constructs and items for the constructs studied were obtained through interviews and document analysis. Semi-structured interviews were conducted, and the findings were analysed using thematic analysis. Two experts agreed upon the formation of themes through the reliability of Cohen's Kappa Coefficient.

Stage 2: The Likert scale was selected for respondents to provide consensus feedback on each item found on the questionnaire. A 6-point Likert scale was used consisting of 1 -Strongly Disagree, 2- Disagree, 3- Slightly Disagree, 4- Moderately Agree, 5- Agree, 6 -Strongly Agree.

Stage 3: Content experts and psychometric experts reviewed the questionnaire to ensure that the content met the psychometric characteristics of the reliability and validity of the items studied. Improvements were made based on comments made by experts before the questionnaire was distributed to the study sample.

The findings were recorded in IBM SPSS version 22 software, Microsoft excel, and Winsteps software.

Results and Discussion

This section describes the demographic distribution of study respondents, item validity analysis consisting of polarity test, fit measure and Principal Component Analysis (PCA) and reliability of study tools for m-learning implementation that have been constructed using Rasch analysis.

a) Respondent Demographics

The final data set comprised a total sample of 55 teachers, with 19 male (35%) and 36 female (65%). Besides that, 1 of the sample has teaching experience less than 5 years, 13 of them with teaching experience between 11 to 15 years (23.6%), 20 sampel (36.4%) with 16 to 20 years teaching experience, 17 sampel (30.9%) with 16 to 20 years teaching experience and 4 sampel with teaching experience more than 20 years. Table *2* provides the demographic characteristics of the respondent.

Table 2: Respondent Demographic

No	Demographic	Item	Frequency	Percentage (%)
1.	Gender	Male	19	35.0
		Female	36	65.0
2.	Teaching Experience	Less than 5 years	1	1.8
		6-10 years	13	23.6
		11-15 years	20	36.4
		16-20 years	17	30.9
		20 years and above	4	7.3

b) Findings of the Rasch Model Analysis

The findings from the analysis of the Rasch measurement model conducted to validate the research instrument are as in the

Table 3 below.

Table 3: Validity Analysis of Research Instrument

Construct	Validation Criteria								
	Infit		Outfit		Pt Mea Corr	Raw Variance explained by measure		Unexplained variance in 1st contrast	Eigen value
	MNSQ Range	ZSTD Range	MNSQ Range	ZSTD Range		Empirical	Model		
Students	1.17 to 0.75	0.8 to -1.3	1.17 to 0.74	1.1 to -1.3	Positive	45.6%	46.1%	8.6%	2.1
Teachers	1.25 to 0.59	1.2 to -2.2	1.38 to 0.56	1.6 to -2.1	Positive	54.8%	54.8%	8.1%	1.9
Technology	1.64 to 0.71	2.3 to -1.2	1.32 to 0.63	1.1 to -1.4	Positive	71.8%	71.5%	8.3%	2.4
Learning environment	1.44 to 0.62	1.4 to -0.8	1.33 to 0.58	1.2 to -2.1	Positive	63.4%	61.8%	11.9%	1.5
Content	1.42 to 0.70	1.6 to -1.3	1.23 to 0.68	0.9 to -1.3	Positive	71.3%	71.0%	9.3%	2.3
Assessment	1.17 to -1.60	0.7 to -1.6	1.49 to 0.55	1.9 to -1.8	Positive	50.8%	51.3%	10.1%	2.2
Learning strategies	1.40 to 0.60	1.7 to -2.1	1.31 to 0.58	1.4 to -2.2	Positive	73.1%	72.7%	9.5%	1.8
Learning activities	1.37 to 0.59	1.4 to -1.8	1.49 to 0.56	1.6 to -1.9	Positive	62.9%	62.7%	10.4%	1.9

Based on this table, it is found that all constructs have MNSQ outfit values in the range of 0.5 to 1.5. The value of the MNSQ outfit is evaluated to determine the suitability of the item (item fit) that measures a construct. Boone et al (2014) stated that the suitability range of productive items is between 0.5 and 1.5, and if it is found that an item exceeds this range, the value of Zstd is also found to exceed the accepted range of -2.0 to +2.0. This means, each item on the construct contributes to the full measurement of the implementation of M-Learning as modelled by the Rasch Model.

In addition, the analysis of item polarity (Point Measure Correlation) or item parallelism found that the PTMEA CORR value is positive. Based on the analysis that has been conducted as in

Figure 2, it is found that the PTMEA CORR values for all constructs have no negative value. This means, all the items present on each construct move in the same direction parallel to the measured domain (Bond & Fox, 2015b). No items should be paid attention to or dropped from the existing list of items. In short, all items in this research instrument measured the construct.

Furthermore, Rasch analysis can also detect the instrument's capability in a unidimension with an acceptable level of interference (Asbulah et al., 2018). From the analysis conducted, all constructs were found to have a value of Raw Variance Explained By Measure exceeding 40%, which exceeds the value recommended by Aziz et al. (2013). In addition, the measured measurement was reported not to be similar to the one modelled due to interference (noise). It is found that the noise value that is the value of variance that is not explained in construct 1 (unexplained variance in 1st contrast) was between 8.1% and 11.9%. This value is below the 15 per cent value as suggested by Fisher (2007). Next, the eigen value was between 1.5 and 2.4. These values are classified as well-controlled (Asbulah et al., 2018). This means that the most significant factor taken from the residual has a strength in the range of one to two items only.

c) Reliability of the Research Instrument
Reliability refers to the similar expectations generated when an individual is given a similar set of questions measuring the same construct (Bond & Fox, 2007). The reliability of this Rasch Measurement Model approach refers to the Cronbach's Alpha value to measure the level of reliability of the items in the instrument. According to (Bond & Fox, 2007), an item reliability value above 0.80 indicates that the item has an excellent level of reliability. The Table *4* below shows the interpretation of Cronbach's alpha score proposed by (Bond & Fox, 2007).

Table 4: Interpretation of Cronbach's alpha score

Cronbach's alpha score	Reliability Interpretation
>0.8	Very good and effective and has a high consistency
0.7- 0.8	Good and acceptable item
0.6 – 0.7	Items are acceptable
<0.6	Items need to be fixed
<0.5	Items need to be dropped

Therefore, based on the pilot study, the reliability value obtained from the Cronbach's Alpha value for each construct are as shown in Table *5*.

Table 5: Cronbach's alpha value for each construct

Construct	No of item	Alpha Coefficient Value	Interpretation
Students	6	0.829	Very good
Teachers	7	0.897	Very good
Technology	6	0.959	Very good
Learning environment	5	0.886	Very good
Content	6	0.950	Very good
Assessment	7	0.880	Very good
Learning strategies	5	0.937	Very good
Learning activities	5	0.940	Very good

Conclusion

The findings demonstrate the instrument has adequate psychometric properties for its validity and reliability value. As mention earlier, the aim of this study is to test the validity and reliability of m-learning implementation instrument based on Competency-Based Education using Rasch Model Analysis. Researchers built this instrument to develop a framework for the implementation of m-learning for students. The results shows that item polarity test and misfit test could contribute to the instrument's validity for the development of the m-learning implementation framework. In addition, dimensionality also indicates that the instrument is one-dimensional in nature that is uniform with the level of interference received. Reliability tests also suggest that items are acceptable to measure the implementation of m-learning. As a result, based on the validity and reliability test in made on this instrument, it indicates that this instrument is verify and fits to be use in the context of the technical and vocational education (TVET). This instrument also can be used by the other researchers for future study. Thus, it was practical for researchers to use this instrument to identify the best practice for mobile learning implimentation in TVET context to improve teaching delivery to enhance student learning achievement.

Implications and Suggestions for Future Research

The Rasch output has created a paradigm shift in measuring perception by producing more meaningful data and a quality instrument. Thus, the implications of this analysis help researchers in developing a good instrument for the TVET. It help the stakeholders by using this instrument as a guideline to identify the best practice for implement mobile learning. Hence, the framework for mobile learning implimentation in TVET context can be develop based on the teachers or students response. In future research, the researcher suggest that might use the variables developed in this study to check for the relationship between the variables by using Confirmatory factor analysis. In addition, it is better for the instrument to go through further validity and reliability test with larger sample size during real study.

Acknowledgements

Corresponding Author

Siti Azwanee binti Abdullah
Faculty of Social Sciences and Humanities, Universiti Teknologi Malaysia.
Email : sitiazwanee@graduate.utm.my / waneeabdullah@gmail.com

References

Adzhar, H., Karim, A. A., & Sahrin, M. U. (2017). Pembangunan Instrumen Penerimaan E-Pembelajaran Pelajar Pascasiswazah Menggunakan Analisis Rasch. *Jurnal Pendidikan Malaysia*, *42*(2), 147–155. https://doi.org/10.17576/jpen-2017-42.02-08

Asbulah, L. H., Lubis, M. A., Aladdin, A., & Sahrim, M. (2018). Kesahan dan Kebolehpercayaan Instrumen Strategi Pembelajaran Kolokasi Bahasa Arab: Analisis Menggunakan Model Rasch. *Jurnal Pendidikan Malaysia*, *1*(1), 131–140. https://doi.org/10.17576/jpen-2018-43.03-14

Aziz, A. A., Masodi, M. S., & Zaharim, A. (2013). *Asas Model Pengukuran Rasch: Pembentukan Skala & Struktur Pengukuran*. Bangi: Penerbit Universiti Kebangsaan Malaysia.

Bond, T. G., & Fox, C. M. (2007). *Applying The Rasch Model: Fundamental Measurement In The Human Sciences* (Second Edi). Mahwah: New Jersey London: Lawrence Erlbaum Associates, Publishers.

Bond, T. G., & Fox, C. M. (2015a). *Applying The Rasch Model: Fundamental Measurement in the Human Sciences* (Third Edit). New York: Routledge.

Bond, T. G., & Fox, C. M. (2015b). *Applying the Rasch Model Fundamental Measurement* (Third Edit). New York: Routledge.

Boone, W. J., Yale, M. S., & Staver, J. R. (2014). *Rasch analysis in the human sciences*. *Rasch Analysis in the Human Sciences*. https://doi.org/10.1007/978-94-007-6857-4

Brentari, E., & Golia, S. (2007). Unidimensionality in the Rasch model: how to detect and interpret. *Statistica*, 253–261. https://doi.org/10.6092/issn.1973-2201/3508

Fisher, J. W. (2007). Rating Scale Instrumen Quality Criteria. *Rasch Measurement Transactions*, *21*(1), 1095.

Harun, N., & Ghani, F. A. (2016). Kesahan dan kebolehpercayaan soal selidik amalan belajar pelajar berpencapaian rendah sekolah berasrama penuh. *Jurnal Kemanusiaan*, *25*(3), 40–52.

Linacre, J. M. (2005). *A User's Guide And Program Manual To Winsteps: Rasch Model Computer Program*. Chicago: MESA Press.

Linacre, J. M. (2012). *A User's Guide to Winsteps Rasch Model Computer Programs*. Chicago: MESA Press.

Linacre, M. (1994). Sample size and item calibration stability. *Rasch Measurement Transactions*, *7*(4), 328.

Rasch, G. (1980). *Weblogs models for some intelligence and Student tests*. Chicago: The University of Chicago Press.

Wu, M., & Adams, R. J. (2007). *Applying the Rasch model to psycho-social measurement: A practical approach. Educational Measurement Solutions*.