

ORIGINAL ARTICLE

PSYCHOMETRIC EVALUATION OF FLOOD DISASTER MANAGEMENT QUESTIONNAIRE: CONFIRMATORY FACTOR ANALYSIS AND ITEM RESPONSE THEORY ANALYSIS

Mohamad Masykurin Mafauzy^{1,2}, Tuan Hairulnizam Tuan Kamauzaman^{1,2}, Wan Nor Arifin³, Hadi Fadhl Mat Said⁴, Fatimah Ismail⁵ and Mohd Najib Abdul Ghani⁵

¹Department of Emergency Medicine, School of Medical Sciences, Universiti Sains Malaysia, Health Campus, 16150 Kubang Kerian, Kelantan, Malaysia.

²Hospital Universiti Sains Malaysia, Health Campus, 16150 Kubang Kerian, Kelantan, Malaysia.

³Biostatistics and Research Methodology Unit, School of Medical Sciences, Universiti Sains Malaysia, 16150 Kubang Kerian, Kelantan, Malaysia.

⁴Hospital Sultan Ismail Petra, 18000 Kuala Krai, Kelantan, Malaysia.

⁵Hospital Raja Perempuan Zainab II, 15586 Kota Bharu, Kelantan, Malaysia.

Corresponding author: Tuan Hairulnizam Tuan Kamauzaman

Email: hairulnizam@usm.my

ABSTRACT

Flood disaster is the commonest natural disaster with huge impact on healthcare services in Malaysia. The FloodDMQ-BM[®] questionnaire was developed as a tool to assess the knowledge, attitude, and practice of healthcare providers regarding patient management during a flood disaster. We aim to further validate the FloodDMQ-BM[®] questionnaire by using Confirmatory Factor Analysis (CFA) and Item Response Theory (IRT). This cross-sectional study involved doctors, nurses and paramedics working in the Emergency Department of Hospital Universiti Sains Malaysia, Hospital Raja Perempuan Zainab II and Hospital Kuala Krai. Respondents were required to complete the FloodDMQ-BM[®] questionnaire. The responses were analysed by using CFA and IRT to establish its validity and reliability. A total of 215 respondents participated in this study. CFA analysis with Maximum Likelihood Robust as the estimation method, on the attitude and practice components resulted in good factor loadings (>0.5) in nearly all items and excellent model fit indices values (CFI = 0.96-0.98, TLI = 0.95-0.96, SRMR = 0.04-0.05, RMSEA = 0.07). Meanwhile, IRT analysis on the knowledge section showed a good two-way marginal fit based on S-X², and a good model fit with RMSEA of 0.08. Based on the 2PL model by using the IRT assessment of the knowledge section, one item in the knowledge section (K3) was removed (chi-squared residual >4) resulting in improved model fit. The included items had well-standardized loadings (>0.3) and marginal reliability of 0.651. Our results confirmed that the FloodDMQ-BM[®] questionnaire displayed valid and reliable psychometric properties.

Keywords: Flood, Disaster, Questionnaire, Validation studies.

INTRODUCTION

Flood disasters are the most frequently occurring natural disaster worldwide over the last 10 years, killing 59,092 people and causing economic losses amounting to 342,836 million USD.¹ Malaysia is one of the countries in Southeast Asia with a tropical climate, which experiences flood almost every year. In the year 2014, Malaysia experienced the worst flood disaster recorded in the eastern states of Peninsular Malaysia. It was considered being a “tsunami-like disaster” in which 202,000 victims were displaced.² As flood disaster causes damage to the health care services, it is important to have a tool to assess the knowledge, attitude, and practice of health care providers in managing patients throughout a flood disaster, specifically during the response phase.

The FloodDMQ-BM[®] is a questionnaire to assess the knowledge, attitude, and practice of health care providers regarding patient management during a flood disaster. It is developed based on four domains, namely alert systems, communication,

standard operating procedure (SOP), and transportation. It has been used and tested for its validity and reliability by using Exploratory Factor Analysis (EFA) and Item Response Theory Analysis (IRT) by Abdul Ghani *et al.* in his study published in 2016. The results have suggested that the FloodDMQ-BM[®] possesses valid and reliable psychometric properties.³

Confirmatory Factor Analysis (CFA) is a multivariate statistical procedure utilized to test how well the measured variables represent the number of constructs.⁴ CFA is optimal for investigating the construct validity of a scale when a strong theoretical hypothesis is present about the structure of the scale. It is often regarded as a stronger source of evidence compared to more exploratory approaches in scale validation.⁵ By performing CFA, the researcher will be able to test the hypothesis that a relationship between the observed variables and their underlying latent constructs exists. Thus, a more conclusive and valid result can be obtained by using CFA.⁶

Item Response Theory allows comparisons to be made between the latent traits of individuals from different populations when subjected to tests with common features. This is only possible because it is focused on the items as a whole, as opposed to the test or questionnaire. Therefore, the level of an individual's response can be compared to specific features of an item, which facilitates the interpretation of the scale created. Furthermore, the researcher can recognize which items on the scale are generating information. It is a model-based method of estimating the parameters for each item included in the scale that separates an individual's responses to the items from their underlying level.⁷

The purpose of this study is to further validate the FloodDMQ-BM[®] by using CFA and IRT with larger sample size. Abdul Ghani *et al.* have recommended recruiting more samples so that the IRT components, standard loading, and marginal reliability are improved.

MATERIAL & METHODS

This cross-sectional study was conducted from 1st June 2018 until 11th February 2019 at the Emergency Department (ED) of Hospital Universiti Sains Malaysia (Hospital USM), Hospital Raja Perempuan Zainab II (HRPZ2), and Hospital Kuala Krai (HKK).

The reference population in this study was the health care workers in the ED of Hospital USM, HRPZ2, and HKK. This included all doctors, nurses, and paramedics, whereby a total of 209 participants were included in the final analysis.

Inclusion Criteria

All health care workers who were currently working in the ED of Hospital USM, HRPZ2, and HKK consented to participate in this study.

Exclusion Criteria

Incomplete questionnaire forms and subjects who had already participated in the previous study.

Technical Information

Structural equation modelling (SEM) research optimally needs a sample size of 200-400. Other research has suggested that the simplified designs can still be optimal with less than 200 samples in the presence of strong factor loadings.⁸ Based on the previous study, the sample size was taken as 200.

The respondents were recruited by non-probability convenient sampling. They were asked to answer the questionnaire with the researcher available at all times to clarify any phrases or terms. Each of the FloodDMQ-BM[®] questionnaires was given a code number, whereby all

respondents were identified via the unique code number given and ensured that their identity remained anonymous. Data obtained from this study were analyzed using the Statistical Package for Social Sciences (SPSS) software version 24.0 and R Studio version 3.6.0.

Tool Description

The FloodDMQ-BM[®] was used as a validated questionnaire to assess the knowledge, attitude, and practice of health care providers regarding patient management during a flood disaster. There were 13 questions in the attitude section, 12 questions in the practice section and nine questions were included in the knowledge section of the FloodDMQ-BM[®].

Statistics

The attitude and practice sections of the FloodDMQ-BM were analyzed using CFA, while the knowledge section was analyzed using a two-parameter logistic model of the IRT analysis.

RESULTS

Out of 250 questionnaires distributed for the study, 215 subjects responded and yielded 86% response rate. However, only 209 subjects were included in the final analysis as six of the questionnaires were incomplete.

Multivariate Normality

Multivariate normality was conducted to check the normality of the data. The results showed that the multivariate normality for both attitude and practice sections of FloodDMQ-BM[®] were not normal (kurtosis > 5, $P < 0.05$). Hence, Maximum Likelihood Robust (MLR) was used as the estimation method for CFA. MLR is generally suitable for complete and incomplete non-normal data.⁹

Fit Indices

Table 1 shows the result of factor analysis and reliability analysis for the attitude and practice sections. All items for both sections generated a factor loading of over 0.5 each, except for item S9 in the attitude section that scored 0.4. A factor loading > 0.5 indicates that the factors are significant. Furthermore, the attitude section of FloodDMQ-BM had a good model fit index with RMSEA of 0.07, SRMR of 0.05, CFI of 0.96, and TLI of 0.95. Meanwhile, the practice section yielded a very good model fit indices too, with RMSEA of 0.07, SRMR of 0.04, CFI of 0.98, and TLI of 0.96. According to Brown¹⁰, fit indices and recommended cut-off values are as follows: RMSEA $\square$ 0.08, SRMR $\square$ 0.08, CFI > 0.05, and TLI $\square$ 0.95.

Table 1: Factor analysis and reliability analysis of the attitude and practice sections for the FloodDMQ-BM®

Section	Items	Factor Loading	RMSEA	SRMR	CFI	TLI
Attitude	S1	0.69	0.07	0.05	0.96	0.95
	S2	0.67				
	S3	0.70				
	S4	0.76				
	S5	0.77				
	S6	0.77				
	S7	0.77				
	S8	0.67				
	S9	0.40				
	S10	0.63				
	S11	0.73				
	S12	0.76				
	S13	0.78				
Practice	A1	0.58	0.07	0.04	0.98	0.96
	A2	0.73				
	A3	0.70				
	A4	0.80				
	A5	0.84				
	A6	0.85				
	A7	0.77				
	A8	0.72				
	A9	0.81				
	A10	0.78				
	A11	0.89				
	A12	0.72				

Modification Indices

Localized areas of a misfit were conducted by using the modification indices (MI). A modification index indicates any expected parameter change if a particular specification in the model is included. Six suggested specifications were highlighted with $MI > 10$ in the attitude section of FloodDMQ-BM®, which were: S1~S2, S2~S3, S4~S5, S5~S7, S9~S10, and S10~S11 (see Figure 1). In contrast, the practice section had 12 suggested specifications with $MI > 10$, which were: A7~A9, A3~A4, A9~A1, A10~A1, A8~A9, A2~A3, A5~A11, A1~A3, A2~A4, A6~A11, A6~A10, and A4~A5 (see Figure 2). Applying the suggested MI improved the fit indices values, as well as the factor loadings and factor correlations on all items.

Item Response Theory Analysis

Based on the 2PL model by using the IRT assessment of the knowledge section, items K3, K8, and K9 had a chi-squared residual > 4 ; this shows a poor fit on the two-way margin.¹¹

Although the removal of items K3, K8, and K9 resulted in an improved model fit, many items were wasted. Besides, the removal of item K3 alone resulted in a balance between model fit. Hence, only K3 was removed in the knowledge section of the FloodDMQ-BM®. The IRT analysis of the remaining items is summarised in Table 2. Two-way marginal fit for the finalized items in the knowledge section had residues less than 4 and was considered a good fit at a 5% significant level.¹² The overall model data fit was good with an RMSEA value of 0.08 as suggested by Kenny and Kaniskan¹³ and was statistically non-significant of S-X2 as suggested by Orlando and Thissen¹⁴. Moreover, all items had well-standardized loadings ranging from 0.42 to 0.77 and marginal reliability of 0.651 as suggested by Dimitrov¹⁵. Finally, eight items were retained in the final model of the knowledge section. A higher score in the knowledge section would indicate a better understanding of the health care provider.

Table 2: Item Response Theory parameters estimation for the items in the knowledge section of FloodDMQ-BM©

Item after removal	Item parameters		S-X2 fit index			
	Difficulty (SE)	Discrimination (SE)	Standardized loading	X2	df	P
K1	0.35	1.09	0.54	1.67	4	0.80
K2	-0.60	2.07	0.77	3.57	3	0.31
K4	1.14	1.90	0.74	4.94	3	0.18
K5	0.91	1.22	0.58	11.08	4	0.03
K6	2.33	1.01	0.51	1.27	4	0.87
K7	0.08	1.31	0.61	6.34	3	0.10
K8	2.20	0.98	0.50	4.18	4	0.38
K9	0.83	0.79	0.42	12.35	4	0.02

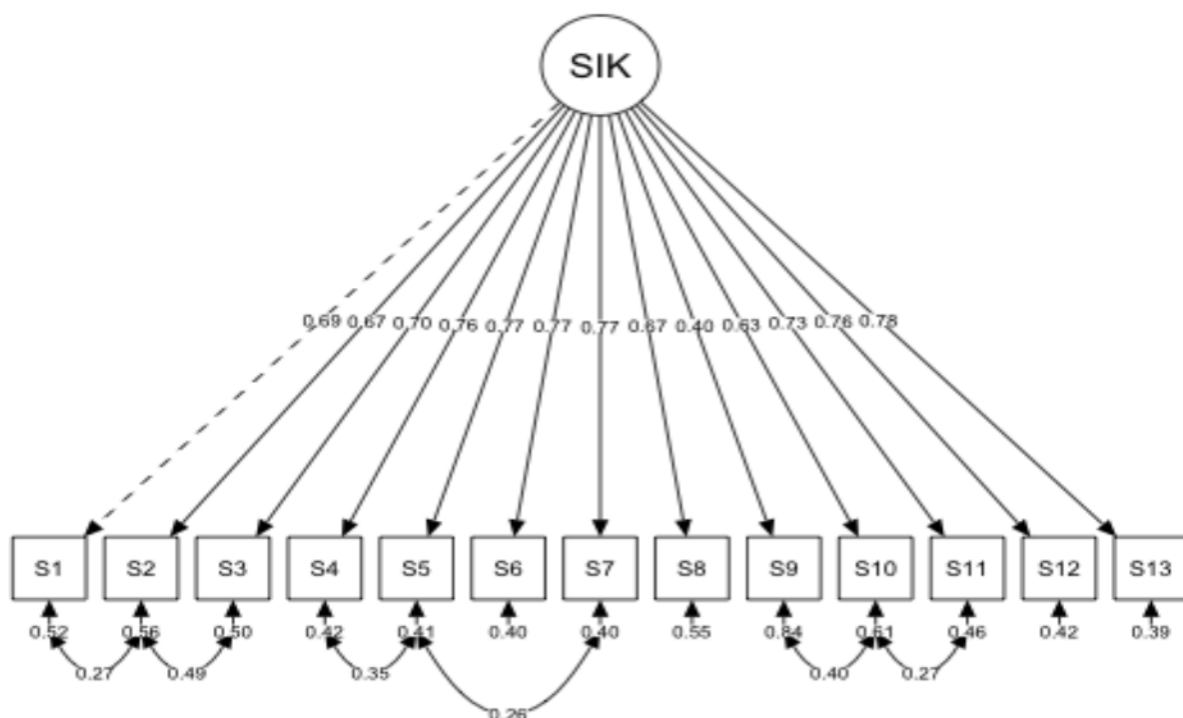


Figure 1: CFA path diagram for Attitude section of FloodDMQ-BM©

Table 3: Summary of the FloodDMQ-BM© items in all sections before and after psychometric analyses

Before		After	
Domain	Item	Domain	Item
Attitude	13 (A1-A13)	Attitude	13 (A1-A13)
Practice	12 (P1-P12)	Practice	12 (P1-P12)
Knowledge	9 (K1-K9)	Knowledge	8 (K1, K2, K4, K5, K6, K7, K8, K9)

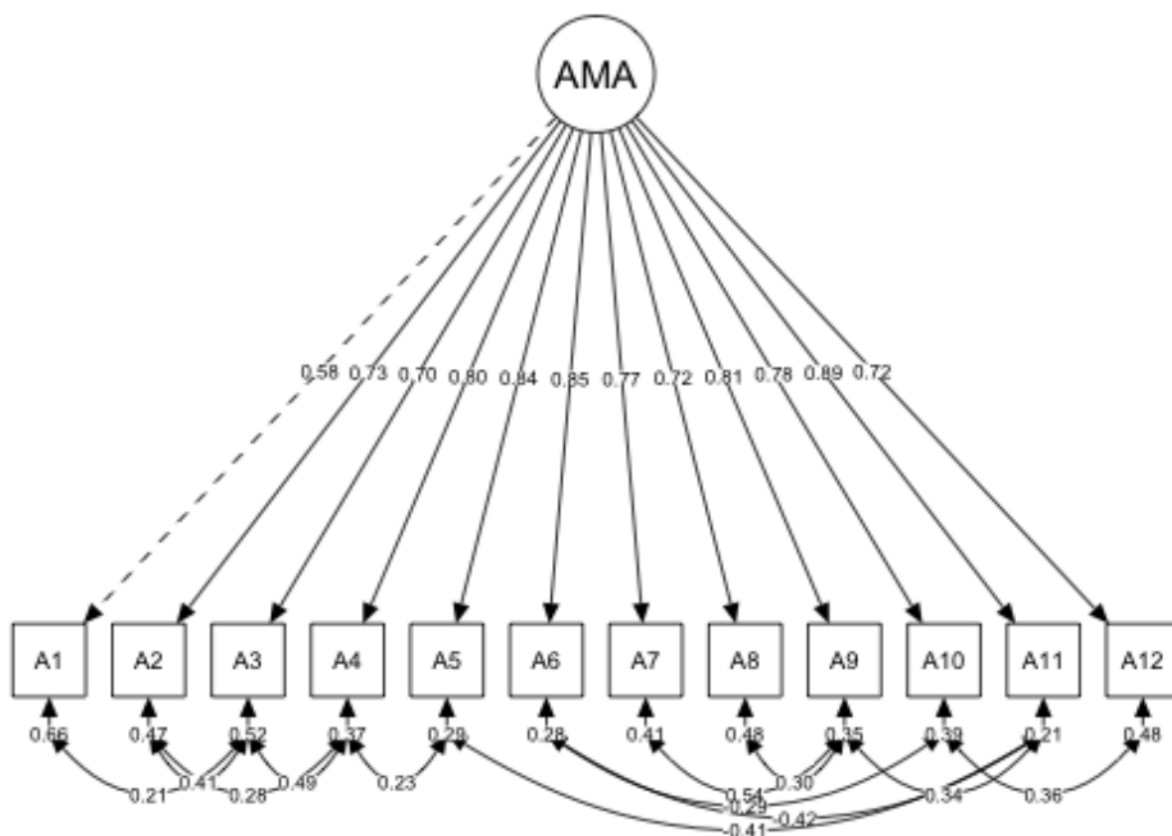


Figure 2: CFA diagram path for practice section of FloodDMQ-BM[®]

Table 3 summarises the items in all sections of the FloodDMQ-BM[®] before and after psychometric analysis. A total of 13 items were retained in the attitude section of the FloodDMQ-BM[®]. The attitude on patient management during flood disaster could be evaluated by the total score, with minimum possible score of 13 points and maximum possible score of 52. All items were positively structured, thereby rendering the reverse scoring to be not applicable to this section. A higher total score in the section indicated a more positive attitude of the health care providers towards patient management during a flood disaster.

Next, a total of 12 items in the practice section that fit the FloodDMQ-BM[®] were retained. The total practice score would assess the practices of good patient management during flood disaster, with a minimum possible scoring of 12 and maximum possible score of 48. All items were positively structured, thereby rendering reverse scoring as not applicable to this section. A higher total score in the section indicated a better practice shown by the health care providers regarding patient management during a flood disaster.

The final model of the knowledge section consisted of eight items in total, which amounted to a minimum score of 0 points and a possible maximum score of 8 points. A higher score in the knowledge section would indicate a better understanding of the health care providers.

DISSECTION

The FloodDMQ-BM[®] is a useful tool to assess the knowledge, attitude, and practice of health care providers towards patient management during a flood disaster. Its delivery in Bahasa Malaysia as the national language for the country allowed the participants to understand the questions well. Furthermore, its validity and reliability have been tested during the previous study by using EFA and IRT. CFA was chosen to analyzed on the attitude and practice section of FloodDMQ-BM[®] as it is optimal for investigating the construct validity of a scale when a strong theoretical hypothesis is present about the structure of the scale. Thus, a more conclusive and valid result can be obtained by using CFA. In this study, the CFA method specifically confirmed its validity, whereas the higher sample size also increased its reliability.¹⁶ Based on the CFA results, the attitude and practice sections both had good validity and reliability.

Meanwhile, IRT was used to analyzed on the knowledge section of this study as it allows comparisons to be made between the latent traits of individuals from different populations when subjected to tests with common features. Therefore, the level of an individual's response can be compared which facilitates the interpretation of the scale created. Furthermore, the researcher can recognise which items on the scale are generating information. After item analysis by using IRT for the knowledge section of

FloodDMQ-BM[®], one item was identified to have a poor fit on a two-way margin. However, other items in the knowledge section had good discrimination and difficulty values, indicating that it was able to differentiate the proportion of participants who scored high and low scores accordingly in the questionnaire. Accordingly, the marginal reliability can estimate the average reliability regarding the respondent's knowledge.³ The exact value of acceptable marginal reliability was not well-documented, but the statistical analysis in this study was underpinned by the studies by Dimitrov¹⁵ and Stoop¹⁷; they have suggested that the value of 0.651 is acceptable. Another point to note is that marginal reliability is influenced by the reduction of some items.¹⁷ In this study, the removal of item K3 significantly improved the marginal reliability score of the knowledge items.

Besides, the items' model fit revealed good RMSEA and non-significant M2 outcomes.¹³ It was noted that CFI and TLI were 0.83 and 0.82, respectively; the acceptable CFI and TLI values should be more than 0.9.¹⁸ However, given a good RMSEA result (less than 0.1), the CFI and TLI values can be ignored.¹³ In this study, the increased sample size resulted in better CFI and TLI compared to the previous study.

By confirming the validity of FloodDMQ-BM[®] using CFA, the researcher may identify the level of understanding and preparedness possessed by health care providers towards patient management during a flood disaster. It can serve as an indicator of the effectiveness in mitigating the risks of the flood disaster, thus ultimately improving health care services and resulting in improved patient care. The limitation of this study was due to the sample population, whereby only health care providers were included, who were from three main hospitals in the Kelantan state, namely Hospital USM, HRPZ2, and HKK. Therefore, only the knowledge, attitude, and practice of health care providers in these three main hospitals in Kelantan were assessed. Meanwhile, many other health care facilities such as government or private general clinics elsewhere also handle various cases during a flood disaster. Thus, the results did not reflect the awareness of health care providers in the Kelantan state as a whole. Another limitation of this study is on the sampling method whereby the author used non-probability convenient sampling method to achieve the optimal number of sample size required for SEM research. This method was chosen as it was not feasible to do probability sampling method in view of the limited number of health care providers in the ED Hospital USM, HRPZ2, and HKK. Hence, the result obtained may not represent the population well.

CONCLUSION

Psychometric evaluation of FloodDMQ-BM[®] questionnaire by using CFA and IRT confirmed its valid and reliable content to evaluate the knowledge, attitude, and practice of health care providers towards patient management during flood disaster. Therefore, it is suggested for this questionnaire be administered to other Malaysian states or translated into the English language to be used in any flood-prone areas worldwide. It is recommended for the future study to include tertiary and district hospitals, and district and rural clinics as well as to do probability sampling technique to produce a result that are more representative of the whole population.

ACKNOWLEDGEMENT

This study was approved by the Human Research Ethics Committee USM (USM/JEPeM/17100473) and National Medical Research Register (NMRR-17-2811-38346). Our special thanks to the School of Medical Sciences of Universiti Sains Malaysia and the National Medical Research Register of the Ministry of Health Malaysia for supporting this study. Our sincere gratitude and appreciation to the staff of Hospital USM, HRPZ2, and HKK who were involved in this study.

REFERENCES

1. International Federation of Red Cross and Red Crescent Societies. World Disasters Report: Focus on local actors, the key to humanitarian effectiveness [Internet]. Geneva: International Federation of Red Cross and Red Crescent Societies; 2015 [cited 2019 June 30]. Available from <https://www.ifrc.org/en/news-and-media/press-releases/general/wdr-press-release/>
2. Baharuddin KA, Abd Wahab SF, Nik Ab Rahman NH, et al. The Record-Setting Flood of 2014 in Kelantan: Challenges and Recommendations from an Emergency Medicine Perspective and Why the Medical Campus Stood Dry. *Malays J Med Sci* 2015; 22(2):1-7.
3. Abdul Ghani MN, Tuan Kamauzaman TH, Mohd Fudzi MFM, et al. Development And Psychometric Evaluation Of Flood Disaster Management Questionnaire-(FloodDMQ-BM): Exploratory Factor Analysis And Item Response Theory Analysis. *International Journal of Public Health and Clinical Sciences* 2016; 3(3).
4. Statistics Solutions. Confirmatory Factor Analysis [Internet]. Florida: Statistics Solutions; 2013 [cited 2019 June 30]. Available from:

- http://www.statisticssolutions.com/academic-solutions/resources/directory-of-statistical-analyses/confirmatory-factor-analysis/
5. Makransky G, Lilleholt L, Aaby A. Development and Validation of the Multimodal Presence Scale for Virtual Reality Environments: A Confirmatory Factor Analysis and Item Response Theory Approach. *Comput Human Behav* 2017; 72.
6. Diana DS. Exploratory or Confirmatory Factor Analysis? [Internet]. Northern Colorado: SAS Institute; 2006 [cited 2019 June 30]. Available from <https://support.sas.com/resources/papers/proceedings/proceedings/sugi31/200-31.pdf>
7. Yang FM, Kao ST. Item response theory for measurement validity. *Shanghai Arch Psychiatry* 2014; 26(3):171-177.
8. Rex BK. Principles and Practice of Structural Equation Modeling. New York: The Guilford Press 2011:427.
9. Rosseel Y. Package 'lavaan': Latent Variable Analysis [Internet]. Lavaan; 2019 [cited 2019 June 30] Available from <https://cran.r-project.org/web/packages/lavaan/lavaan.pdf>
10. Timothy AB. Confirmatory Factor Analysis for Applied Research. New York: The Guilford Press 2015:462.
11. Arifin WN, Yusoff MSB. Item response theory for medical educationists. *Education in Medicine Journal* 2017; 9(3):69-81.
12. Bartholomew DJ, Steele F, Moustaki I, et al. Analysis of Multivariate Social Science Data. Florida: CRC Press 2011:384.
13. Kenny DA, Kaniskan B, McCoach DB. The Performance of RMSEA in Models with Small Degrees of Freedom. *Sociology Method & Research* 2015; 44(3):486-507.
14. Orlando M, Thissen D. Further Investigation of the Performance of S-X2: An Item Fit Index for Use with Dichotomous Item Response Theory Models. *Appl Psychol Meas* 2003; 27(4):289-298.
15. Dimitrov D. Marginal True-Score Measures and Reliability for Binary Items as a Function of Their IRT Parameters. *Appl Psychol Meas* 2003; 27(6):440-458.
16. An GY, Sean P. A Beginner's Guide to Factor Analysis: Focusing on Exploratory Factor Analysis. *Tutor Quant Methods Psychol* 2013; 9(2):79-94.
17. Stoop VD. Using Item Response Theory to Analyze Properties of the Leadership Practices Inventory. *British Journal of Science* 2011; 1(1):140-162.
18. Hooper D, Coughlan J, Mullen M. Structural Equation Modelling: Guidelines for Determining Model Fit. *Electronic Journal of Business Research Methods* 2008; 6(1):53-60.