Journal of Contemporary Issues and Thought
ISSN 2232-0032/ e-ISSN 0128-0481/ Vol 11, Issue 1, 2021 (105-112)

Optimal Reliability and Validity of Measurement Model in Confirmatory Factor Analysis: Different Likert Point Scale Experiment

Mohd Azry Abdul Malik[1*], Muhammad Firdaus Mustapha[2], Norafefah Mohamad Sobri[3], Nor Fatihah Abd Razak[4], Mohamad Nurifaizal Mohd Zaidi[5], Ahmad Aizat Shukri[6], & Muhammad Amir Luqman Zalimie Sham[7]

[1,2,3,4,6,7]*Faculty of Computer and Mathematical Science, Universiti Teknologi MARA, Malaysia*
[5]*Koperasi Muafakat Makmur Berhad, Bangi, Selagor, Malaysia*
**Email: azry056@uitm.edu.my*

Abstract

Designing a questionnaire is one of the most difficult challenges in research design, particularly when deciding which level of Likert point scale is appropriate for the instrumentation. Suitable Likert point scale used in the instrumentation able to reduce the risk of facing potential problems of not achieving reliability (indicator and internal consistency reliability) and validity (convergent validity, discriminating validity, and construct validity) and simultaneously preventing the occurrence of multicollinearity. This study compares the performance of reliability and validity of measurement construct in Confirmatory Factor Analysis (CFA) by using Likert 5-point Likert scale, 6-point Likert scale, 7-point Likert scale, 9-point Likert scale, and 10-point Likert scale. The study uses primary data based on a questionnaire data collection method which involves 100 samples from similar population characteristics for each Likert. The data were analysed using Smart-PLS software. The results suggest that expanding the range of the Likert point scale optimizes the performance of reliability and validity of the measurement model. This study offers an insight to researchers in deciding the best choice of Likert point scale to adapt in instrumentation for a better result in the quantitative analysis process.

Keywords: Convergent; Discriminant; Multicollinearity; SEM; PLS

1. Introduction

Likert scale is the rating scale that measures the impressions of the selected respondents towards the statement in the questionnaire. Likert scale is a widely employed psychometric methods in education and social sciences studies (Joshi et al., 2015). It is a tool that can easily capture the reaction of respondents towards specifics statements either they agree, disagree or neutral (Wuensch & Kar, 2005).

When designing instrumentation, choosing the appropriate level of the Likert scale is very important for the successful analysis process. Recently, there are a variety of suggestions related to the suitable level of the Likert scale to be used in social studies such as 5-point Likert scale (Adelson et al., 2010; Chachamovich et al., 2009), 6-point Likert scale (Chomeya, 2010), 7-point Likert scale (Dawes, 2008; Norman, 2010), 9-point Likert scale (Dawes, 2008; Tarka, 2016), and 10-point Likert scale (Awang et al., 2016; Cummins et al., 2000; Preston et al., 2000). However, few studies have reported on the effect of using different Likert point scale on the analysis process.

The purpose of this study was to ascertain the effect of using different Likert point scale on reliability and validity for measurement model in CFA. Achieving good reliability and validity in CFA determines either the data can be processed for further analysis or not. This topic was identified as being important to researchers in providing them the necessary background to construct good research instrumentation.

2. Materials and Methodology

Study Design

The primary data was obtained through a direct questionnaire distributed to the 100 students of Bachelor Degree in Statistics at Universiti Teknologi MARA (UiTM) Kota Bharu, Kelantan, Malaysia. The questionnaires consist of 5 exogenous and 1 endogenous variable as in Figure 1. There are 5 sets of questionnaires constructed using a 5-point Likert scale, 6-point Likert scale, 7-point Likert scale, 9-point Likert scale, and 10-point Likert scale. Each respondent needs to answer five sets of questionnaire with different Likert point scale based on their point of view. The questionnaire contains 6 sections as in Table 1a.

Table 1a: Summary of the questionnaire

Sections	Variables	No of items	Adapted from
A	Academic Self Efficiency	7	Pintrich et al, 1991
B	Attitude Toward Students	7	
C	Autonomy-Power Sharing	7	Treagust, D. F., & Fraser, B. J. (1986).
D	Peer Relationships	7	
E	Student Interest Motivation	7	
F	Class Organization	7	

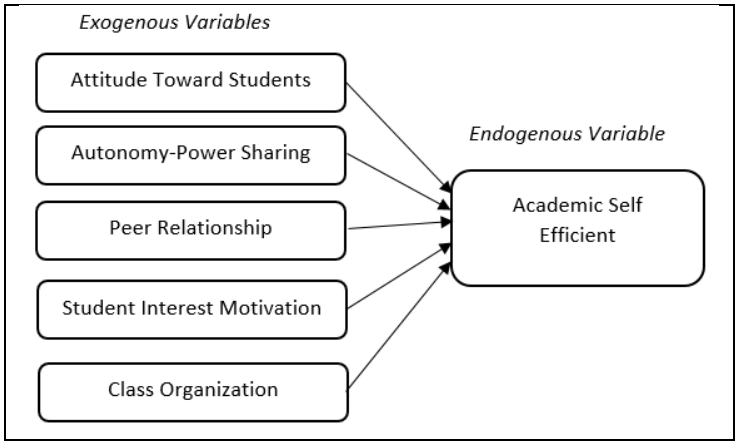

Figure 1: Sample model use in the study

Data Analysis

PLS-SEM is the technique that performs confirmatory factor analysis and regression simultaneously. The study uses Smart-PLS software to run the confirmatory factor analysis on the dataset. The CFA will confirm to either the measurement model achieves good reliability and validity or not.

Reliability is the extent of how consistent the instrument in measuring the latent construct. The criteria that the researcher should consider in measuring the reliability were indicator reliability (factor loading), and internal consistency reliability (Cronbach alpha). Validity is a test of how well an instrument that was developed measures the particular concept it was intended to measure (Sekaran & Bougie, 2010). There are three types of validity; construct validity, convergent validity and discriminant validity. For convergence validity, Average Variance Extracted Estimate (AVE) is used to assess the amount of measurement error variability recorded by a group of objects on a scale and a stringent rate of 0.50 or higher was recommended (Valerie, 2012). For discriminant validity, the Heterotrait-monotrait (HTMT) ratio of correlation is used to identify the discriminating validity when two variables are genuinely empirically distinct from each other and lower HTMT is better (Hamid et al., 2017; Hair et al. 2017). Construct validity uses Absolute Fitness Indexes (model fit) to assess how well a priori model fits the sample data and it should meet the required level of Rooted Mean Squared Error Approximation (RMSEA) less than 0.1 (McDonald & Ho, 2002) as indicate reasonable errors of approximation.

3. Results

Reliability Analysis

Reliability Analysis must achieve the internal consistency and reliability of the indicator. Internal consistency tests the continuity of the homogeneous objects and indicator reliability focuses on the factor loading for the item.

Indicator Reliability: Factor Loading

Indicator reliability identifies how many factors loading can be explained by the latent variable. The appropriate values of factor loading are 0.70 and above (Valerie, 2012). Table 1b shows the value of factor loading for Academic Self Efficient, Attitude Toward Students, Autonomy-Power Sharing, Peer Relationship, Student Interest and Class Organization for 5-point Likert scale, 6-point Likert scale, 7-point Likert scale, 9-point Likert scale and 10-point Likert scale. Overall, the 9-point Likert scale and 10-point Likert scale indicate good indicator reliability across all the variables involves.

Table 1b: Factor loading values

	Construct	Likert 5	Likert 6	Likert 7	Likert 9	Likert 10
Academic Self Efficient	AQ1	0.747	0.702	0.716	0.818	0.752
	AQ2	0.766	0.678	0.727	0.825	0.792
	AQ3	0.427	0.769	0.503	0.787	0.834
	AQ4	0.581	0.636	0.782	0.872	0.831
	AQ5	0.758	0.824	0.752	0.889	0.881
	AQ6	0.802	0.854	0.812	0.831	0.845
	AQ7	0.770	0.808	0.600	0.818	0.741
Attitude Toward Students	BQ1	0.765	0.716	0.782	0.880	0.856
	BQ2	0.845	0.788	0.895	0.851	0.893
	BQ3	0.831	0.812	0.820	0.802	0.880
	BQ4	0.822	0.825	0.722	0.856	0.894

	BQ5	0.270	0.427	0.363	0.559	0.446
	BQ6	-0.398	-0.298	0.141	-0.137	0.035
	BQ7	0.76	0.837	0.222	-0.072	0.827
Autonomy-Power Sharing	CQ1	0.719	0.470	0.659	0.406	0.548
	CQ2	0.670	0.479	0.721	0.353	0.607
	CQ3	0.498	0.697	0.528	0.734	0.626
	CQ4	0.692	0.668	0.520	0.668	0.838
	CQ5	-0.293	0.674	0.605	0.587	0.678
	CQ6	-0.187	0.740	0.742	0.811	0.806
	CQ7	-0.040	0.442	0.732	0.583	0.594
Peer Relationship	DQ1	0.564	0.701	0.776	0.406	0.716
	DQ2	0.730	0.669	0.572	0.353	0.780
	DQ3	0.822	0.815	0.681	0.734	0.807
	DQ4	0.401	0.648	0.683	0.668	0.742
	DQ5	-0.001	0.096	0.280	0.587	0.125
	DQ6	0.763	0.805	0.604	0.811	0.669
	DQ7	-0.333	-0.078	0.019	0.583	0.094
Student Interest	EQ1	0.698	0.696	0.499	0.735	0.802
	EQ2	0.546	0.730	0.509	0.818	0.808
	EQ3	0.683	0.776	0.734	0.709	0.755
	EQ4	-0.107	0.292	0.406	0.596	0.409
	EQ5	0.653	0.683	0.702	0.824	0.830
	EQ6	0.686	0.761	0.469	0.825	0.744
	EQ7	0.815	0.873	0.560	0.862	0.818
Class Organization	FQ1	0.174	0.801	0.620	0.807	0.856
	FQ2	0.633	0.751	0.709	0.798	0.864
	FQ3	0.610	0.791	0.431	0.749	0.718
	FQ4	0.860	0.679	0.668	0.860	0.830
	FQ5	0.823	0.778	0.726	0.767	0.831
	FQ6	0.826	0.745	0.65	0.748	0.796
	FQ7	0.724	0.663	0.666	0.771	0.825

Internal Consistency Reliability: Cronbach Alpha

The study measures internal consistency reliability performance using Cronbach's Alpha. A bigger value of Cronbach's alpha is better for internal consistency reliability. Table 2 and Figure 2 summarise the Cronbach values obtained. The plot in Figure 2 shows that majority of the variables using 9-point Likert scale 10-point Likert scale creates acceptable Cronbach's Alpha values (above 0.7).

Table 2: Cronbach's alpha values

	Academic Self Efficient	**Attitude Toward Students**	**Autonomy-Power Sharing**	**Peer Relationship**	**Student Interest**	**Class Organization**
Likert 5	0.824	0.761	0.459	0.544	0.719	0.812

Likert 6	0.873	0.700	0.732	0.605	0.820	0.869
Likert 7	0.831	0.697	0.775	0.554	0.635	0.762
Likert 9	0.927	0.684	0.735	0.736	0.884	0.897
Likert 10	0.913	0.830	0.817	0.686	0.917	0.864

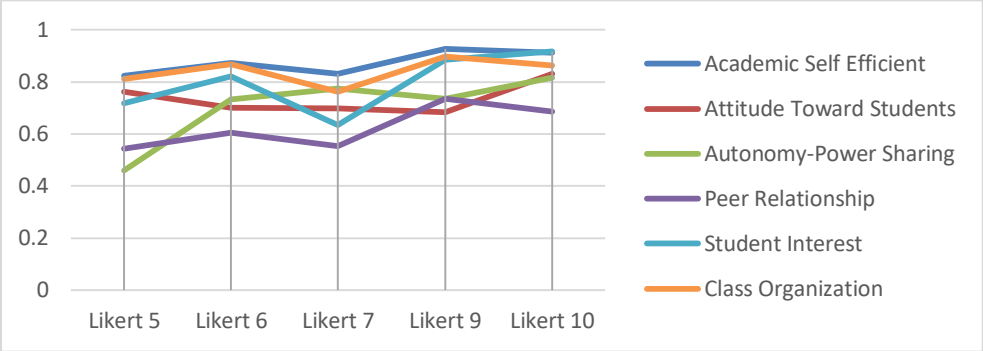

Figure 2: Cronbach's alpha values

Validity Analysis

Validity is a test of how well an instrument that was developed measures the particular concept it was intended to measure (Sekaran & Bougie, 2013).

Convergent Validity: Average Variance Extracted (AVE)

The Average Variance Extracted (AVE) determines the sum of measurement error uncertainty produced by a collection of items in a range. A strict AVE value of 0.50 or higher is suggested to achieve a good convergent validity for the measurement model (Hair et al., 2017). Table 3 and Figure 3 summarise the AVE values for every variable involves in the measurement model when using 5 different points of Likert scale. The results agreed that the 9-point and 10-point Likert scale produce a better AVE value for all the variables involved in the measurement model.

Table 3: Average variance extracted (AVE) values

	Academic Self Efficient	Attitude Toward Students	Autonomy-Power Sharing	Peer Relationship	Student Interest	Class Organization
Likert 5	0.496	0.496	0.259	0.340	0.403	0.489
Likert 6	0.573	0.492	0.366	0.384	0.502	0.559
Likert 7	0.499	0.401	0.422	0.329	0.320	0.416
Likert 9	0.694	0.459	0.374	0.464	0.595	0.618
Likert 10	0.659	0.570	0.461	0.399	0.670	0.564

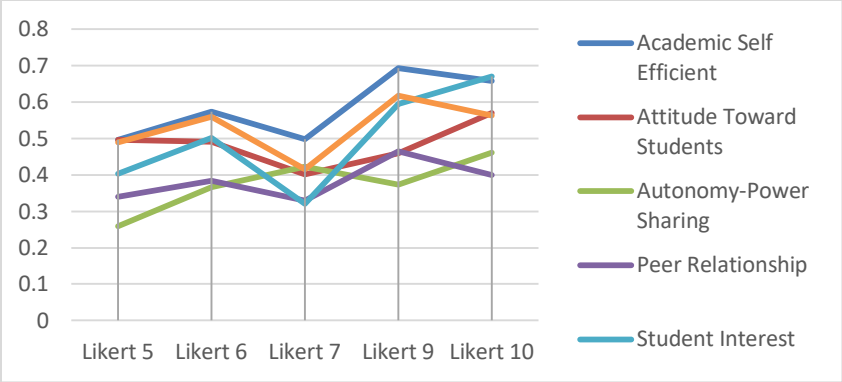

Figure 3: Average variance extracted (AVE) values

Discriminant Validity: Heterotrait-Monotrait Ratio (HTMT)

To achieve the discriminating validity, the diagonal value of HTMT must be lower than 0.9 (Hair et al., 2017). A lower HTMT value gives a better performance in discriminant validity. Figure 4 revealed that only 5-point and 9-point Likert scales contain the acceptable HTMT values across all the variables.

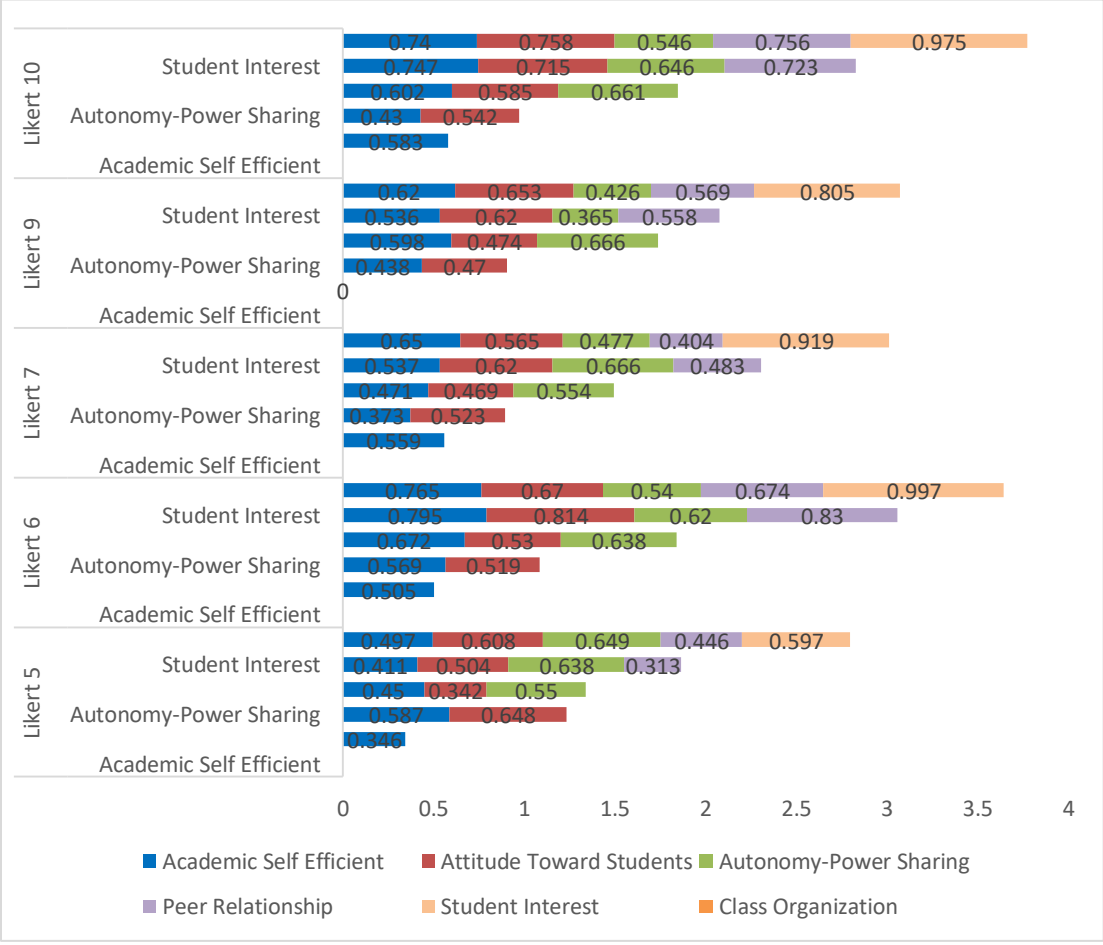

Figure 4: Heterotrait-Monotrait (HTMT) ratio

Construct Validity

Absolute fit indices determine how well a priori model fits the sample data (McDonald & Ho, 2002) and demonstrates which proposed model has the most superior fit. Rooted Mean Squared Error Approximation (RMSEA) value less than 0.1 is acceptable. Figure 5 shows the value of RMSEA for 5-point Likert scale, 6-point Likert scale, 7-point Likert scale, 9-point Likert scale and 10-point Likert scale which is 0.114, 0.119, 0.118, 0.098 and 0.117 respectively. The results indicate that 9-point Likert scale produces a better construct validity since it has the lowest value among all other observation.

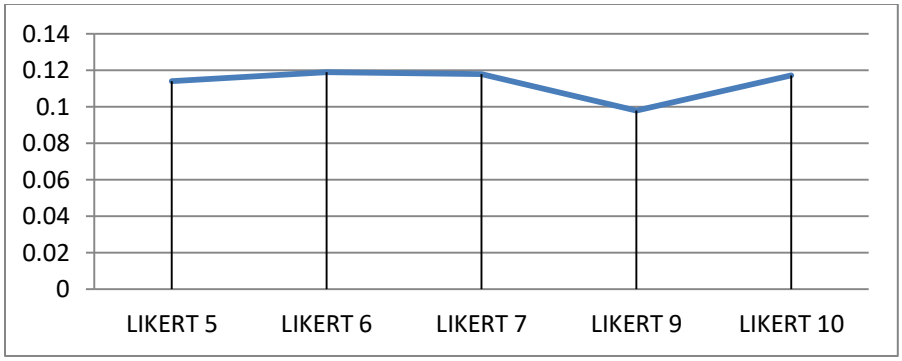

Figure 5: Value of rooted mean squared error Approximation (RMSEA)

Overall, by considering the performance of the indicator reliability, internal consistency reliability, convergent validity, discriminant validity and construct validity, the 9-point Likert scale is the best Likert scale to produce optimal reliability and validity for measurement model in confirmatory factor analysis.

4. Conclusion and Recommendations

The findings from the study show that the 9-point Likert scale is the most suitable Likert point scale to produce optimal reliability and validity for the measurement model in CFA. By expanding the range of the Likert point scale in the instrumentation reduce the risk of facing potential problems of convergent validity, discriminating validity, construct validity and preventing the existence of multicollinearity. Increasing Likert's range creates a bigger variance in the respondents' response. As a result, the level at which the construction indicators reveal the latent variable (composite reliability) and the inter-item consistency of the measuring items (Cronbach's alpha) produce a better result. Good quality of composite reliability and the alpha values of Cronbach increase the outcome as many things that calculate the same concept agree (convergent validity), simultaneously resulted in improving discriminant validity and reduce the probability of multicollinearity occurrence. However, the result obtained only can be generalized in educational setting. Future work is suggested to pay more attention to the other factors to spur effective instrumentation in research design.

References

Adelson, J. L., & McCoach, D. B. (2010). Measuring The Mathematical Attitudes of Elementary Students: The Effects of A 4-Point or 5-Point Likert-Type Scale. *Educational and Psychological measurement*, *70*(5), 796-807. https://doi.org/10.1177/0013164410366694

Awang, Z., Afthanorhan, A., & Mamat, M. (2016). The Likert scale analysis using parametric based Structural Equation Modeling (SEM). *Computational Methods in Social Sciences*, *4*(1), 13. https://doi.org/10.5539/mas.v9n9p58

Chachamovich, E., Fleck, M. P., & Power, M. (2009). Literacy Affected Ability to Adequately Discriminate Among Categories in Multipoint Likert Scales. Journal of Clinical Epidemiology, 62(1), 37-46. https://doi.org/10.1016/j.jclinepi.2008.03.002

Chomeya, R. (2010). Quality of psychology test between Likert scale 5 and 6 points. *Journal of Social Sciences*, *6*(3), 399-403. https://doi.org/10.3844/jssp.2010.399.403

Cummins, R. A., & Gullone, E. (2000, March). Why We Should Not Use 5-point likert scales: The Case for Subjective Quality of Life Measurement. In *Proceedings, second international conference on quality of life in cities*, 74-93.

Dawes, J. (2008). Do Data Characteristics Change According to The Number of Scale Points Used. *International journal of market research*, *50*(1), 61-77. https://doi.org/10.1177/147078530805000106

Hair, J.F., Hult G.T.M, Ringle, C.M., Sarstedt, M. (2017). *A Primer on Partial Least Square Structural Equation Modelling (PLS-SEM)* (2 ed.). Thousand Oaks, CA: Sage. ISBN 9781483377445.

Hamid, M. R. A., Sami, W., & Sidek, M. H. M. (2017). Discriminant Validity Assessment: Use of Fornell & Larcker criterion versus HTMT Criterion. Journal of Physics: Conference Series, 890, 012163. doi: 10.1088/1742-6596/890/1/012163

Joshi, A., Kale, S., Chandel, S., & Pal, D. K. (2015). Likert scale: Explored and explained. *British Journal of Applied Science & Technology*, *7*(4), 396. https://doi.org/10.9734/bjast/2015/14975

McDonald, R.P. and Ho, M.-H.R. (2002). Principles and Practice in Reporting Statistical Equation Analyses. Psychological Methods, 7 (1), 64-82. https://doi.org/10.1037/1082-989x.7.1.64

Norman, G. (2010). Likert scales, levels of measurement and the "laws" of statistics. Advances in Health Sciences Education, 15(5), 625-632. https://doi.org/10.1007/s10459-010-9222-y

Preston, C. C., & Colman, A. M. (2000). Optimal Number of Response Categories in Rating Scales: Reliability, Validity, Discriminating Power, And Respondent Preferences. *Acta psychologica*, *104(1),* 1-15. https://doi.org/10.1016/s0001-6918(99)00050-5

Sekaran, U., & Bougie, R. (2010). *Research methods for business: A skill building approach* (5th ed). Chichester: John Wiley & Sons Ltd.

Valerie, F. (2012). Re-discovering the PLS approach in management science. *Management, 15(1),* 101-123. https://doi.org/10.3917/mana.151.0102

Wuensch, Karl L. (2005). "What is a Likert Scale? and How Do You Pronounce 'Likert?'". East Carolina University.