Discriminant Validity Assessment: Use of Fornell & Larcker criterion versus HTMT Criterion

To cite this article: M R Ab Hamid *et al* 2017 *J. Phys.: Conf. Ser.* **890** 012163

View the article online for updates and enhancements.

Related content

- Generalized shrunken type-GM estimator and its application
  C Z Ma and Y L Du

- Modeling Governance KB with CATPCA to Overcome Multicollinearity in the Logistic Regression
  L Khikmah, H Wijayanto and U D Syafitri

- SEVIRI Cloud mask by Cumulative Discriminant Analysis
  M G Blasi, C Serio, G Masiello et al.

Discriminant Validity Assessment: Use of Fornell & Larcker criterion versus HTMT Criterion

M R Ab Hamid[1,2], W Sami[1,3] and M H Mohmad Sidek[1]

[1]Faculty of Industrial Management, Universiti Malaysia Pahang, 26300 Gambang, Kuantan, Pahang, Malaysia
[2]UMP Advanced, Kompleks Pentadbiran Utama, Universiti Malaysia Pahang, 26300 Gambang, Kuantan, Pahang, Malaysia
[3]Department of Public Health & Community Medicine, College of Medicine, Majmaah University, Kingdom of Saudi Arabia

E-mail: rashid@ump.edu.my

Abstract. Assessment of discriminant validity is a must in any research that involves latent variables for the prevention of multicollinearity issues. Fornell and Larcker criterion is the most widely used method for this purpose. However, a new method has emerged for establishing the discriminant validity assessment through heterotrait-monotrait (HTMT) ratio of correlations method. Therefore, this article presents the results of discriminant validity assessment using these methods. Data from previous study was used that involved 429 respondents for empirical validation of value-based excellence model in higher education institutions (HEI) in Malaysia. From the analysis, the convergent, divergent and discriminant validity were established and admissible using Fornell and Larcker criterion. However, the discriminant validity is an issue when employing the HTMT criterion. This shows that the latent variables under study faced the issue of multicollinearity and should be looked into for further details. This also implied that the HTMT criterion is a stringent measure that could detect the possible indiscriminant among the latent variables. In conclusion, the instrument which consisted of six latent variables was still lacking in terms of discriminant validity and should be explored further.

1. Introduction
The objective of this paper is to apply a new criterion for assessing the discriminant validity as suggested by Henseler [1] in 2015. The assessment of discriminant validity is of utmost importance in research that involves latent variables along with the use of several items or indicators for representing the construct. Therefore, the researcher needs to establish the discriminant validity beforehand. This is to ensure that the latent constructs used for measuring the causal relationships under study are truly distinct from each other. In other words, they are different and not measuring the same thing that would arise the issue of multicollinearity. If the researcher still proceeds with the testing of the hypothesised model without addressing this issue, then the interpretation of the whole model can be misleading or useless. Therefore, the assessment of discriminant validity should be established first. Previously, the researcher used Fornell and Larcker Criterion [2] that was suggested in 1981 for assessing the discriminant validity. However, recently in 2015 Henseler [1] disapproved the Fornell and Larcker criterion. They

found out that the previous criterion is still lacking in establishing the distinctiveness between constructs that prompted them to suggest a more robust approach that could capture the discriminancy among the constructs under study.

2. Literature Review

Despite significant and good reliability coefficient, it does not guarantee an accurately measured construct [3]. Thus, the assessment of validity needs to be carried out. Hair et al. [3] also pointed out that the results of the construct validity of the test will give a better understanding of the quality measures used. After all, prior to testing the significance of the relationships in the model structure, measurement models need to have a level of satisfactory validity and reliability [2]. In this study, data was analysed using the SmartPLS sofware version 2.0.M3 developed by Ringle et al. [4] in 2005.

2.1. Evaluation of outer model

The first step in PLS-SEM analysis is to evaluate the outer model (or measurement model). The purpose is to determine how well the item (questions) load on the hypothetical-defined construct. Analysing the outer model comprises of unidirectional predictive relationships between each of the latent construct that is linked with the observed indicator [5]. Generally, there are two distinct measures of the indicators in PLS-SEM that are reflective and formative outer model [6]. The assessment of reflective outer model involves the examining of reliabilities of the individual items (indicator reliability), reliability of each latent variables, internal consistency (Cronbach alpha and composite reliability), construct validity (loading and cross-loading), convergent validity (average variance extracted, (AVE)) and discriminant validity (Fornell-Larcker criterion, cross loading, HTMT criterion) [7].

2.2. Internal Consistency

The most common measurement used for internal consistency is Cronbach alpha and composite reliability, in which it measures the reliability based on the interrelationship of the observed items variables. In PLS-SEM, the values are organised according to their indicator's individual reliability [7]. The values range from 0 to 1, where a higher value indicates higher reliability level. In exploratory research, values of composite reliability/Cronbach alpha between 0.60 to 0.70 are acceptable, while in more advanced stage the value have to be higher than 0.70 [7]. However, the value that is more than 0.90 is not desirable and the value that is 0.95 or above is definitely undesirable [8].

2.3. Indicator Reliability

Indicator reliability is the proportion of indicator variance that is explained by the latent variable. The values range from 0 to 1. The outer loadings value should be higher than 0.70 and it should be considered for deletion if the removal of the indicator with outer loadings which is between 0.40 and 0.70 if it contributes to an increase in composite reliability and average variance extracted (AVE) [7]. On the other hand, indicators with outer loading below 0.40 should always be removed [5],[9].

2.4. Convergent Validity

Convergent validity is the assessment to measure the level of correlation of multiple indicators of the same construct that are in agreement. To establish convergent validity, the factor loading of the indicator, composite reliability (CR) and the average variance extracted (AVE) have to be considered [7]. The value ranges from 0 to 1. AVE value should exceed 0.50 so that it is adequate for convergent validity [[10],[11],[7],[2].

2.5. Discriminant Validity Test

Discriminant validity is referring to the extent in which the construct is actually differing from one another empirically. It also measures the degree of differences between the overlapping construct [7]. The discriminant validity can be evaluated by using cross-loading of indicator, Fornell & Larcker criterion and Heterotrait-monotrait (HTMT) ratio of correlation. By looking at the cross-loading, the

factor loading indicators on the assigned construct have to be higher than all loading of other constructs with condition that the cut-off value of factor loading is higher than 0.70 [[5],[7].

The second criterion is to assess discriminant validity using Fornell-Lacker criterion [12]. This method compares the square root of the average variance extracted (AVE) with the correlation of latent constructs [7]. A latent construct should explain better the variance of its own indicator rather than the variance of other latent constructs. Therefore, the square root of each construct's AVE should have a greater value than the correlations with other latent constructs [7].

The other measure for discriminant validity is Heterotrait-monotrait (HTMT) ratio of correlation. Henseler et al. [1] proposed the superior performance of this method by means of Monte Carlo simulation study and found that HTMT is able to achieve higher specificity and sensitivity rates (97% to 99%) compared to the cross-loadings criterion (0.00%) and Fornell-Lacker (20.82%). HTMT values close to 1 indicates a lack of discriminant validity. Using the HTMT as a criterion involves comparing it to a predefined threshold. If the value of the HTMT is higher than this threshold, one can conclude that there is a lack of discriminant validity. Some authors suggest a threshold of 0.85 [13]. In addition, Gold et al. [14] argued with it and proposed a value of 0.90.

3. Methodology

Data from previous study was used in this article for establishing the discriminant validity test using Fornell and Larcker 1981 criterion [2] and Heterotrait-monotrait (HTMT) criterion as proposed by Henseler in 2015 [1]. Four hundred and twenty nine (429) data set were available for the data analysis and no missing data were found. The data was collected previously by Ab Hamid [15] in 2012 for assessing the value-based excellence model for higher education institution (HEIs) in Malaysia. There were six latent constructs involved i.e. leadership, culture, productivity, employee, stakeholder and university performance. Each construct was measured by six items. Readers can refer to Ab Hamid et al. 2014 for the items in the questionnaire/instrument used. The results of the analysis using the two criterions were compared and analysed. SmartPLS software version 2 was used in this study.

4. Results

4.1. Evaluation of outer model

Assessment of the reliability of each item was done by checking the cross-loadings and it was found that the values of factor loading was high on their respective constructs i.e. each factor loading was greater than the cut-off value of 0.70. This also shows that the reliability of each item was good and gives reinforcement to the allocation for each item on the specified latent construct. Indirectly, it supported for convergent validity. In other words, there is shared variance between the constructs and the items [[16],[17]. Each factor loading was significant at 5% level of significance.

4.2. Discriminant Validity: Fornell and Larcker criterion

Output from the analysis revealed the composite reliability (CR), the average variance extracted (AVE) and the correlation coefficients between the constructs that are summarized as in Table 1.

Table 1. Composite reliability (CR), the square root of the average variance extracted (AVE) *(in bold)* and correlations between constructs *(off-diagonal).*

Latent constructs	CR	AVE	Latent constructs					
			A	B	C	D	E	F
Leadership (A)	0.969	0.839	**0.916**					
Culture (B)	0.965	0.820	0.759	**0.906**				
Productivity (C)	0.942	0.729	0.763	0.822	**0.854**			
Employee (D)	0.954	0.778	0.743	0.845	0.863	**0.882**		
Stakeholder (E)	0.966	0.826	0.735	0.794	0.861	0.856	**0.909**	
University Performance (F)	0.963	0.814	0.673	0.709	0.825	0.769	0.848	**0.902**

Referring to Table 1, the CR for all constructs are above 0.70 and the AVE values are within 0.729 and 0.839. The discriminant validity was assessed using Fornel and Larcker (1971) by comparing the square root of each AVE in the diagonal with the correlation coefficients (off-diagonal) for each construct in the relevant rows and columns. For the productivity-employee construct and the productivity-stakeholder construct, there are little disputes. However, the difference is too small, each with 0.009 and 0.007 respectively, and can be ignored [18]. Overall, discriminant validity can be accepted for this measurement model and supports the discriminant validity between the constructs.

4.3. Discriminant Validity: Heterotrait-monotrait (HTMT) criterion
Table 2 below showed the output from HTMT analysis. The output can easily be calculated using the formula as in Henseler [1] in 2015.

Table 2. HTMT results.

	Leadership	Culture	Productivity	Employee	Stakeholder	UP
Leadership	-					
Culture	0.79	-				
Productivity	0.81	**0.87**	-			
Employee	0.78	**0.93**	**0.92**	-		
Stakeholder	0.77	**0.89**	**0.91**	**0.90**	-	
UP	0.70	0.79	**0.88**	0.81	**0.88**	-

From the HTMT results, the values (in bold) in Table 2 indicated discriminant validity problems according to the HTMT$_{0.85}$ criterions. This implied that the HTMT criterion detect the collinearity problems among the latent constructs (multicollinearity). The constructs of culture-productivity, culture-employee, culture-stakeholder, productivity-employee, productivity-stakeholder, productivity-university performance (UP), employee-stakeholder and stakeholder-university performance are having problems. Probably most of the items of constructs are measuring the same thing. In other words, it contains the overlapping items from the respondents' perception in the affected constructs.

5. Conclusions

To sum up, the instrument used for the survey in this study must be revised and re-evaluated as it has problems in establishing the discriminant validity. Fornell and Larcker criterion and the assessment of the cross-loadings are inadequately sensitive to detect discriminant validity when compared with Heterotrait-monotrait (HTMT) criterion. Thus, the use of HTMT criterion should be adopted for this purpose so that the interpretation of the causal effect in the modelling analysis is not misleading. Despite its strictest procedure (HTMT compared to Fornell and Larcker criterion), the measurement model would be free from any problems besides creating good quality measurement tool through the items in the developed questionnaire. In conclusion, HTMT criterion has high sensitivity and specificity in detecting discriminant validity problems and more empirical evidence is needed to use this approach. However, to assess discriminant validity, model setup along with how conservative the researcher wants to be in assessment of discriminant validity should also be kept in view.

References

[1] Henseler J, Ringle CM, Sarstedt M 2015 *J. Acad. Market. Sci.* **43** 115-35
[2] Fornell C, Larcker DF 1981 *J. Mark. Res.* **1** 39-50
[3] Hair JF, Black WC, Babin BJ, Anderson RE 2010 *Multivariate Data Analysis Seventh Edition Prentice Hall* (New Jersey: Prentice Hall)
[4] Ringle CM, Wende S, Will A 2005 SmartPLS 2.0 M3. www.smartpls.de. [Last assessed 6 January 2011]
[5] Hair JF, Ringle CM, Sarstedt M 2011 *J. Mark. Theory. Pract.* **19** 139-52
[6] Becker JM, Klein K, Wetzels M 2012 Hierarchical latent variable models in PLS-SEM: guidelines for using reflective-formative type models *Long. Range. Plann.* **45** 359-94
[7] Hair J, Hult GTM, Ringle C, Sarstedt M 2014 *A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM)* (Los Angeles: SAGE Publications, Incorporated)
[8] Nunnally JC, Bernstein IH 1994 *Psychological theory* (New York, NY: MacGraw-Hill)
[9] Hulland J 1999 *Strategic Management Journal* **20** 195–204
[10] Bagozzi R P, Yi Y 1988 *J. Acad. Market. Sci.* **16** 74–94
[11] Henseler J, Ringle, Sinkovics RR 2009 *J. Acad. Market. Sci.* **20** 227–319
[12] Fornell C, Cha J 1994 *Advan. Meths. Market. Res.* **407** 52–78.
[13] Kline RB 2011 *Principles and Practice of Structural Equation Modeling Third Edition* (New York: The Guilford Press)
[14] Gold AH, Arvind Malhotra AH 2001 *J. Manage. Inform. Syst.* **18** 185-214.
[15] Ab Hamid MR, Mustafa Z, Suradi NR, Idris F, Abdullah M, Omar M 2012 *Journal of Statistical Modeling and Analytics* **3** 52-7.
[16] Barclay D, Higgins C, Thompson R 1995 *Technology Studies*. **2** 285-309
[17] Sang S L, Lee JD, Lee J 2010 E-government adoption in Cambodia: A partial least squares approach *Transforming government: People, Process and Policy*. **4** 138-57
[18] Rahim MA, Magner NR 1995 *J. Appl. Psychol.* **80** 122-32