Available online at www.elixirpublishers.com (Elixir International Journal)

Statistics

ISSN: 2229-712X

Elixir Statistics 80 (2015) 31035-31039

Moderated mediation using partial least square structural equation modeling (PLS-SEM)

Ahmad NazimAimran[1], AsyrafAfthanorhan[2] and NurHidayahMohdRazali[3]

[1]Malaysia National Population and Family Development Board, Ministry of Women, Family and Community Development, Kuala Lumpur, Malaysia.

[2]Department of Mathematics, Universiti Malaysia Terengganu, Kuala Terengganu, Malaysia.

[3]Department of Computer and Mahematical Science, UniversitiTeknologi MARA, Seremban, Malaysia.

ARTICLE INFO

Article history:
Received: 30 January 2015;
Received in revised form:
25 February 2015;
Accepted: 10 March 2015;

Keywords
Moderation, Mediation,
Moderated Mediation,
Structural Equation CB-SEM,
PLS-SEM.

ABSTRACT

Moderated mediation has been proven to be one of the useful techniques in providing powerful analysis in many research areas such as social science, statistics, marketing, health science and others. By using secondary data obtained from Trends In Mathematics and Science Study (TIMSS), moderated mediation analysis is used to determine the significance difference of direct effect and total effect including indirect effect of exogenous latent constructs toward endogenous latent construct through mediator latent construct between moderator; male and female samples. From the Moderation Analysis, it is found that there is no significant difference between male and female samples in the direct effect of all exogenous latent constructs toward endogenous latent construct. The same result obtained in Moderated Mediation Analysis where there is no significant difference between male and female samples in the total effect including indirect effect from all exogenous latent constructs toward endogenous latent construct through mediator latent construct.

Introduction

Structural Equation Modeling or also known as SEM has gained popularity among researchers, academicians and students nowadays. It is due to its flexibility and generality besides can generate an accurate and precise estimation in making prediction. SEM analysis goes through the steps of model specification, data collection, model estimation, model evaluation and also model modification (ZainudinAwang, 2012). SEM is a unique method because the researcher can modify the structural model in order to increase the model fitness.

The assumption that data is collected from a single homogeneous population is often unrealistic (Esposito Vinzi, 2007). Assuming that data in empirical studies are homogeneous and represent a single population is often unrealistic in the social and behavioral sciences, such as information systems, management, and marketing (Rust and Verhoef 2005; Wedel and Kamakura 2000). The failure to account for heterogeneity leads to ambiguous PLS path modeling results and, thus, to conclusions that are incomplete and ineffective (Esposito Vinzi, 2007).

Analyzing the moderating effect for the model with latent constructs is very complicated. The normal modeling procedure using interaction terms is not practical with latent construct since it would cause problem with model convergence as well as distortion of standard errors. In the end, it result in model misfit and the procedure stops (ZainudinAwang, 2012).

Modeling the moderating effect for latent constructs also known as Multi-Group Confirmatory Factor Analysis (CFA). It was used for assessing the effect of moderator; in this case gender of respondents to the constructed model. Every path may have its' own moderation effect and the path(s) is/are selected in parallel to the research objective. Every path selected will be estimated separately; constrained and unconstrained model. Another method used is the critical ration difference tests which consist of z-test and z-score test. The independent sample z-test for a difference in proportion test the same null hypothesis as does the chi-square test with 1 degree of freedom, these test produce the same result in terms of reject or fail to reject decision (R. Clifford et.al, 2008).

According to Byrne, B.M. (2001), critical ratio consists of z-test and z-score. This test has can be obtained by dividing the regression weight estimates to its standard error. However, this study will be using the t-statistics suggested by Chin et.al. (2010) since researcher is using SmartPLS which is appropriate for the nonparametric assumption study. The latest application of SmartPLS suggests this method since its ability to help researchers to analyze data even with small sample size and nonparametric assumption.

In this study, authors are interested to apply the confirmatory factor analysis (CFA) using Covariance Based Structural Equation Modelling (CB-SEM) since its character methodology fits with the parametric test. Then after, this model will be executed in Partial Least Square Structural Equation Modeling (PLS-SEM) or Composites Based Structural Equation Modeling in obtaining of hypothesis testing. Indeed, hypothesis testing is applicable for the confirmatory purpose as deployed by the character of CB-SEM (theory driven), but, this study also paves away to better understanding the real relationship of Mathematics achievement with the corresponding factors. Plus, multi-group analysis was carry out to understand whether the observed variable based on socio demographic (gender) would provide an effect on mathematics achievement. If so, the observed homogeneity may influence on the parameter estimation of Mathematic achievement. Therefore, the study that involving of gender variable can be a pioneer to drive the practitioners conduct their research in future research due to the capability of PLS-SEM is more practical to predict the relationship of each factor that involves in the study in view of

Tele:
E-mail addresses: ash_18raft@yahoo.com

the gender variable also be tested. In order to investigate this issue, t-statistics was used for helping the researchers to obtain the family wise error rate (p-value) for hypothesis testing.

Materials and methods

The population of the study is defined as the eighth grades (form two) students in Malaysia. Our target population is eighth grades students in Malaysia. A value of 5733 respondents was randomly chosen from 180 randomly chosen schools in Malaysia. In this case study, eighth grades student's attitude towards Mathematics acts as mediating variable, gender acts as moderator variable and achievement in Mathematics examination acts as dependent variable.

Data were obtained from the Trends in International Mathematics and Science Study (TIMSS) international database. Respondents were normally selected through a two stage stratified cluster sampling technique in which consist of cluster sampling for the first stage, school sampling for the second stage and class sampling for thethird stage. The questionnaire consists of four independent variables which are; school environment, teacher's characteristics, student's self-confidence in Mathematics and student's motivation in Mathematics besides student's attitude in Mathematics (mediation variable), student's gender (moderation variable) and student's Mathematics achievement (dependent variable). The data mining software SPSS AMOS was used. Several analysisusedin this study are Confirmatory Factor Analysis (CFA), Discriminant Validity, Path Analysis, Structural Equation Modeling (SEM) and Chi Square test.

Confirmatory Factor Analysis(CFA)is a special form of factor analysis, most commonly used in social research.It is the extended analysis of Exploratory Factor Analysis (EFA) and used to test whether measures of a construct consistent with a researcher's understanding of the nature of that construct (or factor). As such, the objective of confirmatory factor analysis is to test whether the data fit a hypothesized measurement model.Model fit measures could then be obtained to assess how well the proposed model captured the covariance between all the items or measures in the model (Zainudin, 2012). All redundant items exist in a latent construct will be either removed or constrained. Model fitness estimations are as follow:

Table 1: Fitness Indexes Acceptance Level

Name of Category	Name of Index	Level of Acceptance	Literature
Factor Loading	Standardized Regression Weight	Weight > 0.6	Hair et. al.(2006)
Absolute Fit	Chisq	P > 0.05	Wheaton et. al. (1977)
	RMSEA	RMSEA < 0.08	Browne and Cudeck (1993)
	GFI	GFI > 0.9	Joreskog and Sorbom (1984)
Incremental Fit	AGFI	AGFI > 0.9	Tanaka and Huba (1985)
	CFI	CFI > 0.9	Bentler (1990)
	TLI	TLI > 0.9	Bentler and Bonett (1980)
	NFI	NFI > 0.9	Bollen (1989)
Parsimonious Fit	Chisq/df	Chisq/df< 5.0	Marsh and Hocevar (1985)

DISCRIMINANT VALIDITYis the degree to which scores on a test do not correlate with scores from other tests that are not designed to assess the same construct. Correlation coefficients between measures of a construct and measures of conceptually different constructs are usually given as evidence of discriminant validity. If the correlation coefficient is high (>0.85), then the discriminant validity is considered as weak, depending on the theoretical relationship and the magnitude of the coefficient. On the other hand, if the correlations are low to moderate, this demonstrates that the measure has discriminant validity. However, this threshold may meaningless if the correlation matrix and square root Average Variances Extracted (AVE) do not meet the requirement especially during the implementation of second order construct CFA. This method is usually quite restrictive and difficult to handle. The formula of AVE is given by:

$$AVE_J = \frac{1}{p}\sum_{h=i}^{p_j} Cor^2(x_{jh}, \xi)$$

MEDIATION or an indirect effect is said to occur when the causal effect of an independent variable (X) on a dependent variable (Y) is transmitted by a mediator (M) (Preacher et.al., 2007). In other words, X directly affects Y and indirectly affects Y through M . For further explanation, mediation variable can be estimated without the direct effect of exogenous on endogenous construct since the application used is partial least square algorithm (Ramayah, 2013). In this case, the causal effect of exogenous is imposing on mediator and endogenous construct simultaneously.

MODERATION according to (Preacher et.al., 2007), when the strength of the relationship between two variables is dependent on a third variable, moderation is said to be occurring. The third variable, or moderator (W), interacts with X in predicting Y if the regression weight of Y on X varies as a function of W . Moderation is typically assessed with the regression equation:

$$Y = a_0 + a_1X + a_2W + a_3XW + r$$

Where W is considered the moderator, the above equation may be expressed as

$$Y = (a_0 + a_2W) + (a_1 + a_3W)X + r$$

Clarifying how the simple slope of Y regressed on X , $(a_1 + a_3W)$, is the function of the moderator.

Formally, moderated mediation occurs when the strength of indirect effect depends on the level of some variable, in other words, when the mediation relations are contingent on the level of a moderator.

T-STATISTICS in path analysis aims to calculate the difference in paths between groups of sample. Chin (2000) suggests a quick fix to treat the estimates of the bootstrap re-sampling in a parametric sense via t-test. In particular, studenditized t-test has two types of proportions namely equal and unequal variances that should be suit with the characteristics of sampling in population. However, Ott (1984) put forth that the result between equal and unequal variance is almost same as long the discrepancy of each group not higher than of 1.5. In this study, the authors claims that equal variance of t-test can be implemented since the required level of proportion is meet.

$$t = \frac{Path_{sample1} - Path_{sample2}}{\left[\sqrt{\frac{(m-1)^2}{(m+n-2)} * STERR^2_{sample1} + \frac{(n-1)^2}{(m+n-2)} * STERR^2_{sample2}}\right] * \left[\sqrt{\frac{1}{m} + \frac{1}{n}}\right]}$$

where:

m = number of samples 1

n = number of samples 2

$Path_{sample(i)}$ = sample mean for i group(s)

$STERR^2_{sample(i)}$ = the square of standard error for i groups(s)

To perform the moderated mediation analysis, a best fit model for all samples must be first done by using CFA, Discriminant Analysis and Path Analysis. In this study, researcher use Pooled CFA since it has been proven by (Eugenie Eugene Chong et.al.,2014) can obtain equal result as in individual CFA. Then, the data must be separated and saved into different files according to the groups of interest variable (i.e.: in this study, researcher separates the data into male and female groups). After that, execute the Bootstrap Analysis for overall, male and female samples to analyze the difference of before and after the data separation is done. To seek the moderation between (i.e.: male and female) samples, researcher must determine sample size, sample mean (M) and sample standard error (STERR) of path coefficient for male and female samples and calculate the difference in paths between groups of sample by using the t-statistics equation provided by Chin (2000). To proceed with the moderation mediation analysis, the sample size, sample mean (M) and sample standard error (STERR) of Total Effect for male and female samples must be determined before inserted into the t-statistics provided by Chin (2000).

Findings

POOLED CONFIRMATORY FACTOR ANALYSIS (CFA) AND DISCRIMINANT VALIDITY

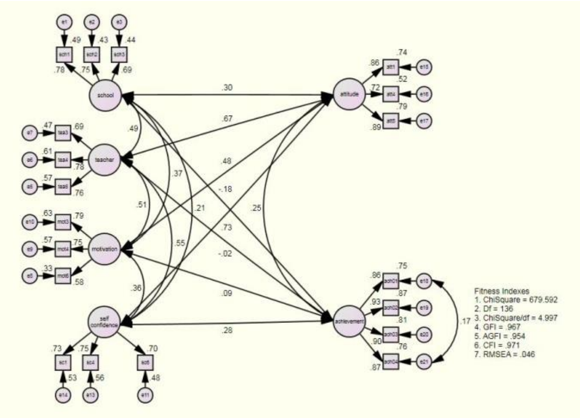

Figure 1: Pooled CFA and Discriminant Validity

Table 2. Table of Summary

Construct	**Item**	**Factor Loading**	**Cronbach Alpha**	**CR**	**AVE**
School	sch1	.781	0.712	0.787	0.552
	sch2	.751			
	sch3	.694			
Teacher	tea3	.685	0.784	0.786	0.551
	tea4	.782			
	tea5	.757			
Confidence	sc1	.729	0.770	0.768	0.525
	sc4	.748			
	sc6	.695			
Motivation	mot3	.794	0.743	0.755	0.511
	mot4	.755			
	mot6	.577			
Attitude	att1	.858	0.855	0.863	0.680
	att4	.720			
	att5	.886			
Achievement	ach01	.864	0.942	0.940	0.798
	ach02	.933			
	ach03	.901			
	ach04	.873			

Table above shows the Factor Loading, Cronbach Alpha, Composite Reliability (CR) and Average Variance Extracted (AVE) values for all latent constructs after Pooled CFA has been performed. All constructs have achieved the minimum estimation required; 0.70(Cronbach Alpha), 0.60 (CR) and 0.50 (AVE). Therefore, it can be concluded that Convergent Validity (AVE ≥ 0.5), Internal Reliability (Cronbach Alpha ≥ 0.6) and Construct Reliability (CR ≥ 0.60) of all constructs had been achieved. Therefore, the model is good enough for the analysis. Besides that, all latent exogenous constructs are correlated with the correlation strength of less than 0.85. Therefore, the discriminant validity is achieved and all latent exogenous constructs are kept in the full model.

Table 3: Discriminant Validity Index Summary

Constructs	**Sch**	**Tea**	**Sc**	**Mot**	**Att**	**Ach**
Sch	**.743**					
Tea	.495	**.742**				
Sc	.205	.547	**.725**			
Mot	.372	.515	.352	**.715**		
Att	.298	.666	.714	.484	**.825**	
Ach	-.178	-.023	.290	.088	.249	**.893**

The diagonal value (in bold) is the square root of AVE, while other values are the correlations between the respective latent construct. The discriminant validity is achieved when a diagonal value (in bold) is higher than the values in its row and column. Referring to the above table, it can be concluded that discriminant validity for all constructs are achieved.

Table 4: Model Fitness Indexes

Name of Category	**Name of Index**	**Index value**	**Required level**
Absolute Fit	RMSEA	0.046	< 0.08
	GFI	0.967	> 0.90
Incremental Fit	CFI	0.971	> 0.90
Parsimonious Fit	Chisq/df	4.997	< 5.0

Table 4 shows that all fitness indexes values had achieved the required level except for Parsimonious fit. Therefore, the model is good enough for the analysis.

Moderation Analysis

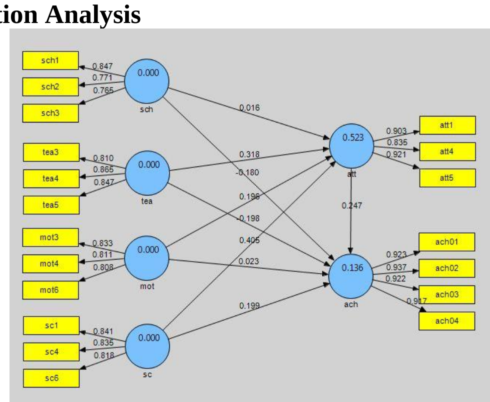

Figure 2: Structural model for overall sample

Table 5: Bootstrap analysis for overall sample

Overall sample	Original Sample (O)	Sample Mean (M)	Standard Error (STERR)
Mot → ach	0.0227	0.0358	0.1178
Sc→ Ach	0.1991	0.2152	0.1255
Sch→ Ach	-0.1798	-0.2018	0.1445
Tea → Ach	-0.1982	-0.1984	0.1335

Table 5 shows the original sample (O), sample (M) and standard error (STERR) for overall sample. The score in the table describe the path for the overall sample without include the heterogeneity factor existing in the model.

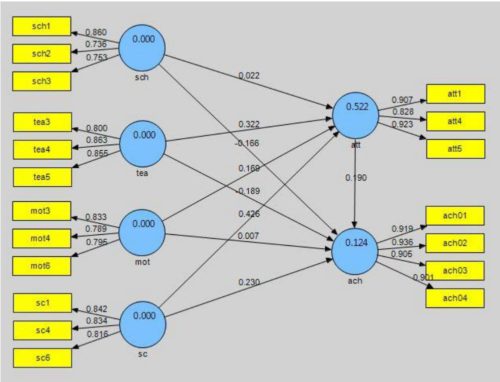

Figure 3: Structural model for male

Table 6: Path Coefficient for male

Male	Original Sample (O)	Sample Mean (M)	Standard Error (STERR)
Mot → Ach	0.0070	0.0050	0.1162
Sc→ Ach	0.2298	0.2421	0.1335
Sch→ Ach	-0.1660	-0.1798	0.1140
Tea → Ach	-0.1095	-0.1928	0.1168

Table 6 and shows the original sample (O), sample (M) and standard error (STERR) for male samples. From the table, it can be seen that the scores obtained are different from the overall sample. Taking sample mean (M) scores into consideration for prior analysis, scores for all exogenous latent constructs for male sample compared to overall sample are0.0050 < 0.0358 (Mot → Ach), 0.2421 > 0.2152 (Sc→ Ach), -0.1798 < -0.2018 (Sch→ Ach) and –0.1928 < -0.1984 (Tea → Ach).

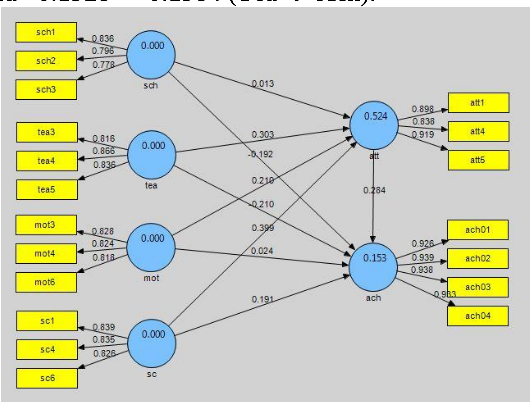

Figure 4: Structural model for female

Table 7: Path Coefficient for female

Female	**Original Sample (O)**	**Sample Mean (M)**	**Standard Error (STERR)**
Mot → Ach	0.0241	0.0177	0.1091
Sc→ Ach	0.1912	0.2096	0.1228
Sch→ Ach	-0.1920	-0.1814	0.1087
Tea → Ach	-0.2104	-0.2070	0.1170

Table 7 shows the original sample (O), sample (M) and standard error (STERR) for male samples. From the table, it can be seen that the scores obtained are different from the overall sample. Taking sample mean (M) scores into consideration for prior analysis, scores for all exogenous latent constructs for male sample compared to overall sample are 0.0177< 0.0358 (Mot → Ach), 0.2096 < 0.2152 (Sc→ Ach), -0.1814< -0.2018 (Sch→ Ach) and –0.2070> -0.1984 (Tea → Ach).

By obtaining Sample Mean (M) and Standard Error (STERR) of Path Coefficient for male and female samples, insert the values into following formula:

$$t = \frac{Path_{sample1} - Path_{sample2}}{\left[\sqrt{\frac{(m-1)^2}{(m+n-2)} * STERR^2_{sample1} + \frac{(n-1)^2}{(m+n-2)} * STERR^2_{sample2}}\right] * \left[\sqrt{\frac{1}{m} + \frac{1}{n}}\right]}$$

For male = 2,610 samples and female = 2,431 samples

Table 8: t-statistics for path coefficient

	t-statistics	Conclusion
Mot → Ach	0.079	There is no significant difference between male and female in the effect of Mot towards Ach.
Sc→ Ach	0.178	There is no significant difference between male and female in the effect of Sc towards Ach.
Sch→ Ach	0.010	There is no significant difference between male and female in the effect of Sch towards Ach.
Tea → Ach	0.086	There is no significant difference between male and female in the effect of Tea towards Ach.

From table 8, it is found that the effect of all exogenous latent constructs toward endogenous latent construct; Mot → Ach, Sc→ Ach, Sch→ Ach and Tea → Ach have no significant difference between male and female samples since all t-statistics value obtained for these paths are lower than 1.96.

Moderated Mediation

By obtaining Sample Mean (M) and Standard Error (STERR) of Total Effects for male and female samples, insert the values into following formula:

$$t = \frac{Path_{sample1} - Path_{sample2}}{\left[\sqrt{\frac{(m-1)^2}{(m+n-2)} * STERR^2_{sample1} + \frac{(n-1)^2}{(m+n-2)} * STERR^2_{sample2}}\right] * \left[\sqrt{\frac{1}{m} + \frac{1}{n}}\right]}$$

For male = 2,610 samples and female = 2,431 samples

Table 9: Total Effects for male

Male	Original Sample (O)	Sample Mean (M)	Standard Error (STERR)
Mot → Ach	0.0391	0.0386	0.1116
Sc→ Ach	0.3107	0.3217	0.1110
Sch→ Ach	-0.1619	-0.1731	0.1137
Tea → Ach	-0.1238	-0.1357	0.1177

Table 10: Total Effects for female

Male	Original Sample (O)	Sample Mean (M)	Standard Error (STERR)
Mot → Ach	0.0837	0.0793	0.1101
Sc→ Ach	0.3044	0.3164	0.1084
Sch→ Ach	-0.1883	-0.1752	0.1096
Tea → Ach	-0.1242	-0.1232	0.1110

From table 11, it can be concluded that the total effect including indirect effect from all exogenous latent constructs toward endogenous latent construct; Mot → Ach, Sc→ Ach, Sch→ Ach and Tea → Ach through mediator latent construct; Att are statistically not different between male and female since all t-statistics value obtained for these paths are lower than 1.96.

Table 11: t-statistics for total effect

	t-statistics	Conclusion
Mot → Ach	0.259	There is no significant difference between male and female in the effect of Mot towards Ach through Att.
Sc→ Ach	0.034	There is no significant difference between male and female in the effect of Sc towards Ach through Att.
Sch→ Ach	0.013	There is no significant difference between male and female in the effect of Sch towards Ach through Att.
Tea → Ach	0.077	There is no significant difference between male and female in the effect of Tea towards Ach through Att.

Conclusion

From the results obtained, it is found that there is no significant difference between male and female samples in the direct effects and total effect including indirect effects of exogenous latent constructs toward endogenous latent construct. Based on the result obtained, both groups can be considered as a homogenous group even though there are some differences in the sample mean scores. However, it can't be concluded that the samples are from homogenous group because they might be distinguish in other observed and unobserved heterogeneity. For further study, researcher may consider other observed heterogenous groups or unobserved heterogeneity groups to be included in the study so that the path models for all heterogenous groups can be estimated.

References

Ahmad Nazim&Sabri Ahmad (2013). Assessing The Unidimensionality, Reliability, Validity And Fitness Of Influential Factors Of 8th Grades Student's Mathematics Achievement In Malaysia. International Journal Of Advanced Research Mathematics (IJOARM) ISSN 2320-9186 Volume 1, Issue 5.

Bentler, P.M. (1990). Comparative fit indexes in structural models. Psychological Bulletin 107: 238-246.

Bentler, P.M. &Bonett, D.G. (1980). Significance Tests And Goodness Of Fit In The Analysis Of Covariance Structures. Psychological Bulletin. 88: 588-606.

Bollen, K.A. (1989). A new incremental fit index for general structural equation models. Sociological Methods and Research 17: 303-316.

Byrne, B. M. (2001), Structural Equation Modeling with AMOS: Basic Concepts, Applications, and Programming, Mahwah, New Jersey: Lawrence Erlbaum Associates.

Chin, W. W., and Dibbern, J. (2010). A Permutation Based Procedure for Multi-Group PLS Analysis: Results of Tests of Differences on Simulated Data and a Cross Cultural Analysis of the Sourcing of Information System Services between Germany and the USA, in Handbook of Partial Least Squares: Concepts, Methods and Applications, V. Esposito Vinzi, W.W. Chin, J. Henseler and H. Wang (eds.), Berlin: Springer, pp.171-193.

Eugenie Eugene Chong, Ahmad Nazim&Sabri Ahmad (2014). A Comparison between Individual Confirmatory Factor Analysis and Pooled Confirmatory Factor Analysis: An Analysis of Library Service Quality, a Case Study at a Public University in Terengganu. International Journal of Engineering Science and Innovative Technology (IJESIT) ISSN 2319-5967Volume 3, Issue 1, January 2014. Pg. 110-116.

Esposito Vinzi, W.W Chin, J. Henseler, H.Wang (2010). Handbook of Partial Least Squares Concepts, Methods and Applications. Springer Handbook of Computational Statistics.DOI 10.1007/978-3-540-32827-8.Springer Heidelberg Dordrecht London New York.

Esposito Vinzi, V., Ringle, C. M., Squillacciotti, S., &Trinchera, L. (2007). Capturing and Treating heterogeneity by Response Based Segmentation in PLS Path Modeling: A Comparison of Alternative Methods by Computational Experiments. Essec Research Centre, DR-07019 juil. 07, Retrieved from http;//www45.essec.eudu/pwr/servlet/showRef?bibID=7153.

Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E. and Tatham, R. L. (2006).Multivariate Data Analysis. 6th edition. Upper Saddle River, New Jersey: Prentice Hall.

Preacher, K.J., Rucker, D.D., & Hayes, A.F. (2007). Addressing moderated mediation hypothesis: Theory, methods and prescription. Multivariate Behavioral Research, 42, pp.185-227.

Rust, R.T., and Verhoef, P.C. (2005).Optimizing the Marketing Interventions Mix in Intermediate-Term CRM, Marketing Science (24:3), pp.477-489.

Tanaka, J.S. and Huba, G.J. (1985). A fit index for covariance structure models under arbitrary GLS estimation. British Journal of Mathematical and Psychology 38: 197-201.

Wedel, M., and Kamakura, W. (2000). Market Segmentation: Conceptual and Methodological Foundations (2nd ed.), New York: Kluwer Academics Publisher.

Zainudin, A. (2012). Research Methodology and Data Analysis 5th Edition. Shah Alam: UniversitiTeknologi MARA Publication Centre (UiTM Press).