RICHTMANN
PUBLISHING

Research Article

Testing the Construct Validity and Reliability of Curiosity Scale Using Confirmatory Factor Analysis

Heronemus Sujati[1]

Sajidan[2]

Muhammad Akhyar[2]

Gunarhadi[2]

[1]Doctoral Program of Educational Science,
Universitas Sebelas Maret, Surakarta, Indonesia
[2]Faculty of Teacher Training and Education,
Universitas Sebelas Maret, Surakarta, Indonesia

Abstract

Curiosity is one of the important characters that should be acquired by every elementary student in Indonesia. Currently, a standardized instrument to measure the student's curiosity has not yet available. This study aims to test the construct validity and reliability of the developed curiosity scale. The research involved 300 random samples from 565 students of the fifth grade of elementary school in Ngaglik district of Sleman region, in Yogyakarta, Indonesia. The data was analyzed with the second order confirmatory factor analysis using LISREL 8.80 software. The results show that the developed curiosity scale had fulfilled the criteria of goodness of fit, convergent validity, discriminant validity, construct reliability and omega composite reliability. Therefore, the developed scale was feasible to use.

Keywords: curiosity scale, second order confirmatory factor analysis, convergent validity, discriminant validity, construct reliability and omega composite reliability

1. Introduction

Education is an important foundation for every nation's development (Mustapha & Bolaji, 2015). Indonesia's national education system has a goal to produce intelligent, skillful and characterful citizens. Regarding to the character, the government has recently developed a character education from the early childhood to the higher education level. It is expected that character education would produce educated citizens with good personalities and values of hard-working, discipline, ethics, religiousness, honesty, politeness, tolerance, confidence, patriotism, curiosity, and many more. This idea is in line with Lickona (1991) who believes that character education aims to create a student's personality. Similarly, Berkowitz & Bier (2005) state that character education encourages the student's ethics development through universal values teaching.

The aforementioned description shows that curiosity has become one of the goals of character education for several reasons. According to Gülten, Yaman, & Özsari (2011), curiosity serves some

benefits like strengthening social relationship, preventing Alzheimer, and supporting learning activities. Therefore, a curious person is prone to be happy and innovative. Curiosity encourages a student to gain new experiences and to learn from his surroundings. Engel (2015) argues that curiosity is an important component in learning. Thus, stimulating the student's curiosity should be an essential part of education and learning.

To obtain data on the quality of the students' curiosity, an assessment is needed. Assessment is a set of procedure that is designed to receive information about the students' growth, development and achievement to be compared with a certain standard (Shermis & Di Vesta, 2011). Clements & Cord (2013) state that an assessment is an important component in learning process. The assessment could identify and evaluate the students' strengths and weaknesses, assess the students' learning process, monitor and provide feedback on the students' progress. Thus, an assessment instrument for teachers is urgently required.

In reality, an assessment instrument on curiosity based on psychometrics criteria is currently not available. Consequently, teachers apply simple instrument that does not reflect scientific results. This research is aimed to develop a curiosity scale that meets two psychometric criteria which are validity and reliability.

2. Literature Review

Worthington & Whittaker (2014) state that factor analysis is a statistical procedure that can be used to reduce the number of observed variables to a small number of latent variables by examining the covariance between observed variables. According to Morata-Ramírez & Holgado-Tello (2013), the use of factor analysis would allow researchers to obtain empirical evidence about the internal structure of measurement instruments, namely the relationship between latent variables and their latent factors to observable variables. McCoach, Gable & Madura (2013) added that the use of factor analysis would enable the researcher to identify number of constructs contained in an instrument and explain the pattern of relationships between observational variables with their constructs and between constructs and latent variables.

When it is viewed from the usage, there are two kinds of factor analysis called as exploratory and confirmatory (Kerlinger, 1996). Exploratory analysis factor is the use of factor analysis to identify any factors underlying a group of variables. This analysis does not hypothesize the number of factors from the items that construct the variables. The items are left alone to form some particular patterns and inform how many factors existing in one variable. In other words, the researcher does not know much about any theories underlying the variables which were measured (Matsunaga, 2011). On the contrary, confirmatory factor analysis aims at examining the validity of the number of factors as they are stated in one variable and the connection between each item and its factor (Furr, 2011).

The difference between exploratory and confirmatory analysis factor located on the perspective on how a factor is formed. On the exploratory analysis factor, theoretical hypothesis is not employed when using analysis factor. The formation of the factors is based on analysis results. Meanwhile, the factor formation on confirmatory analysis factor is based on the hypothesis developed on the prior stage, where the researcher has already known the structure of latent x and Larcker (1981) and Agarwal (2013) that the construct validity test can be carried out by confirmatory factor analysis which has been considered as an appropriate way to see the relationship between observed variables and latent variables (Jackson, Gillaspy & Stephenson, 2009). The CFA examines the relationship between the observed variables and the unobserved variables, in which the researcher uses a hypothetical model to estimate the population covariance matrix compared to the observed sample covariance matrix (Schreiber, Nora, Stage, Barlow & King, 2006).

CFA assumes that the observed variables are imperfect indicators of certain latent variables or underlying constructs (Wijanto, 2008). McCoach, Gable and Madura (1986) argue that in the CFA researchers have postulated a priori specific models and then tested whether the data collected was in accordance with the hypothesized model. Therefore, in CFA the number of factors has been

determined, as well as the hypothesized items to be measured. Associated with many items for each factor, Tabachnick & Fidell (2012) proposed at least three items. In this way, CFA provides information about how well the factor model is hypothesized.

3. Purpose of the Study

The purpose of this study is to test a curiosity scale that fulfills two psychometric criteria which are validity and reliability. The construct validity includes convergent validity and discriminant validity, while the reliability includes construct reliability and composite reliability.

4. Methodology

4.1 Research Approach

This study employed a research and development (R&D) approach which is aimed to develop and validate an educational product (Borg & Gall, 1983). The product of this study was a curiosity scale that met the validity and reliability criteria. The research procedure was adapted from the theory developed by McCoach, Gable & Madura (1986) involving 13 steps as follows: (1) identifying research problems and needs, (2) determining variables, (3) reviewing a comprehensive study on curiosity, (4) writing conceptual and operational definition, (5) constructing indicators, (6) choosing instruments, (7) designing blueprint, (8) writing the items based on the blueprint, (9) conducting content validation qualitatively and quantitatively by some experts, (10) revising the items based on the experts' suggestions, (11) conducting an empirical field test, (12) conducting validity and reliability analysis with confirmatory factor analysis (CFA) using LISREL 8.80 software and (13) designing the final instrument.

4.2 Sample Size

Sample size requirement for a structural equation modelling (SEM) has become a concern for the researchers. An advance in statistical modelling approach and the easiness of software operation resulted in various sample size requirements in testing a model. Schumacher & Lomax (2010: 42) argue that to achieve a precise calculation with CFA, a researcher needs 250 to 500 respondents, while Hoetler (1983) suggests 500 respondents. Comrey & Lee (1992) determine sample size of 50 - very poor, 100 - poor, 200 – fair, 300 – good, 500 – very good and 1000 – excellent. Besides, Anderson, & Gerbing propose sample size of 150 or more to achieve a minimum standard error. Conclusively, 300 respondents were randomly chosen from 565 elementary school students in Ngaglik district of Sleman region in Yogyakarta, Indonesia.

4.3 Data Analysis

The two main requirements in an instrument development are the validity and reliability of a research model (Bacharach, 1998). Hence, this research used the second order confirmatory factor analysis (2nd order CFA). Brown (2015), O'Rouke & Hatcher (2013) state that CFA is suitable for determining construct validity and instrument reliability. Moreover, Hair, Black, Babin,, & Anderson (2019) express that CFA could be used to test not only construct validity, but also construct reliability. Hill & Hughes (2007) state that CFA allows factors, variance and relationships between latent constructs to be reviewed. In this case, it permits the establishing of both convergent validity and discriminant validity.

5. Findings and Discussion

5.1 Goodness of fit test

A goodness of fit test aims to see how well a theoretical model fits the empirical model. In fact, there is no single rule in testing modeling fit. Hu & Bentler (1998) propose two criteria in assessing modeling fit: comparative fit index (CFI) and roots-mean-square error of approximation (RMSEA). Meanwhile, Gerbing & Anderson (1992) suggest three criteria: p-value of χ_2, normed fit index (NFI) and relative fit index (RFI). Accordingly, this research used the p-value of χ_2, CFI, RMSEA, NFI, and RFI for the assumption testing. The goodness of fit test result showed the value of $\chi_2 = 210.02$; p-value = 0.091; CFI = 1,0; RMSEA = 0.022; NFI = 0,98 and RFI = 0,98. The results indicated that the empirical data fit the developed theoretical model. Thus, it could proceed to an analysis stage.

5.2 Factor Loading

Two results of the second order confirmatory factor analysis are the determination of factor loading (λ) of each item and factors that develop scale. Initially, the researchers provided four factors that involve 28 items, but after the preliminary analysis it was found that there were only 21valid items. Table 1 shows the factor loading of every factor and its items.

Table 1. Factor Loadings Factor and Item

Factor	Item number	Statement	λ	λ
Interested in new things	1	Search for new information on various media	0,74	
	2	Try something new	0,75	
	3	Investigate new things	0,79	0,87
	4	Observe new things	0,76	
	5	Use new ways to solve problems	0,79	
	6	Wish for new atmosphere	0,79	
Question something	9	Clarify unclear idea	0,66	
	10	Ask something to anyone who knows more	0,74	0,84
	11	Look for answers to my own question	0,74	
	12	Question strange things	0,79	
Courage to try	13	Dare to face failure in trying new things	0,77	
	16	Never give up in difficult situation	0,79	
	17	Tackle problems by thinking out of the box	0,69	0,66
	18	Solve problems which are different from the job sheet	0,78	
	19	Try to search by myself than told by others	0,77	
Enthusiastic to solve problems	21	Reject friends' invitation while doing task	0,84	
	23	Keep spirited even if I failed	0,77	
	24	Search for a more difficult task after completing a certain task	0,70	0,78
	25	Use my spare time to work on a challenging task	0,72	
	26	Reduce my break time for completing my task	0,84	
	27	Feel anxious when the task is incomplete	0,82	

Factor loading is a correlation of each variable and the factor (Field, 2009). Factor loading is a way to interpret the role of every variable in defining its factor. Therefore, a higher factor loading would make the variable more representative for the factor. Concerning this theory, the number of factor loading should be carefully considered. Hair et al. (2019) claim that factor loading of ≥0.50 is practically significant. According to the result of the second order confirmatory factor analysis as presented on table 1, it was found that all items and factors shown factor loading of > 0.5. As a result,

all items and factors were practically significant and feasible to be used in data collecting.

5.3   Construct Validity

According to Furr & Bacharach (2003) construct validity refers to an extent to which the measurement score reflects latent construct to be measured. Meanwhile, Hair et al. (2019) define construct validity as an approach to make sure that a set of variables represents the theoretical latent construct which is being measured.

Fornell & Larcker (1998), Agarwal (2013) note that construct validity of confirmatory factor analysis includes two main tests, namely convergent validity test and discriminant validity test. Campbell & Fiske (1959) describe that convergent and discriminant validity are essential requirements on every instrument development to obtain accountable data psychometrically. Accordingly, this research reported both convergent and discriminant validity.

5.3.1 Convergent Validity

Convergent validity refers to the degree to which similar constructs are measured with different variables (Hill & Hughes, 2007; Kenny & Kashy, 1992). In other words, convergent validity ensures that variables belong to the latent construct to be measured (Wang, French & Clay, 2015). Convergent validity is based on the correlation between responses of different variables in measuring the same construct (Peter, 1981: 136). Subsequently, variables should be highly correlated with the latent construct (Engellant, Holland & Piper, 2016).

The amount of factor loading is a fundamental consideration in determining convergent validity (Hair et al., 2019). Igbaria, Zinatelli, Cragg, & Cavaye (1997) demonstrate that a variable is good if the latent variable shows the factor loading of ≥0.50. Hair et al. (2019) recommend an average variance extracted (AVE) as convergent validity measure since AVE could explain the degree to which items are shared between the construct in structural equation modelling (SEM) where AVE 0.5 or more are acceptable as convergent validity.

The scale development in this study involved four constructs, namely interested in new things, question something, courage to try and enthusiastic to solve problems. The result shown that the AVE values for the four constructs respectively were: 0.59, 0.54, 0.58 and 0.61. As all constructs exceeded the threshold AVE value of >0.50, it is concluded that they could measure the latent variables. Hence, they fulfilled the convergent validity criteria.

5.3.2 Discriminant Validity

A discriminant validity test is a requirement in an instrument development that involves latent variable (Ab Hamid, Sami & Sidek, 2017). Discriminant validity which also refers to divergent validity (DeVelis, 2017) means that two concepts should show significant differences conceptually. A discriminant validity test aims to prove that one construct is highly different from the other one (Voorhees, Brady, Calantone, & Ramirez, 2015). Discriminant validity reveals the extent to which a construct is distinguished from other constructs in a model (Hair et al., 2019; Barclay, Higgins, & Thompson, 1995).

Discriminant validity is shown by the correlation between latent constructs which is not too high (Peter, 1981), or low covariance factor (Kenny & Kashy, 1992). Discriminant validity confirms that every latent construct is unique. In other words, one latent construct should not be highly correlated with the other one (Henseler, Ringle, & Sarstedt, 2014). It is fulfilled when the two latent constructs are not correlated theoretically and empirically proven from the scores showing one construct is higher than the other (Bagozzi & Dholakia, 2002).

Hair et al. (2019) state that discriminant validity could be established by correlating one construct to another. If the correlation value of both constructs is lower than 0.85, it means that the

E-ISSN 2240-0524
ISSN 2239-978X

Journal of Educational and Social Research
www.richtmann.org

Vol 10 No 4
July 2020

discriminant validity exists. Besides, Fornell, & Larcker (1981) argue that discriminant validity exists if latent variable shows more variances on related indicator variable rather than share with other construct in the same model.

Table 2. Discriminant Validity

	Interest	Question	Courage	Enthusiasm
Interest	**0,77**			
Question	0,637	**0,73**		
Courage	0,496	0,485	**0,76**	
Enthusiasm	0,623	0,554	0,484	**0,78**

The results presented on Table 2 inform that the four latent constructs respectively had square roots of AVE: 0.77, 0.73, 0.76, and 0.78. The square roots of AVE of the four latent constructs were greater than the inter-construct correlation. Conclusively, the four latent constructs had fulfilled the criteria of discriminant validity.

5.4 Reliability

Beside validity, reliability is also required for a scale. Brown (2015) refers reliability to the consistency of the measurement results. Likewise & Robert (2006) states that a reliable instrument could maintain the consistency of the measurement results in a certain range. The degree of consistency is determined by reliability coefficient. Margono (2015) suggests that an instrument is reliable if it could measure the same phenomenon repeatedly, yet it yields relatively consistent results. This study reported two types of reliability, namely constructs and composite reliability.

5.4.1 Construct reliability

According to Retnawati (2017), construct reliability (CR) is a measure of internal consistency of the variables that represent latent construct to be measured. Construct reliability is used to measure to what extent variables underlying the constructs served in structural equation modelling (Zinbarg, Revelle, Yovel, & Li, 2005). Construct reliability could be estimated after the construct validity is proved using confirmatory factor analysis. Based on the factor loading analysis of the factor, construct reliability is estimated (Geldhof, Preacher, & Zyphur, 2014). Gefen, Straub, & Boudreau (2000) suggest that construct reliability coefficient higher than 0.70 is acceptable. A high coefficient indicates high internal consistency. It would only be possible if every variable consistently measure the same latent construct.

The result of this study shown that the four constructs: interested in new things, question something, courage to try and enthusiastic to solve problems have CR coefficient of 0.90, 0.82, 0.87, and 0.90. Referring to the CR coefficient threshold of 0.70 by Gefen, Straub, & Boudreau (2000), it is concluded that every variable in this study was reliable and feasible to use.

5.4.2 Composite Reliability

Composite reliability, known as internal consistency, (Fornell & Larcker, 1981) is the combination reliability of latent constructs underlying a scale (Hair et al. 2019; Geldhof, Preacher, & Zyphur, 2014). Alpha and omega are two popular methods of composite reliability (Bacon, Sauer, & Young, 1995; Peters, 2014; Padilla & Divers, 2015). Dunn, Baguley and Brunsden (2013) suggest that if a researcher is sure that the scale had fulfilled the unidimensional principle, he might use alpha. Conversely, if the researcher is not sure whether the scale is unidimensional, he might use omega. McDonald (1981) clarifies that an instrument is unidimensional if the covariant between the item in the instrument is

zero (0) when it is used by test takers with similar proficiency. Furthermore, omega reliability could estimate the scale reliability more accurately and higher than alpha (Revelle & Zinbarg, 2008). Based on this consideration, omega (ω) reliability was applied in this study. Viladrich, Angulo-Brunet & Doval (2017) claim that the threshold coefficient of composite reliability higher than 0.70 is reliable. The result of this study shown that the scale had coefficient omega of 0.96. Thus, the scale was reliable.

6. Conclusion

Based on the discussion, it is concluded that the four indicators: interested in new things, question something, courage to try and enthusiastic to solve problems which were designed theoretically to develop curiosity scale were proven to meet the criteria of goodness of fit, validity and construct reliability. Furthermore, the scale had fulfilled reliability criterion comprehensively. Therefore, the developed scale was feasible to be used in collecting data to measure the student's curiosity.

References

Agarwal, V. (2013). Investigating The Convergent Validity of Organizational Trust. *Journal of Communication Management*, 17(1), 24–39. doi:10.1108/13632541311300133

Anderson, J. C. & Gerbing, D. W. (1988). Structural Equation Modeling in Practice: A Review and Recommended Two-Step Approach. *Psychological Bulletin*, Vol. 103, No. 3, p. 411-423. DOI: org/10.1037/0033-2909.103.3.411

Bacharach, S. B. (1989). Organizational Theories: Some Criteria for Evaluation. *Academy of Management Review*, 14(4), 496–515. doi:10.5465/amr.1989.4308374

Bacon, D. R., Sauer, P. L., & Young, M. (1995). Composite Reliability in Structural Equations Modeling. *Educational and Psychological Measurement*, 55(3), 394–406. doi:10.1177/0013164495055003003

Barclay, D., Higgins, C. & Thompson, R. (1995). The Partial Least Squares (PLS) Approach to Causal Modelling: Personal Computer Adoption and Use as an Illustration. *Technology Studies*, Vol. 2(2), pp. 285-309.

Berkowitz, M. W. & Bier, M.C. 2005. *What Works in Character Education: A Research-Driven Guide for Educators*. Washington DC: University of Missouri St Louis.

Brown, T. A. (2015). *Confirmatory Factor Analysis for Applied Research*. Second Edition. New York: The Guilford Press.

Carlson, K. D., & Herdman, A. O. (2014). Understanding the Impactof Convergent Validity on Research Results. *Organizational Research Methods*, 15(1), 17–32. DOI: 10.1177/1094428110392383

Clements, M. D. & Cord, B. A. (2013). Assessment Guiding Learning: Developing Graduate Qualities in An Experiential Learning Program. *Assessment and Evaluation in Higher Education*, 38 (1), 114-124

DeVellis, R. F. (2017). *Scale development: Theory and applications*, Fourth Edition. Los Angeles: Sage.

Dunn, T. J., Baguley, T., & Brunsden, V. (2013). From Alpha to Omega: A Practical Solution to The Pervasive Problem of Internal Consistency Estimation. *British Journal of Psychology*, 105(3), 399–412. doi:10.1111/bjop.12046

Engel, Susan. (2015). *The Hungry Mind: The Origins of Curiosity in Childhood*. London: Harvard University Press.

Field, A. (2009). *Discovering Statistics Using SPSS*, 3rd Edition. New Delhi: SAGE Publications Ltd.

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, 18, 39-50. DOI: 10.2307/3151312

Furr, R. M. & Bacharach, V. R. (2013). *Psychometrics: An Introduction*. Singapore: SAGE Publications, Inc.

Gerbing, D. W., & Anderson, J. C. (1992). Monte Carlo Evaluations of Goodness of Fit Indices for Structural Equation Models. *Sociological Methods & Research*, 21(2), 132–160. doi:10.1177/0049124192021002002

Gefen, D., Straub, D. W. & Boudreau, M.-C. (2000), 'Structural Equation Modeling and Regression: Guidelines for Research Practice', *Communications of the Association for Information Systems*, vol. 4, no. 7, pp. 1-77. doi: 10.17705/1CAIS.00407

Kerlinger, F.N. (1996). Asas-asas Penelitian Behavioral. Yogyakarta: Gadjah Mada University Press.

McCoach, D. B., Gable, R. K., & Madura, J. P. (2013). Instrument Development in the Affective Domain, 3rd Edition. New York: Springer.

Geldhof, G. J., Preacher, K. J., & Zyphur, M. J. (2014). Reliability estimation in a multilevel confirmatory factor analysis framework. *Psychological Methods*, 19(1), 72–91. doi:10.1037/a0032138

Ghozali, I. (2014). *Strutural Equation Modeling: Teori, Konsep, dan Aplikasi dengan program Lisrel 9.10.* Semarang: Badan penerbit Universitas Diponegoro.

Gülten, D. Ç, Yaman, Y., Deringöl, Y., & Özsari, I. (2011). Investigating the relationship between Curiosity Level and Computer Self Efficacy Beliefs of Elementary Teacher Candidates. *The Turkish Online Journal of Educational Technology*, 10(4), p. 248-154.

Hair, J.P., Black, J.P., Babin, J.P., & Anderson, R.E. (2019). *Multivariate Data Analysis*, Eighth Edition. Harlow: Cengage Learning.

Hamid, M. R., Sami, W., & Mohmad Sidek, M. H. (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

Henseler, J., Ringle, C. M., & Sarstedt, M. (2014). A New Criterion for Assessing Discriminant Validity in Variance-Based Structural Equation Modeling. *Journal of the Academy of Marketing Science*, 43(1), 115–135. doi:10.1007/s11747-014-0403-8

Hill, C.R., & Hughes, J.N. (2007). An Examination of the Convergent and Discriminant Validity of the Strengths and Difficulties Questionnaire. *Sch Psychol Q*, 22(3), p. 380–406. doi: 10.1037/1045-3830.22.3.380

Hoelter, J.W. (1983). The analysis of covariance structures: Goodness-of-fit indices. *Sociological Methods and Research*, 11: 325344. doi: org/10.1177/0049124183011003003

Hu, L., & Bentler, P. M. (1999). Cutoff Criteria for Fit Indexes in Covariance Structure Analysis: Conventional Criteria Versus New Alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1–55. doi:10.1080/10705519909540118

Igbaria, M., Zinatelli, N., Cragg, P., & Cavaye, A. L. M. (1997). Personal Computing Acceptance Factors in Small Firms: A Structural Equation Model. *MIS Quarterly*, 21(3), 279. doi:10.2307/249498

Jackson, D. L., Gillaspy, J. A., & Purc-Stephenson, R. (2009). Reporting practices in confirmatory factor analysis: An overview and some recommendations. Psychological Methods, Vol. 14, No. 1, p. 6–23. DOI:10.1037/a0014694

Kenny, D. A., & Kashy, D. A. (1992). Analysis of The Multitrait Multimethod Matrix By Confirmatory Factor Analysis. *Psychological Bulletin*, 112(1), 165–172. doi:10.1037/0033-2909.112.1.165

Kevin A. Engellant, K. A., Holland, D. D. & Randy T. Piper, R. T. (2016). Assessing Convergent and Discriminant Validity of the Motivation Construct for the Technology Integration Education (TIE) Model. *Journal of Higher Education Theory and Practice*, Vol. 16(1), p. 37-50. doi.org/10.33423/jhetp.v16i1.1935

Latan, H. (2012). *Structural Equation Modeling Konsep dan Aplikasi Menggunakan LISREL 8,80*. Bandung: Alfabeta.

Lickona, T. 1991. *Educating for Character, How Our Schools Can Teach Respect and Responsibility*. New York: Bantam Books.

Margono, G. (2015). Multidimensional Reliability of Instrument for Measuring Students' Attitudes Toward Statistics by Using Semantic Differential Scale. *American Journal of Educational Research*, 2015, Vol. 3, No. 1, 49-53, doi: 10.12691/education-3-1-10

Matsunaga, M. (2011). How to Factor-Analyze Your Data Right: Do's, Don'ts, and How-To's. International Journal of Psychological Research, 3 (1), 97-110.

McCoach, D. B., Gable, R. K., & Madura, J. P. (1986). *Instrumen Development in the Affective Domain*, 3rd Edition. New York: Springer, doi: 10.1007/978-94-015-7259-0

McDonald, R. P. (1981). The Dimensionality of Tests and Items. *British Journal of Mathematical and Statistical Psychology*, 34(1), p. 100–117. doi:10.1111/j.2044-8317.1981.tb00621.x

Morata-Ramírez, M. A., & Holgado-Tello, F. B. (2013). Construct validity of likert scales through confirmatory factor analysis: a simulation study comparing different methods of estimation based on pearson and polychoric correlations. International Journal of Social Science Studies, Vol. 1(1), p. 54-61.

Mustapha, M. & Bolaji, B. Y. (2015). Measuring Lecturers Commitment Scales: A Second order Confirmatory Factor Analysis (CFA). *International Journal of Education and Research*, Vol. 3, No. 3, p. 505-516.

O'Rourke, N. & Hatcher, L. (2013*). A Step-by-Step Approach to Using SAS for Factor Analysis and Structural Equation Modeling*. Cary: SAS Institute Inc.

Padilla, M. A., & Divers, J. (2015). A Comparison of Composite Reliability Estimators. *Educational and Psychological Measurement*, 76(3), p. 436–453. doi:10.1177/0013164415593776

Peter, J. P. (1981). Construct Validity: A Review of Basic Issues and Marketing Practices. *Journal of Marketing Research*, 18(2), 133–145. doi:10.1177/002224378101800201

Peters, G. Y. (2014). The Alpha and The Omega of Scale Reliability and Validity. *The European Health Psychologist*, Vol. 16, Issue 2, p. 56-69.

Retnawati, H. (2017). Validitas dan Reliabilitas Konstruk Skor Tes Kemampuan Calon Mahasiswa. *Jurnal Ilmu Pendidikan*, 23(2), p. 126 -135.

Revelle, W., & Zinbarg, R. E. (2008). Coefficients Alpha, Beta, Omega, and the glb: Comments on Sijtsma. *Psychometrika*, 74(1), 145–154. doi:10.1007/s11336-008-9102-z

Roberts, P., Priest, H., & Traynor, M. (2006). Reliability and validity in research. *Nursing Standard*, 20(44), 41–45. doi:10.7748/ns2006.07.20.44.41.c6560

Schreiber, J. B., Nora, A., Stage, F. K., Barlow, E. A., & King, J. (2006). Reporting structural equation modeling and confirmatory factor analysis results: A review. The Journal of Educational Research, 99(6), 323–338. doi:10.3200/joer.99.6.323-338

Schumacher, R. E., & Lomax, R. G. (2010). *A beginner's Guide to Structural Equation Modeling* (3rd ed.). New York: Routledge.

Shermis, M. D. & Di Vesta, F. J. 2011. Classroom Assessment in Action. New York: Roman & Littlefield Publishers, Ink.

Tabachnick, B. G., & Fidell, L. S. (2012). Using multivariate statistics (7th ed.). New York: Harper & Row.

Viladrich, C., Angulo-Brunet, A., & Doval, E. (2017). Un Viaje Alrededor de Alfa Y Omega Para Estimar la Fiabilidad de Consistencia Interna. *Anales De Psicología*, 33(3), 755. doi:10.6018/analesps.33.3.268401

Voorhees, C. M., Brady, M. K., Calantone, R., & Ramirez, E. (2015). Discriminant Validity Testing in Marketing: An Analysis, Causes for Concern, and Proposed Remedies. *Journal of the Academy of Marketing Science,* 44(1), 119–134. doi:10.1007/s11747-015-0455-4

Wang, X., French, B. F., & Clay, P. F. (2015). Convergent and Discriminant Validity with Formative Measurement: A Mediator Perspective. *Journal of Modern Applied Statistical Methods*, Vol. 14(1), pp. 83-106. doi: 10.22237/jmasm/1430453400

Worthington, R. L. & Whittaker, T. A. (2014). Scale Development Research: A Content Analysis and Recommendations for Best Practices. The *Counseling Psychologist*, Vol. 34 No. 6, p. 806-838. DOI: 10.1177/0011000006288127

Wijanto, S.H. (2008). Strutural Equation Modeling dengan LSREL 8.8: Konsep dan Tutorial. Yogyakarta: Graha Ilmu.

Zait, A. (2014). Methods for testing discriminant validity. *Management & Marketing*, Volume IX, issue 2, pp. 217-224

Zinbarg, R. E., Revelle, W., Yovel, I., & Li, W. (2005). Cronbach's α, Revelle's β, and Mcdonald's ωH: Their Relations with Each Other and Two Alternative Conceptualizations of Reliability. *Psychometrika*, 70(1), 123–133. doi:10.1007/s11336-003-0974-7