International Journal of Psychosocial Rehabilitation, Vol. 24, Issue 05, 2020
ISSN: 1475-7192

Validation of Developed Instruments on Study Habits, Creativity, and Self Concept on College Students in Nigeria

Sulaiman Adamu Mayanchi, Aqeel Khan and
Adibah Binti Abdul Latif

Abstract--- *The role of validating a questionnaire is to ensure that the questionnaires for use in research settings are psychometrically sound, valid and reliable as well as efficient and effective. This article will therefore provide the validation of existing resources on study habits, creativity and self-concept through the use of the Rasch Measurement approach. A pilot study was randomly selected from each of the two Colleges in Zamfara State, with a sample of 180 respondents. The employed Rasch-based Winsteps code produces the required parameter estimate automatically. The results of the reliability coefficients showed the reliability of 0.83, 0.96 and 0.87 items, with 0.96, 0.91 and 0.94 corresponding to Cronbach Alpha. The unidimensional construct of the test measures supported by the raw variance explained by measurements of 48.3 percent, 53.8 percent and 53.0 percent closely match the variance expected. Investigation of the map of the item person revealed that all items fell within the respondents ' ability level. Likewise, fitness indices showed that 5 items were listed to delete items for study habits, and sub-scale creativity reveals that all 23 items in the scale have reasonable fitness while 5 items have poor self-concept sub-scale fitness indices. The results confirm the accuracy of the explanations and inferences of the scores on the objects and subscales of the instruments.*

Keywords--- *Validation, Reliability, Rash Measurement Approach, Study Habit, Creativity and Self-concept.*

I. INTRODUCTION

Validity shows how well an instrument can calculate what it is supposed to measure and function as well as how it is supposed to do, since it is virtually impossible for a device to be accurate at 100% (Yue Li, 2016). Therefore, validity refers to how effectively an instrument calculates what it aims to measure (Zelt et al.2018); this article will provide how to validate established tools on study habits, creativity and self-concept by using the Rasch Measurement approach hence, it has been found that having high expectation-value values is correlated with the continuity of a student in good study habit, positive self-concept, and creativity in improvising new innovation from local materials within the society (Alana Unfried & Malinda et, al, 2015; Khan et al., 2020; Khan, 2013). The Rasch Measurement framework was used to estimate the accuracy or reliability and validity of the' Study Habit Creativity and Self-Concept Sub-Scales to assess sufficient internal quality reliability and construct alignment using unidimensional test component characteristics (Linacre, Aziz et al., 2019 and 2013). However, it will be considered three validity areas that include: face validity, content validity, and construct validity.

Sulaiman Adamu Mayanchi, School of Education, Faculty of Social Science & Humanities, Universiti Teknologi Malaysia (UTM), Johor, Malaysia.
Aqeel Khan, School of Education, Faculty of Social Science & Humanities, Universiti Teknologi Malaysia (UTM), Johor, Malaysia.
Adibah Binti Abdul Latif, School of Education, Faculty of Social Science & Humanities, Universiti Teknologi Malaysia (UTM), Johor, Malaysia.

DOI: 10.37200/IJPR/V24I5/PR2020261
Received: 16 Mar 2020 | Revised: 30 Mar 2020 | Accepted: 08 Apr 2020

Face validity is an external evaluation to determine if a measure appears to subjectively measure the framework of each instrument (Connell et al., 2018). If accurate representation of the variable is provided, a test is said to show face validity (Richard Huan Xu,et, al 2018), therefore, face validity of the instruments in this study is determined by the team of experts who validate the instruments from the Faculty of Education UTM Malaysia and recommended that the instruments of this study has face validity after modification and delineation of certain items from the sub-scales of the study

Several changes have been made to the concept of validity. The reliability of the calculation of what a test aims to measure was described by Kelley (1927:14). Guilford (1946: 429) argued that "the experiment extends to anything it corresponds to" (Baghaei, 2008). Construct validity is, therefore, the extent to which the measure ' behaves ' in a manner consistent with theoretical hypotheses and shows how well the theoretical construct indicates scores on the instrument (Morris,et, al., 2017). The reliable tool is critical to measuring what is meant to be measured in addition to using the correct measuring tool (Gholamabbas Shirali, et, al, 2018). To ensure proper measurements of performance are carried out (Fuchs et al., 1999; Figlio & Lucas, 2004; Banta, 2007; Tahir, 2014) and valuable inferences can be made (Wright & Mok, 2004), it is crucial that the measuring tool is reliable and valid (Baghaei, 2008) cited in (Azrilah Abdul Aziz, Muhammad Shahar Jusoh, Omar, Mohd Haris Amlus, & Tuan Salwani Awang, 2014). The confidentiality of the score's meaning and interpretation is an important aspect of the validity of the construct (Shroff, & Ting, et., al 2019). The scientific research aimed at establishing this aspect of validity is called the validity evidentiary basis of the test items (Baghaei, 2008).

II. RASCH MEASUREMENT MODEL

The Rasch model is called a prescriptive model and has two essential internal scaling and invariance properties (Wu, M. & Adams, R. 2007). The model is called a prescriptive model because it prescribes specific data fulfilment conditions. This means that the whole research process must be consistent with the design requirements from the very beginning when unidimensionality is assumed (i.e. when test data match the model) is met (Reckase, Ackerman, & Carlson,1988). Unidimensionality (one function at a time should be tested by the experiment) is one of the basic assumptions of the Rasch measurement method. One of the basic assumptions of the Rasch measurement method is unidimensionality: the experiment will test one element at a time (Embretson, & Reise, 2000). Although the hypothesis is theoretically valid, it is practically impossible to build tests that evaluate only one attribute or prevent the test from affecting external influences (Baghaei 2008).

Rasch analysis is primarily designed to fulfil the specified construct validity and Rasch item analysis also focuses on measuring examinee ability and item complexity; model fit approximation, unidimensionality assessment and distractor analysis (Kathy Green and Catherine Frantom, 2002**)**. These are the metric used to measure the quality of the test item and the relevance to the evaluated trait taking into account the person's ability (Baghaei, 2008). Complexity, model fit approximation, unidimensionality assessment and distractor analysis. Rasch's use in education after the introduction of Georg Rasch in 1960 resulted in better learning outcomes and applied to medicine, public health, and other disciplines respectively (Wolfe, & Chiu, 1999).

DOI: 10.37200/IJPR/V24I5/PR2020261
Received: 16 Mar 2020 | Revised: 30 Mar 2020 | Accepted: 08 Apr 2020

Therefore, this research uses the Rasch measurement method to evaluate the instrument's construct validity; hence, unidimensionality was used to determine the construct validity through the Rash measurement model (Linacre, (2002) & Kline, (2005). One of the Rasch model's strengths is that it constructs a theoretical one-dimensional line along which objects and individuals are based on their measures of complexity and ability (Baghaei 2008). The items falling as close to the theoretical line as possible help calculate the single dimension specified in the theory of construction (Smith, 1992). The factor similar to the main dimension of Rash is determined by those who fell out of it (Levine, & Saintonge, 1993). Long distances between the items on the line suggest that there are significant differences between item difficulties so that people who fall in skill near this part of the line are not assessed as accurately by the test (Azrilah Abdul Aziz, Muhammad Shahar Jusoh, Omar, Mohd Haris Amlus, & Tuan Salwani Awang, 2014).

III. METHODOLOGY

Participant/Respondent

The participants in the present study comprise 180 students that are randomly selected from the two colleges of education in Zamfara State, Nigeria. One of the colleges of education is mixed which includes male and female students, while the second one comprises only female students. In the selection of participants, the students ' stratum (gender) was recognized to ensure adequate representation of the target population intended for the test being created in the mixed college.

Instruments

This research will use the questionnaire with series of items to collect data from the respondents. Questionnaires are more advantageous compared to other instruments because they collect data from a large sample and have organized and standardized responses (Leal Filho & Kovaleva, 2015), And that most researchers and respondents are familiar with the general methodology, especially in the field of educational research. (Leal Kovaleva & Filho, 2015). In this study, the questionnaire to be used consists of four different sections. Section A contains demographic information on the brief biography of the respondents, while section B contains 6 constructs of the respondents with 34 items measuring study habits (SHI), Section C consists of 3 dimensions with 23 items measuring creativity (RIBS), Section D consists of 6 dimensions with 30 items measuring the respondents ' self-concept (ASCSS). In addition, the expert team reviewed and checked the test items and recommended that certain items in the instruments be changed and removed based on the order of the expert examination and decreased the items from 93 test items to 87 test items.

Administration of the Instruments

Prior to the administering of the test, permission was sought and obtained from the school authorities and venues where the test items will be conducted were also provided in the two colleges respectively. The instructions and purpose of the test was explained to the respondents and their attention and consents were obtained before they start the test.

DOI: 10.37200/IJPR/V24I5/PR2020261
Received: 16 Mar 2020 | Revised: 30 Mar 2020 | Accepted: 08 Apr 2020

IV. DATA ANALYSIS

Using Rasch Analysis WINSTEPS 3.72.3, the data are analyzed using the Rasch Approach (Linacre, & Wright, 2000). In the Rasch test, three different assessment scales are considered, (i) Calibration of the skill of the applicant and difficulties of the element (ii) Assessment of fit (iii) Evaluation of unidimensionality using Rasch residuals main component analysis (PCA) (Bond and Fox, 2015). Using person-item maps the relationship between the skill of students and the difficulties of items were presented (Wright, & Stone, 1979*).* To test the fit statistics, the mean square values (MNSQ) and the standard Z values (ZSTD) were tested.

V. RESULTS AND DISCUSSIONS

It is also important to measure the validity of the content; to what extent the set of items covers the different components to be measured (Connell et al., 2018). There are various ways or methods to assess content validity such as test specification table or panel of experts (Gholamabbas Shirali, et, al, 2018). Therefore, based on the previous psychological assessments, the content expert review is the best method to check an instrument's content validity (Connell et al., 2018). However, team of experts who are specialists in the field of counselling, educational psychology and educational measurement under faculty of education UTM Malaysia were selected to assess the content of the instruments.

The content validity index was obtained using Kappa coefficient generated from the content validity ratings of the items by the team of experts. The Kappa values for the three main research instruments based on the expert's assessment are 0.843 for study Habit, 0.698 for Creativity and 0.910 for Self-Concept as presented in table 1 below:

Table 1: Interpretation of Content Validity Indices (Kappa)

Construct	**Number of Items**	**Kappa Value**	**Kappa Interpretation**
Study Habit Inventory	34	0.843	Almost perfect agreement
Runco Ideational Behaviour Scale (Creativity Measure)	23	0.698	Substantial agreement
Adolescents Self-Concept Short Scale	30	0.910	Almost perfect agreement
Total	**87**	**0.815**	**Almost perfect agreement**

Validation of interpretation of study habits, creativity and self concept sub scales amendments from the team of experts is illustrated in table 2 below:-

Table 2: Validation Interpretation of Independent Variable Amendment

Independent Variables Measures	**Experts' Agreement and Comment**	**Amendment**
Study Habit	Two experts identified two dimensions with less than 3 items (2 dimensions should be discarded)	Recording and Language were removed as dimensions
	Two experts recommend item 38, 39 and 40 to be discarded, while 1 expert recommend for Modification or discard	Dimension (Support) with items 38-40 were removed.
	Two out of the three experts recommend for the modification of items 1, 2, 16, 25, 26, 27, 30, and 37 while one examiner recommend item 1 and 30 to be discarded but items 2, 16, 25, 26, 27 and 37 are ok	Items 1,2, 16, 25, 26, 27, 30 and 37 were modified
Creativity	Two experts recommend items 22 and 23 to be modified while one expert feel the items are ok	Items 22 and 23 were Modified.
Self-concept	One expert recommends items 6 and 7 to be modified while the two experts feel they are ok.	No modification required

DOI: 10.37200/IJPR/V24I5/PR2020261
Received: 16 Mar 2020 | Revised: 30 Mar 2020 | Accepted: 08 Apr 2020

Following the process of evaluating reliability, face and content validity of the instruments as well, a study was conducted as a preliminary experiment on a smaller scale of 90 respondents from each school totalling 180 respondents for the study. Through the study, (Wright & Mok, 2004) researcher can be able to identify unforeseen problems that may include unnecessary items from the questionnaires and to collects feedback from study samples to refines the length of the scale to improve the instruments during the real study (Roberta Heale, & Alison Twycross, 2015)

VI. Rasch Analysis Measurement Results

Reliability (Internal Consistency)

The Rasch Measurement approach was used to estimate the reliability of the 'Study Habit, Self-Concept and Creativity Sub-Scales. The reliability coefficient of the scales in both person and item based on relevant standards has excellent person, item as well as adequate internal consistency reliability (Linacre, 2019 & Aziz et al., 2013). Summary presented on Table 1 showed that, person sample used in the Test is large enough to confirm the item difficulty hierarchy of the test items.

Table 3: Summary Statistics

SN	Reliability	Study Habit	Self-Concept	Creativity
1	Item Reliability	0.83	0.96	0.87
	Item Separation	*2.14*	*5.04*	2.64
2	Person Reliability	0.94	0.91	0.91
	Person Separation	*4.02*	*3.14*	*3.18*
3	Cronbach Alpha	0.96	0.91	0.94

The reliability coefficients presented in Table 1 revealed the item reliabilities of 0.83, 0.96 and 0.87, with corresponding Cronbach Alpha of 0.96, 0.91 and 0.94. These parameters were considered satisfactory reliability because according to Vandenberg & Lance, (2000), Hair, Anderson, Tatham & Black (1995) and Cohen et al. (2013), a Cronbach's alpha scale greater than 0.70 is acceptable for the internal consistency reliability of the items and can therefore be accepted for research purpose. These criteria served as the guidelines in interpreting the internal consistency-reliability coefficients in the study (Wright, 1977 & Hu & Bentler 1999)

VII. Construct Validity

Unidimensionality Analysis

In validating the research instrument, the'Study Habit, Self-Concept and Creativity Sub-Scales'. The alignment of the items to the constructs can be identified using unidimensional characteristics (Reckase, Ackerman, & Carlson, 1988). Two (2) indicators; the raw variance explained by measures (RVEM) and unexplained variance in the first contrast (UVFC) can be used to detect unidimensionality using PCA (Linacre, 2019 and Aziz et al., 2013). The results of the Principal Component Analysis (PCA) of Rasch residuals to check unidimensional characterises of the test items is presented in Table 4:-

Table 4: Summary Statistics

SN	Sub-Scale	Raw variance explained by measures	UVFC	Expected
1	Study Habit,	48.3%	9.6%	49.0%
2	Self-Concept	53.8%	12.9%	53.4%
3	Creativity	53.0%	8.6%	51.9%

DOI: 10.37200/IJPR/V24I5/PR2020261
Received: 16 Mar 2020 | Revised: 30 Mar 2020 | Accepted: 08 Apr 2020

The result presented in Table 4 showed that the unidimensional nature of the scales and confirmed further whether the items are measuring study habit, self-concept and creativity. The raw variance explained by measures are 48.3%, 53.8% and 53.0% closely match the expected variance of 48.0%, 53.4% and 51.9% respectively. The rule as provided by Linacre, (2019), Aziz et al., (2013) & Gorsuch, (1983), is that, for an instrument to achieved unidimensionality, it requires the minimum of raw variance explained by measures (RVEM) of 40% and unexplained variance in the first contrast (UVFC) to be less than 15% to fully established that, the Scale is unidimensional and no substantial and meaningful secondary dimension seem to exist within the data (Aziz et al., 2013, & Brown, & Moore, 2012; Alshemmeri et al 2011).

Model Fitness Statistics

Based on the Rasch model analysis various tests were used to identify error responses labelled as 'outliers' or 'misfit' as these are commonly referred terms used in psychometrics (Karabatsos, G. 2000). As the name suggests 'misfit' refers to the estimates that do not fit into the overall model fit. According to the IRT framework the model fit is determined by examining the misfit indices revealed by the *Outfit Mean Square* (MNSQ) measures, the *Outfit Z Standard* (ZSTD) estimates as well as the Point Measure Correlation indices (Smith, Schumacker, & Bush, 1995).

According to Bond and Fox (2015) the variance is larger for well-targeted observations and smaller for extreme observations. In this study the *Outfit MNSQ* (+1.05) shows acceptable variances within the responses because the accepted value of this fit statistic is close to 1. It is also important to examine and consider the existence of misfit items in a questionnaire which could arise if items are either too easy or too difficult for respondents (Wright, & Panchapakesan, 1969). This could result in a situation where the items do not exactly measure the desired latent variable. *Outfit MNSQ* measures are more susceptible to extreme responses compared to *Infit MNSQ* measures (Abdullah & Lim, 2013). Thus, to maintain any item in a test is should satisfy the following conditions as provided by Linacre (2019):

i) PTMEA CORR is positive and not 0 or close to it
ii) The INFIT and OUTFIT MNSQ index fall within the acceptable range for Multiple choice Questions, i.e., $0.7 \leq MNSQ \leq 1.3$
iii) The Z standard (ZSTD) values fall within acceptable range of $-2.0 \leq Z \leq$

Misfit Items in Study Habit

Based on the item map (Appendix), all the items fall within the ability level of the respondents. Further, the investigation of fitness indices revealed that, 5 items (SH 8, 9, 10, 12 and 30) have out of range fitness statistics. To decide whether to omit them from the scale or maintain to be use in the next administration, the indicators of fit were investigated i.e., Point Measure Correlation (PTMEA CORR), INFIT Mean Square (INFIT MNSQ) and OUTFIT Mean Square (OUTFIT MNSQ) (Hair & Black, et, al, 2006). The result shows that, all these 5 items have unacceptable Infit and Outfit parameters (see appendix) and therefore cannot be used to collect the relevant data in measuring study habit in this study (Linacre, 2012). Thus items 8, 9, 10, 12 and 30 were deleted from the final instrument.

DOI: 10.37200/IJPR/V24I5/PR2020261
Received: 16 Mar 2020 | Revised: 30 Mar 2020 | Accepted: 08 Apr 2020

Misfit Items in Self-Concept

Based on the item map (Appendix), there are 4 items lower than the less able respondents (SC 12, 13, 16 and 25) but all these items have satisfactory fitness indices. However, there are 5 items that have out of range fitness indices (SC 10, 19, 20, 27 and 28). After careful investigation of the fit statistics in terms of Point Measure Correlation (PTMEA CORR), INFIT Mean Square (INFIT MNSQ) and OUTFIT Mean Square (OUTFIT MNSQ). The result showed that, all these 5 items have poor fitness indices (Linacre, 2012). Thus items SC 10, 19, 20, 27 and 28 were completely eliminated from the final instrument.

Misfit Items in Creativity

The investigation on the Creativity sub-scale shows that, all the 23 items in the scale have acceptable fitness statistics (Linacre, 2012). Thus all the 23 items are retained in the final draft of the scale. Depending on these results, various criteria are used to pick a psychometric protocol and Professionals should explain their appraisal strategy and think about the target group. In view of the Rasch results analysis instruments sub scales can be used to test college students.

VIII. CONCLUSION

Determining the criteria of performance or problematic and better items is an important step in creating valid and reliable test items to assess students ' true ability (Lawshe, 1975). This article therefore, provides how developed instruments on study habits, creativity and self-concept are validated through the use of Rasch Measurement approach to ascertain its construct validity and reliability evidences to the college students. According to Cheung, & Rensvold, (2002), the Rasch frameworks provided a scientific insight on how different test items performed in the developed test by identifying several poor or problematic using Person-Map-Items, Item Fit Statistics (MNSQ, ZSTD and PTMEA CORR). Ten items were deleted from the instruments through Rasch analysis from the final instruments.

Interpretation using Rasch, however, offers more detailed information on the item structure required to validly determine the student's ability and suitability of the items to calculate the intended outcome. Likewise, in validating sub-scale instruments in study, consideration should be given to the magnitude of the decision to be made from the responses obtained from the rasch analysis and experts evaluation.

REFERENCES

[1] Abdul Aziz, A., Masodi,M.S. & Zaharim, A. Asas(2013). Model pengukuran Rasch. *Bangi: Universiti Kebangsaan Malaysia*.

[2] Abdullah, N. & Lim, B. K.(2013). Parallel circuit conceptual understanding test (PCCUT). *Procedia—Social and Behavioral Sciences*, *90* (InCULT 2012), 431– 440.

[3] Alana Unfried & Malinda et, al, (2015). *Journal of Psycho-educational Assessment* 1–18 © 2015 SAGE Publications Reprints and permissions: sagepub.com/journalsPermissions.nav DOI: 10.1177/0734282915571160 jpa.sagepub.com.

[4] Alshemmeri, F., Putih, A., Siraj, S., Khan, A., Abdallah, N. (2011). Art Ability and Academic Achievement in the Kingdom of Saudi Arabia: Role of Age and Sex. *New Educational Review*, *26* (4), 238-247.

[5] Azrilah Abdul Aziz, Muhammad Shahar Jusoh, Omar, A. R., Mohd Haris Amlus, & Tuan Salwani Awang (2014). Construct Validity: A Rasch Measurement Model Approaches. *Journal of Applied Science and*

Agriculture, *9*(12), 7–12.

[6] Baghaei P. (2008). The Rasch model as a construct validation tool. *Rasch Measurement Transactions*, 22 (1): 1145-1146.

[7] Banta, T.W. A (2007). Warning on Measuring Learning Outcomes. *Inside Higher Education*.

[8] Bond T., Fox C. M. (2015) *Applying the Rasch Model: Fundamental Measurement in the Human Sciences*, 3rd Edn New York, NY: Taylor & Francis10.4324/9781315814698 [CrossRef] [Go ogle Scholar]

[9] Brown, T. A., & Moore, M. T. (2012). Confirmatory factor analysis. In R. H. Hoyle (Ed.), Handbook of structural equation modeling (pp. 361-379). *New York, NY: Guilford Press.*

[10] Cohen, A. S., Morrison S. C., Callaway, D. A. (2013). Computerized facial analysis for understanding blunted affect: Initial feasibility, reliability, and validity data. *Schizophrenia research.* 148(1-3):111-6.

[11] Connell, J., Carlton, J., Grundy, A., Taylor Buck, E., Keetharuth, A. D., Ricketts, T. Brazier, J. (2018). The importance of content and face validity in instrument development: lessons learnt from service users when developing the Recovering Quality of Life measure (ReQoL). *Quality of Life Research*, *27*(7), 1893–1902.

[12] Cheung, G. W. & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, *9*, 233-255.

[13] Embretson, S. E., & Reise, S. P. (2000). Item response theory for psychologists. *Mahwah, NJ: Lawrence Erlbaum Associates.*

[14] Figlio, D.N. and M.E. Lucas,. Do High Grading Standards Affect Student Performance? *Journal of Public Economics,* 88: 1815-1834 2004.

[15] Fuchs, L.S., D. Fuchs, K. Karns, C.L. Hamlett and M. Katzaroff 1999. Mathematics Performance Assessment in the Classroom: Effects on Teacher Planning and Student Problem Solving. *American Educational Research Journal,* 36(3): 609-649.

[16] Gorsuch, R. L. (1983). Factor analysis (2nd ed.). *Hillsdale, NJ: Lawrence Erlbaum*

[17] Guilford, J.P. (1946). New standards for test evaluation. *Educational and Psychological Measurement,* 6, 427-439.

[18] Hair, J. F. Anderson, R. E., Tatham, R. L., & Black, W. C. (1995). Multivariate data analysis. *Englewood Cliffs, NJ: Prentice-Hal.*

[19] Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., & Tatham, R. L. (2006). Multivariate data analysis (6th ed.). *Upper Saddle River, NJ: Prentice Hall.*

[20] Hu, L., & Bentler, P. M. (1999). Cut-off criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*, 1-55.

[21] Karabatsos, G. (2000). A critique of Rasch residual fit statistics. *Journal of Applied Measurement, 1*, 152-176.

[22] Kathy E. Green and Catherine G. Frantom, (2002.) Survey Development and Validation with the Rasch Model. *A paper presented at the International Conference on Questionnaire Development, Evaluation, and Testing, Charleston, SC, November 14-17.*

[23] Khan, A., et al. (2020). Mediating Effect of Positive Psychological Strength and Study Skills on Examination Anxiety among Nigerian College Students. *Sustainability, 12*, 1479.

[24] Khan, A. (2013). Predictors of Positive Psychological Strengths and Subjective Well-being among North Indian Adolescents: Role of Mentoring and Educational Encouragement. *Social Indicators Research, 114*, 3, 1285-1293.

[25] Kelley T.L. (1927). Interpretation of educational measurements. *Yonkers, NY, World Book Company*

[26] Kline, R. B. (2005). Principles and practice of structural equation modeling (2nd ed.). *New York, NY: Guilford Press.*

[27] Lawshe, C. H. (1975). A quantitative approach to content validity. *Personnel Psychology*, *28*, 563-575.

[28] Linacre JM. (2012). A user's guide to Winsteps-Ministep: Rasch model computer programs. Program manual 3.68.0.

[29] Linacre, J.M. (2019). Winsteps (Version 4.4.2) [Computer Software]. *Beaverton, Oregon: Winsteps.com.*

[30] Linacre, J. M., & Wright, B. D. (2000). WINSTEPS: A Rasch computer program. *Chicago: MESA Press.*

[31] Linacre, J. M. (2002). Understanding Rasch measurement: Optimizing rating scale category effectiveness. *Journal of Applied Measurement, 3*, 85-106.

[32] Li, Y. (2016). How to Determine the Validity and Reliability of an Instrument. *Miami University Discovery Center*.

[33] Leal Filho, W., & Kovaleva, M. (2015). Research methods. *Environmental Science and Engineering (Subseries Environmental Science), 5*(9783319109053), 81–82.

DOI: 10.37200/IJPR/V24I5/PR2020261
Received: 16 Mar 2020 | Revised: 30 Mar 2020 | Accepted: 08 Apr 2020

[34] Levine, J. B., & Saintonge, S. (1993). Psychometric properties of the Separation- Individuation Test of Adolescence with a clinical population. *Journal of Clinical Psychology, 49*, 492-507.

[35] Rasch, G. (1960). Probabilistic models for some intelligence and attainment tests. *University of Chicago Press, Chicago.*

[36] Reckase, M. D., Ackerman, T. A., & Carlson, J. E. (1988). Building a unidimensional test using multidimensional items. *Journal of Educational Measurement, 25*, 193-203.

[37] Smith, R. M. (1992). Application of Rasch measurement. *Chicago: MESA Press.*

[38] Smith, R. M., Schumacker, R. E., & Bush, J. M. (1995, April). Using item mean squares to evaluate fit to the Rasch model. *Paper presented at the annual meeting of the American Educational Research Association, San Francisco.*

[39] Vandenberg, R. J. & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods, 3*, 4-69.

[40] Wu, M. & Adams, R. (2007). Applying the Rasch model to psycho-social measurement: A practical approach. *Educational Measurement Solutions, Melbourne*

[41] Wolfe, E. W., & Chiu, C. W. T. (1999). Measuring change across multiple occasions using the Rasch rating scale model. *Journal of Outcome Measurement, 3*, 360-381.

[42] Wright, B. D., & Stone, M. H. (1979*). Best test design*. Chicago: MESA Press.

[43] Wright, B. D. (1977). Solving measurement problems with the Rasch model. *Journal of Educational Measurement, 14*, 97-116.

[44] Wright, B.D. and M.C. Mok, (2013). An Overview of the Family of Rasch Measurement Models. *A chapter in the Introduction of Rasch Measurement. Chapter,* 1: 1-24.

[45] Tahir L, Mohd Said, MNH., Daud, K, Vazhathodi, SH, Khan A. (2014). The Benefits of Headship Mentoring: An Analysis of Malaysian Novice Head teachers' Perceptions. *Educational Management Administration & Leadership, 44*, 420-450.

DOI: 10.37200/IJPR/V24I5/PR2020261
Received: 16 Mar 2020 | Revised: 30 Mar 2020 | Accepted: 08 Apr 2020