Universiti Teknologi MARA
Cawangan Kedah
Kampus Sungai Petani

Voice of Academia

Academic Series of Universiti Teknologi MARA Kedah

BUSINESS MANAGEMENT AND ENTREPRENEURSHIP

SPECIAL ISSUE

VoA
Volume 15 Issue 1
August 2019

COMMITTEE PAGE

e-ISSN: 2682-7840

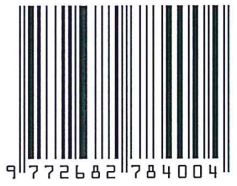

TABLE of CONTENTS

Voice of Academia
e-ISSN: 2682-7840

Kabiru al. / Voice of Academia 15 Special Issue August (1) 2019, 62-72

Voice of Academia
Academic Series of Universiti Teknologi MARA Kedah
Available online at http://voa.uitm.edu.my

Revalidation of Student Evaluation of Teaching Effectiveness Rating Scale in Ahmadu Bello University Zaria

Kabiru Jinjiri Ringim, Bello Sabo, Mukhtar Yusuf Abubakar & Nasir Abdullahi

ABU Business School, Ahmadu Bello University, Zaria, Kaduna State, Nigeria

ARTICLE INFO

Article history:
Received June 2019
Received in revised form
Accepted June 2019
Published August 2019

Keywords:
SETERS, Teacher' Effectiveness, Factor Analysis, Construct Validation, ABU Zaria

Corresponding Author:
Jinjiri Ringim kabirujinjiri@yahoo.com

ABSTRACT

The growing interest in moving beyond traditional measures of teachers' performance and effectiveness in the classroom as the basis for making decisions for promotion and selection for leadership roles in Ahmadu Bello University (ABU), Zaria led to the adoption of Student Evaluation of Teaching Effectiveness Rating Scale (SETERS). Therefore, the objectives of this study were to determine the suitability and validity of SETERS adoption in the Business Administration Department, Ahmadu Bello University, Zaria, Nigeria. The study used descriptive research design. A proportionate stratified random sampling technique was used for a sample selection of the respondent from undergraduate students of semester one (1) to semester eight (8) respectively. Eight hundred (800) questionnaires were returned out of the 1,000 questionnaires distributed by hand delivery to the students with the help of four Research Assistants. The data collected was processed using SPSS for descriptive and inferential analysis. Results emanating from principal component analysis suggested five factors: 1) Clear delivery of course information; 2) Knowledge and Experience sharing; 3) Clear course outline and direction; 4) Lecturer Interaction and Student participation; 5) Provision of progress report to the student, the most important. The components revealed an Eigenvalue greater than one explaining the variance respectively and accounted for 65.46% of the total scale variance. The correlation matrix showed the presence of many significant coefficients of 0.3 and above. The Kaiser-Meyer-Olkin value reached statistical significance, supporting the factorability of the correlation matrix. In an attempt to establish construct validity, convergent validity was examined using Average Variance Extracted (AVE) that shows how indicators of SETERS construct converged and share common variance. Finally, reliability test for each dimension of SETERS emerged with a Cronbach's alpha ranges from 0.692 to 0.826 and a composite reliability coefficient of 0.907 for the major variables in the questionnaire used for the study implied that the instrument was reliable. Hence the instrument had excellent reliability concerning the internal consistency. The outcomes also confirmed SETERS as demonstrating strong convergent va-

lidity. It was concluded that SETTERS in its present form (i.e., in this study) is suitable for use in the evaluation of teaching effectiveness of lecturers in Ahmadu Bello University, Zaria Nigeria.

1. Introduction

In recent years there has been growing interest in moving beyond the traditional measures of assessing academics qualifications, publications, and community service for promotion in most of the Nigerian universities, as all these criteria are not sufficient predictors of teaching effectiveness (Mordi, 2003). The key to improving public education and teaching effectiveness is to place highly skilled and effective teachers in all classrooms. In Nigerian Universities, the major problem is that current measures for assessing academics for promotion in most of the Nigerian Universities are not linked to their capacity to teach. To evaluate teachers' actual performance and effectiveness in the classroom, previous research findings have demonstrated that effective teaching would lead to good academic performance (Abdulkadir, 2006). Parents and policymakers agree that the key to improving public education and teaching effectiveness is by placing highly skilled and effective teachers in all classrooms (Sanders and Rivers, 1996). In a practical situation, existing Nigerian University practices for measuring teacher effectiveness rely on perception by Heads of Departments or focus on teachers' course-taking records and paper-and-pencil tests of basic academic skills and subject matter knowledge. This is in addition to the criteria for assessing academics for qualifications, publications and community service at the local, national or international levels (Adomi and Mordi, 2003). All these are poor predictors of teaching effectiveness. The measurement of teaching effectiveness using SETER Scale has not been validated for adoption. This means, its suitability for use in Nigeria has not been established. A past study in Europe indicated that Students' evaluation of their teachers had been suggested as a solution for assessing the teaching effectiveness at the Universities of Medieval Europe (Central, 1993). It has shown that there has been a tremendous increase in interest regarding students' ratings of instruction and this topic has been the subject of a substantial body of research spanning approximately 70 years (Areola, 1995). Students' ratings were used in the North American Universities in the mid-1920s as the basis for promoting their teachers (D'Apollonia & Abrami, 1997).

Students' Evaluations of Teaching Effectiveness Rating Scale (SETERS) has been useful in assessing teaching effectiveness in Western Countries (Marsh, 1987). The SETERS was developed in the United States of America, whose culture, and the environment is different from Nigeria's and therefore there is a need for cross-cultural validation of SETERS. As recommended by many authors that scales be cross-validated before use once such scales are to be used in different cultures and national boundaries (Tschannen-Moran, Woolfolk Hoy & Hoy, 1998; Henson, 2001 and Brouwers, 2003). Moreover, previous researchers, Toland and De Ayala (2005) saw a need for additional empirical research on the SETERS before widespread use. Hence, there is a need to revalidate the SETER SCALE dimensions (Instructor's Delivery of Course Information, Teacher's Role in Facilitating Instructor/Student Interaction and Instructor's Role in Regulating Student's Learning) for proper evaluation of teachers' effectiveness in Business Administration of ABU Zaria.

Also, the study explored the inter-relationship between SETER Scale factors and Business Admin students of ABU Zaria. Given the research problem that is presented above, specifically

in the Business Administration Department ABU Zaria, and the Nigerian context, there is a dearth of literature on revalidation of a research instrument to the knowledge of the researcher, this study seeks to address the following three (3) research questions. Firstly, how reliable is SETERS Scale items for measuring teaching effectiveness in the Business Administration department in ABU Zaria? Secondly, is the factor structure of SETERS significantly valid in measuring the teaching effectiveness? Thirdly, to what extent does SETERS Scale factors are inter-related? Therefore, the three (3) objectives of the study are: to investigate the reliability (the internal consistency and stability) coefficients of SETERS; to determine the factor structure for convergent and divergent validity of SETER Scale in teaching effeteness in Business administration department of ABU Zaria; and to determine the inter-relationships between the SETERS Scale factors.

2.0 Methodology

2.1 Research Design

This study focuses on descriptive and inferential statistical research analysis. The research setting was a cross-sectional study design. It involves gathering the data only once or at one point in time to meet the research objectives (Cavana, Dalahaye, & Sekaran, 2001). Therefore, non experimental design or survey using the quantitative method of administering the questionnaire is employed in this research.

2.2 Population and Sampling design and determination of sample size

The population of the study is 3,000 students of level one, two, three and four in the Business administration department. This study employed a probability sampling method because it gives each respondent an equal chance of being selected as the object (Sekaran, 2003). Furthermore, simple random is used in this study because it ensures equal and independent representation of the chosen data. This method is exceptionally appropriate as the selection or probability that one person was chosen does not affect the probability of another person being chosen thereby avoiding bias (Salkind, 2003). It also ensures high generalizability (Cavana et, al., 2001). Krejcie and Morgan (1970) greatly simplified the determination of sample size decision by providing the table that ensures a good precision for margin of error and confidence level. Based on the table, the sample size for this study is approximately 910. For increased representation, 10% of the population sample size was added to the minimum sample size given by the table to make it 1,000. This will also take care of other unavoidable errors such as incorrect filing and failure of some respondents to return questionnaire (Israel, 2013).

2.3 Data collection strategy

The questionnaire is perhaps the most widely used survey instrument to get first-hand information from the respondents (Osuala, 2005). In an attempt to get the completed questionnaire returned as quickly as possible, the hand delivery and collection method was chosen; which is expected to give a high response rate. Hand delivery and collection are an efficient method in an environment where a research culture is not sufficiently developed, such as Nigeria. For instance, research has shown that the rate of return of mailed questionnaires is between 3 percent and 4 percent (Asika, 1991). The questionnaire was distributed based on the proportion of the population in each class.

2.4 Measurement and validation test of construct

The study adapted the Students Evaluation of Teaching Effectiveness Rating Scale' (SETERS) developed by Toland & DeAyala (2005). The questionnaire design for this study consist of two main sections. Section A consist of questions relating to the personal profile of the students. Section B consists of 34 items measuring three dimensions of teaching effectiveness, namely: Instructor's Delivery of Course Information (12 items), Teacher's Role in Facilitating Instructor/ Student Interaction (10 items) and Instructor's Role in Regulating Student's Learning (12 items). The response is of the Likert-type format ranging from Strongly Disagree to Agree Strongly, on a rating scale of 1 to 5. The information gathered from the survey was processed using the Statistical Package for the Social Sciences (SPSS) for descriptive and inferential statistical analysis. The result of the analysis was used to answers the research questions of the study. Factor analysis was done to validate the constructs (SETERS) and reduce the dimensionality of a data set in which there are a large number of interrelated items while retaining as much as possible of the variation present in the data set. Exploratory factor analysis was conducted using principal component analysis, face validity, convergent, discriminant and Nomological and construct reliability of both composite and Cronbach alpha coefficient analysis would be done to validate the constructs.

2.5 Pre-Test/Face Validity

A pre-test of the questionnaire was conducted to enable testing of the alternative wordings and question sequences to determine which format best suits the respondents. The purpose of the pre-test was to alert the researcher to potential problems that may be caused by the questionnaire. Thus, pre-tests were conducted to answer questions on the questionnaire, such as the following: 1) Can the questionnaire format be followed by the researcher 2) Does the questionnaire flow naturally? 3) Can respondents answer the question easily? 4) Which alternative form of questions works best? Pre-testing also provides the means to test the sampling procedure, whether efficient or not. Therefore, the benefit of conducting a pre-test of the questionnaire is to improve the validity and reliability of the instrument measures (Zikmund, 2000).

Face validity, according to Sekaran and Bougie (2010), indicates that the items that are intended to measure a construct seem to have measured it. With regards to the measurement scale of this study, six experts - Senior Lecturers, Associate Professor, and Professor in ABU Zaria - were consulted, and their observations noted and affected. Similarly, for the meaningful and logical understanding of variables, the same people were consulted.

3.0 Data Analysis and Discussions

3.1 Descriptive Statistics

3.1.1 Lecturer/ Student Interactions

The descriptive analysis of lecturer/student interaction in the class was shown in Table 1.0 in appendix 1.0. The finding revealed that all the items had mean values of more than 2.47. This generally indicated that there is a cordial relationship between the lecturers and the students in terms of encouraging students to ask questions, encouraging active participation in class discussions, encouraging class discussions, encouraging the expression of opinion on course materials, encouraging sharing of knowledge, respecting individual opinion, encouraging individual class interac-

tions, goodness in answering questions, contacting lecturers outside class and provision of conducive learning atmosphere. In addition, Lecturers spend time encouraging students to ask questions, encourages active participation of students in class discussions. Lecturers encourage students' academic interaction in classes by sharing knowledge and personal experience that are relevant to the course curriculum.

3.1.2 Lecturer Regulating Student Learning

The finding of the result shown in Table 2.0 in Appendix 1.0 demonstrated that all the items measuring the construct had mean values of more than 2.40. This generally indicated that the students agreed that lecturers regulate their learning through clarity of course syllabus, practical application, rating their learning, acknowledging achievements, being knowledgeable on course content, helping them learn, grading their assignment and exams, providing feedback on progress, giving them valuable feedback, covering of course content, and providing result outcome for students.

3.1.3 Lecturers Delivery of Course Information

The descriptive analysis of lecturer/delivery of course information to the student in the class was shown in Table 3.0 in appendix 1.0. The finding revealed that all the items had mean values of more than 2.67. This generally indicated that the lecturers were motivated to deliver effective course information to the student in terms of content appropriate and related material for current and future needs, effective presentation strategies for student understanding of the course content, periodic review of course content for effective performance. In addition, the result revealed that the lecturers are knowledgeable in their areas of specialization with adequate knowledge of the course they handle. Those course syllabi provided by lecturers were simple, clear and provided a direction for students not only on the theory aspect but also on practical application. This implies that lecturers try to carry everybody along when teaching a course and acknowledge students' performance and achievements. By issuing feedback on the assessment. This is valuable to students as they can learn from their mistakes and take corrective measures.

3.2 Factor Analysis

Factor analysis: this technique takes a large set of variables and looks for a way the data may be reduced or summarized using a smaller set of factors or components. It does this by looking for groups among the intercorrelations of a set of variables. The sample size guideline by Coakes and Steed, (2003); Hair et al., (2010) indicates that a minimum of five subjects per variable is needed for factor analysis. A sample size of more than 350 requires a factor loading of 0.30 to assess statistical significance (Hair et al., 2010). Hence, the minimum requirement for factor analysis was fulfilled. The original SETERS questionnaire administered contains 34 items with three constructs namely Instructor's delivery of course information; Lecturer's role in facilitating Instructor/Student Interaction and Lecturer's role in regulating student's learning (Toland and Ayala, 2005). The collected data were subjected to exploratory factor analysis, intercorrelation of the items using principal component analysis.

The result of factor analysis is shown in appendix 1 revealed that the correlation matrix showed the presence of many coefficients of .3 and above. The Kaiser-Meyer-Oklin value was 0.87 this value exceeded the recommended value of 0.6 (Kaiser 1970, 1974) and Bartlett's Test of Sphericity (Bartlett 1954) reached statistical significance, supporting the factorability of the correlation

matrix. Also, implies that the sample size is adequately meritorious for factor analysis to be conducted. Principal components analysis revealed the presence of five components with eigenvalues exceeding 1, explaining 36.8%, 8.9%, 8.2% 6.4% and 5.2% of the variance respectively. An inspection of the scree plot revealed a sharp and clear break after the third component. The first component Instructor's Delivery of Course Information was defined by six items relating to Student understanding of course content, Interest on the subject, course materials, Lecturer's presentation of course content and material that would motivate student learning. Thus, this dimension was named Lecturer's clear delivery of course information factor. The second component Instructor's Role in Regulating Student's Learning split into two factors, the first was dominated by four items relating to the provision of a progress report, valuable feedback to the students assignment and quiz/ test examination which the dimension was named Provision of student progress report to student factor. The second dimension emerged with three items related to the clear outline of the syllabus, course content, practical application, hands-on desk for the student was named as a Clear course outline and direction factor. The third component relating to Teacher's Role in Facilitating Instructor/Student Interaction also give rise to two dimensions named interaction and participation factor that consisted of three items pertaining to active participation in the class discussions and encouragement of Student to ask questions as well as the second factor consists of three items related to expression of opinion by student on course materials, share Knowledge, and personal experiences by both the lecturer's and students were named share knowledge and experience factor.

3.3 Construct Cronbach's Alpha Analysis

The construct reliability must be assessed before examining its validity (Hair et al., 2010). To this end, the reliability of all the items was examined through the Cronbach's Alpha, factor loadings and the index of composite reliability. Although there is a lot of debate concerning the best method to estimate reliability, coefficient alpha remains the commonly used method even though it may underestimate reliability (Hair et al., 2010). The different methods of assessing reliability produced similar results. The values of Cronbach's alpha and composite reliability are shown in Table 5. In appendix 1.0.

Reliability analysis test is best carried out using the Cronbach's alpha statistic. It measures the internal consistency of a research instrument. The extent to which the results are consistent with time can be best measured using reliability test also it could act as the best representation of the population under study (Joppe, 2000). Cronbach's alpha is a consistency test of whether all items within the instrument measure the same thing. It measures the reliability of the questionnaire items. Its values vary between 0 and 1. When a negative value occurs, it implies that a scale in which some items measure the opposite of what other items measure. The closer the alpha is to 1.00, the greater the internal consistency of items in the research instrument. At a more conceptual level, the coefficient of Cronbach's alpha may be considered as the coefficient between a sincere response and all other sincere responses of the same item that are drawn randomly from the same population of interest.

$$\alpha = \frac{kr}{1+(k-1)r}$$

The reliability test for each dimension emerged after factor analysis was conducted. Flynn, Schroeder, and Sakakibara (1994) argued that a Cronbach's alpha of 0.6 and above was considered effective reliability for judging a scale. The summary of the result reliability analysis of the major variables of SETERS demonstrated in appendix 1 revealed that the Cronbach's alpha ranges from 0.692 to 0.826 for the five (5) dimensions of SETERS that emerged after factor analysis. It implied

that the adapted questionnaire instruments for this study were reliable. Hence the instrument had excellent reliability concerning the internal consistency.

3.4 Construct Composite Reliability

Fornell and Larcker, (1981) argued that composite reliability is more robust than Cronbach's alpha. From the table, it is obvious that each of the indexes of construct reliability (composite reliability) is greater than the threshold of 0.7 (Fornell & Larcker, 1981). The composite reliability values range between 0.746 and 0.907. This result means that the constructs have internal consistency and that all the measures consistently represent the same latent construct. Although composite reliability is stronger than the Cronbach's alpha, in this study, the latter was also assessed to complement the former. Flynn, Schroeder, and Sakakibara (1994) argued that a Cronbach's alpha of 0.6 and above was considered effective reliability for judging a scale. The agreed lower limit for Cronbach's alpha may decrease to 0.60 in exploratory research (Hair et al., 2010). Again, from Table 6 in the appendix, the results presented the factor loadings for all the items ranging from 0.625 to 0.800, confirming that the indicators are strongly related to their various constructs. Hence, it indicates good construct validity (Hair et al., 2010).

Where: CR = composite reliability; Li = standardized factor loading; ei = error variance

3.5 Convergent Validity

In an attempt to establish construct validity, convergent validity was examined using Average Variance Extracted (AVE) as recommended by Hair et al. (2010). The AVE shows how indicators of construct converged and how they share common variance. In other words, the indicators should converge and share a high proportion of variance on a common point, the latent construct. AVE is computed as the mean of variance extracted for the items loading on a construct. This computation can be done using the formula below with the standardized loadings:

Where: AVE =average variance extracted; Li = standardized factor loading; i = number of items

The Average Variance Explain results in Table 6 of the appendix indicated the AVE of the three dimensions of SETERS Provision of progress report 0.561, Knowledge & Experience 0.582 and Lecturer/Student interaction &Participation 0.618 are higher than the required value of 0.5 (Fornell & Larker, 1981). The results indicate that each dimension has the capability to explain more than half of the variance to its measuring items on average. Thus convergent validity is confirmed. The other two dimensions Clear course outline & Direction 0.495 and Lecturer clear delivery of course information 0.457are also to too far away from the threshold limit of 0.5.

3.6 Validity

Discriminant validity assesses the extent to which a construct is truly different from other constructs (Hair et al., 2010). Consequently, a high level of discriminant validity suggests that a latent construct is unique and captures some phenomena that other constructs do not. Although, there are several ways to compute discriminant validity, a more rigorous method is to compare the AVE values for any two constructs and with the square of the correlation estimate between these two constructs. The AVE should be greater than the square correlation estimate (Hair et al., 2010). Another way of

doing this test is to compare the square-root of AVE for a given construct with the absolute correlations of that construct and all other constructs (Tang, Luo, & Xiao, 2011). For either, however, the AVE must be greater than the construct correlation to establish discriminant validity (Fornell & Larcker, 1981). Table 7 in the appendix indicates all the square roots of AVE ranging between 0.676 and 0.786 are greater than the values of the constructs in the corresponding matrices. This indicates that each construct shares more variance with its items than with other constructs, and supports discriminant validity.

3.7 Nomological Validity

To assess the Nomological validity correlation matrix was used based on the suggestion of Hair et al. (2010). They stress that Nomological validity is supported to the extent that a construct relates to other constructs in a theoretically consistent manner. This enables us to examine the extent of relationships among the constructs under investigation based on the established literature. Previous studies have established some form of relationships among the constructs in this study, namely, Clear delivery, of course, Provision of Progress Report, Interaction & Participation, Knowledge & Experience, and clear course outline and direction. Hence, to confirm this relationship, correlation analysis was run. From Table 8, it could be seen that the constructs have a significant positive correlation, and, therefore, Nomological validity is supported.

3.8 Bivariate Relationship Between SETERS Dimensions

Correlation analysis was conducted during this study to explore the strength and direction of the linear relationship between five variables of SETERS. Specifically, this analysis determined 1) the inter-relationship between the variable of SETERS. In determining the strength to the relationship, Pallant (2001) noted that a correlation of 0 signifies no relationship, a correlation of 1.0 signifies a perfect positive correlation and a value of -1.0 signifies a perfect negative correlation. In interpreting the values between 0 and 1, the following guideline was suggested by Cohen (1998): r=0.10 to 0.29 or r=-0.10 to be -0.29 small; r= 0.3 to 0.49 or r= -0.30 to -0.49 medium; r=0.50 to 1.0 or r=-0.50 to -1.0 large. The result of the Pearson correlation is presented in Table 8. The results of the inter-correlation between variables. The correlation analysis was subjected to a one-tailed test of statistical significance at two different levels; significant ($p<0.01$) and significant ($p<0.05$). Overall, the results indicate that all the variables were significant at $p<0.01$. The strongest positive correlation was the relationship between Clear course delivery by the lecturers and their knowledge and experience in teaching ($r=0.496$@ $p<0.01$) with a high level of knowledge and experience associated with a clear level of delivery of course information by the lecturers. The next strongest positive correlation was between clear course outline and clear course delivery ($r=0.488$, $p<0.01$), followed by the provision of progress report to student and course delivery ($r=0.470$, $p<0.01$). The participation and interaction of student and lecturers in the class ($r=.383$, $p<0.01$).

4.0 Conclusion

The analysis of data collected on the Students' Evaluation of Teaching Effectiveness Rating Scale (SETERS) among undergraduate students of the Business administration department, ABU Zaria demonstrated the following:

The correlation analysis of the study showed that there is the strongest positive correlation between Clear course delivery by the lecturers and their knowledge and experience in teaching

with a high level of knowledge and experience associated with a clear level of delivery of course information by the lecturers. The next strongest positive correlation was between clear course outline and clear course delivery followed by the provision of progress report to student and course delivery. The participation and interaction of student and lecturers in the class. The significant positive correlation between the variables revealed that Nomological validity is supported to the extent that the variable relates to other variables in a consistent manner

The reliability analysis using the Cronbach's Alpha and composite reliability also revealed that the high values of the Cronbach's Alpha coefficient demonstrated that SETERS is highly reliable, stable and measuring teaching effectiveness. In other words, this scale is suitable for use in the Business administration department, ABU Zaria. This is in agreement with the findings of Toland & De Ayala (2005).

The confirmatory factor analysis of all the original SETERS Scale containing 34 items with three constructs namely Instructor's delivery of course information; Lecturer's role in facilitating Instructor/Student Interaction and Lecturer's role in regulating student's learning conducted using principal component analysis, varimax rotation, correlation matrix and the scree plot, revealed that the correlation matrix showed the presence of many coefficients of .3 and above. The Kaiser-Meyer-Oklin value exceeded the recommended value of 0.6 (Kaiser 1970, 1974) and Bartlett's Test of Sphericity (Bartlett 1954) reached statistical significance, supporting the factorability of the correlation matrix. Also, implies that the sample size is adequately meritorious for factor analysis to be conducted. Principal components analysis revealed the presence of five components with eigenvalues exceeding 1, of the variance respectively. An inspection of the scree plot revealed a sharp and clear break after the third component. These results appeared to be in agreement with previous studies conducted by Toland & Ayala (2005). The assessment of the construct discriminant validity by comparing the square-root of AVE for a given construct with the absolute correlations of that construct and all other constructs indicated that each construct shares more variance with its items than with other constructs, and supports discriminant validity of the SETTER Construct.

Therefore, based on the findings of the analysis (Bivariate correlation, Confirmatory factor analysis, Cronbach's Alpha and Composite reliability analysis, Convergent, discriminant, and Nomological validity analysis) of the construct SETTERS scale in its present form revealed that it is suitable for use in Ahmadu Bello University, Zaria, Nigeria.

Acknowledgment

We would like to express our deepest appreciation to the Head of Department Business Administration and Dean Administration, Ahmadu Bello University Zaria, who provided the support and resources to complete the research and publication of the study's findings in an eminent University based journal.

References

Abdulkadir, H. S. (2006). Strategies for improving the efficiency of teachers in Jigawa State. Unpublished masters' thesis, University of Nigeria, Nsukka, Nigeria.

Adomi, E. E. & Mordi, C. (2003). Publication in foreign journals and promotion of academics i n Nigeria. Learned Publishing, 16(4), 259-63

Areola, R. A. (1995). Developing a comprehensive faculty evaluation system: A handbook for college faculty and administrators on designing and operating a comprehensive faculty evaluation system. Bolton: Anker Publishing

Asika, N. (1991). Research methodology in the behavioral sciences. Lagos: Longman Nigeria Plc

Bartlett, M. (1954). A note on multiplying factors for various chi square approximations. Journal of the Royal Statistical Society, 16(series B), 296-298.

Cavana, R. Y., Dalahaye, B., & Sekaran, U. (2001). Applied Research: Qualitative and quantitative methods. Australia: John Wiley and sons

Centra, J. A. (1993). Reflective faculty evaluation: Enhancing teaching and determining faculty effectiveness. San Francisco, CA: Jossey-Bass.

Coakes, S. J., & Steed, L. G. (2003). SPSS: Analysis without anguish. Sydney: John Wiley & Sons

D'Apollonia, S. & Abrami, P. C. (1997). Navigating student ratings of instruction. American Psychologist, 52(11), 1198-1208

Flynn, B., Schroeder, R., & Sakakibara, S. (1994). A framework for quality management research and an associated measurement instrument. Journal of Operations Management, 11, 339-366.

Fornell & Larcker, (1981) Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. Journal of Marketing Research 18(1), 39-50

Hair, J., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). Multivariate data analysis (7th ed.). Upper Saddle River, New Jersey: Pearson Education International

Henson, R. K. (2001). Teacher efficacy scale: Substantive implications and measurement dilemmas. Keynote Address, Annual Meeting of the Educational Research Exchange, Texas A & M University, College Station, Texas. Retrieved December 2009, from http://www.css.edu/users/dswenson/web/scaledevstat.html

Isreal, G. D. (2013). Sampling the Evidence of Extension Program Impact. Programs Evaluation and Organizational Size, University of Florida. IEAS Extension Development, IFAS, University of Florida. Journal of Personality and Social Psychology, 51 (6), 1173-1182

Kaiser, H. (1970). A second generation little jiffy. Psychometrika, 35, 401-415.

Kaiser, H. (1974). An index of factorial simplicity. Psychometrika, 39(31-36).

Krejcie, R. V., & Morgan, D. W. (1970). Determining sample size for research activities. Educational and Psychological Measurement, 30, 601-610

Marsh, H. W. & Hocevar, D. (1994). The factorial invariance of student evaluations of college teaching. American Educational Research Journal, 21, 341-366.

Marsh, H. W. & Roche, L. A. (1997). Making students' evaluation teaching effectiveness effective: The critical issues of validity, bias, and utility. American Psychologist, 52, 1187-1197

Marsh, H. W. (1977). The validity of students' evaluations: Classroom evaluation of instructors independently nominated as best and worst teachers by graduating seniors. American Educational Journal, 14, 441- 447

Marsh, H. W. (1983). Multidimensional ratings of teaching effectiveness by students from different academic settings and their relation to student/course/instructor characteristics. Journal of Educational Psychology, 75(1), 150-166.

Marsh, H. W. (1984). Students' evaluations of university teaching: Dimensionality, reliability, validity, potential biases, and utility. Journal of Educational Psychology, 76(5), 707-754.

Marsh, H. W. (1987). Students' evaluations of university teaching: Research findings, methodological issues, and directions for future research. International Journal of Educational Research, 11(2), 253-388

Pallant (2001) Pallant, J. (2001). SPSS Survival Manual: A Step by step guide to data analysis using SPSS for windows (3rd ed.). England: McGraw Hill Open University Press. Research, 18(1), 39-50

Salkind, N.J. (2003). Exploring Research. ISBN-10: 0130983527 USA

Sanders, W. L. & Rivers, J. C. (1996). Cumulative and residual effects of teachers on future student academic achievement. Knoxville: University of Tennessee Value-Added Research and Assessment Centre. Research Progress Research Report 1-11 Retrieved December 2016 from https://www.beteronderwijsnederland.nl/files/cumulative%20and%20residual%20effects%20of%20teachers.pdf

Sekaran and Bougie (2010). Research methods for business (4th ed.). New York: John Wiley & Sons, Inc.

Tschannen-Moran, M., Woolfolk Hoy, A., & Hoy, W. K. (1998). Teacher efficacy: Its meaning and measure. Review of Educational Research, 68(2), 202-248.

Zikmund, G. W. (2000). Business research methods (7 ed.). Published by the Dryden Press, 2000 ISBN 10: 0030258170 / ISBN 13: 9780030258176.

e-ISSN: 2682-7840

UNIVERSITI TEKNOLOGI MARA
Cawangan Kedah
Kampus Sungai Petani