

Development and Validation of Lecturer Teaching Assessment Scale for Nigerian Universities

Amuche Christian Igomu Amuche^{1*} Mahmood Hamid Umar²

1.Faculty of Education, Taraba State University, Jalingo, PMB 1167, Jalingo Taraba State Nigeria

2.Faculty of Agriculture, Taraba State University, Jalingo, PMB 1167, Jalingo Taraba State Nigeria

This study is part of a research funded through the Tertiary Education Trust Fund (TETFund) research grant, Abuja Nigeria. The authors are therefore grateful to the TETFund office Abuja and Taraba State University Management for facilitating the grant.

Abstract

The present study developed and examined the factorial validity of scores on the Lecturer Teaching Assessment Scale across lecturers of universities in North-east Nigeria. This study presents a validation report of the Lecturer's Teaching Assessment Scale (LECTAS) developed for the assessment of lecturer's teaching effectiveness in Universities. It also examined the factor structure of the LECTAS, its logical validity, and internal consistency reliability coefficients. The study adopted the Developmental and survey research designs. A total of 2600 students that completed the developed LECTAS constituted the sample for the study. The 34-item LECTAS was used to collect data for the study. Collected data were subjected to reliability and factor analyses. Results showed that the scale is valid and reliable in assessing lecturers teaching output. The LECTAS was adjudged to possess construct validity. The results revealed that the LECTAS was reliable (Cronbach Alpha reliability coefficient of 0.932). Thus, the LECTAS possessed adequate psychometric qualities that make it suitable to assess lecturers' teaching effectiveness in Nigerian universities.

Keywords: Development, Factor Analysis, Teaching Assessment, Validation, Scaling

DOI: 10.7176/JEP/10-5-05

1. Introduction

The call for teaching accountability in universities initiated teaching effectiveness research and its scales development. Attention in many institutions of higher learning has been diverted recently to the improvement of teaching performance as another way besides academic research to promote the higher institutions. The diversity of attention is a response to external calls for accountability in teaching as a result of the under-estimation of the significance of the teaching process compared to research activities. As research on teaching effectiveness has increased, so has the number of different measures of teaching effectiveness.

Evaluating teachers can be approached from three different but related angles: measurement of inputs, processes and outputs (Goe, Bell & Little, 2008). This study considered the process approach because it bothers on the interaction that occurs in the classroom between teachers and students. Students are in a unique position to comment on their satisfaction with a course and the impact of the instruction on their own learning. However, they are not subject matter experts, and therefore are not in a position to make judgments about the currency or accuracy of course content. In addition, research has shown that ratings by students are sometimes influenced by their level of motivation for taking the course, attitude toward the course or the instructor, and needs or contextual variables (e.g., whether the course is required).

Clarifying the way teacher effectiveness is defined is important for two main reasons: first, what is measured is a reflection of what is valued and as a corollary, what is measured is valued. This revolves around process and output of teaching. Thus, when policy discussions revolve around classroom practices, particular approaches to teaching becomes the focus. To give further enlightenment to the concern of this study, it is but necessary to define teacher effectiveness. Hunt as cited by Calagaus (2015) posit that the term teacher effectiveness is used broadly, to mean the collection of characteristics, competencies, and behaviors of teachers at all educational levels that enable students to reach desired outcomes. Given that teacher characteristics are relatively stable traits that are related to and influence the way teachers practice their profession, and because effective teachers are a must in Nigerian universities, evaluating teachers to help them be effective is necessary. Evaluating the performance of teachers in universities is equivalent to evaluating the learning of students.

An extremely useful and increasingly common approach to evaluating teaching effectiveness is to measure students' knowledge or skills at the beginning of a course or unit of the course and again after some body of material has been covered in class. Instructors can then observe and quantify the amount of improvement and draw inferences about the instructor's effectiveness in helping students learn the subject matter. For measures of student learning to be considered valid and reliable, however, considerable effort is required to develop pre- and post-learning tests that actually measure the kind of learning desired. In addition, changes observed in students' learning and performance cannot be attributed solely to the effectiveness of an individual instructor. Many factors, including students' ability and motivation to learn and even their health status when taking either

examination, can also influence the outcomes.

As a major component of the education, teacher evaluation is at the core. Gathering teacher evaluation is undeniably helpful in identifying exemplary teacher and teaching in universities (Feldman, 2007). It is also helpful for on-going self-monitoring of one's teaching, evaluating one's professional development needs, and preparing a case for promotion or tenure, providing information for students to use in the selection of courses and instructors, and providing an outcome for research (Casey, Gentile, & Bigger, 1997). The feedback from such evaluation helps to improve teaching and learning.

There are generally two fundamental aspects of teacher evaluation which include improvement function which relates to formative nature and accountability function which relates to summative nature (Tucker & Stronge, 2005). Classroom observations, principal evaluations, analysis of classroom artifacts, portfolios, self-report of teacher practice, and value-added models are methods of evaluating teacher effectiveness (Goe, Bell, & Little, 2008). In addition, one of the key components of teacher evaluation can be thought as student evaluation of teachers as a recipient of service provided and affected. Implementing student questionnaires which include different dimensions of teaching to evaluate teaching effectiveness and quality is a fairly common procedure and it aims the improvement of teaching quality (Dresel & Rindermann, 2011).

Given that students are in a unique position to comment on their satisfaction with a course and the impact of the instruction on their own learning, they are not subject matter experts, and therefore are not in a position to make judgments about the currency or accuracy of course content. In addition, research has shown that ratings by students are sometimes influenced by their level of motivation for taking the course, attitude toward the course or the instructor, and needs or contextual variables (e.g., whether the course is required). Our study had two objectives, both of which addressed validity considerations. The first was to support the external validity of a questionnaire used for student evaluations of lecturers. The second objective was to provide evidence for the validity of the questionnaire.

Freeman (1979) found common dimensions which students often use as yardstick in perceiving and rating their teachers to include the instructors' subject matter competence, ability to relate materials and quality, fairness of feedback, evaluation procedure and the degree of instructor-student rapport. He further found that some students place relatively greater importance in associating with a teacher because of the teacher's personality characteristic, which is an important value in the whole process of teacher evaluation, and that it affects students' academic achievement. How to define operationally and promote instructional effectiveness is still a problem for educators and researchers. There is also the problem of instrumentation (nature of instrument, reliability and validity). However, instrument validation and reliability could be guided by some basic criteria: first the instrument should provide a good coverage of instructional effectiveness of teachers' and concerns of the students; second, the instrument's structure should be consistent with general psychometric principles in that it should possess several internally consistent, mutually exclusive scales; thirdly, individual scale items should be sensitive to different levels of concerns of students in the study; and fourthly, the instrument should be relatively economical to administer, answer, score and analyse. In order to operationalize these criteria, it is necessary to employ both intuitive-rational and factor analytic approaches to scale development. This involves identification of salient dimensions, writing tentative scale items, and conducting field-testing and applying factor analysis to group items into scales (Harse & Goldberg, 1967). This was essentially followed in this study. Hence, in this article, the researchers examined the psychometric properties of the developed lecturers' teaching assessment scale

2. Statement of the problem

University faculty members are evaluated in many ways as a means of determining whether they should be promoted or rewarded, and to potentially improve their performance. A convenient measure of the research productivity of faculty members that is frequently employed is the number and quality of technical papers and reports published. A similar metric for teaching effectiveness is not so readily available. Attempts to establish metrics for teaching effectiveness have assumed many forms but normally have focussed on student responses to written questions. Despite some strong opposition to incorporating such student ratings in faculty evaluation, they are widely used. A number of student questionnaires have been developed with the difference arising primarily because of the intent of the questionnaire.

A key problem is that current measures for assessing academics for promotion in most Nigerian Universities are not often linked to their capacity to teach effectively. Existing Federal University policies for measuring teacher effectiveness either rely almost exclusively on perception by heads of departments, or focus on teachers' course-taking records and on paper-and-pencil tests of basic academic skills and subject matter knowledge. Also, criteria for assessing academics for promotion in most Nigerian universities include qualifications, publications and community service at the local, national or international levels, all these, research as exemplified by research, are poor predictors of teaching effectiveness (Faleye & Awopeju, 2012). It is evident that such evaluation practices do not provide feedback for teaching improvement.

Even though, there is still a dearth of comparable validated instruments for assessing the effectiveness of lecturers' teaching using Nigerian samples, research and theory have shown that teaching assessment, as measured by students' rating of teaching, is multidimensional in nature. Arguably, empirically derived factor structure of lecturer teaching effectiveness scale seems to have varied agreement. Variations exist in the factor structure of exploratory as well as confirmatory factor analyses of Teaching Effectiveness Scale across researchers as approaches of preparing teachers in countries differ. Also, student evaluation of teaching most often comes in the form of questionnaire that asks students to rate teachers on a Likert-type scale. Students have the ability to assess various aspects of teaching from course content to specific teaching practices and behaviours (Little, Goe & Bell, 2009). This may be because students have the most contact with their lecturers and are the most direct consumers of their services. Therefore, there is a need to develop instrument in Nigeria to measure teaching effectiveness of lecturers in tertiary institutions. Therefore, the thrust of this study is to develop and validate an instrument capable of measuring lecturers' teaching effectiveness in Nigerian universities.

3. Objective of the Study

This study aimed to prepare a scale that will assess lecturers' effectiveness in universities. Specifically, it sought to answer the following questions:

- i. How can an instrument which will measure teacher effectiveness in universities be developed?
- ii. How can the scale be validated?
- iii. what is the reliability of the developed tool?

4. Hypothesis

The Lecturer Teaching Effectiveness Scale in universities (LECTAS) is a psychometrically sound scale that measures the multidimensional aspects of teaching effectiveness in universities.

5. Literature Review

The core of education is teaching and learning, and the teaching-learning connection works best when we have effective teachers working with every student every day. Scholars have argued there is little consensus within the education community about how to define effective teaching (Popp, Grant, & Stronge, 2011). The clearest and potentially most useful example identified in my review of the literature comes from the Center for High Impact Philanthropy (2010). A quality teacher is one who has a positive effect on student learning and development through a combination of content mastery, command of a broad set of pedagogic skills, and communications/interpersonal skills. Quality teachers are life-long learners in their subject areas, teach with commitment, and are reflective upon their teaching practice. They transfer knowledge of their subject matter and the learning process through good communication, diagnostic skills, understanding of different learning styles and cultural influences, knowledge about child development, and the ability to marshal a broad array of techniques to meet student needs. They set high expectations and support students in achieving them. Students may assess various aspects of teaching, from course content to specific teaching practices and behaviours; although students have the most contact with their teachers and are the direct consumers of teachers' performance delivery in teaching; their ratings may be susceptible to bias because they lack adequate knowledge about the full context of teaching. However, studies validate that student rating to be considered as part of teacher performance evaluation method (Spooren, Brockx, & Mortelmans, 2013).

Student evaluations have become routine at most colleges and universities. Evidence from many studies indicates that most universities and colleges throughout the world use student ratings of instruction as part of their evaluation of teaching effectiveness (Hobson & Talbot, 2001). Student evaluations of teaching effectiveness are commonly used to provide: (1) formative feedback to faculty for improving teaching, course content and structure; (2) a summary measure of teaching effectiveness for promotion and tenure decisions; (3) information to students for the selection of courses and teachers (Marsh & Roche, 1993). Research on student evaluations of teaching effectiveness often examines issues like the development and validity of an evaluation instrument (Marsh, 1987)

The process of student evaluation of teaching is used extensively in a growing number of tertiary institutions around the world (Taheri, Ryasi, Afshar, & Mofatteh, (2014). This form of evaluation have several advantages which includes cost and time effective, requires little or no preparation, data can be gathered cautiously and most importantly monitor changes over time (Worrell & Kuterbach, 2001). Research literature has been almost exclusively concerned with the construction of ratings instruments and the reliability and validity of student ratings. However, through a better understanding of students' needs and behavioural intentions, the results of this study can aid in creating evaluation systems that truly respond to the needs of those who evaluate teaching performance.

This study has been modelled on the output of a teacher production function (Harris, 2001). The production function expresses teaching effectiveness as an output determined by three primary input characteristics: student,

course and teacher. The model is expressed as follows:

$$E = f(S, C, T) \dots \dots \dots \text{equation (1)}$$

Where, E = Evaluation measure of teacher effectiveness, S = vector of students' characteristics, C = vector of course characteristics and, T = vector of teacher characteristics. Mathematically, the marginal effect of any one input of teacher evaluation is simply the partial derivative of equation 1 with respect to that input. Investigation into students' evaluation of teaching effectiveness has been the subject of research in Nigeria and a number of instruments have well-defined factor structures and provides measures of distinct components of teaching effectiveness scales (SETS) but are often criticized as lacking validity (Awofala, 2011).

In Africa, particularly in Nigeria, Okpala (1999) developed the first indigenous model for evaluating teaching effectiveness. Okpala opines that evaluation of teaching effectiveness is an integral part of teaching-learning process and that each stage of the teaching-learning process is subject to evaluation, and that the evaluation data from each Stage could be used to influence decision-making at other stages. Okpala's model suggests eight components of teaching-learning process, which were adopted in this study as the eight dimension of teaching effectiveness. Research evidence currently justifies observing and judging the following teacher behaviours; clarity of presentations and explanation, enthusiasm, variety in use of instructional materials and techniques, task orientation and "business like" behaviour in provision of ample learning opportunities. Each characteristic can be precisely defined and can be reliably measured through training, and using observational judgmental method most of the time. This method involves the collection of students' evaluation of their teachers. The data collected from students are of several types but include information to corroborate outside observers' rating of their teacher behaviours. Involvement of students in the observational-judgmental model may form an important contribution in the teacher evaluation process.

Muhammed and Sally (2015) developed and validated a Self-assessment Instrument for Teacher Evaluation (SITE II) based on five National Professional Standards for Teachers developed by the Ministry of Education, Pakistan: subject matter knowledge, instructional planning and strategies, assessment, learning environment, and effective communication. The data were collected from 279 English and mathematics teachers of grade 10 in 40 public boys' and girls' high schools in district Okara who self-evaluated their performance on five teacher performance components. The overall reliability of the questionnaire was found high ($\alpha=.94$). The SITE II factor structure was discovered through exploratory factor analysis. Confirmatory factor analyses provided evidence of construct validity of the questionnaire. Significant positive relationship was found between teachers' scores on self-evaluation questionnaire and their students' achievement ($n= 7245$) in English as well as in mathematics. The findings suggest that the questionnaire is valid and efficient tool for measuring components of teacher self-assessment. The self-assessment approach is prone to bias hence, the present study considered student assessment of teaching effectiveness due to its efficacy. Also, Estu and Franciscus.(2016) developed an observation instrument for assessing English teaching at vocational secondary schools (VSSs), examined the validity of the developed instrument, examined the inter-rater reliability, and assessed the effectiveness of the developed instrument. The subjects of the three-phase field-tests were English teachers and students of four VSSs. Observation sheets, questionnaires, and interviewes were employed to collect the data. The instruments were field-tested in three phases. Expert judgment, Kappa Coefficient, and Kane's (2006) argument-based validation and Bachman & Palmer's intepretive and use argument were used to determine the instrument validity and reliability. The field-test which involved stakeholders proved that the observation instrument could be implemented effectively to assess the teaching of English at VSSs.

In Nigeria, education literature shows efforts at developing and validating scales for measuring teaching effectiveness in universities have not been sufficiently harnessed (Awofala, 2011). He argued that the questions about the validity, reliability, utility, generalizability, interpretability as well as acceptability of students rating of teachers especially when such a result of such evaluation exercises are to serve some summative purposes have been a source of concern.

6. Research Methodology

The study adopted Research and Development design (R & D) as well as the survey designs. Research and Development design (R & D) according to Gall, Gall and Borg (2007), is a field of trial of Research and Development of an instrument to assess teaching effectiveness. The population of this study comprised all undergraduate students of Universities in North-eastern Nigeria. The universities have a total of 73 faculties with a population of 106,000 students (2017 estimate). A sample of 2600 would be drawn using stratified random sampling technique. The selected sample was made of 1320 males and 1280 female students. The LECTAS was completed (by every student) for each lecturer that participated in the teaching of each of their courses of study. This means that each student completed the form once irrespective of the number of courses taken by each student as well as the number of lecturers that taught each of such courses. Thus, for the purpose of validation, only the 2600 of responses to the LECTAS was for the study. The students were undergraduate students of the universities in the study area and the sample is inclusive of a mix of students across the departments in

universities. The students are those in second to the fifth year of their programme, depending on duration of programme. The items on LECTAS are positively stated and there is no need for reversal of scores for any item. The response format for the LECTAS is the Likert type will comprise; Always, Sometimes, Rarely and Never; and the scoring relating to each of the four response options are 4, 3, 2 and 1 respectively. The dimensions of instructional effectiveness used for this study were: Knowledge of subject matter, classroom communication, classroom management, motivation and reinforcement, use of variety of teaching methods, effective use of instructional materials, teacher's technique of evaluating students, and students/teacher relation. Data collected were analysed using factor and reliability analyses.

6.1 Search for content domain

Review of literature published between the years 2013 to 2018 was carried out. Such undertaking provided the framework that served as guide in writing the items in the initial scale on teacher effectiveness. Reviewed literature was limited to teaching effectiveness in Nigerian universities.

6.2 Item writing

The review of literature steered to the writing of items in the initial scale which were based on the identified seven dimensions of teacher effectiveness, namely: Personality, Subject- matter proficiency, Interactive Proficiency, Professional proficiency, Teaching style, Classroom management, Evaluation. A total of 78 items were written with the following breakdown: personality 15 items, subject matter proficiency 10 items, Interactive proficiency 15 items, professional proficiency 10 items, teaching style 12 items, classroom management style 10 items and, evaluation 6 items.

6.3 Administration of the initial scale

The administration of the initial scale was done to evaluate the psychometric properties of the scale. A total of 2600 students chosen via systematic random sampling were originally included in this study. These students were second year students and above. Only second year students and above were chosen because of the assumption that these students were already very much aware of how learning is best facilitated especially by lecturers in universities. However, out of the 2600 students originally included in the study, only 2570 completely responded to the scale and therefore served as participants.

6.4 Descriptive statistics

With the responses of the participants in the administration of the initial scale as bases, the means and standard deviations of the items included in the initial scale were computed. This was done because the means and standard deviations could serve as guide in deciding what items to include in the final scale.

6.5 Evaluation of reliability

The reliability of the scale was evaluated using Crombach's Coefficient Alpha formula. This method examined the internal consistency of the items in the scale.

7. Results and discussion

The purpose of this study was to develop and validate a scale that measures teacher effectiveness in universities from the perspective of students. Following this aim, the study started with a review of literature on teacher effectiveness. The ideas from the reviewed literature served as bases in writing the 78 items included in the initial scale and which were distributed as follows: personality - 25 items, subject matter expertise - 10 items, relational competence with students - 18 items, professional competence - 10 items, teaching style - 32 items, and classroom management style - 12 items.

7.1 Descriptive statistics

The 78 item initial scale was administered to 2600 universities students chosen via systematic random sampling but only 2570 completely responded and therefore considered as participants. With the responses of the participants in the administration of the initial scale as bases, the means and standard deviations of the items included in the initial scale were computed. Descriptive statistics come first whenever data is examined and the most common of these are the means and standard deviations.

The means and standard deviations of items can provide clues about which items will be useful and which ones will not be. If the variance of an item is low, this mean that there is little variability on the item and it may not be useful. While it is not common to examine item level descriptive statistics in most research applications, in creating and validating tests it is a crucial first step (Kline, 2005). The descriptive statistics of the items in the final scale is presented in Table 1 with their standard deviations.

Table1: Descriptive statistics of the items in the final scale is presented with their standard deviations.

| S/N | CRITERIA/STATEMENTS | N | Min | Max | Mean | SD |
|-----|--|------|-----|-----|------|------|
| 1 | Lecturer's Adherence to time (My lecturer arrives in the classroom on time) | 2570 | 1 | 4 | 3.73 | .565 |
| 2 | Utilization of lecture time. (The lecturer teaches for the number of hours required for teaching the course per week e.g. 3 hours for a 3-unit course; 2 hours for a 2-unit course) | 2570 | 1 | 4 | 3.21 | .659 |
| 3 | Provision of lecture materials. (My lecturer provides lecture notes and/or refers students to available texts and other sources of information) | 2570 | 1 | 4 | 3.63 | .723 |
| 4 | Presentation indicators (She/he provides useful presentation indicators e. g. gives outlines, states objectives, reviews main points) | 2570 | 1 | 4 | 3.21 | .552 |
| 5 | Reception (She/he welcomes and encourages ideas, comments, questions and students' active participation) | 2570 | 1 | 4 | 3.78 | .812 |
| 6 | Mode of presentation. (She/he delivers the lecture without excessive reliance on reading lecture notes.) | 2570 | 1 | 4 | 3.22 | .722 |
| 7 | Communication (She/he speaks fluently without distracting verbal and/or non-verbal mannerisms) | 2570 | 1 | 4 | 3.55 | .343 |
| 8 | Student Respect (She/he commands respect and treats students with respect.) | 2570 | 1 | 4 | 3.67 | .660 |
| 9 | Discipline (She/he maintains discipline in and out of the lecture room.) | 2570 | 1 | 4 | 2.91 | .802 |
| 10 | Attendance Register (She/he ensures that students attend lectures regularly by taking weighted attendance) | 2570 | 1 | 4 | 3.88 | .744 |
| 11 | Evaluation (She/he grades and returns tests and assignments in time) | 2570 | 1 | 4 | 2.87 | .581 |
| 12 | Lecturer's preparedness to teach (She/he arrives in the classroom prepared) | 2570 | 1 | 4 | 2.65 | .878 |
| 13 | Lecturer's Accessibility in Class (She/he is accessible in class) | 2570 | 1 | 4 | 2.55 | 1.08 |
| 14 | Lecturer's availability outside the classroom (She/he is available during office hour) | 2570 | 1 | 4 | 2.65 | .753 |
| 15 | Asking question in Class (The lecturer feel comfortable asking question during class) | 2570 | 1 | 4 | 3.77 | .812 |
| 16 | Explanation of Concepts (She/he explains the concept and material well, and displays mastery of content. | 2570 | 1 | 4 | 3.54 | .662 |
| 17 | Lecturer's Teaching Consistency (She/he is content in his/her teaching) | 2570 | 1 | 4 | 3.34 | .562 |
| 18 | Use of Digital instructional Technology (She/he uses digital instructional technology such as e-book, e-journal, journals articles, YouTube, computer programs, videos, online resources, social media, mobile speakers etc.) | 2570 | 1 | 4 | 2.18 | .557 |
| 19 | Learning Compliance with the Syllabus (I am learning what is in the course description/syllabus) | 2570 | 1 | 4 | 3.21 | .546 |
| 20 | Use of Digital Skills (I am using digital and other online skills in this course) | 2570 | 1 | 4 | 2.11 | .655 |
| 21 | Relevance of Assignment (My lecturer's assignments are relevant to the course) | 2570 | 1 | 4 | 3.34 | .487 |
| 22 | Grading Policies (She/he grading policies are fair and consistent to the syllabus) | 2570 | 1 | 4 | 3.87 | .774 |
| 23 | Relevance of Course content to future career prospect (She/he relates course content and skill to my future career) | 2570 | 1 | 4 | 3.34 | .629 |
| 24 | Lecturer's Openness to multiple perspective (She/he is open to a variety of perspectives on a topic) | 2570 | 1 | 4 | 3.20 | .570 |
| 25 | Developing writing Skills (She/he gives writing assignments in class to help develop my writing skills) | 2570 | 1 | 4 | 3.76 | .542 |
| 26 | Motivation/Inspiration (She/he motivates and inspires me) | 2570 | 1 | 4 | 2.78 | .662 |
| 27 | Due Process (She/he has respect for due process) | 2570 | 1 | 4 | 2.55 | .701 |
| 28 | Mode of Dressing (She/he is lecturer is always well dressed) | 2570 | 1 | 4 | 3.17 | .920 |

| | | | | | | |
|---------------------|---|------|---|---|------|------|
| 29 | Students Friendliness (She/he is students friendly) | 2570 | 1 | 4 | 2.56 | .658 |
| 30 | Leadership skills (Ability of the lecturer to lead students) | 2570 | 1 | 4 | 3.18 | .587 |
| 31 | Assessment (She/he encourages students to give feedback in the course of the lecture) | 2570 | 1 | 4 | 3.67 | .671 |
| 32 | Teaching Environment (She/he creates friendly and supportive lecture environment) | 2570 | 1 | 4 | 3.12 | .562 |
| 33 | Communication (She/he use correct vocabulary and grammar in speaking & writing) | 2570 | 1 | 4 | 3.11 | .712 |
| 34 | Lecturer's presentation capability (She/he presents challenging and thought-provoking ideas to stimulate reflective and critical thinking) | 2570 | 1 | 4 | 3.08 | .578 |
| TOTAL RATING | | | | | | |
| MEAN RATING | | | | | | |

Table 1 shows the mean and Standard deviations of the final instrument. Apart for items 18 and 20 with a mean below the 2.5 benchmark, the other items had acceptable mean values. This could be deduced from the low use of technology by lecturers for teaching.

7.2 Evaluation of validity

Three experts validated the instrument, from Modibbo Adama University of Technology Yola, Taraba State University Jalingo and University of Maiduguri. The suggestions of the validates were used to improve the instrument and consensus of opinion of the validators were sieved to obtain a logical validity index of 0.80 for the LECTAS . Factor analysis was primarily used to evaluate the scale's validity. Specifically, Exploratory Factor Analysis (EFA) was used with Principal Components Analysis (PCA) and Varimax Rotation (with Kaiser Normalization). Principal component analysis, Varimax Rotation and forced choice solution of the seven factors were executed.

The 34 items with seven factors and unequal subscales were selected from correlation matrix generated by item-total correlation analysis were the final 34 items made up the scale. Items selected for factor analysis were those with Crombach's alpha coefficient of 0.42 to 0.80. before factor analysis was conducted, the instrument was subjected to Kaiser-Meyer-Othion (KMO) measure of sampling adequacy and the Bartlett's test of Sphericity. Keiser recommended a coefficient of 0.60 to be considered good enough for factor analysis (Kolawole & Fabunmi, 2014). The seven subscales have values between 0.65 and 0.86. The Bartlett's test produced chi-square values significant at $p = 0.000$.

7.3 Development of final scale

After analysing the descriptive statistics (means and standard deviations of items), reliability (Cornbach's Alpha Formula), and validity (PCA with Varimax Rotation), items to be included were finally determined. Out of the 78 items in the initial scale, 34 were retained. The 34 items retained were the items that loaded on the seven identified dimensions (Personality, Subject- matter proficiency, Interactive Proficiency, Professional proficiency, Teaching style, Classroom management, Evaluation) through PCA. Table 2 presents the psychometric properties of final scale and its subscales.

Table 2: Factor Analysis of Final Items Selected

| Item | R | Level of Significance | Selection | Factor | |
|------|------|-----------------------|-----------|---------|--------|
| | | | | Loading | Number |
| 1 | .667 | .000 | Selected | .402 | 1 |
| 2 | .702 | .000 | Selected | .545 | 3 |
| 3 | .504 | .000 | Selected | .713 | 7 |
| 4 | .492 | .000 | Selected | .403 | 1 |
| 5 | .560 | .003 | Selected | .541 | 3 |
| 6 | .603 | .000 | Selected | .509 | 3 |
| 7 | .511 | .001 | Selected | .401 | 1 |
| 8 | .490 | .000 | Selected | .560 | 4 |
| 9 | .673 | .000 | Selected | .701 | 6 |
| 10 | .602 | .000 | Selected | .549 | 3 |
| 11 | .641 | .000 | Selected | .492 | 2 |
| 12 | .509 | .000 | Selected | .703 | 6 |
| 13 | .587 | .000 | Selected | .633 | 5 |
| 14 | .612 | .000 | Selected | .564 | 4 |
| 15 | .670 | .000 | Selected | .615 | 5 |
| 16 | .508 | .001 | Selected | .551 | 4 |
| 17 | .552 | .001 | Selected | .492 | 2 |
| 18 | .496 | .000 | Selected | .509 | 3 |
| 19 | .506 | .000 | Selected | .676 | 6 |
| 20 | .631 | .003 | Selected | .750 | 7 |
| 21 | .689 | .000 | Selected | .515 | 3 |
| 22 | .504 | .000 | Selected | .500 | 2 |
| 23 | .661 | .000 | Selected | .750 | 7 |
| 24 | .578 | .014 | Selected | .758 | 7 |
| 25 | .545 | .000 | Selected | .721 | 6 |
| 26 | .653 | .000 | Selected | .701 | 6 |
| 27 | .515 | .001 | Selected | .611 | 5 |
| 28 | .569 | .000 | Selected | .515 | 3 |
| 29 | .599 | .000 | Selected | .412 | 1 |
| 30 | .616 | .000 | Selected | .602 | 5 |
| 31 | .616 | .002 | Selected | .551 | 4 |
| 32 | .562 | .000 | Selected | .477 | 2 |
| 33 | .691 | .001 | Selected | .491 | 2 |
| 34 | .588 | .000 | Selected | .611 | 5 |

The Table 2 shows the result of Principal Component Analysis with Varimax Rotation. The factor analysis with Varimax Rotation produced 7 factors accounting for 58.3 per cent of the variance. Table 2 in addition indicates the factor loading ranged from 0.400 to 0.450 for factor I, from 0.451 to 0.500 for factor II, from 0.501 to 0.550 for factor III, from 0.551 to 0.600 for factor IV, from 0.601 to 0.650 for factor V, from 0.650 to 0.700 for factor V and 0.701 to 0.750 for factor VII. Table 3 shows the Internal Consistency Reliability of Effectiveness Subscale. The summary of the result presented in Table 2 showed that out of 78 items of the instrument, 34 items were factorially pure and are acceptable as suitable for use in the study.

Table 3: Psychometric properties of the final scale and its subscales Factors Number of Items

| | Teaching Effectiveness Sub- Scale | Number of Items | Internal Consistency (α) |
|---|-----------------------------------|-----------------|-----------------------------------|
| 1 | Subject- matter proficiency | 4 | .87 |
| 2 | Interactive Proficiency | 5 | .82 |
| 3 | Professional proficiency | 6 | .82 |
| 4 | Teaching style | 5 | .88 |
| 5 | Classroom management | 5 | .81 |
| 6 | Evaluation | 3 | .83 |
| 7 | Communication | 3 | .84 |
| 8 | Personality | 3 | .74 |
| | TES | 34 | .93 |

The reliability coefficient obtained for the instrument was 0.93 and the coefficients of its sub-scales ranged from 0.74 to 0.88, which were judged to be high and adequate (Fraenkel and Wallen, 1993). The estimate further

shows that the items on the instruments were quite homogeneous and reliable.

As seen in Table 3, the Alpha of the final scale was 0.930 which signifies high internal consistency while the Alpha of its subscales ranged from 0.740 to 0.870. Thus, the final scale has been proven reliable (Alpha value of 0.930 indicating high internal consistency) and valid (from 0.507 to 0.748 factor loadings based on the 0.500 cut-off for screening of items).

8. Conclusion

The purpose of this study was to develop and validate a Teaching assessment scale for lecturers in Nigerian Universities. The study was motivated by dearth of valid and reliable scales and persistent poor performance of students in Nigerian universities over the years. Principal Components Analysis (PCA) was conducted to confirm the number of factors to be extracted. Seven factors were extracted for the instrument and each of the factors was labelled based on the description of items that loaded on them. The index of logical validity was 0.90 while the reliability coefficient of the instrument was 0.93. The LECTAS is a psychometrically sound scale that measures the multidimensional aspects of lecturer effectiveness in Nigeria universities. This is reflected in the results of its reliability evaluation (Cronbach's Alpha value of 0.93 showing high internal consistency) and validity evaluation (from 0.507 to 0.748 factor loadings based on the 0.500 cut-off for screening of items). The final version of the LECTAS is designed with the 34 valid items and 7 dimensions. This scale is a Likert Type four point rating scale (Always, Sometimes, Rarely and Never; and the scoring relating to each of the four response options are 4, 3, 2 and 1 respectively) for positively worded items and reverse for the negatively worded items. The maximum possible score is 136 and the minimum is 34. The highest score indicates high teaching effectiveness in lecturers. Based on the findings of the study, it was concluded that the LECTAS is a valid and reliable instrument for assessing teaching of academic members of staff in Nigerian Universities. The data collected would be useful to lecturers and students in making necessary adjustments in their behavior and practices in order to improve students' performance.

9. Recommendations

The LECTAS has gone stages of testing its reliability (Cronbach's Coefficient Alpha formula) and validity (PCA with Varimax Rotation) in the process of establishing its psychometric properties. The LECTAS is recommended to be included as component part for assessment of lecturers in Nigerian universities for promotion and teaching improvement. Periodic feedback from the instrument should be communicated to lecturers to derive the formative function of the instrument. However, to have a more detailed property of LECTAS, further study is still recommended. It must be reviewed again to determine redundant items to further establish its reliability and validity.

References

- Abdulkadir, H. S. (2006). Strategies for improving the efficiency of teachers in Jigawa State [dissertation]. [Nsukka]: University of Nigeria, Nsukka.
- Ajzen, I. & Fishbein, M. (1980). *Understanding Attitudes and Predicting Social Behavior* (Englewood Cliffs, NJ, Prentice Hall)
- Awofala, A. O. (2011). Development and Factorial structure of students' evaluation of teaching effectiveness scale in mathematics. *Cypriot Journal of Educational sciences*. 7(1): 33-44.
- Calagaus, G. M. (2012). Teacher effectiveness scale in higher education: Development and psychometric properties. *International Journal of Research Studies in Education*. 7(1):45-51.
- Casey, R. J., Gentile, P., & Bigger, S. W. (1997). Teaching appraisal in higher education: An Australian perspective. *Higher education*, 34(4):459-482.
- Dresel, M. & Rindermann, H. (2011). Counseling university instructors based on student evaluations of their teaching effectiveness: A multilevel test of its effectiveness under consideration of bias and unfairness variables. *Research in Higher education*. 52(7):717-737.
- Estu W & Franciscus, X. S.(2016). Developing an observation instrument for assessing the effectiveness of english teaching at vocational secondary schools. *Research and evaluation in education*, 2(2), 2016, 135-154 available online at: <http://journal.uny.ac.id/index.php/reid>
- Faleye, B. A., & Awopeju, O. A. (2012). A Re-Validation of the Students' Evaluation of Teaching Effectiveness Rating Scale. *Ife Psychologia*, 20(2), 220-233.
- Feldman, K. (2007). Identifying exemplary teachers and teaching: evidence from student rating. In R. P. Perry, & J. C. Smart, *The Scholarship of teaching and learning in higher education: an evidence-based perspective* (p. 93-130). Dordrecht, The Netherlands: Springer
- Fraenkel, J.R., & Wallen, N.E. (1993). *How to design and evaluate research* (2nd ed.). New York.
- Freeman, R. D. (1979). *Course faculty instrument*. New York: New York University Press.
- Goe, L., Bell, C. & Little, O. (2008). *Approaches to Evaluating Teacher Effectiveness: A research synthesis*.

- National Comprehensive Center for Teacher Quality. NY.
- Harris, B. D. (2001). An aggregate approach to measuring teaching performance. Unpublished Doctoral Dissertation. San Diego State University.
- Hobson, S. M. & Talbot, D. M. (2001) Understanding student evaluations. *College Teaching*, 49 (1), pp. 26–31
- Jegade, O. P., Faleye, B. A. & Adeyemoo, E. O. (2015). Factor Analytic Study of Lecturer’s Teaching Assessment Scale in Obafemi Awolowo University, Nigeria. *World Journal of Education*. 5(3): 123-130.
- Kolawole, E. B. & Fabunmi, J. I. (2014). Construction and Validation of Career Interest Scale for Senior Secondary Schools Students in Nigeria. *Advances in Social Sciences Research Journal*. 1(3): 35-47.
- Marsh, H. W. (1987) Students’ evaluations of university teaching: research findings methodological issues and directions for future research. *International Journal of Educational Research*, 11 (3): 253–388.
- Muhammad A. & Sally J. Z. (2015). Journal of Research and Reflections in Education December 2015, 9(2):134-148
- Popp, P. A., Grant, L. W., & Stronge, J. H. (2011). Effective teachers for at-risk or highly mobile students: What are the dispositions and behaviors of award-winning teachers? *Journal of Education for Students Placed at Risk*, 16, 275-291
- Spooren, P., Brockx, B., & Mortelmans, D. (2013). On the Validity of Student Evaluation of Teaching: The State of the Art. *Review of Educational Research* 83(4): 598–642. doi: 10.3102/0034654313496870
- Tucker, P. D. & Stronge, J. H. (2005). Linking teacher evaluation and student learning. Danvers, USA: Association of Supervision and Curriculum Development.

APPENDIX 1
LECTURER TEACHING ASSESSMENT SCALE (LECTAS)
STUDENT EVALUATION OF TEACHING

(To be completed by students who have met 75% attendance on the course)

INSTRUCTIONS:

This is an instrument for Assessment of Lecturers by the students during the current semester. This evaluation instrument is aimed at providing useful information for quality improvement of teaching and learning in the University.

Students are expected to honestly give their opinion on the courses and the lecturers who taught the courses without prejudice or bias.

SECTION A: LECTURER INFORMATION

Please fill as appropriate.

| | |
|-------------------------|--|
| NAME OF LECTURER | |
| COURSE TAUGHT | |
| FACULTY | |
| DEPARTMENT | |
| SEMESTER | |
| SESSION | |

SECTION B: Criteria/Statements on the performance of lecturer

INSTRUCTION: Please select the option that directly represents your view on the performance of the Academic Staff with name above. Indicate your rating by ticking (√) against one of the options from: Always, Sometimes, Rarely and Never

| S/N | CRITERIA/STATEMENTS | Always | Sometimes | Rarely | Never |
|-----|--|--------|-----------|--------|-------|
| 1 | Lecturer’s Adherence to time (My lecturer arrives in the classroom on time) | | | | |
| 2 | Utilization of lecture time. (The lecturer teaches for the number of hours required for teaching the course per week e.g. 3 hours for a 3-unit course; 2 hours for a 2-unit course) | | | | |
| 3 | Provision of lecture materials. (My lecturer provides lecture notes and/or refers students to available texts and other sources of information) | | | | |
| 4 | Presentation indicators (She/he provides useful presentation indicators e. g. gives outlines, states objectives, reviews main points) | | | | |
| 5 | Reception (She/he welcomes and encourages ideas, comments, questions and students’ active participation) | | | | |
| 6 | Mode of presentation. (She/he delivers the lecture without excessive reliance on reading lecture notes.) | | | | |

-
- 7 **Communication** (She/he speaks fluently without distracting verbal and/or non-verbal mannerisms)
 - 8 **Student Respect** (She/he commands respect and treats students with respect.)
 - 9 **Discipline** (She/he maintains discipline in and out of the lecture room.)
 - 10 **Attendance Register** (She/he ensures that students attend lectures regularly by taking weighted attendance)
 - 11 **Evaluation** (She/he grades and returns tests and assignments in time)
 - 12 **Lecturer's preparedness to teach** (She/he arrives in the classroom prepared)
 - 13 **Lecturer's Accessibility in Class** (She/he is accessible in class)
 - 14 **Lecturer's availability outside the classroom** (She/he is available during office hour)
 - 15 **Asking question in Class** (The lecturer feel comfortable asking question during class)
 - 16 **Explanation of Concepts** (She/he explains the concept and material well, and displays mastery of content.
 - 17 **Lecturer's Teaching Consistency** (She/he is content in his/her teaching)
 - 18 **Use of Digital instructional Technology** (She/he uses digital instructional technology such as e-book, e-journal, journals articles, YouTube, computer programs, videos, online resources, social media, mobile speakers etc.)
 - 19 **Learning Compliance with the Syllabus** (I am learning what is in the course description/syllabus)
 - 20 **Use of Digital Skills** (I am using digital and other online skills in this course)
 - 21 **Relevance of Assignment** (My lecturer's assignments are relevant to the course)
 - 22 **Grading Policies** (She/he grading policies are fair and consistent to the syllabus)
 - 23 **Relevance of Course content to future career prospect** (She/he relates course content and skill to my future career)
 - 24 **Lecturer's Openness to multiple perspective** (She/he is open to a variety of perspectives on a topic)
 - 25 **Developing writing Skills** (She/he gives writing assignments in class to help develop my writing skills)
 - 26 **Motivation/Inspiration** (She/he motivates and inspires me)
 - 27 **Due Process** (She/he has respect for due process)
 - 28 **Mode of Dressing** (She/he is lecturer is always well dressed)
 - 29 **Students Friendliness** (She/he is students friendly)
 - 30 **Leadership skills** (Ability of the lecturer to lead students)
 - 31 **Assessment** (She/he encourages students to give feedback in the course of the lecture)
 - 32 **Teaching Environment** (She/he creates friendly and supportive lecture environment)
 - 33 **Communication** (She/he use correct vocabulary and grammar in speaking & writing)
 - 34 **Lecturer's presentation capability** (She/he presents challenging and thought-provoking ideas to stimulate reflective and critical thinking)

TOTAL RATING

MEAN RATING

SCORING = MAXIMUM SCORE (136); MINIMUM SCORE = (34) SCORE ARE CONVERTED TO PERCENTAGE

INTERPRETATION OF MEAN RESPONSE

| |
|--|
| 70 & Above: Excellent (exceptional, exemplary) |
| 60-69: Very good (high quality, better than average) |
| 50-59: Good (reasonable, well done, acceptable) |
| 40-49: Marginal (highly below average, needs improvement) |
| 0-39: Poor (far below average, needs improvement) |