DEVELOPMENT AND VALIDATION OF INSTRUMENT FOR STUDENTS' APPRAISAL OF TEACHING EFFECTIVENESS OF LECTURERS IN COLLEGES OF EDUCATION IN NIGERIA

By

EZUGWU, G. G.
Department of Educational Psychology
Federal College of Education,
Zaria, Nigeria
gregorygodwin33@yahoo.com

Abstract
This study was designed to develop and validate an instrument for Students Appraisal of Teaching Effectiveness of Lecturers (SATEL) in Colleges of Education in Nigeria. The study was purely an instrumentation study. From literature, lecturers, students, and the researcher initially identified 20 major constructs' attributes from where 85 items that constituted the elements of teaching effectiveness came. These were re-grouped into 9 clusters, which are: good organisation and planning; demonstration skills; teacher-students interaction; clarity and communication skills. Others are flexible approach to teaching; display of objectivity; academic role model; intellectual and personality disposition; and grading and evaluation of students. A 40-item instrument was structured from the sections, and presented to five specialists in instrument development, for the purposes of validation, after which 27 items survived the exercise. The items were put into a modified Likert Scale scored on a 7-point grid and trial-tested using students from two colleges of education. A test of reliability using trial-test result shows that the instrument was reliable at 0.96. The 27 items were administered to 33,165 NCE III students in the 64 Colleges of Education from the North, East and South zones. Data from the research subjects were analysed using mean scores. The result shows among other things, that the items are stable across the variables. The researcher recommended that the instrument (SATEL) be used on the lecturers before any promotion exercise, and at the end of every course contact.

Keywords: Development of instrument; Students appraisal; Teaching effectiveness; Lecturers in Colleges of Education.
Development and Validation of instrument for Students’ appraisal of teaching effectiveness of Lecturers in Colleges of Education in Nigeria

Introduction

The National Commission for Colleges of Education (NCCE) is charged with the responsibility of fashioning out a policy necessary for the full development of teacher education; seeing to the quality of academic staff; accreditation of NCE courses; certification and other academic awards; disbursement of funds to the Federal, State and privately-owned Colleges of Education based on prescribed rules (Lassa, 1998).

The mandate of the NCCE on face value appears relevant, useful, credible and adequate. Still, there is evidence that NCE graduates have not distinguished themselves as having received quality teaching in the Colleges they attended in this present dispensation. For instance, Ezugwu (1997) observed that many Colleges of Education authorities play down standards in their recruitment of academic staff, which apparently, bring low quality products. To this, they advise that all the institutions producing teachers should make sure that they follow the laid-down minimum standards especially for Colleges of Education and Universities. This may have prompted Lassa (1998) to recommend that in recruiting academic staff for Colleges of Education, it has to be a holder of Masters Degree who is proficient in his field of study. In the same vein Ali, (1998), described academic staff as highly experienced and frontline scholars in their various disciplines and those who may utilize the required essential teaching and learning facilities as well as opportunities for professional and career growth in such system.

Statement of Problem

The Colleges of Education in Nigeria are charged with the responsibility of providing the middle manpower teachers needed for both the primary and junior secondary school system (FRN, 2005). Ultimately, it is the responsibilities of the academic staff of the different Colleges to perform this obvious task creditably. For this goal to be achieved, the lecturers who teach and perform their academic duties in the Colleges must be effective in doing so. Teaching effectiveness serves as a benchmark against which measure the performance in the general preparation of teachers for the task of teaching (Baikie, 2002).

One way of determining if these lecturers are effective is through the appraisal of their teaching effectiveness. Definitely, there is an annual appraisal exercise known as “Annual Performance Evaluation Report” (APER) which is filled by every worker annually for the purposes of promotion. This annual exercise is usually carried out by the superior officers on their subordinates. This is made possible because they work closely with each other. As was the practice, the Heads of department and the Deans do the appraisal of the lecturers or academic staff. Even though they have roles to play in the appraisal exercise, they should not be the principal actors.

It is clear that students who are always with the lecturers in the classroom are, and should be in a better position to correctly do the appraisal of the lecturers' teaching competence. Hence, the opinion of those who eat the dinner should be
considered, if we want to know how it tastes (Arubayi, 2003). According to Centra (1974), the implication of this is that students are in a better place to determine who the effective or ineffective lecturer is. Therefore, the problem of this study put in question form: Is it possible to develop a valid and reliable instrument for use in students' appraisal of lecturers' teaching effectiveness in Colleges of Education in Nigeria?: What are the psychometric properties of such instrument put in another way?

**Purpose of the Study**

The purpose of this study was to develop and validate an instrument for students' appraisal of teaching effectiveness of lecturers (SATEL) in Colleges of Education in Nigeria. Specifically, the study intended to achieve the following:

i. to determine the validity and to establish the reliability of the instrument.

ii. to find out the influence of specialty of lecturer (Arts, Science and Vocational courses) in SATEL.

iii. to establish the influence of gender of lecturers in SATEL.

**Research Questions**

The following research questions were answered in the study:

1. What are the validity and the reliability of the instrument SATEL?
2. What is the influence of lecturers' area of specialization (course type) on the instrument?
3. What is the influence of gender of lecturers in SATEL?

**Research Design and Methods**

An instrumentation research design was adopted for the study. A study belongs to “instrumentation research” if it involves introducing new or modified content procedure, technology or instrument of educational practice (Ogomaka, 1990). Ali (2006) pointed out that a study which is purely geared towards the development and validation of instrument in education is an “instrumentation study” or developmental study.

The study was conducted in Colleges of Education in the six geo-political zones in Nigeria. The six geo-political zones are: North East, North West, South-East, South-West, North Central and South South.

The population of the study consisted of all the thirty-three thousand, one hundred and sixty-five (33,165) NCE III students in the sixty-four Colleges of Education in Nigeria. According to the National Commission for Colleges of Education (NCCE) Statistical Digest (2010), out of the sixty-four Colleges of Education, twenty are federal, thirty-eight are States and six are privately-owned Colleges. The NCE III Students of 2011/12 academic session were used in the study.

A multi-staged sampling technique was used in selecting the sample for this study. The first stage involved the sampling at the geo-political level. The six geo-
political zones of the country were considered, from which three zones were now randomly drawn. The three randomly drawn geo-political zones are North Central, South East and South South. The three randomly sampled geo-political zones equally ensured a broad representation of the geo-political structure of Nigeria.

The researcher identified three strata of Colleges of Education in Nigeria. These are Federal-owned, State owned, and privately owned Colleges of Education. From these three strata, the researcher applied disproportionate sampling to sample two federal colleges and one each from state and privately owned colleges of education. The sample for the study consisted of 840 NCE III Students in the 2011/12 academic session. This was made up of 420 students from the two federal, and 210 students from each State, and a Private Colleges of Education respectively. They were selected through simple random sampling procedure without replacement.

**Instrument Development**

The researcher gathered and listed 85 attributes, skills and traits from literature which were thought to be necessary for competent, good, efficient and effective teaching. The list though long, was later reduced to a more realistic pool of attributes and skills. The following procedures and stages were adopted in the development and validation of the instrument.

i. Stage 1: Pool of Attributes  
ii. Stage 2: Organization of Items  
iii. Stage 3: Assessing the Importance and Relevance of Items  
iv. Stage 4: Assembling of the Initial Draft Instrument  
v. Stage 5: Trial-Testing (pilot study)

**Validation and Reliability of the Instrument**

The items were presented to five specialists in instrument development from University of Nigeria Nsukka, and Ahmadu Bello University, Zaria. These experts were requested to independently make critical observations that could improve each of the identified sections and items. They were asked to examine the items along the following guidelines: to see the relevance of the items in each section; the simplicity of expression; double barreled statements; clarity of items; statement of facts or opinion; vagueness of statements; over-emphasis on any of the teaching effectiveness attributes; any biased statement; and any other relevant considerations or input to the instrument. Inclusion of any item in the instrument was based on its acceptance by at least three out of five validators.

The collected copies from the five experts were used during trial-testing to find the internal consistency of the instrument. Cronbach alpha procedure which is a modified version of Kuder-Richardson formula 21 (Cronbach alpha) was used in testing the internal consistency of the items based on scores derived from the trial-test. The reliability coefficient of 0.9607 was obtained from the items.
Data Collection and Analysis

The researcher administered the instrument on NCE III Students of the sampled Colleges. However, 5 lecturers each in the sampled Colleges helped the researcher to administer and make returns. Out of eight hundred and forty (840) copies of the instrument administered, eight hundred and fourteen (814) copies correctly filled, were returned.

The responses made by the students on the developed instrument, were analyzed based on each research question. For instance, research question one was the result of 5 expert validators and the use of Cronbach Alpha, while others were answered with the use of the mean.

Results

Research Question One: What is the validity and the reliability of the instrument SATEL?

Table 1: The five validators imputes and the result of the Cronbach Alpha test on the 27 items of SATEL

<table>
<thead>
<tr>
<th>Factors</th>
<th>Items</th>
<th>Reliability Coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>a Good organization and planning</td>
<td>1, 2, 3</td>
<td>0.5030</td>
</tr>
<tr>
<td>b Demonstration skills</td>
<td>5, 6, 7</td>
<td>0.3989</td>
</tr>
<tr>
<td>c Teacher-student interaction</td>
<td>8, 9, 10</td>
<td>0.8632</td>
</tr>
<tr>
<td>d Clarity and communication skills</td>
<td>12, 14, 15</td>
<td>0.8565</td>
</tr>
<tr>
<td>e Flexible approach to teaching</td>
<td>17, 18, 28</td>
<td>0.9694</td>
</tr>
<tr>
<td>f Display of objectivity</td>
<td>19, 24, 29</td>
<td>0.5203</td>
</tr>
<tr>
<td>g Academic role model</td>
<td>21, 22, 23</td>
<td>0.8369</td>
</tr>
<tr>
<td>h Intellectual and personality disposition</td>
<td>25, 26, 27</td>
<td>0.5030</td>
</tr>
<tr>
<td>i Grading and Evaluation of students</td>
<td>30, 32, 33</td>
<td>0.7669</td>
</tr>
</tbody>
</table>

The summary of the result presented in tables 1 indicates that out of the 33 items, 27 items were found to be valid and acceptable as suitable. It means that, items 4, 11, 13, 16, 20 and 31 were rejected and discarded by the experts. The reliability coefficient of the 27 items survived following final stage of the actual trial of the instrument was 0.9607. The above Table 1 shows the reliability coefficient of the nine factors, one after the other, as well as the five experts' imputes

Research Question Two: What is the influence of lecturers' area of specialization (Course-type) on the instrument?
Table 2: Mean (x) and Standard Deviation (S) of Clusters or Factors of SATEL by Course Type

<table>
<thead>
<tr>
<th>Arts</th>
<th>Interpretation</th>
<th>Science</th>
<th>Interpretation</th>
<th>Voc. Edu</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>X</td>
<td>S</td>
<td></td>
<td>X</td>
<td>S</td>
</tr>
<tr>
<td>A</td>
<td>Good organization and planning</td>
<td>5.9</td>
<td>3.3</td>
<td>Moderately Effective</td>
<td>5.9</td>
</tr>
<tr>
<td>B</td>
<td>Demonstration skills</td>
<td>5.6</td>
<td>4.0</td>
<td>Moderately Effective</td>
<td>5.4</td>
</tr>
<tr>
<td>C</td>
<td>Teacher-student interaction</td>
<td>6.1</td>
<td>5.1</td>
<td>*Strongly Effective</td>
<td>6.0</td>
</tr>
<tr>
<td>D</td>
<td>Clarity and concrete Skills</td>
<td>5.5</td>
<td>4.2</td>
<td>Moderately Effective</td>
<td>5.4</td>
</tr>
<tr>
<td>E</td>
<td>Flexible approach to teaching</td>
<td>5.4</td>
<td>3.7</td>
<td>Moderately Effective</td>
<td>5.4</td>
</tr>
<tr>
<td>F</td>
<td>Display of objectivity</td>
<td>5.6</td>
<td>3.8</td>
<td>Moderately Effective</td>
<td>5.6</td>
</tr>
<tr>
<td>G</td>
<td>Academic role model</td>
<td>5.7</td>
<td>4.0</td>
<td>Moderately Effective</td>
<td>5.8</td>
</tr>
<tr>
<td>H</td>
<td>Intellectual and personality disposition</td>
<td>5.8</td>
<td>5.6</td>
<td>Moderately Effective</td>
<td>5.7</td>
</tr>
<tr>
<td>I</td>
<td>Grading and evaluation of student</td>
<td>5.7</td>
<td>3.7</td>
<td>Moderately Effective</td>
<td>5.7</td>
</tr>
</tbody>
</table>

The result in Table 2 shows that nine clusters of SATEL were rated moderately effective in all the three course-type. However, in Teacher-student interaction, the students rated Arts and science course lecturers strongly effective and vocational education as moderately effective. They received the mean ratings of 6.19 and 6.03 respectively in both Arts and Science. In general, lecturers in Arts received a higher mean rating than the other two, whereas Science received higher than vocational course lecturers in all the nine clusters of SATEL. **Research Question Three**: What is the influence of the gender of lecturers on their mean scores in SATEL?

Table 3: Mean (x) and Standard Deviation (SD) of clusters of SATEL by Gender

<table>
<thead>
<tr>
<th>Cluster Area</th>
<th>Male</th>
<th></th>
<th>Female</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>X</td>
<td>S</td>
<td>Interpretation</td>
<td>x</td>
<td>SD</td>
</tr>
<tr>
<td>Good organization and planning</td>
<td>5.94</td>
<td>3.24</td>
<td>Moderately Effective</td>
<td>5.96</td>
<td>3.31</td>
</tr>
<tr>
<td>Demonstration skills</td>
<td>5.50</td>
<td>4.15</td>
<td>Moderately Effective</td>
<td>5.55</td>
<td>4.28</td>
</tr>
<tr>
<td>Teacher-Student rapport</td>
<td>6.08</td>
<td>3.33</td>
<td>*Strongly Effective</td>
<td>6.11</td>
<td>3.61</td>
</tr>
<tr>
<td>Clarity and efficient communication</td>
<td>5.45</td>
<td>4.18</td>
<td>Moderately Effective</td>
<td>5.51</td>
<td>4.30</td>
</tr>
<tr>
<td>Flexible approach to teaching</td>
<td>5.40</td>
<td>3.80</td>
<td>Moderately Effective</td>
<td>5.48</td>
<td>3.94</td>
</tr>
<tr>
<td>Display of objectivity</td>
<td>5.53</td>
<td>3.92</td>
<td>Moderately Effective</td>
<td>5.84</td>
<td>4.01</td>
</tr>
<tr>
<td>An academic role model</td>
<td>5.71</td>
<td>4.01</td>
<td>Moderately Effective</td>
<td>5.84</td>
<td>4.02</td>
</tr>
<tr>
<td>Intellectual and personality disposition</td>
<td>5.73</td>
<td>3.74</td>
<td>Moderately Effective</td>
<td>5.80</td>
<td>3.85</td>
</tr>
<tr>
<td>Grading and Evaluation of Students</td>
<td>5.8</td>
<td>3.93</td>
<td>Moderately Effective</td>
<td>5.74</td>
<td>3.90</td>
</tr>
</tbody>
</table>

123
The result in Table 3 shows that with respect to gender, the lecturers received a high mean rating. Even though both the male and female lecturers were rated high, female lecturers received a higher mean rating than their male counterparts. Again, 'teacher-student interaction' in both cases were rated strongly effective as against moderately effective by other items.

Discussion

The 27 items that survived after the five experts exercise were subjected to a test of internal consistency (reliability index) using Chronbach alpha. Reliability of any instrument, according to Anikweze (2009), is the degree of consistency such instrument has in establishing what it intended to continuously. The result of Cronbach alpha indicates a high reliability coefficient of 0.9607 for the entire 27-item instrument. According to Kerlinger (1992), to obtain a high reliability shows that the instrument measures dependability, consistency, predictability and accuracy. This corroborates with Chukwudolue (2002) whose instrument titled: Teacher Motivation Assessment Scale which had a high reliability of 0.74. The result is equally in agreement with Ogbaezi and Okpala (1994) who posited that "greater values indicate greater reliability".

Influence ofCause Type

Data in Table 3 clearly shows the analysis of mean rating on lecturers' area of specialty or course-type (Arts, Science and Voc. Education). In that table, the mean ratings of lecturers based on their area of specialization or course-type, did not vary remarkably. For instance, in Good organization and Planning: Arts, Science and Vocation related course lecturers received mean ratings of 5.92, 5.92 and 5.94 respectively, indicating moderately effective. In Demonstration Skills, Arts lecturers were rated a mean of 5.60, indicating moderately effective, while Science and Vocational courses had 5.47 and 5.39 respectively indicating moderately effective.

A look at table 3 shows that the ratings of the Arts related courses were higher than Science and Vocational education. This could be that very few students offer science courses because of poor understanding of science concepts, and the assumption that science subjects are difficult to conceptualize. This is in consonance with Ngwu (2002) who observed that the major problem of Science discipline is the non-attainment of reasonable depth of knowledge and practical skills required for attainment of science culture. In a survey study conducted by Okeke (2000), it was revealed that out of 445 teachers that taught Introductory Technology in 362 schools in Imo State, none of them was of the right quality. Ngwu (2002) equally observed that Integrated Science has been a part of program of teacher preparation, we have hardly gotten student teachers to the point at which they can conduct an independent science investigation that would show the "integrated-ness of knowledge". From these instances, one would see reasons why there is preference for Arts related course instead of science and vocational courses.

The Influence of Gender of Lecturers

The summary of the result in Table 3 shows the mean ratings of gender of lecturers. A look at the table reveals that in almost all the constructs, both male and
female lecturers obtained moderately effective rating across board. It was in “Teacher-Students Interaction” that received higher ratings of “strongly effective” in both sexes. However, in all, there is a clear evidence that though both of them obtained high ratings, female lecturers received higher ratings than their male counterparts.

This finding is in consonance with the finding of Centra and Geubatz (2000) who found that the sex of lecturers was found to affect students’ ratings, with female lecturers receiving more favorable ratings than male lecturers. Similarly, sex of the students was equally found to affect students’ ratings with female students rating higher than the male students.

**Educational Implication of the Findings**

This study has interesting implications not only for teaching and learning, but also in administration of colleges in general. The study has developed a valid and reliable instrument which measures, assess and appraises teaching effectiveness. The implication here is that the problem of how best to measure teaching effectiveness of lecturers in Colleges of Education can now be addressed. Initially, administrators in Colleges of Education are confronted with the problem of lecturers’ appraisal instrument. With this instrument in place or available, the effectiveness of lecturers could be appraised from time to time in Colleges of Education in Nigeria.

This has additional implication for job efficiency and control. According to Arubayi (2003), in an era of increasing accountability and evaluation of learning outcomes, achieving a more robust understanding of evaluation of teaching effectiveness may be a necessary step towards including the scholarship of teaching, in decision of appointment, tenure and promotion of academic staff in tertiary institutions. Therefore, SATEL is a necessary check and balance with respect to accountability on the job.

Further, this newly developed instrument going by the psychometric attributes and in consonance with the established findings and theories, has been shown to exhibit relative stability across gender, and area of specialty of lecturers. This implies that from now, lecturers’ teaching effectiveness can be efficiently appraised without a sniff of bias.

**Recommendations**

Based on the findings of this study, the following recommendations were made:

i. The instrument: Students’ Appraisal of Teaching Effectiveness of Lecturers’ should be used in all the Colleges of Education in Nigeria.

ii. If the goal of instructional improvement will be achieved, this instrument should be given at the end of each course contact. This is particularly to find out how effective the lecturer fared on the course.

iii. Similar studies should be carried out in Universities and Polytechnics in Nigeria.
Reference


Development and Validation of instrument for Students' appraisal of teaching effectiveness of Lecturers in Colleges of Education in Nigeria


