Development of a Measuring Scale for University Teachers' Assessment Literacy in Osun State, Nigeria

[1]A. A. ADEDIWURA, Ph.D
yemtoy20002000@gmail.com

[2]Bankole Isaac OGUNSAKIN
sakinbamikole@yahoo.com

[3] Yusuf Olayinka SHOGBESAN
yusuflayinka813@gmail.com

[1, 2 & 3]Department of Educational Foundations and Counselling,
Obafemi Awolowo University, Ile-Ife

Abstract

The study developed a scale for measuring assessment literacy of university teachers and determined the validity and reliability of the scale. This is with a view of improving university teachers' assessment literacy. The study adopted the non-experimental research design of the descriptive type. The population comprised all lecturers working in universities within Osun State. The sample consisted of 549 university teachers(lecturer II and above) that were not in training position from different disciplines selected using convenient sampling technique in all the eight accredited universities within Osun state. The study made use of three instruments namely; University Teachers Assessment Literacy Scale (UTALS), Teacher Self-Efficacy Scale (TSES) and Teacher Burnout Scale (TBS). The results showed that the 60- item UTALS was reduced to 43 items after the psychometric procedure of item retention and deletion were applied. Scree plot and eigen value methods showed that the final UTALS construct validity had seven factors of university teachers' assessment literacy that accounted for 86.46% of the total scale variance and also converged and discriminate significantly with the Teacher Efficacy Scale (r = 0.75) and Teacher Burnout Scale (0.52) respectively. The internal consistency reliability of the scale was 0.96 (Guttman), 0.90(Cronbach), and 0.96 (Spearman). The study concluded that the university teachers assessment literacy scale developed in this study had adequate psychometric characteristics and is suitable for measuring assessment literacy of university teachers in Osun State.

Keywords: Assessment, assessment literacy, scale development, university teachers

Introduction

Assessment is considered to be an important integral component in any instructional context and educational setting. In any higher institution of learning, assessment is a major business and plays at least two key roles; first in ensuring institutional quality and

accountability, and second, in improving the students' learning (Ewell, 2009). University lecturers' use of assessment affects the depth and quality of what students learn, their choice of learning strategies, how they manage their study time (Australian National Training Authority, 2002), and quite significantly, their motivation to continue learning (Harlen & Crick, 2003). Thus, the importance of assessment cannot be overstated as no factor influences a learning environment as much as assessment (Anderson, 2004; Hannafin, Hill, Oliver & Glazer, 2003).

Assessment literacy can be defined as the solid and sound educational assessment knowledge and skills that is required by teachers in assessing students' mastery of learning outcome. Thus, it is one of the most important responsibilities of teachers in classroom teaching and learning (Volante & Fazio, 2007). The effectiveness of teaching methods used in classroom may be exposed by the quality of the assessments used. Therefore, the results gathered from classroom assessments should be reliable and valid. Accordingly, the problems of lack of teachers' consensus on deciding the most beneficial and suitable assessment practices, limited assessment training, lack of assessment knowledge and skills may result in the teachers' variation in implementing the classroom assessment (Volante & Fazio, 2007; Suah, 2012).

The conduct of quality assessments require that a teacher carefully consider key features of assessment. Stiggins, (1993) opined that such features include; teacher clarity about the purpose of the assessment, teachers need to start with a clear and appropriate vision of the learning target, designing assessment that is adequate and appropriate for the context and giving feedback of the results to intended user in a timely and understandable way. Aside the mentioned four quality assessment key, another related key feature that ensures classroom effective use of assessments in the classroom is student involvement in the process while they are learning.

The Nigeria educational system calls for the accountability of higher education in promoting student learning. In this regard, universities are expected to give paramount importance to two key components in the structure of academic programmes. They are: (i) clear and measurable programme learning outcomes, and (ii) quality assessment that is well-aligned with the intended outcomes. The relationship between these two components is an intricate one where the provision of clear and measurable learning outcomes is expected to guide the use and practice of assessment among university lecturers. However, many universities in Nigeria do not seem to provide the lecturers with a clear policy on assessment and likely not provide necessary training for lecturers on assessment practice, hence leaving them clueless as to the '*what*' and the '*how*' of student assessment. This is evidence in Nigeria university lecturers' use of assessment that tends to be restricted to paper-and-pencil tests (such as multiple choice, true-false and essay questions) being the most widely and frequently utilized methods. The use of alternative and authentic assessments (e.g. observations, demonstrations, portfolios, e.tc.) seem to be uncommon

among the lecturers. The adherence to traditional formats is known to be a long standing practice, and provides some indication as to what Nigerian university lecturers commonly understand to be student assessment. It also suggests that their assessment knowledge and practice seems not to have improved much over the years thus the assessment literacy (the collection of the teacher's skills in test development, the composition of closed questions, the development of assessment rubrics, and statistical analysis of cumulative data for variety of teaching and learning needs Donoho, 2000; Popham, 2004) of Nigeria universities teachers come to question. In order for assessment to be effective, classroom teachers need to be assessment literate knowledgeable about the key concepts of testing and how they can inform the design of assessments and decisions surrounding their usage. Knowing more about assessment will not only help in the effective assessment of students, but it will also provide a means of evaluating teachers' own teaching and help in the production of tests that will actually motivate students to learn.

Assessment literates mean the ability to detect the differences between sound and unsound assessment (Stiggins, 1995). He specified that assessment-literate educators should be able to identify: (a) the content and learning outcome to be assessed; (b) the purpose of having assessment; (c) the best way to assess the students' skills and knowledge; (d) the development of quality instrument to assess students' performance; (e) the potential problems with the assessment; (f) the prevention of the problem; and (g) the awareness of the potential negative consequences of poor, inaccurate assessment. According to the standards for teacher competence in the Educational Assessment of Students (American Federation of Teachers, National Council on Measurement in Education, National Education Association, 1990), there are seven standards for teacher competence in classroom assessment. Teachers should be skilled in: choosing assessment methods that are suitable for the learning outcomes to be assessed; developing assessment instrument that are suitable for the learning outcomes to be assessed; administering, scoring, and interpreting the assessment findings of both externally-produced and teacher-produced assessment; using assessment results perceptively when making decisions about students' learning, teachers' teaching, school development, programme, or curriculum; developing the reliable and valid grading procedures; communicating and reporting efficiently the assessment results to students, parents, school administrators, and other educators; and recognizing the illegal, unethical, inappropriate uses of assessment information and methods of assessment.

There has been a growing expectation concerning Nigeria universities teachers' assessment knowledge and skills. Universities teachers are required by stakeholders to implement assessment professionally and creatively so as to align with the learning outcomes in the curriculum standards as a means of improving students' learning ability and performance and in turn provide the nation with high level man power that is required for the development of the country. A substantial part of university teachers' professional

time is devoted to assessment-related activities. Optimal implementation of these activities requires strong knowledge and abilities in educational testing. Teachers' knowledge and abilities in educational testing can be equated to assessment literacy. It entails knowing what it is being assessed, why it is assessed, how best to assess it, how to make a representative sample of the assessment, what problems can occur within the assessment process, and how to prevent them from occurring. There is a concern on the adequacy of most university teachers' assessment literacy. The concern most often emanated from the quality and employability of the universities turned out graduates. However, the management of most Nigeria universities appears not to determine the assessment literacy level of their teachers at the point of entry into university teaching and most often do not likely organize in-service training for them on fundamentals of educational assessment. Thus, assessment literacy of most university teachers may be considered to be feeble and the issue of inaccurate and invalid educational assessment may be due to the unacceptably low levels of assessment literacy. This therefore, may have prevented the students from reaching their full potential. Furthermore, to the best of the authors knowledge based on available literature, little or no study is carried out by university management on teachers' adherence to the fundamental principles of educational assessment as recommended by experts when assessing students' learning quality and not much empirical studies have been carried out to validate existing Western teacher assessment literacy inventory for use among Nigeria university teachers and there is no known published locally developed scale in Nigeria with which this important attribute of a university teachers activities could be measured. This study therefore developed and validated a University Teachers' Assessment Literacy Scale (UTLS) for the measurement of university teachers' assessment literacy.

The major objective of this study is to develop a valid instrument that is reliable and validly capable of measuring university teachers' assessment literacy.

Research Questions

The following research questions were raised to guide the study.

1. What items would be adjudged fit to measure university teachers' assessment literacy

2. What is the construct validity of the developed UTALS?

3. What is the convergent validity of the UTALS?

4. What is the discriminant validity of the UTALS?

5. What is the factor structure of the developed UTALS?

6. What is the reliability of UTALS?

Methodology

The design employed for the study is the non-experimental research involving the descriptive survey type. The population for the study comprised all the 2,485 lecturers working in universities within Osun State of Nigeria. A total of 1640 (66.0%) of the teachers works in the two government owned universities in Osun State while the remaining 845(34.0%) works with private owned universities in the State. The sample consisted of 549 university teachers (lecturer II and above) from different disciplines that were not in training position. The university teachers were selected from all the eight National University Commission accredited universities in Osun State using convenient sampling. The sample comprised 339 male and 210 female lecturers. The breakdown of the sample comprised 275 university teachers that were below the rank of Senior Lecturer, 199 Senior Lecturers and 75 university teachers above the rank of Senior Lecturer. A total of 305 of the selected university teachers were from public universities while the remaining 295 were from private universities.

The study made use of self-developed instrument titled: University Teachers Assessment Literacy Scale (UTALS) and two adopted instruments: Teacher Self-Efficacy scale (TSES) and Teacher Burnout Scale (TBS). The first stage of UTALS item development was the generation of initial items on university teachers' assessment literacy. A pool of 85 items was generated from the literature and from ideas of universities teachers with good assessment practices experience. The 85 items covered the seven aspects of Standards for teacher competence in the educational assessment of students as put forward by (American Federation of Teachers, National Council on Measurement in Education, National Education Association, 1990).These items were moderated and reviewed by experts in the field of Tests and Measurement to determine the appropriateness, relevance and adequacy of the items (content validity). This was then reduced to 60 items. The response pattern adopted was Likert (1932) format with four option range from SA= Strongly Agree, A = Agree, D = Disagree and SD = Strongly Disagree. The TSES is a 10 item scale developed by Schwarzer, Schmitz, and Daytner, (1999). The TSES identifies job skills and groups them into four major areas: (a) job accomplishment, (b) skill development on the job, (c) social interaction with students, parents, and colleagues, and (d) coping with job stress.The TBS was developed by Seidman and Zager (1987) to investigate the amount of "burnout syndrome" that might be experienced by teachers in the classroom. The TBS is a twenty-item Likert scale with the following four factors: Teacher Burnout-Career Satisfaction (TB-CS), Teacher Burnout-Administrative Support (TB-AS), Teacher Burnout-Coping with Stress (TB-CWS) and Teacher Burnout-Attitudes towards Students. The University Teachers' Assessment Literacy Scale (UTALS), Teacher Self-Efficacy Scale (TSES) and the Teacher Burnout Scale (TBS) were administered simultaneously on the selected university lecturers by the researcher with the help of 6 trained research assistants.

Results

Research Question 1: What items would be adjudged fit to measure university teachers' assessment literacy?

To answer this question, the initial 60 items first version of University Teachers' assessment Literacy Scale (UTALS) was evaluated with the aim of determining their appropriateness for the scale. The correlation matrix was examined so as to ascertain the first quality, high inter-correlation of the items (DeVellis, 1991). The 60 items of the UTALS were Factor Analyzed. This method was chosen because the researcher was interested in understanding the underlying structure of a set of variables (Conway &Huffcutt, 2003). In the study, initial factor analysis of the 60 items loaded on nine factors and 10 of the 60 items had factor loadings less than 0.32 (Worthington & Whittaker, 2006) and/or loaded on more than one factor. The 10 items were then eliminated from further analyses to obtain a 50 item second version of (UTALS). An investigation into the 50 items after a second round of factor analysis revealed that the 50 items loaded on seven factors with reliability estimate (Cronbach Alpha = 0.888), items mean (3.136) and inter-item correlation (0.155) with seven items that were not satisfactorily inter-correlated and that increased the reliability of the scale if they are deleted. The reduction of the second UTALS version was based on Govaerts and Gregoire(2008) item reduction criteria which stipulated that any item affected by the three or any two of the conditions below should be expunged.

i.   Items with Low Item Mean (LIM) 3.136 or less.

ii. Items with Low Item total Correlation (LITC) of 0.155 and below.

iii. Items having a High Cronbach's Alpha if Item Deleted (HCAID) of 0.888 or more.

After applying the rules, the following items as contained in Table1 were retained in the developed university teachers' assessment literacy scale.

Table 1:
University Teachers' Assessment Literacy Scale (Third and Final Version)

OLD S/N	NEW S/N	Statement	SD	D	A	SA
1	1	Application of what was learnt in class is best assessed using performance test				
2	2	Accurate and consistence grading of students' knowledge depends on developing a scoring rubric after getting a feel for what students can do				
3	3	General impression on a student performance in a standardized test is only acceptable if the reliability of the test does not exceed 0.06				
4	4	Assigning students' grade is an appropriate use of standardized test result				

5	5	Students' prior performance before assigning a final grade should be considered in improving the validity of continuous assessment grading procedure
6	6	A student with 80th percentile in a test is said to have scored 80% of the test items correctly
7	7	To appropriately use assessment information , scores from standardized test should be used to determine teacher instructional effectiveness
8	8	The choice of an assessment that covers single-step problem solving skills is the best for assessing students' problem solving abilities
9	9	Portfolio assessment is the most suitable for accurate assessment of students' performance in anticipation of how students will perform in an external exam
10	10	Receiving a percentile rank 60 in a nationwide assessment, will yield a student's that is of national average
11	11	The result of a test in which a student received a scaled score of 196 with cut off score of 200 and standard error equal 6 should be ignored
12	12	Student's achievement is best reflected when grades are based on daily homework and chapter tests with point deducted for poor efforts
13	13	A student with scores 60th and 56th percentile in two different tests needs a remedial course on the test in which 56th percentile is scored
14	14	Teachers should encourage students to practice with items from an alternate form of a high stake test to improve their performance
15	15	Item analysis should be carried out on each item of a test to examine the validity coefficient
16	16	A student is believed to know 85% of the content covered in an instructional unit, if in a test scored using a 100-percent scale he/she scored 85
17	17	Norm- and criterion-referenced information is used when a teacher adjust his/her instruction based on pretest result
18	18	Grades assigned based on a single test could be biased against some minority students
19	19	The use of traditional assessment is involved when teacher Assigned grades based on classroom observation
20	20	It is unethical for a teacher to use previous grades to adjust current grades

21	21	Teachers should check for clarity of sentence construction when designing story-based mathematics test
22	22	The primary purpose for conducting formative assessment that involve teacher quick "check in" on students understanding is to identify cumulative knowledge
23	23	Utilizing a holistic scoring method minimize teacher subjectivity in scoring
25	24	A test item with difficulty value between 0.50 and 0.75 should be discarded
26	25	Portfolio grading system is used by teachers to detect students level of content mastery
27	26	Authentic assessment should be considered in testing application of what was learnt in the class
29	27	When comparison group is comprised of grade level peers, the general impression on a student's performance in a Standardized test is acceptable
30	28	Standardized test result can be used to plan instruction
33	29	Utilizing information from a variety of assessments when making decision about student learning is an appropriate use of assessment information
34	30	Students' problem- solving abilities is assessed using an assessment that is consistent with the content and skills they were taught
36	31	A student with a percentile rank of 60 on the problem-solving skills subset of a statewide assessment score above average
40	32	In anticipation of an external assessment teachers should plan so that it focuses on concepts and skills to be covered on the test
42	33	In a test scored using a 100-percent correct scale, a score of 85% means answering 85% of the items On the test correctly
44	34	Decisions like grades should be based on more than one piece of information
50	35	To plan classroom instruction, teachers should do a quick "check in" on the students to get an impression of their understanding
51	36	Grading all responses to essay question 1 before grading responses to essay questions 2 remove inconsistency in scoring

52	37	To ensure that standardized test results provide an accurate picture of what students really know, it is recommended that teachers clarify items that are confusing students
53	38	A diagnostic assessment is used to identify students specific difficulties
54	39	A test item with a discrimination value equal to -0.50 should be discarded
55	40	A restricted-response essay items that is concerned with students' demonstrating level of understanding of several specific criteria is best scored using analytic rubric
56	41	To enhance school improvement in standardized exam the teachers must ensure alignment between instruction and what is measured on the test
57	42	Criterion-referenced grading systems reflect each student's respective level of content mastery.
60	43	Raw scores are purely criterion-referenced and percentile Ranks are merely one form of norm-referenced scoring

NB: SA- Strongly Agree, A-Agree, D-Disagree, SD-Strongly Disagree

Research Question 2: What is the construct validity of the developed UTALS?

To answer this question, the Kaiser or eigenvalues greater-than-one criterion (Kaiser, 1960) and scree test, which involves an examination of a plot of the eigenvalues for breaks or discontinuities was used to determine the construct validity. To do this, the Spector (1992) Exploratory Factor Analysis (EFA) technique for studying the dimensionality of a scale was applied so as to explore the dimensionality of UTALS with the aim of determining (a) the number of factors that best represent the items and (b) the interpretation of the factors. From the initial eigenvalues as presented in Table 2, seven factors of university teachers assessment literacy emerged, which accounted for 86.457% of the total scale variance on the UTALS. The factor solution was in line with the initial assumption of the researcher (which was seven).

Table 2:
Eigenvalues and total variance on the UTALS

Component	Total	Initial Eigenvalues % of Variance	Cumulative %
1	9.747	20.668	20.668
2	8.174	17.010	37.678
3	6.583	13.310	50.988
4	6.362	12.796	62.784
5	5.847	11.598	73.382
6	3.868	8.996	81.378
7	2.184	5.079	86.457
8	.199	.463	
9	.032	.073	
10	.003	.007	
11	3.553E-016	8.262E-016	
12	3.237E-016	7.527E-016	
13	2.839E-016	6.603E-016	
14	1.979E-016	4.601E-016	
15	4.052E-017	9.422E-017	
16	2.192E-017	5.097E-017	
17	1.793E-017	4.170E-017	
18	1.187E-017	2.759E-017	
19	9.611E-018	2.235E-017	
20	6.703E-018	1.559E-017	
21	6.475E-018	1.506E-017	
22	4.906E-018	1.141E-017	
23	2.360E-018	5.488E-018	
24	1.156E-018	2.687E-018	
25	-1.308E-033	-3.043E-033	
26	-7.816E-033	-1.818E-032	
27	-2.906E-019	-6.758E-019	
28	-1.152E-018	-2.680E-018	
29	-2.598E-018	-6.041E-018	
30	-3.980E-018	-9.257E-018	
31	-6.596E-018	-1.534E-017	
32	-7.326E-018	-1.704E-017	
33	-9.678E 018	-2.251E-017	
34	-1.189E-017	-2.766E-017	
35	-1.532E-017	-3.564E-017	
36	-1.610E-017	-3.743E-017	

37	-2.117E-017	-4.923E-017
38	-9.122E-017	-2.121E-016
39	-1.628E-016	-3.785E-016
40	-2.603E-016	-6.053E-016
41	-2.727E-016	-6.341E-016
42	-4.292E-016	-9.980E-016
43	-5.879E-016	-1.367E-015

The standardized factor loadings for the 43 items were statistically significant at $p < .05$. Thus, the standardized item loadings of the UTALS items showed that the instrument is valid. Scree plot was also employed to further confirm the number of factors on which the UTALS items would load. The plot is as presented in Figure 1.
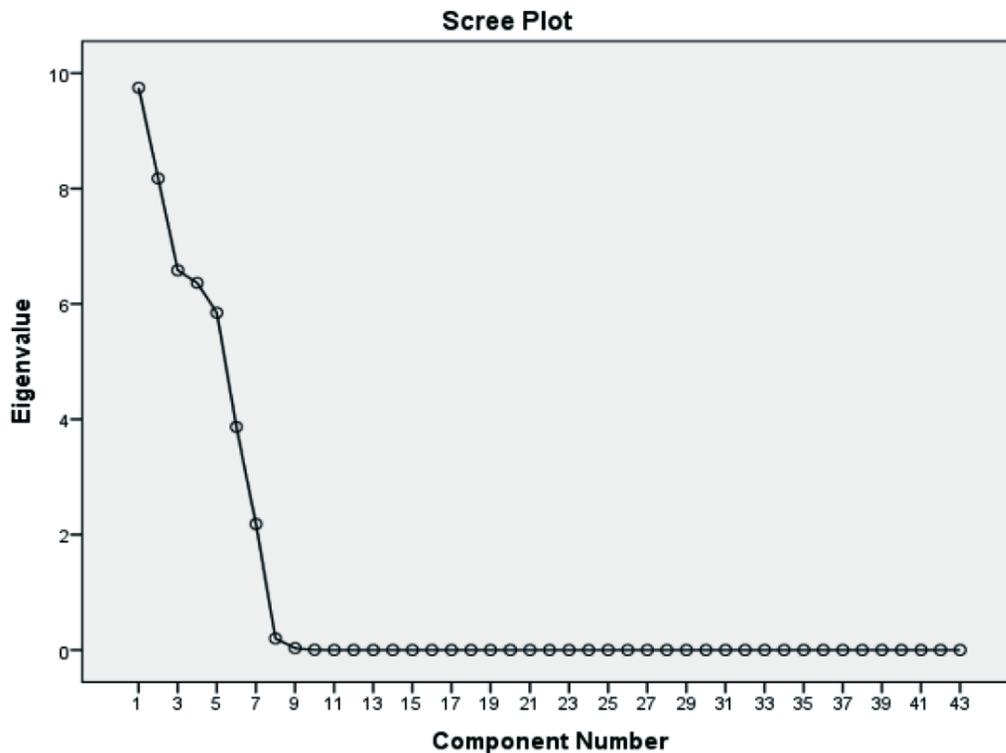

Figure 1: Scree plot showing seven factors on UTALS

The Scree plots in Figure 1 showed also seven factors on the UTALS and thus, confirm the number of factors in Table 2. Thus, there are seven factors on the developed University Teachers' Assessment Literacy Scale (UTALS) for measuring university teachers' assessment literacy.

Table 3:
The UTALS third version subscales and corresponding items

S/N	SUBSCALE	ITEMS
1	Developing Valid Grading Procedures	5, 12, 18, 23, 26, 44, 51, 57
2	Using Assessment Results	4, 11, 17, 22, 30, 50, 56
3	Administering, Scoring and Interpreting the Results	3, 10, 16, 36, 42, 55
4	Choosing Appropriate Assessment Methods	1, 8, 27, 29, 34, 53
5	Developing Appropriate Assessment Methods	2, 9, 15, 21, 25, 54
6	Communicating Assessment Result	6, 13, 19, 52, 60
7	Recognizing Unethical and Otherwise Inappropriate Assessment Methods	7, 14, 20, 33, 40,

From Table 3, it can be observed that Eight (8) of the 43 items of the UTALS loaded on factor 1 (Developing Valid Grading Procedures). It could therefore be concluded that Developing Valid Grading Procedures is the most important of the factors. Seven items loaded on factor 2 (Using Assessment Results), which makes it next most crucial to the first factor on UTALS, Six items loaded on factor 4 and 5(Choosing Appropriate Assessment Methods and Developing Appropriate Assessment Methods) and five items loaded on each of factor 6 and 7 (Communicating Assessment Result and Recognizing Unethical and Otherwise Inappropriate Assessment Methods). Thus, the 43 items on Table 1 were considered suitable and adequate for measuring university teachers' assessment literacy.

Research Question 3: What is the convergent validity of the UTALS?
To answer this question, scores from the Teacher Self-efficacy Scale (TSES)were correlated with those from the newly developed 43-item UTALS a related construct. Table 4 presents the result.

Table 4:
Convergent validity of UTALS

	N	R	p-value
TSES	551	0.747	<.01
UTALS	545		

From Table 4, the correlation coefficient between the two scales, UTALS and TSES, was 0.747, which is significant at 0.01 level of significance. Since the TSES is a widely used scale with a significant Cronbach's alpha that was found to be between 0.76, and 0.82 using three samples for a period of one year and for a period of two years it was found to be .65 (Adewolu, 2006). A high and positive correlation with it by the UTALS thus establishes the validity of the latter. By implication the two instruments are measuring the same traits or features. Hence, the UTALS does measure university teachers' assessment literacy in Osun State.

Research Question 4: What is the discriminant validity of the UTALS?
To answer this question, scores from the Teacher Burnout Scale (TBS)were correlated with those from the developed 43-item UTALS a related construct. Table 5 presents the result.

Table 5:
Discriminant validity of UTALS from Teacher Burnout Scale (TBS)

	N	R	p-value
TBS	523	-0.517	<.01
UTALS	545		

From Table 5, the correlation coefficient between the two scales, UTALS and TBS, was -0.517, which is significant at 0.01 level of significance. Since the TBS is a widely used scale with a significant subscales Cronbach's alpha of 0.89 for career satisfaction, 0.84 for perceived administrator support, 0.80 for coping with job related stress, and 0.72 for Attitudes towards students. A negative correlation with it by the UTALS thus establishes the discriminant validity of the latter. By implication the two instruments are not measuring the same traits or features. Hence, the UTALS does measure university teachers' assessment literacy in Osun State.

Research Question 5: What is the reliability of UTALS?
To answer this question, internal consistency reliability analyses were conducted on the 43-items final version of University Teachers' Assessment Literacy Scale using SPSS (version 20) to ascertain the general reliability of the scale. The results were as presented in Table 6.

Table 6:
Internal consistency reliability estimates of the UTALS

Reliability Type	Coefficient	No of Items
Cronbach Alpha	0.901	43
Spearman Brown (Split-half unequal Length)	0.963	43
Guttman Split-half Coefficient	0.962	43
Guttman	0.909	43

The results as presented in Table 6 showed that Cronbach Alpha coefficient of the UTALS was 0.901, while the Spearman Brown (split-half unequal length) coefficient was 0.963 and the Guttman coefficient was 0.909. These results are psychometrically satisfactory as opined by Devells (1991) cited by Adewolu (2006). Thus the UTALS can be considered generally reliable. However, the reliability of the sub-scales was also ascertained using internal consistency reliability estimates.

Table 7:
University teacher assessment literacy scale sub-scales reliability

Sub-scale	Cronbach's Alpha	No of Items
Developing Valid Grading Procedures	0.894	8
Using Assessment Results	0.799	7
Administering, Scoring and Interpreting the Results	0.814	6
Choosing Appropriate Assessment Methods	0.798	6
Developing Appropriate Assessment Methods	0.883	6
Communicating Assessment Result	0.818	5
Recognizing Unethical and Otherwise Inappropriate Assessment Methods	0.873	5

The result as presented in Table 7 showed that UTALS sub-factors estimated reliabilities are very high, indicating that the items were internally consistent and can be used to measure university teacher assessment literacy consistently.

Discussion

Sixty items were subjected to psychometric properties analyses. The 60-items UTALS were reduced to 43-items final version based on Govaerts and Gregoire (2008) item reduction criteria. This is because the items were selected during development have been

initially scrutinized by tests experts to also ensure they are measuring construct relevant domain which aided meeting the inclusion criterion. The validity of UTALS was examined through construct, convergent and divergent validity while the reliability was examined through internal consistency. A factor analytical technique, principal component analysis, was applied to determine the underlying component structure of the UTALS.

The principal component analysis results with eigenvalues greater-than-one showed that the 43 items that final emerged on the UTAL converge on seven distinct components. The loading of the items on the components were moderate to high, and all loadings but one were above 0.5. With the use of scree plot the seven factors on which UTALS loaded were confirmed. The seven factors are; Developing Valid Grading Procedures, Using Assessment Results, Administering, Scoring and Interpreting the Results, Choosing Appropriate Assessment Methods, Developing Appropriate Assessment Methods, Communicating Assessment Result and Recognizing Unethical and Otherwise Inappropriate Assessment Methods. All items were expected to correlate since the items were meant to represent the overall construct of university teachers' assessment literacy; the hypothesis that seven distinct components would emerge that lined up with each domain of assessment practices addressed in the UTALS was supported. This was in line with other studies that examine similar scales (Zhang, 1995; Burry-Stock & Frazier, 2008). With UTALS 43-items correlating into distinct components, the results suggested that the scale as a whole is a strong measure of university teachers' assessment literacy and thus, it allows researchers to examine and discuss university teachers' assessment literacy as a unidimensional construct. The items on the UTALS showed evidence of construct validity as the initial factor loadings on the data collected using T-CARS were statistically significant. These were good enough for declaring the UTALS usable for measuring the invisible believe that university teachers assessment literacy is capable of enhancing (or impeding) the success with which teaching task would be discharged.

The convergent and discriminant validity of the UTALS were determined by examining the relationship between total score on the UTALS and total score on 10-items TSES and 20-items TBS respectively. A coefficient of 0.747 and -0.517 for showed that the UTALS correlates highly with the TSES and TBS respectively and has acceptable convergent and discriminant validity.The results of the examination of the psychometric properties of the STAP demonstrated internal consistency reliability. Cronbach's alpha, Guttman's and Split-half coefficients respectively were 0.901, 0.909 and 0.962 for all 43 items. This is in line with the findings of Adewolu (2006) which shows that the UTALS has good internal consistency. This implies that the items on the scale have been constructed following the laid down procedures and that they are measuring what they are to measure consistently well among targeted respondents.

Conclusion

The 43-items UTALS, based on the analyses that were carried out could be adjudged reliable and valid for the measuring university teachers' assessment literacy. University teachers' skill in developing valid grading procedures and using assessment results remains the two most important factors. Based on the results, the following recommendations were made:

1. University management should therefore ensure that regular and suitable professional in-service development opportunities foracademic members of staff on measurement and evaluation considering their grade level.

2. University teachers themselves should endeavour to attend seminars, workshops and conferences relating to tests, measurement, assessment and evaluation.

3. a wider application of the UTALS should be explored to ascertain the commonality of the dominant factors among Nigerian university teachers.

4. UTALS could be standardized and published in commercial quantities for easy access to university management and academic staff members.

5. UTALS should be administered on teachers upon their entry into the profession to ascertain their assessment literacy level.

References
Adewolu, B. A. (2006). The development and validation of a teacher efficacy scale for Nigerian secondary school teachers. *Unpublished Ph.D. Thesis, Faculty of Education, Obafemi Awolowo University, Ile-Ife, Nigeria.*

American Federation of Teachers, National Council on Measurement in Education, & National Education Association. (1990). *Standards for teacher competence in educational assessment of students.* Washington, DC: National Council on Measurement in Education.

Anderson, T. (2004) Teaching in an online learning context, in Anderson, & F. Elloumi, (eds.), *Theory and practice of online learning.* Canada: Athabasca University, 273-294.

Australian National Training Authority (ANTA). (2002). *Assessment and online teaching: Australian flexible learning quick guide series.* Available at: http://www.flexiblelearning.net.au/guides/assessment.pdf

Burry-Stock, J. A., & Frazier, C. H. (2008, March)*. Revision of the assessment practice inventory (APIR): A combined exploratory factor analysis and polytomous IRT approach.* Paper Presented at the American Educational Research Association, New York, NY.

Conway, J. M., &Huffcutt, A. I. (2003). A review and evaluation of exploratory factor analysis practices in organizational research. *Organizational Research Methods*, *6,* 147168.

Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology, 78*(1), 98-104.

DeVellis, R. F. (2003). Scale development: Theory and applications (2nd ed.). Newbury Park: Sage Publications

Ewell, P. T. (2009). *Assessment, accountability, and improvement: Revisiting the tension,* NILOA Occasional Paper No.1, Urbana, IL: University of Illinois and Indiana University, National Institute for Learning Outcomes Assessment. Available At:http://www.learningoutcomeassessment.org/documents/ Peter Ewell_005.pdf

Govaerts, S.O. & Gregory J. (2008). Debt and construct validation of academic emotions scale. *International Journal of Testing,8*(1), 34-54.

Hannafin, M., Hill, J. R., Oliver, K.,& Glazer, E. (2003). Cognitive and learning factors in web based distance learning environments in M. G. Moore, and W. G. Anderson, (Eds.) *Handbook of distance education* Mahwah, NJ: Lawrence Erlbaum Associates.

Harlen, W., & Crick, R.D. (2003). Testing and motivation for learning. *Assessment in Education,10*(2), 169-207.

Likert, R. (1932). A technique for the measurement of attitudes*. Archives of Psychology*, *22*(140), 155.

Schwarzer, R., Schmitz, G. S., & Daytner, G. T. (1999). The teacher self-efficacy scale. [Online publication]. Available at: Http://www.fu-berlin.de/gesund/skalen/t_se.htm

Seidman, S. A. & Zager, J. (1987). The teacher burnout scale. *Educational Research Quarterly*,*11,* 26-33.

Spector, P. E. (2006). *Industrial and organizational psychology: Research and practice.* 4th edition. John Wiley & Sons, New Jersey.

Stiggins, R. J. (1995). Assessment literacy for the 21st century. *Phi Delta Kappan,   77.*

Suah, S. L. (2012). Analisis model literasidanamal an pentaksiran guru sekolahsertaka jiantentangjurangantarakeduanya (Analysis of literacy model and assessment practices   of school teachers and study of the culture gap) (Unpublished Ph.D. dissertation, School of Educational Studies, Universiti Sains Malaysia).

Volante, L. & Fazio, X. (2007). Exploring teacher candidates' assessment literacy: Implications for teacher education reform and professional development. *Canadian Journal of Education*, *30*(3), 749-770.

Worthington, R. L., & Whittaker, T. A. (2006). Scale development research: A content analysis and recommendations for best practices. *The Counselling Psychologist*, *34*, 806-838. https://doi.org/10.1177/0011000006288127

Zhang, Z. (1995). *Investigating teachers perceived assessment practices and assessment   competencies on the assessment practices inventory (API).* The University of Alabama. *ProQuest Dissertations and Theses,* 163 p. Http://search.proquest.com/docview/304160100?accountid=10639.