Standardization of Test for Assessment and Comparing of Students' Measurement

Patrick U. Osadebe[1]

[1] Department of Guidance and Counseling, Delta State University, Abraka, Nigeria

Correspondence: Patrick U. Osadebe, Department Of Guidance and Counseling, Delta State University, Abraka, Nigeria. Tel: 803-577-6610. E-mail: drosadebeuzo@gmail.com

Abstract

The study Standardized Economics Achievement Test for senior secondary school students in Nigeria. Three research questions guided the study. The standardized test in Economics was first constructed by an expert as a valid and reliable instrument. The test was then used for standardization in this study. That is, ensuring that the Economics achievement test is standardized. It was administered to 3,000 students using the same guidelines with no case of malpractice. The sex, location and school-type norm of students were considered for standardization. The measurement of students in form of Percentile rank, Z-score, T-score and Stanine statistics were used as derived scores to normalize the students' raw scores, using the knowledge of the normal curve as the theoretical base. This completed the process that made the test a Standardized Economics Achievement Test. The result shows that the test scores were normally distributed through the use of Percentile rank, Z-score, T-score and Stanine. Recommendations were made that the standardized test could be used to asses and compare the measurement of students from year to year.

Keywords: standardization, test, measurement

1. Introduction

Standardization of a test is a major area after test construction and administration. The test may be given to a small or large defined number of testees under the same guidelines or conditions. Then the test scores after administration are transformed using Percentile rank, Z-score, T-score and Stanine, among others. The transformed scores help to confirm or establish that a test is standardized. It could also be said that such test is normalized using the knowledge of the normal curve. Therefore, standardization is a process of ensuring that a test is standardized (Osadebe, 2001). It could also be define as the process of developing and implementing technical standard (Wikipedia, 2013).

There are so many studies on test construction and validation but few on standardization. This is a major problem in Nigeria. Therefore, there was the need to begin the process of standardization to ensure that the tests in Nigeria are standardized. There are lots of advantages when a test is standardized. A standard test is usually produced by experts and it is better than teacher made test. Then standardized test is highly valid, reliable and normalized with Percentile rank, Z-score, T-score among others derived scores to produce age norm, sex norm, location norm and school-type norm. These are in order of usage. Teachers can use standardization test with the manual (Ukwuije & Opara, 2012). It has been advised that teachers should consult experts in terms of test preparation and use (Osadebe, 2012). Generally, a standardized test could be used to assess, and compare students in the same norming group. In specific term, it was pointed out that comparative function is the primary role of standardized test (Nworgu, 2003; Thorndike & Hagen, 1977). Since it is valid and reliable, a standardized test could be used to assess students' behaviour. Similarly as a norm test, it could be used to compare students in the same norming group. Standardized test is not common in Nigeria, in spite of its numerous advantages. Therefore, there was the need to produce one in Economics through standardization process. The first process was to administer the test constructed by an expert to a large defined group under the same conditions or guidelines. The test is highly valid and reliable. The second process was to norm or transforms the raw score of testees using Percentile rank, Z-score, T-score and Stanine. These processes help to produce the standardized test in Economics. It should be noted that the same process of standardization may be applied to other subjects (Osadebe, 2004). Standardization involves the establishment of norms. The standardized test could be used to

compare students' scores or measurement from year to year. The scores for this study include percentile rank, z-score, t-score and stanine.

2. Literature Review

Standardization is required to standardize a test. Nworgu (2003) defines standardization as the uniformity of procedure in administering and scoring a test. It implies that scores obtained from different students under the same testing conditions can be compared. Thus, the process of developing and implementing technical standard is called standardization (Wikipedia, 2012). Osadebe (2001) defines standardization of a test as a process of producing a standardized test, and it evolves the establishment of norms. The first process is to obtain a test that is valid and reliable, and administer to small or large defined number of testees under the same conditions or guidelines. The second process is to transform or norm the raw score of testees using Percentile rank, Z-score, T-score and Stanine among others with the knowledge of the normal curve which forms the theoretical base of the study. These processes help to produce a standardized test. It could be used to assess students' behaviour and compare the students according to the norms. These could be in form of age norm, sex norm, location norm, school-type among others. The study determined sex norm, location norm and school-type norm. Norm is a reference group upon which the test is standardized. This could be achieved through the use of Percentile rank, Z-score, T-score and Stanine (Osadebe, 2001; Ohuche & Akeju, 1988).

In achievement test, when the final test has been composed, the next stage is standardization. Thus, "in standardizing achievement tests, one of the most important steps is the establishment of norms" (Nunnally, 1981, p. 264). Again, Aiken (1979) points out that for a test to be standardized, it must be administered with standard directions under uniform conditions to a sample of examinees representatives of the group for whom the test is intended. The purpose of this standardization procedure is to determine the distribution of raw scores in the standardization group (the norm group). These raw scores are then converted to some form of derived scores or norms such as age equivalent, grade equivalent, percentile rank or standard scores. Here, examinees' standing on a test may be evaluated by referring their raw scores to the norm table appropriate for their particular group. In this way, norms serve as a frame of reference for interpreting raw scores (Aiken, 1979).

In this study norms were expressed both in the form of percentile and standard scores (Z-score, T-scores and Stanine). It is also useful to have local norms such as norms based on samples of students in particular locality and in a particular school. Then the score of a particular student can be compared with scores of students across the country, students in the same locality and students in the same school. Thus, the norms published in test manuals are useful for comparing an examinee's score with those of sample of people from various localities, sometimes across section of the nation. But quite often a test administrator is more interested in demanding how well an examinee has done in comparison with the other students in the school or school system rather than a nationally selected sample. In these instances, the administrator will want to convert the raw scores of the particular school group to local norms. "Local norms are used quite often for selection and placement purposes in schools and classes" (Aiken, 1979, p. 52). It is also pointed out that "it is important to note that standardized tests are valuable when there is need to compare achievement between different schools and classes (Ughamadu, Onwuegbu, & Osunde, 1991, p. 58). Norms are very important in that they reveal how others have performed on a test and also enables the comparison of a student or students who at any time take the test with the reference groups or standardized sample.

Joe (1995) and Ughamadu et al. (1991) uphold the points of Gronlund (1976), Aiken (1979), and Osadebe (2012) on the types of test norm. These include grade norms, age norms, percentile norms and standard score norms. The names of the desired score for the grade norms, age norms, percentile norms and standard norms are grade equivalents, age equivalents, percentile ranks and standard scores respectively. But this study required the use of percentile rank and standard scores. In an empirical study on achievement test, Okonkwo (1995) used percentile and standards score norms of nine clearly defined groups. The groups are: rural boys, rural girls, urban boys, urban girls. Rural (boys & girls), urban (boys & girls), boys (rural & urban), girls (rural & urban) and the overall. This is similar to Osadebe (2001). The essence of many norms is to serve as variety of user's needs. Percentile and standard scores are commonly used in schools. These are the scores used in this study as measurement for comparing students' achievement from year to year. Measurement refers to use of score or number to describe individual behaviour (Osadebe, 2001).

Theory of the normal curve helps to explain the statistics required for standardization. The use of normal curve to describe Z-score, T-score, stanine and percentile rank has been widely supported by scholars. Some of these include Osadebe (2014), Nunnally (1981), Aiken (1979), Owen and Jones (1994), Weiss (1999), Anastasi (1976), Gronlund (1985), Cohen and Swerdlik (2002), Frank and Althoen (1994).

When a standardized test is subsequently administered to students, their scores will be normed and compared with the previous related norming group. In this way, the standardized test serves a comparative purpose. When the test is administered to students to judge their performance then it serves an assessment purpose. The students' performance or measurement could be compared from year to year with their standardization group.

3. Research Questions

The following research questions guided the study:

(1) What is the sex norm of students in the Economics Achievement Test?

(2) What is the location norm of students in the Economics Achievement Test?

(3) What is the school-type norm of students in the Economics Achievement Test?

4. Method

The study was based on standardization of a constructed Economics Achievement Test. The test was constructed by researcher who is an expert in measurement and evaluation, as non-standardized. It has a high content validity and a reliability coefficient of 0.95 estimated with Kuder-Richardson 20 as a measure of internal consistency. It was as a result of the need for standardized test that the researcher developed it through the standardization process. This process requires administration of the non-standardized test and the norming of testees' scores with Percentile rank, Z-score, T-score, and Stanine. The process helped to produce Standardized Economics Achievement Test for senior secondary schools in Nigeria.

There are various steps a researcher should follow to achieve the standardization of a test. The first is to obtain or develop a valid and reliable instrument. Here, the researcher is an expert in test construction had already developed an Economics achievement test as non-standardized. The next stage is to standardized and this became the focus of the study.

The second step is to administer the non-standardized test to a large group of testees or students to obtain their raw scores. The researcher administered the non-standardized test to a large group of students across sex, location and school-type of students and obtained their raw scores.

The third step is the use of statistics to norm the raw scores using Z-score, T-score, percentile rank and stanine. These are statistical norms. The researcher used the statistical norms to obtain sex, location and school-type norms. These norming groups will serve as frames of reference for comparing students' scores or measurement when another group of testees or students take the test in future.

A sample of 3,000 students was randomly selected for the administration or norming of the test. Sex norm, location norm, and school-type norm of students were determined through the use of Percentile rank, Z-score, T-score and Stanine. The Percentile rank was obtained by totaling all the frequencies below the particular score plus half the frequency at the score and dividing by the total number of cases then multiplies by 100. For the Z-score, the mean (62) and the standard deviation (9.86) of the distribution were obtained. The difference between each score and the mean was divided by the standard deviation to get the Z-score for each student. There was the need to convert Z-score to T-score to serve users who may need it. This also removed some negatives and decimals in the Z-score. This was computed as $T = 10z + 50$. Stanine has a mean of 5 and standard deviation of approximately 2. The Stanine used is from 1 to 9 with 1 as the lowest and 9 as the highest; with pre-assigned percentage in each of the nine values such as 4% = 9; 7% = 8; 12% = 7, 17% = 6; 0% = 5; 17% = 4; 12% = 3; 7% = 2; 4% = 1.

5. Results

Research Question One: What is the sex norm of students in the Economics Achievement Test?

Table 1. Sex norm (Male (M) = 1500, Female (F) = 1500). Analysis of percentile rank, Z-score, T-score and Stanine N = 3,000

Raw Score X	f	Sex		Percentile Rank $PR = \dfrac{100}{N}\left[CFb + f\left(\dfrac{X-LL}{i}\right)\right]$	Z-score $Z = \dfrac{x - \overline{X}}{SD}$	T-score $T = 10z+50$	Stanine S (9-1)
		M	F				
85	47	26	21	99	2.3	73	9
84	0	0	0	98	2.2	72	9
83	0	0	0	98	2.1	71	9
82	13	7	6	98	2.0	70	9
81	36	18	18	97	1.9	69	9
80	24	10	14	96	1.8	68	9
79	26	14	12	95	1.7	67	8
78	32	17	15	95	1.6	66	8
77	85	41	44	93	1.5	65	8
76	39	13	26	91	1.4	64	8
75	28	15	13	90	1.3	63	8
74	92	45	47	88	1.2	62	7
73	0	0	0	86	1.1	61	7
72	96	50	46	84	1.0	60	7
71	85	45	40	81	0.9	59	7
70	87	46	41	79	0.8	58	7
69	93	50	43	76	0.7	57	6
68	91	51	40	72	0.6	56	6
67	102	51	51	69	0.5	55	6
66	113	60	53	66	0.4	54	6
65	111	60	51	62	0.3	53	6
64	103	52	51	58	0.2	52	5
63	100	55	45	55	0.1	51	5
62	150	80	70	50	0.0	51	5
61	102	55	47	47	-0.1	49	5
60	145	77	68	42	-0.2	48	5
59	0	0	0	40	-0.3	47	4
58	155	80	75	37	-O.4	46	4
57	156	83	73	32	-0.5	45	4
56	105	56	49	28	-O.6	44	4
55	94	50	44	26	-O.7	43	4
54	95	50	45	21	-0.8	42	3
53	80	41	39	19	-0.9	41	3
52	32	17	15	17	-1.O	40	3
51	52	21	31	15	-1.1	39	3
50	101	41	60	13	-1.2	38	3

49	30	16	14	11	-1.3	37	2
48	82	29	53	9	-1.4	36	2
47	34	10	24	7	-1.5	35	2
46	33	11	22	6	-1.6	34	2
45	31	18	13	5	-1.7	33	2
44	33	10	23	4	-1.8	32	1
43	38	13	25	2	-1.9	31	1
42	0	0	0	2	-2.0	30	1
41	48	16	32	1	-2.1	29	1
40	1	0	1	0	-2.2	28	1

Table 1 indicates the raw scores and corresponding Percentile ranks, Z-score, T-score and Stanine on sex of SS III students in the Economics Achievement Test. The scores are normally distributed in line with the normal curve.

Research Question Two: What is the location norm of students in the Economics Achievement Test?

Table 2. Location norm (urban (U) = 2000, Rural (R) = 1000). Analysis of percentile rank, Z-score, T-score and Stanine. N = 3,000

Raw Score X	f	Location		Percentile Rank $PR = \dfrac{100}{N}\left[CFb + f\left(\dfrac{X - LL}{i}\right)\right]$	Z-score $Z = \dfrac{x - \overline{X}}{SD}$	T-score T	Stanine S (9-1)
		U	R				
85	47	33	14	99	2.3	73	9
84	0	0	0	98	2.2	72	9
83	0	0	0	98	2.1	71	9
82	13	9	4	98	2.0	70	9
81	36	25	11	97	1.9	69	9
80	24	17	7	96	1.8	68	9
79	26	18	8	95	1.7	67	8
78	32	22	10	95	1.6	66	8
77	85	60	25	93	1.5	65	8
76	39	27	12	91	1.4	64	8
75	28	20	8	90	1.3	63	8
74	92	64	28	88	1.2	62	7
73	0	0	0	86	1.1	61	7
72	96	67	29	84	1.0	60	7
71	85	59	26	81	0.9	59	7
70	87	61	26	79	0.8	58	7
69	93	65	28	76	0.7	57	6
68	91	64	27	72	0.6	56	6
67	102	71	31	69	0.5	55	6
66	113	79	34	66	0.4	54	6
65	111	78	33	62	0.3	53	6

Raw Score	f	S	M	Percentile Rank	Z-score	T-score	Stanine
64	103	72	31	58	0.2	52	5
63	100	70	30	55	0.1	51	5
62	150	76	74	50	0.0	51	5
61	102	70	32	47	-0.1	49	5
60	145	102	43	42	-0.2	48	5
59	0	0	0	40	-0.3	47	4
58	155	73	82	37	-O.4	46	4
57	156	75	81	32	-0.5	45	4
56	105	74	31	28	-O.6	44	4
55	94	66	28	26	-O.7	43	4
54	95	67	28	21	-0.8	42	3
53	80	56	24	19	-0.9	41	3
52	32	22	10	17	-1.O	40	3
51	52	36	16	15	-1.1	39	3
50	101	71	30	13	-1.2	38	3
49	30	21	9	11	-1.3	37	2
48	82	57	25	9	-1.4	36	2
47	34	24	10	7	-1.5	35	2
46	33	23	10	6	-1.6	34	2
45	31	22	9	5	-1.7	33	2
44	33	23	10	4	-1.8	32	1
43	38	27	11	2	-1.9	31	1
42	0	0	0	2	-2.0	30	1
41	48	34	14	1	-2.1	29	1
40	1	0	1	0	-2.2	28	1

Table 2 shows the raw scores and corresponding Percentile ranks, Z-score, T-score and Stanine on location of SS III students in the Economics Achievement Test. The scores are normally distributed in line with the normal curve.

Research Question Three: What is the school-type norm of students in the Economics Achievement Test?

Table 3. School-type norm (Single (S) = 1,800 Mixed (M) = 1,200). Analysis of percentile rank, Z-score, T-score and Stanine. N = 3,000

| Raw Score X | f | School Type | | Percentile Rank | Z-score | T-score | Stanine |
		S	M	$PR = \dfrac{100}{N}\left[CFb + f\left(\dfrac{X-LL}{i}\right)\right]$	$Z = \dfrac{x - \overline{X}}{SD}$	T = 10z+50	S (9-1)
85	47	30	17	99	2.3	73	9
84	0	0	0	98	2.2	72	9
83	0	0	0	98	2.1	71	9
82	13	8	5	98	2.0	70	9
81	36	21	15	97	1.9	69	9
80	24	15	9	96	1.8	68	9

79	26	18	8	95	1.7	67	8
78	32	21	11	95	1.6	66	8
77	85	50	35	93	1.5	65	8
76	39	21	18	91	1.4	64	8
75	28	18	10	90	1.3	63	8
74	92	54	38	88	1.2	62	7
73	0	0	0	86	1.1	61	7
72	96	67	39	84	1.0	60	7
71	85	49	36	81	0.9	59	7
70	87	50	37	79	0.8	58	7
69	93	55	38	76	0.7	57	6
68	91	54	37	72	0.6	56	6
67	102	57	45	69	0.5	55	6
66	113	59	54	66	0.4	54	6
65	111	68	43	62	0.3	53	6
64	103	52	51	58	0.2	52	5
63	100	60	40	55	0.1	51	5
62	150	76	74	50	0.0	51	5
61	102	56	46	47	-0.1	49	5
60	145	108	37	42	-0.2	48	5
59	0	0	0	40	-0.3	47	4
58	155	85	70	37	-O.4	46	4
57	156	87	69	32	-0.5	45	4
56	105	70	35	28	-O.6	44	4
55	94	64	30	26	-O.7	43	4
54	95	55	40	21	-0.8	42	3
53	80	50	30	19	-0.9	41	3
52	32	22	10	17	-1.O	40	3
51	52	35	17	15	-1.1	39	3
50	101	69	32	13	-1.2	38	3
49	30	20	10	11	-1.3	37	2
48	82	56	26	9	-1.4	36	2
47	34	24	10	7	-1.5	35	2
46	33	20	13	6	-1.6	34	2
45	31	11	20	5	-1.7	33	2
44	33	21	12	4	-1.8	32	1
43	38	23	15	2	-1.9	31	1
42	0	0	0	2	-2.0	30	1
41	48	30	18	1	-2.1	29	1
40	1	1	0	0	-2.2	28	1

Table 3 indicates the raw scores and corresponding Percentile ranks, Z-score, T-score and Stanine on school-type of SS III students in the Economics Achievement Test. The scores are normally distributed in line with the normal curve.

6. Discussion

The results revealed that Percentile rank, Z-score, T-score and Stanine were suitable for standardization of the Economics Achievement Test. The norms were normally distributed. This is an indication that the Economics Achievement Test has been standardized. The results agreed with that of Osadebe (2001). The use of Percentile rank, Z-score, T-score and Stanine disagreed with Osadebe (2014) who supported Z-score and T-score for the normal distribution of scores.

The percentile rank used in the study is in line with Angoff (1976), Aiken (1979), Joe (1995), and Ukwuije (1996). The use of Z-score and T-score are as approved by many experts including Aiken (1979) and Nunnally (1981). The result shows a normal distribution of scores. The stanine scores used are as recommended by Aiken (1979), Ohuche and Aiken (1988) and Ukwuije (1996). The result shows a normal distribution of SS III students' scores.

7. Conclusion

The study focused on standardization, which is the process of standardizing or norming a test. The test used for standardization was constructed by an expert as non-standardized Economics Achievement Test. It was highly valid and reliable before use. It was as a result of the need to standardize the test that the study was carried out.

The test was administered to 3,000 senior secondary three (SS III) students for standardization. The raw score of students were normed or transformed with Percentile rank, Z-score, T-score and Stanine. The scores were normally distributed using the knowledge of the normal curve. The test was normed into sex norm, location norm and school-type norm. This completed the process of standardization that helps to produce a standardized Economics Achievement Test for senior secondary school students. The test could be used to assess the students and compare their measurement from year to year.

8. Recommendations

It was recommended that the standardized Economics achievement Test could be used from year to year to assess students and compare their achievement with the norming group. Users of the test should derive their local norms with Percentile rank, Z-score, T-score and Stanine. The test should only be administered to students when they have completed the content areas of ordinary level Economics curriculum. The test will help the students to prepare for external examinations. The method used for standardization in this study may be applied to other subject areas. Most examination bodies keep secret the process of standardizing their tests. This may be due to security or for commercial purpose where tests are sold to users. This study has provided a simple procedure for standardization for others to follow. It is recommended that the test should be used with the manual.

Average/Midpoint

Lower High

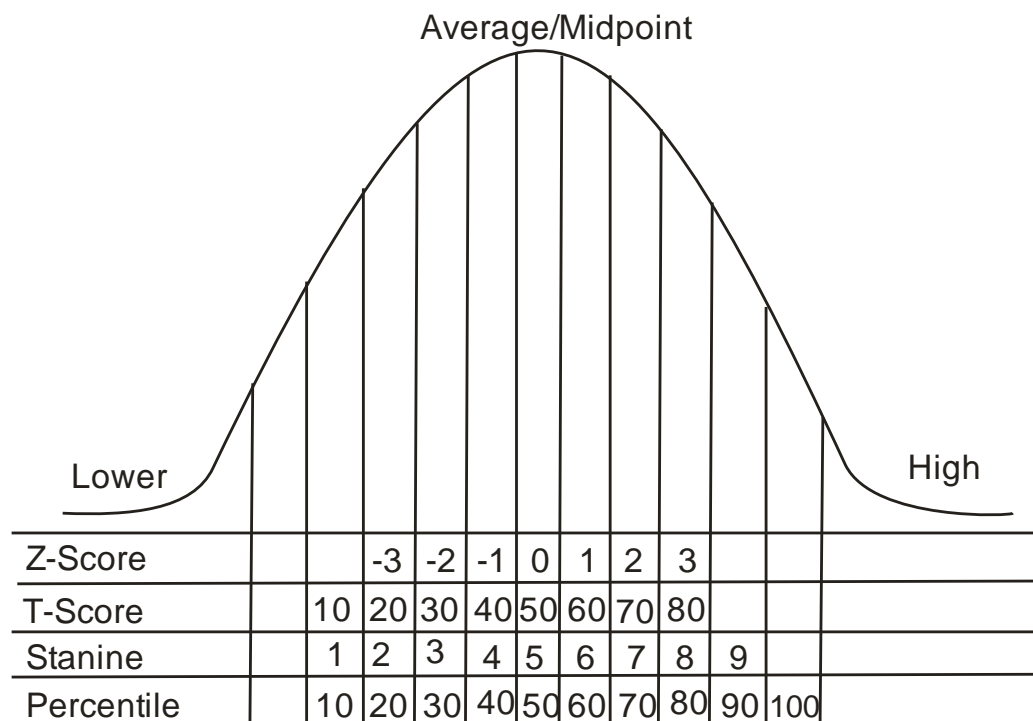|										
Z-Score		-3	-2	-1	0	1	2	3		
T-Score	10	20	30	40	50	60	70	80		
Stanine	1	2	3	4	5	6	7	8	9	
Percentile	10	20	30	40	50	60	70	80	90	100

Figure 1. Normal curve with Z-Score, T-Score, stanine, and percentile

Figure 1 shows the normal curve with Z-score, T-score, stanine and percentile rank. Z-score has a mean of 0 and standard deviation of 1. The T-score has a mean of 50 and standard deviation of 10. Stanine has a mean of 5 and standard deviation of approximately 2. Percentile rank of scores is the percentage of scores below or at the mid-point of a given score. The 50th percentile is the mid score or median based on 100. These were bases upon which raw scores were transformed or normed for the study.

References

Aiken, L. R. (1976). *Psychological Testing and Assessment*. Boston: Allyn and Bacon, Inc.

Anastasi, A. (1976). *Psychological testing*. New York: Macmillan Publishing Co. Inc.

Angoff, W. H. (1976). *Scales, Norms and Equivalent Scores*. Washington: American Council on Education.

Cohen, R. J., & Swerdlik, M. E. (2002). *Psychological Testing and Assessment*. McGraw-Hill companies, Inc.

Frank, H., & Altheon, S. C. (1994). *Statistics: concepts and application. Cambridge*. Cambridge University Press.

Gronlund, N. E. (1976). *Measurement and Evaluation in Teaching* (3rd ed.). New York: Macmillan Publishing Co. Inc.

Gronlund, N. E. (1985). Measurement *and Evaluation in Teaching*. New York: Macmillan Publishing Company.

Joe, A. I. (1975). *Basic Concepts of Educational Measurement and Evaluation*. Port Harcourt: Paragraphics.

Nunnally, J. C. (1981). *Psychometric Theory*. New Delhi: Tata McGraw-Hill Publishing Company.

Nworgu, B. G. (2003). *Educational Measurement and Evaluation Theory and Practice*. Nsukka: University Trust Publishers.

Ohuche, R. O., & Akeju, S. A. (1988). *Measurement and Evaluation in Education*. Ibadan: Africana-Feb Publishers Ltd.

Okonkwo, S. C. (1995). *Development and Standardization of a readiness test for Junior Secondary School Mathematics* (Unpublished Doctoral dissertation, University of Nigeria, Nsukka).

Osadebe, P. U. (2001). *Construction and Validation of Test*. A seminar paper presented at the University of Port Harcourt.

Osadebe, P. U. (2012). *Procedures for Construction, Validation and Standardization of Test*. A seminar paper presented at Delta State University, Abraka.

Osadebe, P. U. (2014). Assessment Score of University Lecturers. *Journal of Education and Practice, 5*(2), 8-14.

Owen, F., & Jones, R. (1994). *Statistics*. London: Longman Group UK Limited.

Thorndike, R. L., & Hagen, E. P. (1977). *Measurement and Evaluation of Psychology and Education*. New York: John Wiley and Sons.

Ukwuije, R. P. I., & Orluwene, G. W. (2012). *Peanuts Educational Statistics*. Port Harcourt: Chadik Printing Press.

Ukwuije, R. P. I. (1996). *Test and Measurement for Teachers*. Port Harcourt: Abe Publishers.

Ukwuije, R. P. I., & Opara, M. I. (2012). *Test and Measurement for Teachers*. Port Harcourt: Chadik Publishers.

Weiss, N. A. (1999). *Introductory statistics*. USA: Addison Wesley Longman Inc.

Wikipedia. (2012). *Standardization*. Retrieved from http://en.wikipedia.org/wiki/Standardized_test