DEVELOPMENT AND VALIDATION OF OPPORTUNITY TO LEARN MATHEMATICS SCALE USING ITEM RESPONSE THEORY

Loretta Chibuzor AUDU.
auduloretta@yahoo.com
+2348063277270.

Georgina. N. OBAITAN
gnobaitan@yhoo.com
+2348061197902.

Joshua O. Adeleke
joadeleke@yahoo.com
+2348033510688

Institute of Education, University of Ibadan

Abstract
This study is a correlation study aimed at applying Item Response Theory (IRT) to validate Opportunity To Learn Mathematics (OTLM) scale. Six (6) research questions were set to guide the study. Opportunity to learn was theoretically defined along five components. A survey approach was employed to collect responses on OTLM. The sample consists of one thousand, nine hundred (1900) students of SS3 from selected 36 secondary schools in Imo and Rivers states of Nigeria. The selection was done through a multi-stage cluster sampling techniques. A pool of one hundred and sixty-two (162) items was generated. The analysis of data was carried out using exploratory factor analysis (EFA), Confirmatory factor analysis (CFA) and Graded Response Model (GRM) of IRT. The reliability of the subscales was ascertained using Item-Total statistics and the estimated reliability coefficient of the subscales were as follows: Teacher effectiveness (0.91), Curriculum content/Learning experience (0.85), Instructional time (0.72), Facilities (0.78) and Learning materials (0.73). Graded response theory model (GRM) of IRT was employed for the IRT analysis and 82 items were selected and calibrated, items and scale properties were estimated and produced. None of the items has less than 0.5 discriminating index. Thus, OTLM scale is both valid and reliable and therefore recommended for use for measuring students' mathematics learning opportunities in secondary schools.

Keyword: Development, Validation, Opportunity To Learn Mathematics, and Item Response Theory.

Background

Mathematics is an indispensable subject in the school curriculum. The importance accorded mathematics in the school curriculum from the primary to the secondary levels reflects accurately the vital role played by the subject in contemporary society. According to Yara (2009), every individual needs the knowledge of mathematics in order to live a useful life and to be an effective member of the society. The learning of mathematics in schools represents first, a basic preparation for adult life and secondly a gateway to a vast array of career choices.

The knowledge of mathematics as a subject affects all aspects of human life in various degrees. The social, economic, political, geographical, scientific and technological aspect of man is centered on numbers. Mathematics is seen as the language used to describe the problems arising in most branches of science and technology (Yara, 2009). The progress of any nation depends upon her scientific and technological advancement which can only be built on a sound mathematical education capable of making the citizens as effectively functional in the natural and applied sciences. In fact, the importance of mathematics cannot be overemphasized, that is why a credit pass in mathematics is a pre-requisite for securing admission into Nigerian universities.

In realization of the importance of mathematics in the school curriculum, many countries have resort to making special comprehensive and well-programmed efforts towards the effective teaching and learning of science and mathematics at all levels of the educational system through the development and implementation of innovative programmes and projects (Odili,2007). Various researches had been undertaken to investigate trends in mathematics achievement and the factors influencing mathematics learning and performance (House and Telese, 2008). Unfortunately, despite all successive efforts to improve the teaching and learning of mathematics, it has been observed that students performance in mathematics is dismal, Obaitan and Rasheed (2010). There is a general public outcry about the poor performance in secondary school mathematics. Various reasons have been linked with this poor performance and various programmes had been carried out in order to solve these problems. Stakeholders in learning outcomes in this area have suggested various instructional strategies that could improve students' performance in mathematics. Improving instruction in mathematics education has been a major topic of interest for teachers, researchers, administrators and public (Popoola 2004).

Unfortunately, the ideal strategies and techniques capable of ameliorating these problems seem to be difficult to attain in our schools today. The poor state of mathematics in the country was brought to a sharp focus and was partly seen as a long-term effect of playing down on the opportunity to learn mathematics provided in our secondary schools such as having access to qualified and competent teachers, adequate facilities, adequate curriculum content, adequate instructional time, adequate learning materials and others provided in our secondary schools. Kurz (2011) reports that our nation's test and standard-base school accountability system is very limited because the people it holds most responsible-students and teachers, have little control over learning opportunities that really matter. For instance, teachers cannot correct overcrowded classrooms and students cannot insist on having mathematics teachers who are fully credentialed and competent. Parents, policy-makers, education officials and other civic leaders can all share the responsibility for guaranteeing opportunities to learn in the schools.

Opportunity to Learn (OTL) refers to equitable conditions or circumstances within the school or classroom that promote learning for all students. It includes the provision of curricula, learning materials, facilities, teachers and instructional experiences that enable students to achieve high standards. OTL also refers to the absence of barriers that prevent learning, (NCTE, 2012). According to Meyers (2010), OTL is a way of measuring and reporting whether students and teachers have access to different ingredients that make up quality schools, school district or even schools across the nation. To him, the more OTL ingredients that are present in an individual school, the more opportunities students have to benefit from high quality education. By measuring and reporting the presence or absence of learning opportunities against a set of standards, OTL can bring to light examples of unfair conditions that limit students' equal access to a high quality education (Kurz, 2011). OTL standards can also help students, parents, communities and school officials to discover and correct problems in schools. Some examples of OTL include students' access to: qualified and competent teachers; clean and safe facilities; up-to-date books and quality learning materials; high quality course work; school conditions that provide students a fair and equal opportunity to learn and achieve knowledge and skills generally and in mathematics in particular.

Opportunity to Learn Mathematics (OTLM) scale when developed would help to look into key issues in the teaching and learning of mathematics: Are students in classes where the Mathematics content is taught? For example, if a school does not have mathematics laboratory, students cannot learn some mathematics concepts. In other words, they do not have the "opportunity to learn some mathematics concept." Do students spend enough time with the content or subject matter for their level? For example, if students are forced to learn in a class that does not spend enough time on a school subject, they do not have the opportunity to gain deep knowledge about the subject. If students are forced to learn in a year-round calendar where the school year is cut short because of school overcrowding, they have less time or opportunity to learn at all. If students do not have current mathematics books or have books with missing pages or have no books at all in the classrooms to take home then, they do not have the opportunity to learn current mathematics. If a school does not have computers available, students cannot do research on the internet. Do the mathematics teachers have the knowledge and re-training to be effective? For example if teachers have only basic knowledge or training, they cannot answer advanced questions or teach certain concepts well. This in turn limits students' opportunities to learn; Are the school facilities adequate, healthy and not jam-packed? For example, if the classrooms are too hot, cold or dirty, or if students are forced to take classes in packed rooms, they cannot learn well. How does OTL affect teachers? Just like the students need opportunity to learn, mathematics teachers should also need good working conditions to do their best job at teaching, i.e. by having basic tools such as books, laboratories, libraries and other adequate facilities, ULCA (2003).

In other words, students and teachers are held accountable for meeting the performance standards, opportunities to learn have not been a part of this standard-based system. As a result, it is difficult, if not impossible, to measure student's performance accurately and fairly if there is no information available about whether they had a chance to learn in their schools. For standard-based school accountability system to be accurate, useful and fair, OTL standards must be included along with performance standards.

The International Association for the Evaluation of Educational Achievement (IEA) studies considers the processes and effects of education using the notion of opportunity to learn (OTL) in order to understand the linkages between: Intended curriculum (what policy require); Implemented Curriculum (what is taught in schools); Achieved Curriculum (what students learn).

The third mathematics and science study (TIMSS) emphasized the fact that the extent and quality of school resources are critical for quality instruction (Lee and Zuze, 2011). It showed that students in schools that are well resourced generally have higher achievement than those in schools where there was shortage of resources. The teacher's methods and materials as well as other components of OTL have a direct causal relationship with school achievement. Therefore, any attempt to ameliorate the dismal performance in mathematics in Nigeria, must first of all address the issue of the OTL mathematics that is provided by our schools.

This of course entails ensuring the quality and quantity of OTL materials present in our schools. To this end, we need valid and reliable instruments for measuring the OTL mathematics. There are two major procedures for developing such instruments - the classical test theory (CTT) and item response theory (IRT).

Classical test theory (CTT) describes how error can influence observed scores or measurement. It is based on the true score theory which introduces three concepts – test scores (often called observed score), true score (T) and error score (E). This is often expressed mathematically as X =T+E. The true score, T reflects whether the examinee's amount of knowledge or ability is the true measurement of the examinee which always contaminated by random errors. According to Ojerinde (2013), these random errors can result from several factors such as guessing, fatigue or stress. The observed score is often called a fallible score because of the error contaminant. The true score is the score that would have obtained if there were no error in measurement. CTT operates under three main assumptions (i) the error and the true scores from the same test have a correlation of zero (ii) the error terms have an expected mean of zero and (iii) the errors from parallel measurements are uncorrelated.

CTT therefore has the following limitations:

1. The item statistics such as item difficulty and item discrimination depend on particular examinee samples from which the test was administered. Item parameters are not invariant characteristics of item, but take on values that depend on who tried the items.
2. The definition of reliability in CTT is established through the concept of parallel tests which is difficult to achieve in practice because individuals are never exactly the same on a second administration.
3. CTT has to do with the assumption that standard error of measurement is the same for all subjects and does not take into account variability in error at the different trait level (Hambleton and Swaminathan, 2010).
4. CTT reflects the focus on test level information to the exclusion of item level information.CTT therefore deals with individual total score and not their ability at the individual item level.

Item Response Theory (IRT) provides an alternative to CTT as a basis for examining the relationship between item responses and the ability of an examinee being measured by the test or the scale (Hambleton and Swaminathan, 2010). IRT attempts to model the ability of an examinee and the probability of answering an item correctly based on the pattern of responses to the items that constitute a scale. IRT is able to estimate the parameters of an item independent of the characteristics of both test takers to which it is exposed and other items that constitute the scale. The individual trait level is often designed by theta (θ), which represents the amount of ability, trait or attribute level possessed by an individual. The three parameters associated with the item are discrimination power (a) the difficulty parameter (b) the guessing parameter (c).IRT operates under these three main assumptions:

1. Unidimensionality which assumes that the scale is measuring only and only one construct or ability.
2. Local independent which assumes that the individual response to items is independent of each other and estimation of item parameters are also independent of each other.
3. Model fit which assumes that the model fits the data.

IRT has a lot of advantages over CTT that warrant its usages. CTT scale score is the sum of all the items in the scale and so it is not accurate measure of an individual's ability. The scale score in IRT has a major advantage over CTT, it estimates individual's latent trait level scores based on all the information in a participants response pattern, it takes into consideration which items were answered correctly and which ones were answered incorrectly, and utilizes the difficulty and discrimination parameters to the items when estimating ability levels. This gives a more accurate estimate of ability.

The problem of reliability in CTT is addressed by the information function in IRT; while CTT yields only a single estimation of reliability and corresponding standard error of measurement, IRT provides a test information function and test standard error function to index the degree of measurement precision across the full range of the latent trait construct. An instrument can be evaluated in terms of the amount of information and precision they provide at specific ranges of test scores that are of particular interest. This feature can be used for selection of quality items for particular purposes. High level of precision is usually in the middle of the scale where information is high and low ends of the scale where information is low (Hambleton & Swaminathan 2010; DeVellis, 2012; Embreton & Reise, 2010; Reeve, 2005). Through information function the test developer can precisely asses the contribution of each item to the precision of the total test and hence choose items in a manner that is not contradictory with the aspect of test construction.

One of the interesting features of IRT is that the item parameters are not dependent upon the ability level of the examinee's responding to the items like the CTT. If two groups of examinee are drown from the same population of examinee with different ability levels, the two groups will yield the same values of the item parameters. Hence, item parameters are group invariance. This item invariance principle has the importance of practical consequence that the parameters of large numbers of items can be estimated, even though each item is not answered by every examinee (item calibration). IRT puts all individual scores on standardized interval level scale while CTT often use ordinal scale of measurements for test.

Ability estimation is also invariance with respect to the items used to determine it; the principal rest on the condition that all items are measuring the same construct and the values of all the item parameters are in a common metric. The practical implication of this principle is that a test located anywhere along the ability scale can be used to estimate examinee ability, whether the test items are easy or hard. This makes it possible to compare individual's results from different versions of a test, which is test equating. This principle also resulted to adaptive testing where different individuals are administered different test according to their ability levels.

Based on these advantages IRT has over CTT and also based on the other issues discussed above, the researcher therefore deemed it fit to apply item response theory (IRT) in the development and validation of opportunity to learn mathematics (OTLM) scale.

Objectives of the Study

The major area to be addressed in the issue of poor performance in mathematics is that of the opportunity to learn mathematics provided by schools. This of course entails assessing the quantity and quality of OTL mathematics present in our schools. This may be a prelude to working out modalities for improving on them. The objective of this research therefore is to carefully construct a standard, valid and reliable instrument that will be used to measure OTL mathematics. The instrument will be validated by applying the procedures of item response theory (IRT) in order to get accurate estimate of ability.

Research Questions

1) What is the construct validity of the OTLM scale?
2) How consistent are the factor loadings of the OTLM scale items?
3) How reliable are the subscales (Teacher effectiveness, Curriculum content/Learning experiences, Instructional Time, Facilities and Learning materials)?
4) (a) What is the spread of the category difficulties? (b) How discriminating is each item?
5) Do the OTLM scale items conform to goodness of fit model?
6) Do the OTLM scale items conform to Differentiate among the respondents based on gender and school type?

Methodology

The study adopts correlation research design which aimed at constructing opportunity to learn mathematics scale for secondary school students using item response theory (IRT). A sample survey was used to collect information from students on mathematics learning opportunities. The target population for the study comprised all senior secondary three (SS3) students in unity, public (Government) and private schools in Imo and Rivers states. Senior secondary three (SS3) students were chosen for this study because being at the final year in the secondary school, they can easily give account of the learning opportunities they had in mathematics during their secondary school period. The sample for the study comprised one thousand, nine hundred (1900) students. The sample was selected using multi-stage cluster sampling procedure. Two states were randomly selected from the eleven states of South-South and South-Eastern zones. Two LGA were purposefully selected from each state. Purposive sampling technique was employed to ensure that all the Federal Government Schools in each state are included in the sample. This is to give a good spread of respondents from all the 36 states in the country. Nine (9) schools were selected from each L.G.A. This comprised one (1) Federal Government School, Five (5) State Government schools selected at random and three (3) private schools also selected randomly. This gave a total of thirty six (36) schools selected for the study. A simple random technique was used to select four hundred and seventy five (475) students from each L.G.A (120 students from 1FGS, 50 students from each of the 5 SGS and 35 students from each of the 3 PS). Thus, a total of one thousand nine hundred (1900) students were selected for the study. Items for the instrument were generated to cover the components of opportunity to learn mathematics (OTLM) which are as follows: curriculum content/Learning experience, teacher effectiveness, instructional time, facilities and learning materials. The items for the OTLM scale were generated from three (3) sources. The students' responses to the open ended questionnaire were based on the five components of OTLM. The selection of the items was done through empirical criterion key. Items were retained if more than twenty percent of the respondents listed it in their response to an open ended questionnaire; review of related literature and from the researcher, based on one of the researchers experience as a secondary school mathematics teacher. A four (4) point Likert scale was developed using these items. A score of four (4) indicated the maximum possible positive score for an item while a score of one (1) was assigned the least possible negative response. The instrument consist of one hundred and seventy-six items generated based on the components of OTLM. The generated items were given to two experts to vet on clarity of words, simplicity of standard, grammatical error and also to ascertain the content validity of the instruments. A3- point rating Likert scale was given to the test experts to ascertain if each item is a measure of OTL; A 3 is assigned if it is a measure, a 2 if it is partially a measure and a 1 if not a measure. The experts also indicated in a 3-point scale if each components and the whole scale is adequately covered and to make general comments on the scale. Based on the ratings and general comments of these experts, some of the statements were modified and some omitted. This reduced the number of instrument to one hundred and sixty two. The instrument for data collection consists of one hundred and sixty two (162) items on opportunity to learn mathematics. It comprises two sections: section A and section B. Section A is designed to elicit personal information from the respondents such as age, gender, school type: government, private or unity school. Section B contains the items with 4-point Likert scale ranging from strongly agree to strongly disagree and as well as instructions on how to respond to questionnaire. The researcher and two trained research assistants administered and collected data in all the selected schools in Imo and Rivers States. The data analysis was in two (2) phases viz: Exploratory factor analysis to (a) verify the assumption of unidimensionality of the items and construct validity (b) ascertain the consistency of the factor loadings and (c) the reliability of the subscale of OTLM scale; IRT analysis in which the statistics for all the calibrated items, total score, theta estimation and overall model fit were estimated using IRT PRO version 3.0 software.

Results and Discussions

Out of 162 items, only 82 items satisfy the requirement for unidimensionality and construct validity using 0.3 as cut off mark for factor loading. Hence, the final OTLM scale consists of 82 items. The estimated reliability coefficient of the new scale was 0.90.

What is the Construct Validity of the OTLM scale?

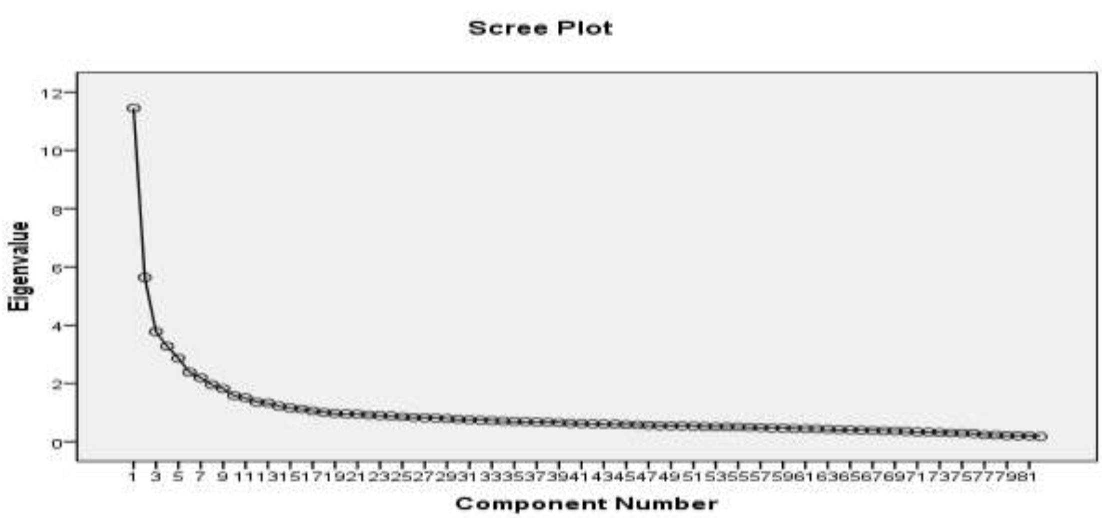

Figure 1: Scree plot indicating Unidimentionality of the items

Table 1: Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	11.449	13.963	13.963	11.449	13.963	13.963
2	5.637	6.875	20.837	5.637	6.875	20.837
.	.	.	.	.	.	.
.	.	.	.	.	.	.
29	0.792	0.966	68.759			

Table 1 and Figure 1 reveals the unidimensionality of the 82 items that constitute OTLM scale. The items meaningfully loaded to components 1and2 measuring Teacher and curriculum dimensions of OTLM scale. The estimated eigen value on component 1 was 11.45 and percentage variance accounted for was 13.96%, while the estimated eigen value on component 2 was 5.64 and percentage variance accounted for was 6.88%. These established the fact that the 82 items measure the construct, opportunity to learn Mathematics meaningfully. Hence, it can be inferred that the scale comprising the 82 items possesses construct validity

How consistent are the factor loadings of the OTLM scale items?

Table 2: Extractions of Principal Component and Verimax Rotation factor Loading

S/N	Principal Component			Verimax Rotation		
	ITEM	Factor1 Loadings	Factor2 Loadings	ITEM	Factor1 Loadings	Factor2 Loadings
1.	TCHR1	0.409		TCHR1	0.592	
2.	TCHR3	0.443		TCHR3	0.456	
3.	TCHR5	0.388		TCHR5	0.586	
4.	LEMAT153	0.352				
.	.	.	.	.	.	.
.	.	.	.	.	.	.
82	LEMAT154	0.413				

Table 2 presents the loadings of the 82 items using two approaches (principal component and verimax rotation) to ascertain the consistency of the loadings. Unrotated principal component exploratory factor analysis approach was found best fit to establish the unidimentionality of OTLM scale. The approach provides the loadings that were highly consistent in estimating the unidimentionality of the OTLM scale.

How reliable are the subscales (Teacher, Curriculum Content/Learning experience, Instructional Time, Facilities and Learning Materials) of OTLM scale?

Table 3a: Item-Total Statistics on Teacher related items

	Item-Total Statistics			
	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted

TCHR1	103.76	327.87	0.43	0.91
TCHR3	103.69	329.05	0.38	0.91
TCHR5	103.82	328.28	0.41	0.91
TCHR6	104.04	327.49	0.41	0.91
.	.	.	.	.
.	.	.	.	.
TCHR49	104.21	328.71	0.32	0.91
Cronbach's Alpha = .91				

Table 3a shows that the coefficient range of the corrected item-total correlation was 0.26 to 0.60. This establishes construct validity of the teacher related items of OTLM scale. The estimated reliability coefficient on the subscale was 0.91. Conclusion can be drawn that 38 items that constitute Teacher related subscale of OTLM scale is valid and reliable to measure the expected dimension of Opportunity to Learn Mathematics scale.

Table 3b: Item-Total Statistics on Curriculum Content/Learning experience related items

		Item-Total Statistics		
	Scale Mean if Item Deleted	**Scale Variance if Item Deleted**	**Corrected Item-Total Correlation**	**Cronbach's Alpha if Item Deleted**
CURRLE52	62.86	209.91	0.38	0.85
CURRLE56	62.97	207.94	0.40	0.85
CURRLE58	63.41	200.97	0.50	0.84
.	.	.	.	.
.	.	.	.	.
CURRLE90	64.43	206.77	0.38	0.85
Cronbach's Alpha= .85				

Table 3b shows that the coefficient range of the corrected item-total correlation is 0.29 to 0.53. This establishes construct validity of the curriculum content/Learning experience related items of OTLM scale. The estimated reliability coefficient on the subscale was 0.85. Conclusion can be drawn that 23 items that constitute curriculum content/learning experience related subscale of OTLM scale is valid and reliable to measure the expected dimension of opportunity to learn mathematics subscale. By measuring and reporting the presence or absence of learning opportunities against a set of standards, OTL can bring to light examples of unfair conditions that limit students' equal access to a high quality education (Kurz, 2011).

Table 3c: Item-Total Statistics on Instructional Time related items

		Item-Total Statistics		
	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
INSTRTIME94	25.04	23.456	0.465	0.677
INSTRTIME95	25.22	24.434	0.368	0.694
INSTRTIME96	25.12	23.668	0.436	0.682
.	.	.	.	.
.	.	.	.	.
INSTRTIME114	24.52	25.569	0.278	0.707
Cronbach's Alpha = .72				

Table 3c shows that the coefficient range of the corrected item-total correlation is 0.28 to 0.47. This establishes construct validity of the instructional time related items of OTLM scale. The estimated reliability coefficient on the subscale was 0.75. Conclusion can be drawn that 10 items that constitute Instructional Time related subscale of OTLM scale are valid and reliable to measure the expected dimension of opportunity to learn mathematics subscale.

Table 3d: Item-Total Statistics on Facilities related items

		Item-Total Statistics		

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
FACI119	14.67	12.03	0.569	0.729
FACI122	14.73	12.98	0.446	0.76
FACI124	14.75	12.125	0.566	0.73
FACI127	14.68	12.371	0.565	0.731
FACI128	14.5	12.424	0.547	0.735
FACI136	14.3	13.36	0.432	0.763
Cronbach's Alpha = .78				

Table 3d shows that the coefficient range of the corrected item-total correlation is 0.43 to 0.57. This establishes construct validity of the Facilities related items of OTLM scale. The estimated reliability coefficient on the subscale was 0.78. Conclusion can be drawn that 6 items that constitute Facilities related subscale of OTLM scale is valid and reliable to measure the expected dimension of opportunity to learn mathematics subscale. Just like the students need opportunity to learn, mathematics teachers should also need good working conditions to do their best job at teaching, i.e. by having basic tools such as books, laboratories, libraries and other adequate facilities, ULCA (2003).

Table 3e: Item-Total Statistics on learning materials related items

		Item-Total Statistics		
	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
LEMAT145	10.46	8.258	0.406	0.684
LEMAT151	10.55	8.511	0.39	0.689
LEMAT152	10.58	7.796	0.525	0.635
LEMAT153	10.87	7.91	0.498	0.646
LEMAT154	10.78	7.737	0.509	0.641
Cronbach's Alpha = .71				

Table 3e shows that the coefficient range of the corrected item-total correlation is 0.39 to 0.53. This establishes construct validity of the Learning Materials related items of OTLM scale. The estimated reliability coefficient on the subscale was 0.71. Conclusion can be drawn that 5 items that constitute Learning Materials related subscale of OTLM scale is valid and reliable to measure the expected dimension of opportunity to learn mathematics subscale. According to Meyers (2010), OTL is a way of measuring and reporting whether students and teachers have access to different ingredients that make up quality schools, school district or even schools across the nation.

(a) What is the spread of the category difficulties? (b) How discriminating is each item?

Table 4: Graded Model Item Parameter Estimates, logit: $a\theta + c$

Item	Label	a		$s.e.$	c_1	
1	TCH1	4	0.9	0.06	1	2.65
2	TCH3	8	1.01	0.06	5	2.63
.	.	.	.	.	.	.

.	.	.	.	.	.	.
82	LEMAT154	293	0.85	0.05	290	1.61

Table 4 provides slopes (a) and intercepts (c) contrasts for the Nominal model. Only one item (CURR 68) has less than 0.5 discriminating index. Conclusion can be drawn that the final OTML scale possesses moderate discrimination and difficulty indices that makes it usable to measure opportunity to learn mathematics in Secondary schools

Do the OTLM items conform to goodness of fit model?

Table 5: S-X^2 Item Level Diagnostic Statistics

Item	Label	X^2	*d.f.*	Probability
1	TCH1	301.31	259	0.0363
2	TCH3	320.00	249	0.0016
3	TCH5	283.23	268	0.2498
.	.	.	.	.
.	.	.	.	.
82	LEMAT154	337.50	284	0.0160

The S- χ^2statistics shows the fit of each individual item as shown on Table 4.5. All the values are very large and almost all are significant at 0.05 level of significant. This implied that there is good/reasonable model-data fit. This is in line with the specification for item goodness of fit that, there may be several items that show misfit but a majority of the items should fit well for the specified IRT model (graded).

Do the OTLM items conform to Differentiate among the respondents based on Gender and School Type?

Table 6: DIF Statistics for Graded Items

Male	Female	Total X^2	*d.f.*	*P*
1	1	26.6	4	0.0001
2	2	17.2	4	0.0018
.	.	.	.	.
.	.	.	.	.
3	3	4.9	4	0.2954
74	74	2.2	4	0.6985

Further, examination of the p-values for the Wald χ^2statistics that tests the difference between reference and focal group item parameters was done. Most items have significant DIF as "anchor items" in the model (alpha =.05). This implies that the final OTML has the potency to differentiate between male and female testees. It is worth noting that 8 items were found to have unequal responses among the groups, hence they were excluded from DIF analysis.

Table 7: DIF Statistics for Graded Items

State	Private	Unity	Contrast	Total X^2	*d.f.*	*p*
1	1	1	1	58.1	4	0.0001

			2	55.3	4	0.0001
2	2	2	1	9.8	4	0.0438
.	.	.	.	.	.	.
.	.	.	.	.	.	.
74	74	74	1	31.7	4	0.0001
			2	59	4	0.0001

Further also, examination of the p-values for the Wald χ^2statistics that tests the difference between reference and focal group item parameters was done based on the type of school (Public, Private and unity) of the respondents. Most items have significant DIF as "anchor items" in the model (alpha =.05).This implies that the final OTML has the potency to differentiate between respondents from different type of school. It is worth noting that 8 items were found to have unequal responses among the groups, hence they were excluded from DIF analysis.

Conclusions

This study is a correlation study aimed at applying Item Response Theory (IRT) to the development and validation of Opportunity to Learn Mathematics (OTLM) scale for secondary school students. Six (6) research questions were set to guide the study. Opportunity to learn was theoretically defined to have five components. A total of 162 items were generated and validated. The sample comprised one thousand, nine hundred (1900) students of SS3 from 36 sampled schools in Imo and Rivers States. A multi-stage cluster sampling techniques was used.

Exploratory factor analysis (EFA), Confirmatory factor analysis (CFA) and Graded response model (GRM) of IRT were employed for data analysis. Eighty-two (82) items were selected and calibrated. The findings are summarized as follows:

(1) The assumptions were not violated.

(2) The OTLM scale was found to possess construct validity with percentage variance of 13.96%.

(3) The items in the OTLM scale will measure students OTL meaningfully.

(4) The factor loading was found to be highly consistent in estimating the unidimensionality of the OTLM scale.

(5) The subscales of OTLM values were found to be reliable with the following reliability indices: Teacher effectiveness (0.91), Curriculum content/Learning experience (0.85), Instructional time (0.72), Facilities (0.78) and Learning material (0.71).

(6) The final OTLM scale possesses moderate discriminating and difficulty indices that make the scale usable to measure OTLM in secondary schools.

(7) The OTLM scale items have good and reasonable model data fit.

(8) The reliability index of the entire scale was found to be very high with a reliability coefficient of 0.90.

(9) Further findings revealed that the final OTLM has the potency to differentiate between male and female respondents and respondents from different types of schools

Based on the findings of this study, the scale is recommended to be used to collect data on opportunity to learn mathematics in Nigerian secondary schools.

Recommendations

a) Researchers should employ IRT procedures to develop and evaluate their survey data in order to get quality result of reliable and valid conclusions.

b) OTLM scale should be used when measuring the quality and quantity of OTLM and presence or absence of equity in resources for teaching and learning of mathematics in secondary schools.

References

DeVellis, R.F.2012. Scale Development Theory and Application; 3rd Edition, 26 Applied Social Research Method Series. SAGE Publication.

Embretson, S.E.& Reise, 2010. Item Response Theory for Psychologist. Lawrence Erbaum Associates Mahwah, NJ.

Hambleton, R.K.& Swaminathan, H. 2010. Item Response Theory: Principles and Applications. Boston: Kluwer-Nijhoff.

.House, J.D. and Telese, J.A. 2008. Relationship between Student and Instructional Factors and an Algebra Achievement of Students in the United States and Japan: An Analysis of TIMSDS 2003. Educ.reg. Evaluation, 14:101-112.

IRTPRO: User guide 2011. Lincolnword, IL Scientific Software International .Inc

Kurz, A. 2011. Access to What should be taught and will be Tested: Students' Opportunity to Learn the Intended Curriculum.

Meyer's, B. 2010. Opportunity to Learn and Equity: Measuring Efforts in Science, Technology, Engineering and Mathematics (STEM) Educational Program. http://www.napedu/catalogue/1067.html.

NCTE 2012. Opportunity-to-Learn Standards, Statement of Principles. National Council of Teachers of English (NCTE) Guideline. Retrieved 1/26/15.

Obaitan, G.N. and Rasheed, M.O. 2010. The Relationship between Selected Learner Characteristic and Performance in Mathematics. M.Ed Thesis Unpublished pages 17 &18.

Odili, G.A. 2007. Mathematics in Nigeria Secondary Schools: A Teaching Perspective. Rex Charles & Patrick Publishers, Lagos.

.Ojerinde, D. 2013. Classical Test Theory (CTT) vs Item Response Theory (IRT): An Evaluation of the comparability of item analysis result. A lecture presented at the institute of education, University of Ibadan (May 23, 2013).

Popoola, A.A. 2004. Effectiveness of Problem Solving Strategy on the Academic Achievement of Secondary School Students in Mathematics Nigerian Journal of Counselling and Applied Psychology. University of Ado Ekiti, Nigeria pgs130-142.

Reeve, B.B 2005. An Introduction to Modern Measurement Theory.

UCLA 2003. Opportunity-to-Learn: UCLA's Institute for Democracy, Education & Access. (www.uda.idea.org), Retrieved 1/26/15.

Yara, P.O. 2009. Students Attitude towards Mathematics and Academic Achievement in some selected Schools in South-Western Nigeria. Kampala International University, Western Campus, Bushenyi-Ishaka Uganda. European Journal of Scientific Research 36(3), 336-341.