An investigation of item biasness of early reading literacy test developed for pre-school children in Oyo State

Olukemi Oluremi Adeosun
Universal Basic Education Board, Oyo State
Assistant Director/ Early childhood care and Development education Desk Officer

Abstract

Fairness and equity of test items across sub-groups of pre-school children is very important as vital decisions on policy, placement, certification, intervention, scholarship award and the likes are arrived at depending on the pre-school children's scores for the test. It is very imperative for test developers to avoid test items that might function differently across sub-groups. This study empirically investigated Differential Item functioning of Developed Early Reading Literacy Test among pre-school children in Oyo State. The study advanced four research questions, where the underlying dimensions and test biasness in terms of gender, school type and location were investigated. The research type is instrumentation while counterbalance design was adopted. Five hundred and sixty pupils were drawn using multi stage sampling procedures from forty (40) pre-schools in twelve municipalities in Oyo State. The instrument used for data collection was self-developed Early Reading Literacy Test (emp_rel =0.84). Stout Test of Essential Dimensionality (STED) was used to assess the dimensionality of the test while Item response theory and DIF method were used to detect gender, school type and location biasness. Findings from this study showed that the traits that underlie questions are unidimensional in nature and 31.0% of the questions were gender biased, 42.0% were biased towards school type while 47.1% were biased with relevance to school location. Thus, the author recommended that test developers should endeavor to determine biasness of their test items so as to avoid it favoring a specific sub – group.

Keywords: Differential Item Functioning (DIF), Early Reading Literacy Test (ERLT), Stout Test Essential Dimensionality (STED)

Introduction

The Early Childhood Care and Development Education (ECCDE) is the care, protection, stimulation and learning promotion of an individual from birth to four years in crèche or nursery, while Pre-Primary Education refers to one year education given to children aged 5 before entering grade school (Federal Republic of Nigeria, 2014). Studies like National Association for the Education of Young Children (2017) and Vidya,T. (2014) have shown that attending ECCDE/Pre-Primary schools does not only increase children's readiness for formal schooling but also enhances positive long-term improvement in participants' school performance and social outcomes. The ECCDE aims at ensuring a smooth transition from home to school and preparing children for future learning activities. Efforts to support and make available quality ECCDE will promote growth in cognitive, language acquisition, motor skills acquisition and other adaptive skills acquisition to enable children function well during other levels of education (Oduolowu, 2011).Thus, ensuring that quality and standard foundation are provided for children is very important since the standard of nurture given to them during early years is crucial to their all-round development.

Moreover, Early Childhood Care and Development Education is the first level of basic education (Federal Republic of Nigeria, 2014) which represents the primary and essential step towards achieving the Millenium Development Goal 2 by 2015 (Federal Republic of Nigeria, 2005). The targeted year, 2015, had come and gone, yet, the MDG 2 goal has not been fully realized in Nigeria and if Sustainable Development Goal 4.2 (2030) would be achieved, there is need for increasing efforts

towards promoting quality Early Childhood Care and Development and Pre-Primary Education programmes. Oduolowu (2011) documented among the benefits of ECCDE as helping in building better critical thinking, better reading and writing skills, ability to handle the demand of formal schooling, higher graduation rates and preparation of youngsters for better future. The Universal Basic Education Commission (UBEC) thus, views the necessity to position Pre-Primary education at the centre of education policy and in response, pays more attention to early reading literacy considering its importance in enhancing academic achievement and nation's development. It thus becomes necessary for stakeholders to harness all measures to ascertain that early reading literacy skills are assessed beginning from the early years. This could be actualized through the construction of early reading literacy test so as to access major pre- skills in early reading literacy like knowledge about print, recognition of signs and symbols and identification of letters, objects, animals, birds and part of human beings and so on.

Achievement Test is an instrument used systematically to gather essential information on the skills and knowledge that the examinees have got on a particular concept and it must be free from biasness (Roever, 2005). The existence of bias is a problem to be addressed because tests serve the role of a gatekeeper for educational opportunities and it is imperative that test items are fair for every examinee. Biased test items refer to items that contain constructs that are irrelevant to the questions resulting in systematically lower or higher scores for identifiable groups of examinees. Under both CTT and IRT frameworks, items on a test that are found to be biased in favour of a group of examinees are considered to exhibit Differential Item Functioning (DIF). This occurs when different groups of examinees show differing possibilities of success on the item after matching on the underlying ability that the item is intended to assess.

Two types of Differential Item Functioning that an item can display are uniform Differential Item Functioning and Non-uniform Differential Item Functioning (Welberg, 2007). Absence of interaction between ability level and group membership, that is, the probability of answering an item correctly over all matched ability level is known as Uniform Differential Item Functioning. Conversely, each item shows Non-Uniform DIF, when interaction between ability level and group membership exists. For an item to display Non-Uniform DIF, the probability of answering an item correctly differs over all matched ability. Detection and estimation of interactions between item difficulties and various subgroups within the population of respondents is the main purpose of DIF, this is frequently applied to interactions with relevant demographic or ethnic groups such as gender, location or race. (Welberg, 2007) Moreover, four major approaches of describing and detecting DIF in a test that contains dichotomously scored items are identified as Logic Regression Analysis, Item Response Theory method, Log Linear Models and Mantel-Haenszel Test . However, only the Item Response Theory methods make use of item parameter estimates of DIF assessment while others are Classical Test Theory based (Oshima,T. & Morris,S. 2008). During this study, IRT method of detecting DIF was emphasized.

Chen, et al., (2009) in a study "using Logistic regression to detect test items in Chemistry Achievement", revealed that there are gender bias and class bias in Chemistry Achievement Test. Also, research findings confirmed that boys and girls have differential abilities. Ijaiya, B. (2007) found that females performed significantly better on word fluency test, while males achieved significantly higher score in spatial test and in general, girls speak earlier than boys, excel in word usage, correctness of sentence structure and comprehensibility of speech more than boys.

Enu,V. (2015) developed and validated item bank in mathematics and geography for Joint Command Schools Promotion Test (JCSPE) of the Nigeria Army Education Corps with a view to ensuring that the items in the bank are calibrated and of top quality. The study sampled 600 students in the two subjects (mathematics and geography). Item Response Theory was adopted. The study found that mathematics

items as well as geography items show DIF in gender, mode of schooling and location. Metibemu (2016) examined the comparability of CTT and IRT frameworks in the development of a 50-item multiple-choice Physics Performance Test from a pool of 100-item multiple-choice Physics Performance Test. The study found that the constructed Physics Performance Test was unidimensional. The items stem and options also gave no clue for answering another item. Some of the test items displayed Differential Item Functioning with reference to gender and learning centre. However, further qualitative analysis revealed the items are not biased. Mantel-Haesznel test - a CTT based method of Differential Item Functioning assessment performed equally as the IRT-based method of DIF assessment. The CTT- based method and IRT- based method of assessing DIF produced similar results in the assessment of DIF in the constructed Physics test items with respect to location. However, the IRT- method out-performed the CTT based-method in assessing DIF with reference to gender.

Nevertheless, items that function differently are of great concern to the developed early reading literacy test, of importance were the observed differences within the ability among the demographic variables such as gender, school type and school location respectively. Thus, this study investigated differential item functioning of early reading literacy test among pre-school children in Oyo State.

Research Questions

Four research questions guided the study. These are:

1. How many dimensions are embedded in the Early Reading Literacy Test (ERLT)?
2. Are Early Reading Literacy Test items biased with respect to gender?
3. Are Early Reading Literacy Test items biased with respect to school type?
4. Are there school location biased test items in Early Reading Literacy Test?

Methodology

The study was anchored on instrumentation research of counterbalance design. This design allows possible order of administering the instrument in to ensure order that none of the questions suffered. It also assisted in reducing errors which may emanate from test administration like random and systematic errors which could unknowingly be taken as variance. The population used for this study consisted of all pre-school children from public and private schools in Oyo State that are in final pre-school classes. Multi stage sampling procedures was used. Oyo State was stratified along the established three (3) senatorial districts while (4) municipalities from each district were selected considering the proportionate to size sample technique. Also, in the sampled LGAs fourteen (14) schools (7 public and 7 private) from Oyo Central, ten (10) schools (six public and four private) from Oyo North and sixteen (16) schools (six public and ten private) from Oyo South senatorial district were selected considering the proportionate to size sampling technique of 40 pre-schools children. An intact final pre-school class from each of the schools was used, totalling the sample size of five hundred and sixty (560). The instrument used was self-developed Early Reading Literacy Test (ERLT) with empirical reliability of 0.84. The test consisted of 226 items each having two response options which were scored as wrong (0) or right (1). Data collected were analysed using Stout Test of Essential Dimensionality (STED) and Item response theory differential item functioning method.

Results

Research question 1: How many dimensions are embedded in the Early Reading Literacy Test (ERLT)?

To answer this question, the null hypothesis posits that the Assessment Subtest (AT) and Partitioning subtest (PT) assess the same dominant underlying dimension, while the alternative hypothesis implies that the items in the AT partition are best represented by a dimension that is distinct from that driving responses to the PT items. The result is presented in Table 1.

Table 1: Stout Test of Essential Dimensionality Statistic of 226 ERLT

TL	TGbar	T	P-value
8.2370	12.7626	-3.5032	0.0820

The result of Stout's test of essential unidimensionality showed that the null hypothesis was not rejected (T=-3.5032, p = 0.0820). This leads to the conclusion that the AT is dimensionally distinct from PT. Thus, the test is viewed as essentially one-dimensional.

Research question 2: Are Early Reading Literacy Test items biased with respect to gender?

To answer this question, the DIF option of Differential Item Functioning Analysis System (DIFAS) software version 5.0 was conducted in order to establish items that function differentially between male (reference group) and female (focal group) pre-school children. The results are presented in Table 2 while Figure 3 depicts the line graph of items that indicated DIF in favour of males and females.

Table2: DIF Statistics with respect to Gender

Item number	MH-CHI	MH-LOR	LOR-SE	LOR Z	BD	ETS
Item 1	0.00	0.00	0.00	0.00	0.00	NO DIF
Item 2	0.01	-0.60	0.99	-0.60	1.54	NO DIF
Item 3	0.00	-0.35	0.81	-0.43	0.34	NO DIF
Item 4	0.05	0.39	0.69	0.56	0.06	NO DIF
Item 5	0.03	0.08	0.66	0.12	1.41	NO DIF
Item 6	0.82	-0.57	0.51	-1.12	2.61	NO DIF
Item 7	0.00	0.09	0.46	0.19	0.49	NO DIF
Item 8	0.68	-0.44	0.42	-1.04	0.38	NO DIF
Item 9	0.82	-0.54	0.47	-1.14	2.84	NO DIF
Item 10	0.02	0.03	0.44	0.07	0.43	NO DIF
+	+	+	+	+	+	+
+	+	+	+	+	+	+
+	+	+	+	+	+	+
Item 216	2.21	-0.38	0.24	-1.58	3.47	NO DIF
Item 217	0.02	0.01	0.33	0.04	0.01	NO DIF
Item 218	0.28	0.28	0.38	0.73	0.17	NO DIF
Item 219	0.08	-0.11	0.26	-0.42	0.02	NO DIF
Item 220	0.02	0.01	0.30	0.02	0.02	NO DIF
Item 221	1.21	0.32	0.26	1.24	0.01	NO DIF
Item 222	0.33	0.45	0.54	0.82	0.01	NO DIF
Item 223	0.55	-0.53	0.51	-1.04	0.48	NO DIF
Item 224	0.01	0.17	0.51	0.33	1.96	NO DIF
Item 225	0.00	0.02	0.27	0.08	0.01	NO DIF
Item 226	0.31	-0.25	0.34	-0.75	0.54	NO DIF

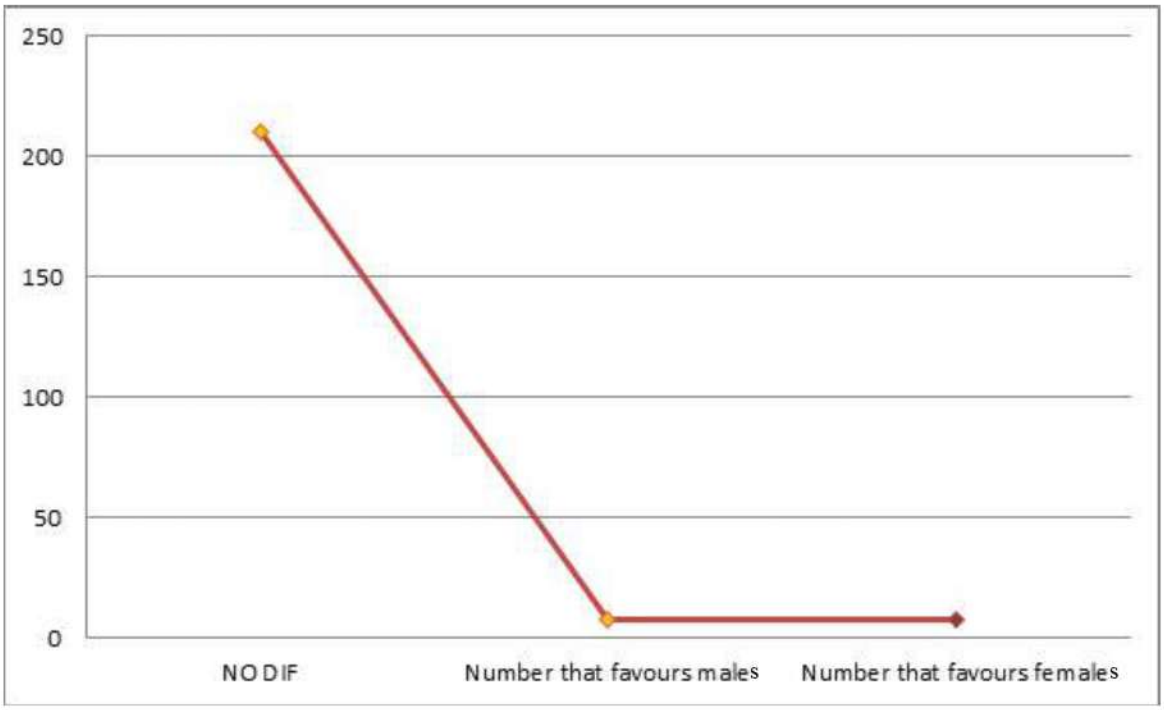

Figure 3: Line graph of items that indicates DIF in favour of males and females

Table 2 revealed the DIF statistics on pre-school children item performance with respect to gender. Column 2 of Table 2 is the Mantel Haenszel Chi-square (MH-CHI). This is distributed as chi-square with one degree of freedom. Column 3 are Mantel Haenszel Common Logs - odds Ratio (MH-LOR), it is asymptotically normally distributed, where negative values mean the reference groups are favoured in the DIF, while positive values indicate presence of DIF in favour of the groups. Column 4 is the standard error of the Mantel-Haenszel Common Log - Odds Ratio (LOR-SE). Non-symmetric estimator presented by Robins, Breslow& Greenland (1988) was computed as the standard error.

Column 5 is the Standardised Mantel-Haenszel Log- Odds Ratio (LOR Z). It is the division of Mantel-Haenszel Log-odds ratio by the estimated standard error (Camili & Shepard, 1994). Column 6 is Breslow -Day Chi-square (BD). The BD chi-square of trend in odds ratio heterogeneity is distributed as chi-square with one degree of freedom, while column 7 shows the ETS, which is used to bring out those items that exhibit DIF. More importantly, Table 2 revealed that 16 items have standardised Mantel-Haenszel Log-Odds Ratio (LOR Z) values greater than ± 2.0. These items are; 24, 39, 44, 70, 84, 107, 118, 125, 132, 133, 142, 146, 149, 172, 173. Out of 226 items, only 16 (31.0%) items were biased with respect to gender (8 favoured males, 8 favoured females).

Research question 3: Are Early Reading Literacy Test items biased with respect to school type?

To answer this question, the DIF option of DIFAS (Differential Item Functioning Analysis System) software version 5.0 was conducted in order to establish items that function differentially between private (reference group) and public (focal group) pre-school children. The results are presented in Table 4 while Figure 5 depicts the line graph of items that indicated DIF in favour of private and public school children.

Table 4: DIF Statistics with respect to School Type

Item number	MH-CHI	MH-LOR	LOR-SE	LOR Z	BD	ETS	Remarks
Item 1	0.00	0.00	0.00	0.00	0.00	NO DIF	
Item 2	0.37	-0.05	1.05	-0.05	3.72	NO DIF	
Item 3	3.05	0.00	0.00	0.00	0.00	NO DIF	
Item 4	4.13	-1.59	0.76	-2.11	0.39	DIF	Favours private examinees
Item 5	1.88	-0.92	0.59	-1.55	0.27	NO DIF	
Item 6	0.98	-0.54	0.47	-1.16	1.38	NO DIF	
Item 7	0.82	-0.47	0.43	-1.09	1.19	NO DIF	
Item 8	2.69	-0.65	0.39	-1.65	0.01	NO DIF	
Item 9	2.44	-0.85	0.50	-1.72	11.90	NO DIF	
Item 10	2.11	-0.89	0.54	-1.65	10.07	NO DIF	
+	+	+	+	+	+	+	+
+	+	+	+	+	+	+	+
+	+	+	+	+	+	+	+
Item 216	1.93	0.43	0.30	1.43	0.46	NO DIF	
Item 217	2.39	-0.59	0.38	-1.55	5.11	NO DIF	
Item 218	0.00	0.06	0.39	0.15	0.00	NO DIF	
Item 219	39.12	1.55	0.29	5.38	24.88	DIF	Favours public examinees
Item 220	7.77	-1.10	0.44	-2.54	8.56	DIF	Favours private examinees
Item 221	10.39	0.92	0.29	3.22	5.49	DIF	Favours public examinees
Item 222	4.78	-1.21	0.59	-2.05	10.36	DIF	Favours private examinees
Item 223	5.17	-1.41	0.61	-2.31	3.13	DIF	Favours private examinees
Item 224	6.06	-1.03	0.57	-1.80	43.51	NO DIF	
Item 225	0.04	-0.11	0.31	-0.35	1.43	NO DIF	
Item 226	0.16	-0.20	0.38	-0.53	0.92	NO DIF	

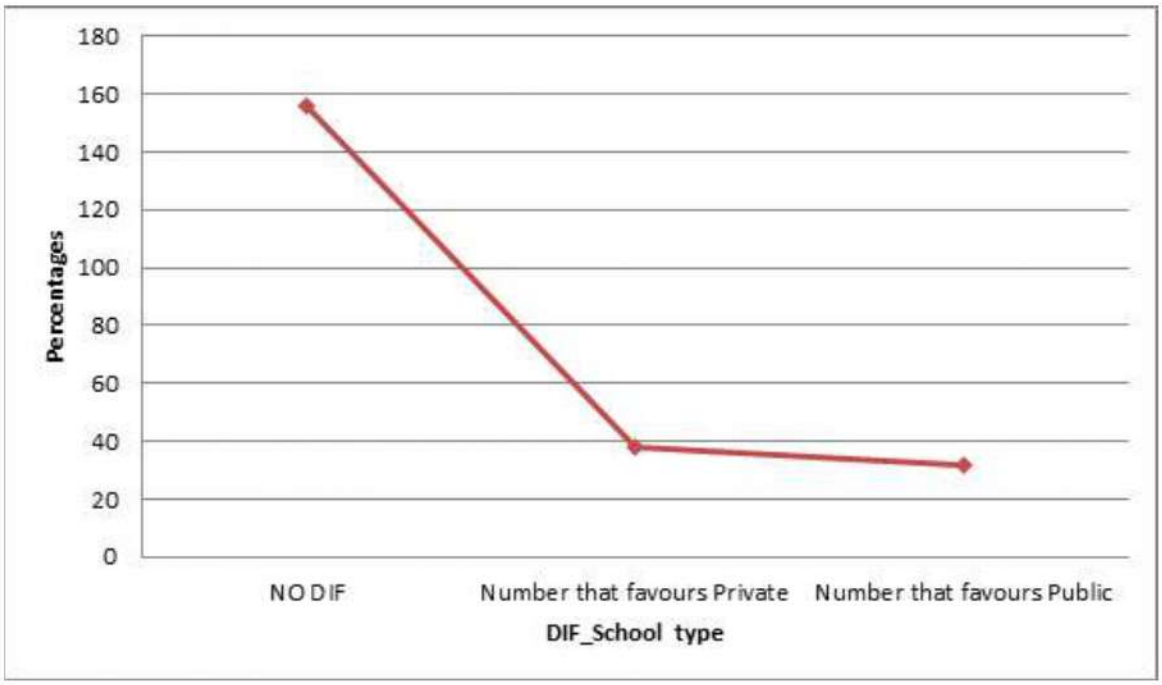

Figure 5: Line graph of items that indicates DIF in favour of Private and Public

Table 4 showed the DIF statistics on examinees item performance with respect to school type. This indicated that 70 items have standardised Mantel-Haenszel Log-Odds Ratio (LOR Z) values greater than ± 2.0. These items are; 4, 23, 24, 31, 52, 83, 85, 90, 105, 106, 107, 108, 113, 115, 116, 119, 122, 123, 126, 133, 134, 137, 139, 140, 141, 142, 145, 146, 147, 148, 149, 155, 156, 157, 158, 159, 160, 161, 162, 164, 166, 167, 171, 174, 176, 178, 180, 181, 186, 187, 191, 193, 194, 197, 198, 199, 200, 201, 202, 209, 210, 211, 212, 213, 214, 219, 220, 221, 222 and 223. Out of 226 items, only 70 (42.0%) items were biased with respect to school type.

Research question 4: Are Early Reading Literacy Test items biased with respect to school location?

To answer this question, the DIF option of DIFAS (Differential Item Functioning Analysis System) software version 5.0 was conducted in order to establish items that function differentially between urban (reference group) and rural (focal group) pre-school children. The results are presented in Table 6 while Figure 7 depicts the line graph of items that indicated DIF in favour of urban and rural.

Table 6: DIF Statistics with respect to School location

Item number	MH CHI	MH LOR	LOR SE	LOR Z	BD	ETS	Remarks
Item 1	0.00	0.00	0.00	0.00	0.00	NO DIF	
Item 2	2.44	0.00	0.00	0.00	0.00	NO DIF	
Item 3	7.98	0.00	0.00	0.00	0.00	NO DIF	
Item 4	2.85	-1.21	0.65	-1.86	0.41	NO DIF	
Item 5	0.85	-0.74	0.64	-1.15	0.26	NO DIF	
Item 6	0.98	-0.81	0.63	-1.29	0.00	NO DIF	
Item 7	9.94	-1.66	0.56	-2.95	4.25	DIF	Favours urban examinees
Item 8	11.47	-1.77	0.57	-3.10	1.76	DIF	Favour urban examinees
Item 9	3.85	-1.12	0.56	-2.02	3.66	DIF	Favour urban examinees
Item 10	4.86	-1.59	0.71	-2.24	0.53	DIF	Favour urban examinees
+	+	+	+	+	+	+	+
+	+	+	+	+	+	+	+
+	+	+	+	+	+	+	+
Item 216	12.35	1.06	0.33	3.18	13.83	DIF	Favour rural examinees
Item 217	3.64	0.76	0.43	1.77	5.81	NO DIF	
Item 218	0.02	-0.07	0.51	-0.13	0.23	NO DIF	
Item 219	0.00	0.03	0.28	0.13	1.51	NO DIF	
Item 220	0.05	0.29	0.57	0.50	0.00	NO DIF	
Item 221	18.53	-1.57	0.39	-4.08	0.08	DIF	Favour urban examinees
Item 222	0.19	0.43	1.60	0.27	0.01	NO DIF	
Item 223	0.08	0.07	0.77	0.09	0.12	NO DIF	
Item 224	0.04	-0.57	0.93	-0.62	1.55	NO DIF	
Item 225	4.24	-0.77	0.38	-2.04	0.67	DIF	Favour urban examinees
Item 226	0.60	0.57	0.58	0.98	0.22	NO DIF	

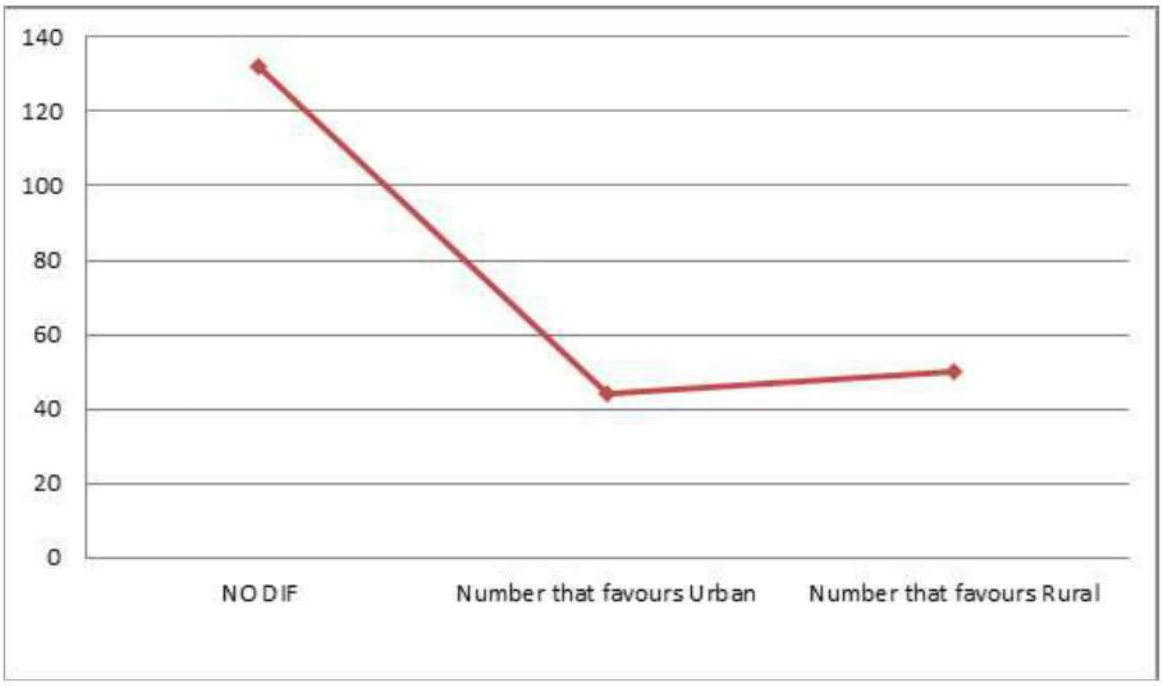

Figure 7: Line graph of items that indicate DIF in favour of Private and Public

Table 6 showed the DIF statistics on examinees item performance with respect to school location. However, table 6 indicated that 94 items have standardised Mantel-Haenszel Log-Odds Ratio (LOR Z) values greater than ± 2.0. These items are; 7, 8, 9, 10, 11, 14, 20, 23, 25, 29, 31, 35, 36, 37, 40, 41, 42, 44, 46, 47, 49, 51, 52, 59, 60, 61, 62, 69, 70, 71, 76, 81, 83, 84, 91, 92, 95, 96, 98, 99, 100, 102, 105, 110, 111, 112, 113, 117, 118, 120, 121, 125, 130, 132, 133, 134, 137, 139, 142, 144, 146, 150, 152, 159, 162, 164, 166, 169, 170, 171, 175, 176, 178, 179, 180, 182, 186, 190, 191, 196, 203, 204, 205, 206, 207, 208, 209, 209, 210, 213, 214, 215, 216, 221 and 225. This implies that 94 (47.1%) of the 226 items exhibited DIF with respect to school type.

Discussion of findings

The findings revealed that there was only one factor underlying the Early Reading Literacy Test. This shows that only one trait accounted for the variation observed in the pre-school children performance in the test. Thus, the unidimensionality assumption of item response theory was not violated. This result lends credence to the findings of Metibemu (2016); Enu (2015) that items of mathematics achievement test fulfilled unidimensionality assumption.

Also, the results revealed that test items had not favoured the males against female pre-schoolers. This supports Igbokwe (2004) who found that there were no significant differences in mathematics performance between boys and girls when the researchers developed item bank in mathematics for NECO common entrance test. The findings corroborated David et al, (2018) result which reviewed that females out performed males on tasks of verbal and language abilities. The evidence is from the National Assessment of Educational progress of 2018. Conversely, the findings contradict the work of Ijaiya (2007) who found that females performed significantly better on word fluency test. More importantly, it is necessary to operationalize the DIF result in the study that is the items that appeared to favour one gender than the other. For example, identification of a bag appeared to favour female pre-school children because the object is gender specific; females are more attached to the use of bags than males. Carrying of bags is seen as part of females' dressing while males rarely carry bags, many of them are only used to their school bags. In the same vein, identification of a tree favoured male children. This could be as a result of the cultural belief that only males are capable of climbing, playing around and on the trees. Most of the time, males take cognisance of issues such as recognizing more trees of different sizes and shapes. Girls are not always

allowed to participate in the activities that will require climbing trees.

Moreover, the results revealed that thirty- seven (37) out of 70 items favoured private schools while 33 items favoured public schools. This implies that 42.0% of the items exhibited biasness with respect to private and public pre-school children. Researcher operationalised the items that were found to exhibit DIF among the examinees in private and public centres. For example, identification of tyre favoured public school children. This might be because some male children in the public schools do play with used tyres, thereby giving them better opportunity to identify tyres than their mates in the private schools. On the other hand, identification of a jug favoured the examinees in the private schools. The reason might be that some of the pre-school children in public schools are from family backgrounds that may not give them the opportunity of having or using jugs at home which might make it a little bit difficult for them to identify the item. These items should not be discarded but it implies that the teachers should ensure these pre-school children are exposed to as many objects as possible that could be found at home and in their environment irrespective of their families' economic status, academic backgrounds and locations.

Also, the results showed that 44 out of 94 items that exhibited DIF were in favour of urban pre-school children examinees while 50 items favoured rural pre-school children. For example, item on the identification of an egg favoured urban examinees. This might be because many pre-school children in the urban setting often see eggs around them and take eggs as part of their meals unlike many examinees in the rural setting who rarely take eggs as part of their meals considering the fact that the free meals given by the federal government to public school children where egg is inclusive do not include pre-primary children. On the other hand, identification of a ladder favoured rural examinees; this might be because examinees in the urban setting may not have had the opportunity of seeing a ladder before unlike many children in the rural areas who, from time to time, see people climbing ladders probably to adjust any damage on the roof of their houses or for some other purposes. Implication of this result is that pre-school teachers should ensure that they expose children to all these objects that could be found in the rural and urban settings irrespective of the pre-schooler's location. Teachers should try as much as possible to balance the examples they use in the learning centres by stimulating the pre-schoolers to learn, not minding the extra efforts doing that would take.

Conclusion and recommendations

Based on the findings, the test items administered on pre-school children revealed biasness with respect to demographic variables in the study, although the test measured one trait of pre-school children. The authors recommend that test developers should always endeavour to establish biasness of their test item in order to avoid favouring a particular sub-group.

Also, considering the uniqueness of the sample of the study, the teachers handling pre-primary children should endeavour to expose them to different examples of items, objects, fruits, and other materials that could assist in their oral language development irrespective of their family, background, economic status and academic location.

Reference

David Reilly, David L. Nuemann & Gleannda Andrews., (2018). Gender differences in reading and writing achievement; Evidence from the National Assessment of Educational Progress. (NAEP). American psychologist association 74 (4) 2018

Enu, V. O. (2015). Validation and calibration of mathematics and geography items for joint command schools promotion examination in Nigeria using Item Response Theory. Ph.D. thesis, International Centre for Educational Evaluation, University of Ibadan, Nigeria.

Falayajo,W.Makoju, G.A.E. Okebukola, P., Onugha, D.C.&Olubodun, J.O. (1997). Assessment of learning performance of

primary four pupil in Nigeria: National report. Abuja: Federal Ministry of Education/UNICEF. UNESCO.

Federal Republic of Nigeria 2005.National Policy on Education.3rd Edition.

Federal Republic of Nigeria 2014.National Policy on Education.6th edition p. 5.

Huifang Chen &Kuan – Yu Jin. (2018) Applying logistics regression to detect differential item functioning in multidimensional data. Applying psychological measurement. 33, 202 - 214

Igbokwe, E. (2004). Development of mathematics item bank for placement of students into federal unity schools in Nigeria. Unpublished Ph.D. Thesis, International Centre for Educational assessement, University of Ibadan, xiv+157pp.

Ijaiya B.S. (2007). A psycho-pragmatic description of performance in the English of selected nursery school children in Kwara State. Unpublished Ph.D. Thesis, University of Ibadan, Ibadan, Nigeria.

Lynn, O. R. (1994). Sex differences on the differential intelligence FEM. test in British and American Adolescents. Educations psychology 12: 10 – 100.

Metibinu, M. A (2016).Comparison of classical test and item response theories in development and scoring of Senior secondary school physics tests in Ondo state, Nigeria. A PhD research post field seminar paper presented at the institute of education, University of Ibadan.

Oduolowu, E. (2011). Contemporary issues in early childhood education. Ibadan: Franco-Ola Publishers.

Oshima, T.C. & Morris, S.B. (2008). Raju's differential functioning of items and tests (DFIT). Educational Measurement: Issue and practice 27, 43-50.

Roever, C (2005). "That's not fair!" Fairness, bias, and differential item functioning in language testing. Retrieved November 18, 2006, from the University of Hawai'i System Web site: http://www2.hawaii.edu/~roever/brown bag.pdf

Vidya, T. (2014). Parental involvement in children education. Journal of the Association for Childhood Education International 10.3. Washington, DC.

Welberg, M. (2007). Measuring and detecting differential item functioning in criterion-referenced licensing test: a theoretical comparison of methods. Educational measurement No. 60 ISSN 1103-2685