European Journal of Education Studies
ISSN: 2501 - 1111
ISSN-L: 2501 - 1111
Available on-line at: www.oapub.org/edu

doi: 10.5281/zenodo.1301272 Volume 4 | Issue 9 | 2018

ASSESSMENT OF DIFFERENTIAL ITEM FUNCTIONING IN SOCIAL STUDIES MULTIPLE CHOICE QUESTIONS IN BASIC EDUCATION CERTIFICATE EXAMINATION

Patrick U. Osadebe[1i],
Bright Agbure[2]
[1] Department of Guidance and Counselling
Delta State University Abraka, Nigeria
[2]Delta State University Abraka, Nigeria

Abstract:
The study examined Differential Item Functioning in Social Studies multiple choice questions in Basic Education Certificate Examination. Five research questions and five hypotheses guided the study. An ex-post facto research design was adopted. The population of this study consisted of all Junior Secondary class three students in Delta central senatorial district. The proportionate stratified random sampling technique was employed. The instruments used for collecting data were the 2014 BECE Social Studies multiple choice questions and socio-economic status whose validity and reliability was ensured. Data were analyzed using SPSS and WINSTEPS. Descriptive statistics was used to answer the research questions while Independence Chi-square test was used to test the hypotheses. The finding revealed that there is incidence of gender, location, socio-economic, school type and school ownership differential functioning in 2014 BECE Social Studies multiple choice test. It was recommended that examination bodies should be mindful of the disparities that exist between gender, location, socio-economic status, school type and school ownership. Some recommendations were made.

Keywords: assessment, DIF, gender, location, school ownership

1. Introduction

Over the years test has been used to analyse the academic performance of students. Educational institutions have used test on students' performance to examine how well they compare to other students. In educational system, test conducted are expected to establish the characteristics of examinees. In Nigeria, there exist a number of national examination bodies who are responsible for conducting test for various reasons such as promotion, selection, placement among others. The examination bodies cater for

[i] Correspondence: email drosadebeuzo@gmail.com

candidates of various backgrounds in Nigeria, hence for personal and environment reasons, these examinations are differently toned (Emaikwa, 2012). As a result of this, there are some instances where an item in these examinations could be more difficult for a particular group of examinees who are on the same ability level but from different subgroups to perform differently, such item can be said to be showing differential item functioning (DIF). Ability is the quality of being able to do something. Hence, the recorded level of accomplishment an individual reaches is referred to as ability level. There is perhaps, no issue more visible among national examinations conducted in a heterogeneous country like Nigeria than differential item functioning (DIF). The problem that necessitated this study centres around the effect of a test differentially functioning.

Differential item functioning can simply be said to occur when test takers from different groups that have been matched on ability levels are performing differently in test items. The effect is that some examinees will be doing well while some will not be doing well. This has created the problem of equal opportunity among the examinees. According to the Federal Government of Nigeria (FGN)(2004) in the national policy on education which states that every Nigerian child shall have a right to equal educational opportunities irrespective of any real or imagined disabilities each according to his or her ability and there shall be the provision of equal access to educational opportunities for all citizens of the country at the primary level, secondary level, and tertiary levels both inside and outside the formal school system. The problem of testing not providing equal opportunity for examinees has been created as a result of the test items functioning differentially. There is no question that education is a key ingredient in improving the lives of children, families, communities and nations. When some examinees are failing while some are passing as a result of the difficulty posed by the test, it has distorted the chance of those who failed to be promoted. This has proved the notion of how DIF could be harmful and dangerous.

There is also the problem of class differentiation. Class differentiation in this context deals with dividing people in the society into groups such that they have certain economic or/and social characteristics in common. The focus of education is to bridge the gap of classes. This could lead to the purpose of education being defeated as a result of the threat being posed by the effect of DIF. The presence of DIF as also led to differential drop out in schools, since the test items are proving difficult to the examinees thereby causing failure not because of their ability to answer the items correctly but because of the unfairness of the test, it has results to many withdrawing out of school. According to Odili (2010), he revealed that interest in analysis of differential item functioning in test derives from the consideration that education is perceived as instrument for achieving egalitarianism among persons. In achieving this demands test items should measure traits which are taught in school subjects and not those that are alien to it. Examinees are failing as a result of the test unfairness and not because of their ability to answer the test correctly, thereby leading to drop out. The consequence of school drop outs has led to increase in crimes and vices in the society. According to the Bureau of Labour statistics, high school drop outs are having a harder

time finding and keeping jobs than individuals with higher levels of education. In fact the national unemployment rate for high school drop outs in July 2009 was 15.4 percent compared to 9.4 percent for high school graduates. This statistics tend to show the effect of a test differentially functioning and as a result of not providing equal opportunity for students, increased class differentiation, differential drop out leading to different crimes and vices by the students who are not in school. The implication of this is that items in the test should be fair to all subgroups in the population. There should be nothing in the test items that would make the items to favour one group above the other group that are of equal ability level.

Atar (2006) explained that it is critical that test items do not differentiate among examinees based on their gender, race, and ethnic background but rather differentiate between them based on their abilities. A fair test is one that is comparably valid for all groups and individuals and that affords all examinees an equally opportunity to demonstrate the skills and which are relevant to the test purpose. The presence of large number of items with DIF is a serious threat to the validity of a test and any inference made from such test scores may not be valid.

There are two types of DIF, which are uniform and non-uniform DIF. Uniform DIF is said to occur when differences in correct response probability are found across all ability levels for a particular item. Non-uniform DIF on the other hand occurs when there is interaction between the ability and group membership such that an item may seem difficult for those at the higher level in one group and after a particular point, it becomes more difficult for those at lower level in the other group. Odili (2003) identified some of the sources of systematic error, which could lead to detection of DIF in a test to be test wiseness, culture, language ability.

Gender refers to specific central patterns attributed to both males and females in terms of behaviour and mechanism (Okoro, 2011). It is the psychological and socio-cultural interpretation of male and female. Gender is the socio-cultural interpretation of male and female based on their expected role, contributions and assigned duties (Ija, 2009). The literature review about gender and differential item functioning conducted nationally have revealed that test items contained differential item functioning whether the test is meant for classification, admission, recruitment, or placement purposes (Abiam & Odok,2006).

According to Alordiah (2015) socio-economic status (SES) is the way people are divided into groups in a society such that they may have certain economic or/and social characteristics in common. Socio-economic status of a family is usually linked with the family's income, parents' educational level, parents' occupation, and social status (Okafor, 2007). Socio-economic status of parents is said to be related to high expectation. The location (urban/rural) which a child finds himself/ herself goes a long way to determine one's academic achievement than those from poor location (Ndifon, Umoinyang & Idiku, 2013). Urban areas are well equipped with learning facilities, qualified teachers, good roads and good communication networks which puts them at advantage position when compared to their rural counterparts where such

opportunities are inadequate or somehow lacking. Iyang (1991) investigated items in the 1986 common Entrance Examination Mathematics items do exhibits location DIF.

The type of school a student attends (single–sex school or mixed- sex school) could to a large extent, influence one's academic performance. Single-sex school refers mostly generally to education at the elementary, secondary, or post- secondary level in which males or females attend school exclusively with members of their own sex. Mixed-sex school on the other hand refers generally to education in which both male and female attend.

School ownership comprises public and private owned schools. Public schools are owned and founded by government while private schools are owned and founded by individuals or organizations. Research has shown that private schools do perform better than their public counterparts. This study will therefore, contribute meaningfully to address the issue posed by differential item functioning to ensure the balance of content reflecting the intent ability of candidate in Basic Education Certificate Examination (BECE).

2. Research Questions

The following questions guided the study;

1. What is the incidence of DIF for male and female students in Social Studies multiple choice test used in BECE in Delta State in 2014?
2. What is the incidence of DIF for urban and rural students in social studies multiple choice test used in BECE in Delta State in 2014?
3. What is the incidence of DIF for students with high and low socio-economic status in Social Studies multiple choice test in BECE in Delta State in 2014?
4. What is the incidence of DIF for students in single-sex and mixed-sex schools in Social Studies multiple choice test in BECE in Delta State in 2014?
5. What is the incidence of DIF for students in public and private schools in Social Studies multiple choice test in BECE in Delta State in 2014?

2.1 Hypotheses

The following hypotheses were formulated;

1. The number of items that function significantly different for male and female students is not significantly different from the number that did not differentially function.
2. The number of items that function significantly different for urban and rural students is not significantly different from the number that did not differentially function.
3. The number of items that function significantly different for high and low socio-economic status students is not significantly different from the number that did not differentially function.

4. The number of items that function significantly different for students in single-sex and mixed-sex schools is not significantly different from the number that did not differentially function.
5. The number of items that function significantly different for students in public and private schools is not significantly different from the number that did not differentially function.

3. Method

This study adopts ex-post-facto correlation design. This is because the data was gathered after the events of interest have occurred. The independents variables cannot be manipulated. It was used also to collect data that enabled the researcher to determine the incidence of Differential Item Functioning in Delta State Ministry of Education, Basic Education Certificate Examination Social Studies multiple choice tests. The population of this study comprises the entire junior secondary school three (JS3) students in all 8 local government areas of Delta Central Senatorial Districts. This population was chosen because they must have covered the syllabuses. There are 179 junior secondary schools in Delta central senatorial district with a population of 22,467 junior secondary school three students at the time of this study (Ministry of Education, Primary and secondary, Asaba). Of this number, 13481 are students in the urban areas while 8986 are students in the rural areas.

The sample size of this study is five hundred (500) students of the population who are JS 3 students from schools in Delta Central Senatorial District. A proportionate stratified random sampling technique was adopted for the study. The population was stratified according to location (urban/rural). The junior secondary school students were randomly selected from schools in public and private schools in the Local Government Areas in Delta Central Senatorial District of Delta State. The instruments used for data collection are 2014 BECE Social Studies multiple choice test and Socio-Economic Status Questionnaire. The 2014 BECE Social Studies multiple choice test contains 60 choice type questions, which cover the JSSCE Social Studies syllabus of Junior Secondary Schools in Delta State. The Socio-Economic Status questionnaire has 20 items on a 4-point scale, Strongly Agree (SA), Agree (A), Disagree (D) and Strongly Disagree (SD).

In order to ensure the validity of the instruments, the 2014 BECE Social Studies multiple choice test items were examined to see whether they cover the Social Studies syllabus for JS3 and it was found to be so. Content validity for the Socio-economic Status questionnaire was established by ensuring that the yardstick for classification of individuals into high and low Socio-economic Status. To ensure further content and face validity, the SES questionnaire was given to the researcher's supervisor and two experts in the Department of Guidance and Counselling for, opinion, comments, recommendations, and modification.

The BECE Social studies 2014 multiple choice test was administered to forty (40) JS III students in order to ensure reliability. Their responses were scored and analysed using Cronbach Alpha formula. A value of 0.82 was obtained, thus, making it suitable

for the study.The reliability for SES questionnaire was also established using Cronbach Alpha. A value of 0.68 was obtained.

The BECE social studies multiple choice test for 2014 was used to collect data from the sample students. Social studies teachers in the schools visited assisted the researcher in the administration of the 2014 BECE Social studies test. The test was administered within the time limit specified by the examination body. The correct response for each item was assigned 1, while the incorrect response was assigned 0. The SES questionnaire has 4-point scale, Strongly Agree (SA), Agree (A), Disagree (D) and Strongly Disagree (SD). The maximum score of the SES questionnaire is 80 while the minimum is 20 if all the items are responded to. The student with 41 and above was grouped as high SES while student with 40 and below was group as low SES. The method used for analyzing the data is Rasch model. The technique was used to answer the research questions and Independence Chi-square was used to test the hypotheses using SPSS and WINSTEPS statistical package at 0.05 levels of significance.

4. Results

Research Question One: What is the incidence of DIF for male and female students in Social Studies multiple-choice test used in BECE in Delta State in 2014?

Table 1: Rasch Model DIF Statistics for Gender

Items	Bm	Bf	Δb	Probability	DIF Incidence	Favoured group
1	-2.78	-2.55	-0.22	0.57	1	F
2	-2.63	-2.01	-0.62	0.69	1	F
3	0.56	0.27	0.29	0.14	1	M
4	1.02	1.28	-0.26	0.20	1	F
5	0.78	1.26	-0.47	0.02	2	F
6	-1.92	-1.47	-0.45	0.11	1	F
7	-0.97	-0.52	-0.46	0.04	2	F
8	1.00	1.26	-0.26	0.20	1	F
9	0.88	0.32	0.55	0.01	2	M
10	-1.46	-1.87	0.41	0.14	1	M
11	-0.61	-0.38	-0.22	0.29	1	F
12	1.14	1.45	-0.32	0.13	1	F
13	-1.10	-1.37	0.27	0.26	1	M
14	0.81	0.81	0.00	1.00	1	M
15	0.41	-0.03	0.44	0.03	2	M
16	0.01	0.25	-0.23	0.24	1	F
17	0.82	1.15	-0.33	0.10	1	F
18	-0.72	-0.65	-0.06	0.78	1	F
19	0.05	-0.01	0.06	0.76	1	M
20	0.25	-0.21	0.47	0.02	2	M
21	-2.05	-1.92	-0.13	0.66	1	F
22	-1.84	-1.74	-0.10	0.72	1	F
23	-1.20	-1.58	0.38	0.14	1	M
24	1.12	0.67	0.44	0.03	2	M

25	2.24	1.92	0.32	0.19	1	M
26	0.81	0.81	0.00	1.00	1	M
27	-1.20	-1.74	0.54	0.04	2	M
28	1.30	1.41	-0.11	0.61	1	F
29	-0.16	-0.36	0.20	0.32	1	M
30	-1.43	-1.82	0.39	0.15	1	M
31	0.84	0.95	-0.11	0.59	1	F
32	1.15	1.15	0.00	1.00	1	M
33	-0.32	-0.07	-0.12	0.22	1	F
34	0.63	0.63	0.00	1.00	1	M
35	1.64	1.76	-0.12	0.58	1	F
36	0.13	0.21	-0.08	0.68	1	F
37	-1.02	-1.15	0.13	0.57	1	M
38	0.18	-0.01	0.19	0.33	1	M
39	1.52	1.52	0.00	1.00	1	M
40	-0.83	-0.78	-0.05	0.81	1	F
41	1.81	1.52	0.29	0.19	1	M
42	0.09	0.23	-0.14	0.48	1	F
43	-1.34	-0.83	-0.52	0.03	2	F
44	-1.15	-1.09	-0.06	0.81	1	F
45	-0.02	-0.02	0.00	1.00	1	M
46	0.94	0.94	0.00	1.00	1	M
47	0.09	0.21	-0.12	0.55	1	F
48	-0.74	-0.74	0.00	1.00	1	M
49	2.37	2.11	0.25	0.31	1	M
50	1.61	1.52	0.09	0.67	1	M
51	0.03	0.67	-0.64	0.00	2	F
52	1.66	2.05	-0.39	0.09	1	F
53	-0.47	-0.49	0.02	0.92	1	M
54	0.32	0.32	0.00	1.00	1	M
55	-1.02	-1.12	0.10	0.66	1	M
56	-1.18	-1.62	0.44	0.08	1	M
57	-0.16	-0.30	0.14	0.49	1	M
58	-0.44	-0.21	-0.22	0.27	1	F
59	0.54	0.36	0.17	0.37	1	M
60	-0.24	-0.47	0.23	0.25	1	M

bm= measure for male, bf = measure for female, Δb = DIF contrast

Table 1 shows the DIF statistics of the Rasch model method for each 0f the 60 items for gender. There is incidence of DIF if the probability is less than 0.05. There is incidence of DIF in 9 items at 0.05 levels of significance. That is 15% of the 2014 BECE Social Studies multiple choice test items functioned differentially for male and female examinees. The DIF items are 5, 7, 9, 15, 20, 24, 27, 43 and 51.

Research Question Two: What is the incidence of DIF for Urban and Rural Students in Social Studies multiple Choice test used in BECE in Delta State in 2014?

Table 2: Rasch Model DIF statistics for Location

Item	Bu	br	Δb	Probability	DIF incidence	Favoured group
1	-2.34	-2.84	0.50	0.20	1	U
2	-2.55	-2.22	-0.33	0.40	1	R
3	1.05	-0.04	1.09	0.00	2	U
4	1.07	1.21	-0.13	0.51	1	R
5	0.88	1.13	-0.25	0.22	1	R
6	-1.46	-1.82	0.36	0.21	1	U
7	-0.29	-1.02	0.72	0.00	2	U
8	1.12	1.12	0.00	1.00	1	U
9	0.79	0.47	0.32	0.10	1	U
10	-1.51	-1.70	0.19	0.19	1	U
11	-0.62	-0.43	-0.20	0.38	1	R
12	1.40	1.19	0.22	0.29	1	U
13	-1.36	-1.16	-0.20	0.44	1	R
14	0.79	0.81	-0.02	0.17	1	R
15	0.03	0.30	-0.28	0.17	1	R
16	0.66	-0.23	0.89	0.00	2	U
17	1.52	0.56	0.95	0.00	2	U
18	-0.59	0.74	0.15	0.52	1	U
19	-0.21	0.16	-0.37	0.07	1	U
20	-0.21	0.18	-0.39	0.06	2	R
21	-1.80	-2.07	0.27	0.38	1	U
22	-2.16	-1.65	-0.51	0.12	1	R
23	-1.80	-1.20	-0.60	0.04	2	R
24	1.12	0.71	0.40	0.04	2	U
25	1.84	2.39	-0.54	0.03	2	R
26	0.55	1.03	-0.49	0.02	2	R
27	-1.68	-1.33	-0.35	0.22	1	R
28	1.25	1.46	-0.21	0.32	1	R
29	-0.38	-0.19	-0.19	0.36	1	R
30	-1.22	-1.82	0.60	0.03	2	U
31	1.16	0.68	0.48	0.01	2	U
32	1.03	1.27	-0.24	0.25	1	R
33	-0.24	-0.17	-0.07	0.74	1	R
34	-0.03	1.13	-1.15	0.00	2	R
35	1.58	1.83	-0.25	0.27	1	R
36	0.46	-0.03	0.48	0.01	2	U
37	-1.27	-1.00	-0.27	0.29	1	R
38	-0.08	0.19	-0.27	0.19	1	R
39	1.82	1.25	0.57	0.01	2	U
40	-0.47	-1.00	0.53	0.18	1	U
41	1.94	1.39	0.55	0.01	2	U
42	-0.21	0.38	-0.59	0.00	2	R
43	-1.01	-1.12	0.10	0.67	1	U
44	-0.90	-1.24	0.34	0.16	1	U
45	-0.62	0.32	-0.94	0.00	2	R
46	0.88	1.00	-0.12	0.56	1	R
47	0.32	0.04	0.28	0.16	1	U

48	-1.05	-0.60	-0.46	0.06	2	R
49	2.32	2.13	0.19	0.46	1	U
50	1.61	1.53	0.08	0.72	1	U
51	0.55	0.21	0.34	0.86	1	U
52	2.66	1.19	1.47	0.00	2	U
53	-0.98	-2.23	-0.74	0.00	2	R
54	0.05	0.50	-0.45	0.03	2	R
55	-1.27	-0.98	-0.29	0.26	1	R
56	-2.01	-1.14	-0.87	0.01	2	R
57	-0.08	-0.31	0.24	0.25	1	U
58	-0.79	-0.09	-0.70	0.00	2	R
59	0.10	0.70	-0.60	0.00	2	R
60	-0.32	-0.35	0.02	0.91	1	U

bu = measure of urban, br = measure of rural, Δb =DIF contrast

Table 2 shows the DIF statistics of the Rasch model method for each of the 60 items for location. There is incidence of DIF if the probability is less than 0.05. There is incidence of DIF in 24 items at 0.05 levels of significance. That is 40% of the 2014 BECE Social Studies multiple choice test items functioned differentially for urban and rural examinees. The DIF items are 3, 7, 16, 17, 20, 23, 24, 25, 26, 30, 31, 34, 36, 39, 41, 42, 45, 48, 52, 54, 56, 58 and 59.

Research Question Three: What is the incidence of DIF for Students with high and low Socio-economic Status in Social Studies multiple choice test in BECE in Delta State in 2014?

Table 3: Rasch Model DIF Statistics for Socio-economic Status

Items	Bh	bl	Δb	Probability	DIF incidence	Favoured group
1	-2.52	-2.79	0.27	0.48	1	H
2	-2.16	-0.43	0.26	0.44	1	H
3	0.60	0.26	0.34	0.08	1	H
4	0.93	1.35	-0.42	0.04	2	L
5	0.85	1.16	-0.32	0.12	1	L
6	-1.65	-1.73	0.08	0.76	1	H
7	-0.77	-0.74	-0.02	0.92	1	L
8	1.16	1.09	0.08	0.70	1	H
9	0.66	0.56	0.10	0.59	1	H
10	-1.52	-1.73	0.21	0.44	1	H
11	-0.49	-0.49	0.00	1.00	1	H
12	1.06	1.52	-0.46	0.03	2	L
13	-1.26	-1.20	-0.06	0.80	1	L
14	0.81	0.81	0.00	1.00	1	H
15	0.27	0.14	0.13	0.52	1	H
16	0.41	-0.11	0.52	0.01	2	H
17	1.01	0.95	0.06	0.76	1	H
18	-0.71	-0.68	-0.03	0.90	1	L
19	0.05	0.00	0.05	0.79	1	H

20	-0.01	0.07	-0.08	0.68	1	L
21	-1.78	-2.14	0.36	0.23	1	H
22	-1.94	-170	-0.24	0.42	1	L
23	-1.40	-1.37	-0.03	0.89	1	L
24	1.30	0.56	0.74	0.00	2	H
25	2.00	2.14	-0.13	0.58	1	L
26	0.68	0.93	-0.25	0.20	1	L
27	-1.29	-1.55	0.25	0.32	1	H
28	1.25	1.45	-0.20	0.33	1	L
29	-0.31	-0.21	-0.09	0.64	1	L
30	-1.29	-1.87	0.58	0.03	2	H
31	1.04	0.77	0.27	0.18	1	H
32	1.23	1.09	0.14	0.48	1	H
33	-0.08	-0.29	0.21	0.31	1	H
34	0.27	0.95	-0.69	0.00	2	L
35	1.43	1.99	-0.56	0.01	2	L
36	0.35	0.02	0.33	0.09	1	H
37	-1.03	-1.12	0.10	0.68	1	H
38	0.14	0.05	0.09	0.66	1	H
39	1.52	1.52	0.00	1.00	1	H
40	-0.45	-1.10	0.65	0.00	2	H
41	1.62	1.71	-0.09	0.69	1	L
42	-0.04	0.31	-0.35	0.08	1	L
43	-1.00	-1.15	0.15	0.52	1	H
44	-0.94	-1.27	0.33	0.16	1	H
45	-0.19	0.12	-0.31	0.12	1	L
46	0.85	1.03	-0.18	0.36	1	L
47	0.39	-0.05	0.44	0.02	2	H
48	-0.97	-0.59	-0.38	0.09	1	L
49	2.18	2.30	-0.12	0.64	1	L
50	1.62	1.52	0.10	0.65	1	H
51	0.27	0.42	-0.15	0.44	1	L
52	1.98	1.73	0.25	0.28	1	H
53	-0.74	-0.29	-0.45	0.03	2	L
54	0.20	0.42	-0.21	0.28	1	L
55	-1.29	-0.92	-0.38	0.12	1	L
56	-1.44	-1.32	-0.12	0.64	1	L
57	-0.26	-0.20	-0.06	0.75	1	L
58	-0.33	-0.33	0.00	1.00	1	H
59	0.31	0.57	-0.27	0.18	1	L
60	-0.19	-0.47	0.28	0.17	1	H

bh= measure of high SES, bl= measure of low SES, Δb =DIF contrast

Table 3 shows the DIF statistics of Rasch model method for each of the 60 items for Socio-Economic Status. There is incidence of DIF if the probability is less than 0.05. There is incidence of DIF in 10 items at 0.05 levels of Significance. That is 17% of the 2014 BECE Social Studies multiple choice test items functioned differentially for high

and low Socio-Economic Status examinees. The DIF items are 4, 12, 16, 24, 30, 35, 40, 47, and 53.

Research Question Four: What is the Incidence of DIF for Students in Single-sex and Mixed-sex Schools in Social Studies multiple choice test in BECE in Delta State in 2014?

Table 4: Rasch Model DIF Statistics for School Type

Items	Bm	bs	Δb	Probability	DIF incidence	Favoured group
1	-2.67	-2.67	0.00	1.00	1	M
2	-2.19	-2.40	0.21	0.54	1	M
3	0.22	0.50	-0.28	0.18	1	S
4	1.63	0.99	0.64	0.01	2	M
5	1.04	1.01	0.03	0.90	1	M
6	-1.94	-1.56	-0.38	0.20	1	S
7	0.80	-0.71	-0.09	0.68	1	S
8	0.50	1.38	-0.89	0.00	2	S
9	0.86	0.50	0.36	0.10	1	M
10	-1.40	-1.80	0.40	0.15	1	M
11	-0.57	-0.46	-0.12	0.59	1	S
12	1.04	1.38	-0.34	0.13	1	S
13	-1.05	-1.33	0.28	0.26	1	M
14	1.08	0.72	0.36	0.11	1	M
15	0.44	0.09	0.35	0.10	1	M
16	-0.54	0.43	-0.97	0.00	2	S
17	0.50	1.18	-0.68	0.00	2	S
18	-0.61	-0.73	0.12	0.59	1	M
19	0.34	-0.13	0.47	0.03	2	M
20	-0.08	0.09	-0.17	0.42	1	S
21	-1.73	-2.17	0.44	0.14	1	M
22	-1.54	-1.97	0.44	0.13	1	M
23	-1.49	-1.31	-0.19	0.48	1	S
24	0.63	1.00	-0.37	0.08	1	S
25	1.95	2.11	-0.16	0.56	1	S
26	0.93	0.77	0.16	0.47	1	M
27	-1.36	-1.48	0.12	0.65	1	M
28	1.08	1.46	-0.38	0.10	1	S
29	-0.48	-0.15	-0.33	0.12	1	S
30	-1.73	-1.54	0.19	0.49	1	S
31	0.76	0.95	-0.19	0.39	1	S
32	1.73	0.97	0.75	0.00	2	M
33	0.10	-0.34	0.44	0.04	2	M
34	1.00	0.49	0.52	0.02	2	M
35	1.63	1.72	-0.10	0.70	1	S
36	-0.11	0.30	-0.41	0.05	1	S
37	-1.49	-0.88	-0.61	0.02	2	S
38	0.09	0.09	0.00	1.00	1	M
39	0.97	1.74	-0.77	0.00	2	S
40	-0.94	-0.73	-0.21	0.36	1	S

41	1.31	1.79	-0.48	0.05	1	S
42	0.31	0.09	0.23	0.29	1	M
43	-1.12	-1.06	-0.07	0.79	1	S
44	-1.16	-1.12	-0.04	0.87	1	S
45	0.40	-0.22	0.62	0.00	2	M
46	0.94	0.94	0.00	1.00	1	M
47	0.28	0.09	0.19	0.36	1	M
48	-0.39	-0.96	0.57	0.01	2	M
49	2.07	2.28	-0.22	0.46	1	S
50	1.19	1.71	-0.52	0.03	2	S
51	0.35	0.35	0.00	1.00	1	M
52	0.69	2.39	-1.70	0.00	2	S
53	-0.11	-0.68	0.56	0.01	2	M
54	0.76	0.13	0.63	0.00	2	M
55	-0.80	-1.24	0.43	0.07	1	M
56	-1.09	-1.56	0.48	0.06	1	M
57	-0.11	-0.28	0.17	0.43	1	M
58	0.04	-0.52	0.56	0.01	2	M
59	0.53	0.42	0.11	0.60	1	M
60	-0.29	-0.38	0.08	0.71	1	M

bm= measure for mixed-sex, bs= measure for single-sex,Δb= DIF contrast

Table 4 shows the DIF Statistics of the Rasch model method for each of the 60 items for school type. There is incidence of DIF if the probability is less than 0.05. There is incidence of DIF in 17 items at 0.05 levels of Significance. That is 28% of the 2014 BECE Social Studies multiple choice test items functioned differentially for Single-sex and Mixed-sex schools examinees. The DIF items are 4, 8, 16, 17, 19, 32, 33, 34, 37, 39, 45, 48, 50, 52, 53, 54, and 58.

Research Question Five: What is the incidence of DIF for Students in public and private schools in Social Studies multiple choice test in BECE in Delta State in 2014?

Table 5: Rasch Model DIF Statistics for School Ownership

Items	Bpu	bpv	Δb	Probability	DIF incidence	Favoured group
1	-2.83	-2.67	-0.16	0.80	1	Pv
2	-1.93	-2.38	0.45	0.30	1	Pu
3	1.13	0.17	0.96	0.00	2	Pu
4	0.55	1.39	-0.84	0.00	2	Pv
5	0.70	1.13	-0.44	0.06	1	Pv
6	-1.78	-1.69	-0.09	0.82	1	Pv
7	-0.71	-0.74	0.03	0.90	1	Pu
8	1.49	0.93	0.52	0.02	2	Pu
9	0.81	0.53	0.27	0.22	1	Pu
10	-1.33	-1.70	0.38	0.27	1	Pu
11	-0.36	-0.53	0.17	0.52	1	Pu
12	1.35	1.27	0.08	0.72	1	Pu
13	-0.91	-1.29	0.38	0.21	1	Pu

14	0.73	0.84	-0.11	0.63	1	Pv
15	0.23	0.20	0.04	0.87	1	Pu
16	0.81	-0.09	0.90	0.00	2	Pu
17	1.20	0.90	0.31	0.17	1	Pu
18	-0.65	-0.68	0.04	0.89	1	Pu
19	-0.26	0.10	-0.36	0.15	1	Pv
20	-0.07	0.06	0.13	0.59	1	Pv
21	-1.93	-1.89	0.05	0.90	1	Pu
22	-1.65	-1.82	0.17	0.66	1	Pu
23	-1.53	-1.34	-0.19	0.60	1	Pv
24	1.24	0.76	0.48	0.03	2	Pu
25	1.96	2.13	-0.17	0.50	1	Pv
26	0.92	0.78	0.14	0.53	1	Pu
27	-1.53	-1,43	-0.10	0.78	1	Pv
28	1.42	1.33	-0.09	0.68	1	Pu
29	-0.31	-0.25	-0.06	0.82	1	Pv
30	-1.33	-1.66	0.34	0.33	1	Pu
31	1.35	0.73	0.62	0.01	2	Pu
32	1.02	1.21	-0.18	0.42	1	Pv
33	-0.42	-0.14	-0.28	0.28	1	Pv
34	-0.12	0.87	-0.99	0.00	2	Pv
35	1.72	1.70	0.02	0.93	1	Pu
36	0.73	-0.01	0.74	0.00	2	Pu
37	-1.33	-1.04	-0.29	0.39	1	Pv
38	-0.16	0.16	-0.32	0.19	1	Pv
39	1.76	1.42	0.34	0.15	1	Pu
40	-0.42	-0.90	0.48	0.07	1	Pu
41	2.17	1.45	0.72	0.00	2	Pu
42	-0.31	0.28	-0.59	0.02	2	Pv
43	-1.15	-1.08	-0.06	0.84	1	Pv
44	-1.15	-1.12	-0.02	0.94	1	Pv
45	-0.77	0.16	-0.93	0.00	2	Pv
46	0.81	0.99	-0.18	0.41	1	Pv
47	0.06	0.17	-0.11	0.65	1	Pv
48	-1.65	-0.60	-1.05	0.00	2	Pv
49	2.08	2.32	-0.24	0.37	1	Pv
50	1.92	1.42	0.50	0.04	2	Pu
51	0.19	0.40	-0.20	0.38	1	Pv
52	2.66	1.53	1.12	0.00	2	Pu
53	-0.84	-0.40	-0.44	0.12	1	Pv
54	0.02	0.41	-0.39	0.11	1	Pv
55	-2.29	-0.91	-1.38	0.00	2	Pv
56	-2.53	-1.24	-1.24	0.02	2	Pv
57	-0.36	-0.19	-0.19	0.50	1	Pv
58	-1.06	-0.18	-0.89	0.00	2	Pv
59	0.02	0.58	-0.56	0.02	2	Pv
60	0.02	-0.45	0.47	0.05	1	Pu

bpu= measure for public, bpv= measure for private,Δb=DIF contrast

Table 5 shows the DIF statistics of the Rasch model method for each of the 60 items for school ownership. There is incidence of DIF if the probability is less than 0.05. There is incidence of DIF in 18 items at 0.05 levels of Significance. That is 30% of the 2014 BECE Social Studies multiple choice test items functioned differentially for public and private examinees. The DIF items are 3, 4, 8, 16, 24, 31, 34, 36, 41, 42, 45, 48, 50, 52, 55, 56, 58, and 59.

Hypothesis One: The number of items that function significantly different for male and female students is not significantly different from the number that did not differentially function.

Table 6: Chi-square Statistics for Gender DIF

Gender	DIF-Gender		Total	X^2	Df	sig	Decision
	Non DIF Items	DIF Items					
Male	28	5	33				
Female	23	4	27	0.10	1	0.97	Not significant
Total	**51**	**9**	**60**				

Table 6 shows a chi-square value of 0.10 and a P-value of 0.97. Testing at an alpha level of 0.05, the p-value is greater than the alpha level. Hence, the null hypothesis is accepted. Consequently, the number of items that function significantly different for male and female students is not significantly different from the number that did not differentially function.

Hypothesis Two: The number of items that function significantly different for urban and students is not significantly different from the number that did not differentially function.

Table 7: Chi-square Statistics for Location DIF

Location	DIF-Location		Total	X^2	Df	sig	Decision
	Non DIF Items	DIF Items					
Urban	18	11	29				
Rural	18	13	31	0.19	1	0.75	Not significant
Total	**36**	**24**	**60**				

Table 7 shows a chi-square value of 0.19 and a p-value of 0.75. Testing at an alpha level of 0.05, the p-value is greater than the alpha level. Hence, the null hypothesis is accepted. Consequently, the number of items that function significantly different for urban and rural students is not significantly different from the number that did not differentially function.

Hypothesis 3: The number of items that function significantly different for high and low socio-economic status students is not significantly different from the number that did not differentially function.

Table 8: Chi-square Statistics for SES DIF

SES	DIF-SES		Total	X^2	Df	sig	Decision
	Non DIF Items	DIF Items					
High	27	5	32				
Low	23	5	28	0.15	1	0.82	Not significant
Total	**50**	**10**	**60**				

Table 8 shows a chi-square value of 0.15 and a p-value of 0.82. Testing at an alpha level of 0.05, the p-value is greater than the alpha level. Hence, the null hypothesis is accepted. Consequently, the number of items that function significantly different for high and low socio-economic status students is not significantly different from the number that did not differentially function.

Hypothesis 4: The number of items that function significantly different for students in Single-sex and Mixed-sex schools is not significantly different from the number that did not differentially function.

Table 9: Chi-square Statistics for School Type DIF

School Type	DIF- School Type		Total	X^2	Df	sig	Decision
	Non DIF Items	DIF Items					
Mixed	23	10	33				
Single	20	7	27	2.14	1	0.71	Not significant
Total	43	17	60				

Table 9 shows a chi-square value of 2.14 and a p-value of 0.71. Testing at an alpha level of 0.05, the p-value is greater than the alpha level. Hence, the null hypothesis is accepted. Therefore, the number of items that function significantly different for students in single-sex school and mixed-sex schools is not significantly different from the number that did not differentially function.

Hypothesis 5: The number of items that function significantly different for students in public and private schools is not significantly different from the number that did not differentially function.

Table 10: Chi-square Statistics for School Ownership DIF

School Ownership	DIF- School Ownership		Total	X^2	Df	sig	Decision
	Non DIF Items	DIF Items					
Public	19	9	8				
Private	23	9	32	0.12	1	0.74	Not significant
Total	42	18	60				

Table 10 shows a chi-square value of 0.12 and a p-value of 0.74. Testing at an alpha level of 0.05, the p-value is greater than the alpha value. Hence, the null hypothesis is accepted. Therefore, the number of items that function significantly different for

students in public and private schools is not significantly different from the number that did not differentially function.

5. Discussion

Incidence of Differential item functioning in BECE Social Studies multiple choice test. The analysis of the student's response to Social Studies multiple choice test used in BECE in 2014 revealed that there is incidence of DIF in 9 items. This implies that BECE 2014 social studies multiple choice test items function differentially for male and female students. The finding of this study is in line with research study by Abedalaziz (2010) who reported incidence of gender DIF in mathematics. Also, Odili (2003) revealed that there was evidence of gender DIF in WAEC/SSCE Biology paper 2 for 1999, 2000 and 2001.

The result of the students' responses to 2014 BECE Social Studies multiple choice test shows that there is incidence of location DIF in 24 items. That is 40% of the 2014 BECE Social Studies multiple choice items functioned differentially for urban and rural examinees. This finding is in consonance with the result of the study carried out by Odili (2003). Consequently, Umoinyang (1991) analysed Mathematics multiple choice test used by West African Examination Council (WAEC) in the 1990 General Certificate Examination (GCE). His study revealed 29 items that differentially function in favour of candidates from educationally advantaged area.

Also, the analysis of the students' response to social studies multiple choice test used in BECE in 2014 revealed that there is incidence of Socio-economic Status DIF in 17 items. That is 17% of the items functioned differentially for high and low Socio-economic status examinees. The result corroborates with the findings of Alordiah (2015). Also, Odili (2003) showed evidence of the presence of SES differential item functioning items in WAEC/SSCE Biology paper 2 for 1999 and 2001.

The result of the students' response to 2014 BECE Social Studies multiple-choice test also shows that there is incidence of School type DIF in 28 items of the 2014 BECE Social Studies multiple choice test for Single-sex and mixed-sex schools examinees. This result agrees with Pahlke, Hyde and Allison (2014) who published a meta-analysis comparing achievement and attitudes in single-sex versus mixed sex schools. The study concluded that there is a little evidence of an advantage of single-sex schooling for boys or girls. This finding however, is in conflict with Cornelius (2015) revealing that there is no advantage to single-sex school or mixed-sex schools.

Consequently, the analysis of the students' response to 2014 BECE Social Studies multiple choice test items revealed the incidence of school ownership DIF in 18 items of the 2014 BECE Social Studies multiple choice test for public and private schools examinees. That is 30%of the 2014 BECE social studies multiple choice test items functioned differentially for school ownership. This finding is in line with Ogbebor and Onuka (2013) in a study titled differential item functioning method as an item bias indicator. They revealed that the items in relation to private and public school using

logistic regression method. Out of 60 items in NECO Economics questions, DIF was present in 10 items.

6. Conclusion

The study investigated differential item functioning in Social Studies multiple choice questions in Basic Education Certificate Examination in Delta State. Based on the findings there are incidence of gender, location, socio-economic, school type and school ownership differential item functioning in 2014 BECE Social Studies multiple choice questions. The study concluded that differential item functioning exist in 2014 Basic Education Certificate Examination in Delta State.

6.1 Recommendations

Based on the findings the study, the researchers made the following recommendations:

- Examination bodies should be mindful of the disparities that exist between gender, location, socio-economic status, school type and school ownership.
- Governments should set up a review panel national/state wide examinations to review DIF items.
- Test developers should write social studies test items that would not favour one group against the other. They should be sensitive to the heterogeneous nature of Nigeria.
- BECE should analyze students' responses to test items for differential functioning before administration on examinees.
- Test developers should be trained and retrained on how to identify DIF items and how to write DIF- items.

References

1. Adebalaziz, N. (2011). Detecting DIF using item characteristic curve approaches. *The international Journal of Education and Psychological Assessment, 8(2),1-15*
2. Adedoyin, O.O. (2010). Using IRT Approach to Detect Gender Based Items in Public Examination. *Education Research and Reviews Academic Journals.* 5(7), 385-399.
3. Alordiah, C. O. (2015). Comparison of index of Differential Item functioning under the methods of Item Response theory and classical test theory in Mathematics. An unpublished Ph.D thesis of Delta State University, Abraka, Delta State, Nigeria
4. Amina, J. (1995). Quality as Stress in Nigeria Education Kaduna, Northern Nigeria Publishing Company.
5. Anastasia A. & Urbina S. (2006). Psychological Testing. New Delhi: Prentice Hall.

6. Arnkelsson, G.A. (2000). Stability of Related Attributes and the Inference of Ability through Social Comparison. Personality and Social Psychology Bulleting, 26, 937-938.
7. Atar, B. (2006). Differential Item Functioning Analysis for mixed response data Using IRT-likelihood ratio test, logistic regression, and GLLAMN Procedures. Electronic thesis, treatises and dissertations (ETDs). Paper 248 http://diginole.fsu.edu/etd/248
8. Augemberg, K.E; & Morgan, D.L. (2008). Differential performance of test Items by geographical regions. Paper presented at the annual meeting of the National Council on measurement in Education in New York, NY
9. Australian Bureau of Statistics Retrieved: 11-11-2016.
10. Baker, F. B. (2001). The Basics Item Response Theory. USA. ERIC clearing-house on assessment and evaluation
11. Berk, T. (2007). Item Bias Detection Methods for Small Samples. Dissertation Abstract International.
12. Blanton, H. (2001). Evaluating the Self in the Context of Another: *The Three Selves Model of Social Comparison Assimilation and Contrast. Journal of Applied Social Psychology*. Bulleting 27 (12) 1628-1629.
13. Bolt, D. M. (2002). A Monte Carlo comparison of parametric and non-parametric polytomous DIF detection methods. Applied psychological measurement, 15,133-141
14. Braun, H., Jenkins, K. & Griggs, W.(2006). Comparing Private Schools and Public Schools using Hierarchical Linear Modelling U.S Department of Education, Institute of Education Sciences, National Centre for Education Statistics, NCES 2006 461.
15. Camilli, G. & Shepherd L.A. (1994). Methods for Identifying Bias Test: Items Vol. 4. Thousand Oaks, CA: Sage Publication.
16. Chong, H. Y. (2010). A simple guide to the item response theory (IRT) and Rasch modelling. http://www.creative-wisdom.com
17. Considine, G. &Zappala, G. (2002). Influence of Social Land Economics Disadvantaged in the Academic Performance of School Students in Australia. *Journal of Sociology, 38, 129-148. Retrieved 11-11-2016 from* http://www.jos.sage.pub.
18. Cronbach, L.J. (1971). Test validation. Psychological Bulletin 52,281-302.
19. Egbule, J.F. (2002). Principles of Psychological Testing, Owerri. Whyte and Whyte Publishers.
20. Emaikwa, S.O. (2012). Issues in Test Item Bias in Public Examinations in Nigeria and Implications for Testing. International Journal of Academic Research in Progressive Education and Development. 1(1), 175-185.
21. Embretson, S. & Reise, S.P. (2000). Item Response Theory for Psychologists. Mahwah Hew Jersey Lawrence Erbaum Associates Publishers.

22. Emanovwe, A. (2009). The Impact of Gender, Socio-Economic Status and School Location on Students' Academic Achievement in Basic Technology. Unpublished M.Ed. Project, Delta State University, Abraka.
23. Federal Republic of Nigeria (2004). National policy on Education. Lagos. NERDC press.
24. Ferder, R.M., Mohr, P.H. & Dietz, E. J (1994). A Longitudinal Study of Engineering Student Performance and Retention. 11 Differences between Students from Rural and Urban Backgrounds. *Journal of Engineering Education, 83 (3).*
25. George, A. & Goethais, V. (2001). Peers Effects, Gender and Intellectual Performance among Students at a Highly Selective College. Comparison of Ability in Analysis: Research in Higher Education. 26, 427-446.
26. Gieris, M.J. (1999). Test bias, Sage Publications, Inc., California.
27. Goux, D. & Maurin, E. (2008). Public School Availability for Two Years old and Mothers' Labour Supply. American Economic Association. 104(1) 252-276
28. Greenward, A.G. (2009). The Brief Implicit Applications and Data Analysis Methods (2nded) Thousand Oaks, CA. Saga Publication.
29. Hambleton, R. K., Swaminathan, H., & Rogers, J. H. (1995). Fundamentals of Item response theory. Newbury park, CA. Sage
30. Ija, C.N. (2009). Towards promoting gender equality and women Empowerment for sustainable development in River state. *Journal Of International Gender studies,4,225-235*
31. Jorgensen, A.R. (2005). Cumulative Deficit in Compensatory Education. *Journal of Psychology, 4, 37-47.*
32. Joshua, M.T. (2005). Test/item Bias in Psychological Testing: Evidence on Nigeria System. A Paper Presented at the Annual Conference of NigerianAssociation of Educational Psychologists held at Ahmadu University Zaria from 24th – 28th.
33. Kolcic, I. (2006). Academic Performance and Scientific Involvement of Final Year Medical Students Coming From Urban And Rural Background of Andjija Stamper School of Public Health, Medical School, University of Zagreb, Croatia http://rrh.Deakia.rdu.au/
34. Lord, F.M. (2003). Application of Item Response Theory to Practical Testing Problems. Hilldale, New Jersey: Erbaum Press.
35. Lee, J. & Melritire, W. (2001). Interstate Variation in the Mathematics Achievement of Rural and Non-Rural Students. *Journal of Research in Rural Education, 16(3).*
36. Malouff, J.M. (2008). Bias in Grading. College Teaching, 58(3), 191-192.
37. Mercer, J.R. (1979). In Defence of Racially and Culturally Non-Discriminatory Assessment. School Psychology Digest, 8, 89-115.
38. Odili, J.N. (2003). Effect of language manipulation on differential item functioning of test items in Biology multiple choice test. An unpublished Ph.D thesis. University of Nigeria Nsukka.

39. Odili, J.N. (2010). Effect of language manipulation on differential item functioning of test items in Biology in a multicultural setting. *Journal of Educational assessment in Africa, 4,268-286*
40. Ogbebor, U.C. (2012). Differential item Functionary Economics Question Paper of National Examination Council in Delta State Nigeria. Unpublished M.Ed. Thesis. University of Ibadan.
41. Owoeye, J.S. (2000). The Effect of Integration of Location Facilities and Class Size on Academic Achievement of Secondary School Students in Ekiti State, Nigeria Ph.D. Thesis, Unpublished. University of Ibadan.
42. Pahike, E., Hyde, J.S & Allison, C.M. (2014). The Effect of Single Sex Compared with Coeducational Schooling On Students' Performance and Attitudes: A Meta-Analysis. Psychological Bulleting. 140 (4), 1042.
43. Park, J.H., Barhem, R.J. & Choi, J. (2012). Causal Effect of Single-Sex School on College Entrance Exams and College Attendance: Random Assignment in Seoul High Schools. PSC Working Paper Servers.
44. Perrone, M. (2006). Differential Item Functioning and Test Bias: Critical Consideration in Test Fairness. Applied linguistics,
45. Rauschemberg, S.W. (2012). Hierarchical Linear Models: Applications and Analysis Methods (2nd .ed) Thousand Oaks, CA. Sage Publication.
46. Regner, I., Huguet, P., & Monteil, J.M. (2002). Effects of Socio Economic Status (SES) Information on Cognitive Ability Inferences: When Low SES Students Make Use of a Self-Threatening Stereotype. Social psychology of Education, 5, 259-269.
47. Rudner, L. A. (2001). Item response theory. http://echo.edres.org.8080/irt/baker/chapter/.pdf.
48. Schemo, D.J. (2006). Most Students in Big Cities Lag Badly in Basic Science. New York Times, November 16 2006, Section A, 22.
49. Skiba, R.J., Knesting, K. & Bush, L.D. (2002). Culturally Competent Assessment: More than Nonbiased Tests. *Journal of Child and Family, 11, 61-78.*
50. Swaminathan, H; & Rogers, H.J. (1990). Detecting differential item Functioning using logistic regression procedures. *Journal of Educational measurement,27,311-370*
51. Teresi, J. (2004). Differential item functioning and health assessment. Colombia University Stroud centre and faculty of medicine, New York State psychiatric institute, research division, Hebrew home for The aged at Riverdale
52. Trembley, S., Ross, N.A. & Berthelot, J.M. (2001). Individual and Community Factors Affecting Grade Three Academic Performance, a Multi-Level Analysis. Education Quarterly Review, 7(4), 25-36.
53. Umoinyang, I.E. (2011). Method of detecting item functioning (DIF) as a Consistent error in achievement test. *Journal of the Science Teachers Association of Nigeria (JSTAN). Stanonline.org/imo2011.pdf*
54. University of Pennsylvania (2012). University Archives and Research Centre. University of Pennsylvania. Retrieved November 25 2016.

55. U.S. Department of Education (2005). Single Sex versus Co-Educational Schooling: A System Review Department of Education, Washington, D.C.
56. Van de Vijiver, F. & Tanzer N.K. (2004). Towards and Integrated Analysis of Bias in Cross-Cultural Assessment. *European Journal of Psychological Assessment, 13, 29-37.*
57. Webster's collegiate Dictionary (2000). An American Dictionary of English Language, 2nd Edition. Springfield Massachusetts.
58. Zumbo, B.D. (1999). A handbook on the theory and methods of modelling as a unitary framework for binary and likert (ordinal) item scores. Ottawa, Canada. Directorate of human resources research and Evaluation department on national defence. Retrieved from: http://www.educ.uba.ca/faculty/zumbo/DIF/index,html