VOLUME 12, NUMBER 1, 2013 ISSN 0795-3607

NIGERIAN JOURNAL OF EDUCATIONAL RESEARCH AND EVALUATION

A PUBLICATION OF

THE NIGERIAN ASSOCIATION OF EDUCATIONAL RESEARCHERS AND EVALUATORS

DIFFERENTIAL ITEM FUNCTIONING OF ECONOMICS QUESTION PAPERS OF NATIONAL EXAMINATIONS COUNCIL IN DELTA STATE, NIGERIA

By

OGBEBOR,Uruemu and Adams ONUKA
Institute of Education, University of Ibadan, Nigeria
+234-803-702-7063/ogbeborremu@yahoo.com
+234-803-356-4064/adamonuka@yahoo.com

Abstract

This study attempted to detect item bias using differential item functioning approach in relation to gender (male and female) and school type (private and public schools).Using National Examinations Council (NECO) Economics questions for 2010. The research design employed in this study was a comparative research type of design. The study sample comprised students in Ethiope East Local Government Area of Delta State. Four hundred and forty seven (447) students were used. The test contains 60 items which was administered to the students. To detect item bias in relation to gender and school type, logistic regression was used; the result showed that there were item biases in terms of gender and school type. It was then recommended that items should be constructed in future taking cognizance of gender and school type.

INTRODUCTION

The issues of educational measurement in research pointing towards enhancing the fairness of test or examination across sub groups of examinees is very essential because important decisions are made based on scores of the examinees. A fair test is one that enables all examinees to have an equal chance to demonstrate the skills and knowledge which they have acquired and which are vital to the test's purpose. Test fairness can be viewed as any test given to a set of testees with an equal chance to demonstrate what they know. Various aspects of fairness in testing have been highlighted in literature, including fairness in regards to standardization, test consequences/score use, and item bias (Kunnan, 2000; Shohamy, 2000).

A fair test is one that affords all examinees an equal opportunity to demonstrate the skills and knowledge which they have acquired and which are relevant to the test's purpose (Roever, 2005). The existence of bias is an issue to be addressed because tests are used as a gatekeeper for educational opportunities and it is a very important issue that test items are fair for every examinee. Bias is the existence of some characteristics of pools of items that cause differential performance for individuals of the same ability but from different ethnic, sex, cultural or religious groups. An examination item is said to be bias if it functions differently for a specified subgroup of test takers. Bias test measures characteristics that are not necessary or items that are irrelevant to the test.

Frequently, examination items are considered biased because they contain sources of difficulty that are not relevant to the construct being measured and these extraneous sources affects test-takers' performance (Zumbo, 1999).

Item bias and differential item functioning (DIF) has critical political, social and ethical implications for test developers, policy makers and examiners. The study of item bias and DIF is critical as such research helps to provide an empirical foundation for the identification and subsequent elimination of examination items that appear to be relatively more difficult for one group of test-takers than another. Further research on these issues will allow us to comprehend more fully the possible substantive interpretation that can be made by focusing on test items considered to be biased. Investigations made in recent years on students' score in economics have shown that students or test takers of particular subgroups perform better than other subgroups. The difference in the performance of students in economics by male and female, public and private or other subgroups have long been studied

Findings, from studies show that differences in test of student achievement and low test scores in some subject areas such as mathematics and economics could be attributed to social and cultural influences that create sex role stereotypes that reduces female interest and achievement in traditionally male-dominated subjects (Williams, Waldaver and Duggal 1992; Anderson, Benjamin, and Fuss 1994; Hirschfeild, Moore, and Brown 1995). Studies have been conducted in this area of item bias and DIF. Pedrajita (2009), in a study "using Logistic regression to detect test items bias in Chemistry Achievement", the result from the study revealed that there are gender bias and class bias in Chemistry Achievement test. Nworgu (2011), revealed that current research evidence has implicated ode used in national and regional examinations as functioning differently with respect to different subgroups. This means that students' scores in such examinations are determined largely by the group to which an examinee belongs and not by ability. The researcher has it in mind that there is the possibility that cognitive differences between male and female, may result in students performance differences, which may limit the economics test, instructional differences may also limit the economics understanding of female. Among these are classroom climate for women based educational materials and poor teachers role model (Ferber 1990, Harvath, Begudin and Wright 1992), also, the constructed response format of a test may influences test results. Which implies that the differences in male and female, private and public school student performances could be as a result of items odelling a set of subgroups, as such bias could be view as a systematic error that distort the inferences made in classification and selection of students. Students who have similar conditions of learning, similar knowledge of the material on a test taken should perform similarly on some examinational items regardless of their gender differences (male and female), and school type (Private and public). An item is bias if it measures attributes that are irrelevant to the tested construct or tracts and these extraneous sources have impact on test takers performance.

Recently researchers have used logistic regression analysis methods to detect item

		Gender		
		Male	female	Total
ITEM 7	NC	97	129	227
	C	117	103	220
	Total	214	232	447

N.B NC = NOT CORRECT C = CORRECT

Item 7 states:

Which of the following is not a primary production ?

a. Driving,
b. Fishing
c. Mining
d. Quarying

The correct answer is A, while the distracters are B, C, and D. The female examinees may have been confused in choosing the right answer because the stems are words that are familiar to the male. Thus the model or equation showed significance at (X^2 =4.031, df=1, N= 447, p< 0.05).

ITEM 26 : Item 26 favoured female and male were the disadvantaged group. And this is represented in

Table 3 Cross tabulation showing group performance on item 26

ITEMS 26 * gender Crosstab

Count

Count

		Gender		
		Male	Female	Total
ITEM 26	NC	155	144	299
	C	59	88	148
Total		214	232	447

NB NC =NOT CORRECT C = CORRECT

Item 26 states:
The consumers price index is calculated as
a. Base year price index 100/Initial year index
b. Base year index 100
c. Current year price index 100/Base year price index
d. current year price index / 100
e. 100 100/Current year index

The correct answer is A while the distracters are B, C, and D. The male examinees may have being confused in choosing between option A and C due to understanding of the words i.e. current and base year index. And this could be caused by teacher incompetence of teaching the subject content with different text books. Thus the model or equation showed significance at (X^2 =4.420 ,df=1, N= 447, p< 0.05).

SCHOOL TYPE

Table 2.0 shows the items that possesses DIF in relation to school type (private and public), identified by logistic regression method using SPSS version 18. Cross tabulation was use to detect the subgroups that are disadvantaged. The biased items were, item 9, item 13 and item 14, item 16, item 43, item 44, item 46, item 49, item 50, and item 59

Table 4: Items that showed DIF in relation to school type Identified by logistic regression method.

Item	df	chi square	p	DIF against
9	1	4.573	.033	public
13	1	11.043	.001	public
14	1	9.370	.002	private
16	1	14.090	.000	private
43	1	13.471	.000	private
44	1	4,498	.034	public
46	1	5.309	.021	public

49	1	4.694	.030	public
50	1	5.035	.027	private
59	1	7.456	.007	public

Critical $X^2 = 3.84, p < 0.05$

Out of sixty items in NECO economics questions DIF in relation to school type was present in ten (10) items. Item 9, item 13, item 44, item 46, item 49, and item 50 favouredprivate school examinees and public school examinee were of the disadvantaged group. Table 2.1 is crosstab table showing group differences in their performance

Table 5: **ITEM 9 *school_type Crosstab**
Count

		school_type		
		Public	private	Total
Item 9	NC	160	113	273
	C	84	90	174
Total		244	203	447

Item 9 states:
The over-all success or failure of a business mainly depends on
A. Amount of capital involve
b. Nearness to the market
c. Nearness to a similar
d. Quality of managerial skills
Size of the business

The correct answer is D, while the distracters are A, B, C AND E, the public school examinees are being disadvantages it may be as a result of construct of the item because the item is constructed in a double barrel form which meant confuse those student from public schools. Thus the model or equation showed significant difference at (X^2= 4.573, df=1, N=447, p< 0.05)

Table 6: Item 13 Crosstab

Count

		school_type		
		Public	private	Total
Item 13	NC	175	115	290
	C	69	88	157
Total		244	203	447

Item 13 states:
Which of the following is a weapon used by trade unions during trade dispute?

A. Black list
b. Lock-out
c. Strike breaker
d. Sack

Work-to-rule

The correct answer is C while the distracters are A, B, D, and E.The distracters meant be seeming confusing to the public school examinees. Thiscould make them select the distracter A or E. Thus the model or equation showed significance at (X^2 = 4.420, df =1, N= 447, p< 0.05).

Table 7: **Item 44 Crosstab**

		school_type		
		Public	Private	Total
Item 44	NC	191	141	332
	C	53	62	115
Total		244	203	447

Item 44 states:
The following are ways by which the government can encourage industrialization EXCEPT by
a. Decongesting urban areas
b. Providing medical facilities
c. Reducing company tax
d. Reducing import duties
Relaxing industrial laws

The correct answer is B while the distracters are A, C, D, and E. The distracters is seemingly confusing to the public school examinees they may have being deceptively made to select distracter A which is negatively stated . This item showed significance at ($X^2 = 4.498$, df=1, N= 447, p< 0.034).

Table 8: **ITEM 46 Crosstab**

Count

		school_type		
		public	Private	Total
Item 46	NC	172	122	294
	C	72	81	153
Total		244	203	447

Item 46 states:
The upward sloping of the supply curve implies that the
A. Price does not affect the quality supplied
b. Quantity supplied has a direct relationship with price
c. Quantity supplied has an inverse relationship with price
d. Price has a negative relationship with the quantity supplied
Two variables do not move in positive directions

The correct answer is B. This item has a cue to identify the correct answer i.e the word "upward" and "direct" this cue meant not have been noticed by most of the public school examinees who were less likely familiar with the movement of the supply curve. These examinees may have being given less exposure on the movement of the supply curve. Thus the item showed significance at ($X^2 = 5.309$, df = 1, N = 447, p<

0.021)

Table 9: **Item 49 Crosstab**

Count

		school_type		
		public	Private	Total
Item 49	NC	144	99	243
	C	100	104	204
Total		244	203	447

Item 49 states:
If Y = C + I + G and C = a + by , then Y would be equal to
A. Y = C + a + by l + G
b. Y = a + by + C + I + G
c. Y = a + by + I + G
d. Y = a + b + I + G
Y = C + I + G

The correct answer is C. The public school examinees may not be familiar with the concept of national income and the use of the mathematical equation and its transposition. The item showed significance at $X^2 = 4.694$, df = 1, N = 447 p > 0.030.

Table 10: Item 59 Crosstab

Count

		school_type		
		Public	private	Total
Item 59	NC	183	128	311
	C	61	75	136
Total		244	203	447

Item 59 states:
Which of these is NOT a feature of the third world nation?]
a. Dual economy
b. Low agricultural productivity
c. Low savings and investment
d. Low population growth
High dependency

The correct answer is A while the distracters are B, C, D, and E. Examinees lacking background knowledge on the history of the third world nation meant not be able to pick the correct response to the item. Thus the model or equation showed significance at ($X^2 = 7.456$, df =1, N= 447, p< 0.007).

Biased Items Against The Public School Examinees

The four biased items against the public school examinees were items 14, 16, 43, and 50.

Table 11: **Item 14 school_typeCrosstabulation**

Count

		school_type		
		Public	private	Total
Q14	NC	34	11	45
	C	210	192	402
Total		244	203	447

Item 14 states:
In the chain of distribution, which of the following arrangement is correct
a. Consumer → manufacturer → wholesaler → retailer
b. Manufacturer → retailer → consumer → wholesaler
c. Manufacturer → wholesaler → retailer → consumer
d. Retailer → manufacturer → consumer → wholesaler
e. Wholesaler → agent → retailer → consumer

The correct answer is C while the distracters are A, B, D, and E. Option E seemed plausible for a great number of private school examinees. However the low performance of the examinees was due to the limited exposure to the topic on chain of distribution. Thus the model or equation showed significance at ($X^2 = 9.370$, df =1, N=

447, p< 0.002).

Table 12**item16 * school_type Crosstab**

Count

		school_type		
		Public	private	Total
Q16	NC	183	179	362
	C	61	24	85
Total		244	203	447

Item 16 states :
Which of the following is NOT a representative of money
A. Cheque
b. Coin
c. Money order
d. Postage stamp
Postal order

The correct answer is D while the distracters are A, B, C, and E. Examinees lacking background knowledge on representative of money meant not be able to pick the correct response to the item. Thus the model or equation showed significance at (X^2 = 14.090, df=1, N= 447, p< 0.000).
Table 13

Item 43 * school_typeCrosstabulation

Count

		school_type		
		Public	private	Total
Q43	NC	211	195	406
	C	33	7	40
Total		244	202	447

Item 43 states:
Institutions involved in the capital market include the following EXCEPT
a. Development bank
b. Insurance company
c. Stock exchange
d. Issuing houses
e. Merchant banks

The correct answer is C while the distracters are A, B, D, and E. Examinees lacking background knowledge on capital market meant not be able to pick the correct response to the item. Thus the model or equation showed significance at ($X^2 = 13.471$, df=1, N=447, p< 0.000).

Table 14
Item 50 * school_typeCrosstabulation

Count

		school_type		
		Public	private	Total
Q50	NC	169	162	331
	C	69	39	108
Total		238	201	439

Item 50 states:
National income statistics as a measure of the living standard of a country is done on the basis of
a. Income earned by the working populace
b. Income per annum
c. Per capital income
d. The amount of aids contribution by the country
The amount that the country has in foreign reserve

The correct answer is C while the distracters are A, B, D, and E. This item has unequal length and complexity of alternatives apparently option B and C seem plausible for the private school examinees this could have miss lead them choosing the correct answer .Thus the model or equation showed significance at ($X^2 = 5.035$, df =1, N= 447, p< 0.027).

DISCUSSION

Logistic regression statistics showed detected items that has DIF against subgroups such as male and female, private and public school examinees. The result indicated that these items that showed DIF were relatively more difficult for the affected groups. Also both the advantaged and disadvantage groups had not gotten equal opportunity to learning on the subject economics. The structure of the questions may be poorly worded or may be unfamiliar to the affected group. However the affected groups could have received limited lessons on the subject being taught or the teacher who taught the subject did not explain the concept of those topic that showed presence of DIF in the NECO economics examination questions. Thus these could be characteristics that affected the test takers response to getting the item correctly. Literature has shown that items used in assessing student ability has element of biasness.

Gierl's (1999), a study on DIF in Alberta examined 30 education Social Studies Diploma students, the study evaluated the effects of DIF between male and female, the results indicated that the majority of multiple choice items did not display DIF using the three-tiered ratings. 65 of 70 item displayed negligible effects, 5 items with moderate DIF, three favoured male and two favoured female, this indicate that the test contained items that functioned differently for male and female. Pedrajita (2009), in a study "using Logistic regression to detect test items in Chemistry Achievement", the result from the study revealed that there are gender bias and class bias in Chemistry Achievement test. Nworgu (2011), revealed that current research evidence has implicated test used in national and regional examination as functioning differently with respect to different subgroups. This means that students' scores in such examinations are determined largely by the group to which an examinee belongs and not by ability.Adedoyin (2010) in his study on investigating gender baised items in public examinations, he found that out of 16 test items that fitted the 3PL item response theory statistical analysis, 5 items were gender biased.

From this study the analysis indicated presences of DIF in two items in relation to gender, while ten items showed DIF in relation to school type. Thus there are presences of gender bias and school type bias in national examinations council economics question for 2010. And this could be attributed to

- Learners not being familiar with the content of the items which leads to examinees to be attracted to the wrong answer.
- Exposure to vocabularies, concepts that reflects on items
- The examinees where unable to comprehend and understand concepts that were present on the item

Conclusion

Based on the forgoing findings the following conclusion were made

There were presences of gender and school type DIF and item bias in NECO economics questions. Despite the presences of gender bias items in NECO economics questions it is not enough to say that gender bias was actually present because the items

that indicated gender biasness is not significant.

Recommendation

On the basis of the findings and conclusion, the following recommendations are made:

1. Test experts and developer should explore the use of differential item functioning method, particularly the use of logistic regression to detect both uniform and no uniform biased items.
2. A study of this should be conducted to provide further empirical evidence on the validity of the method in detecting biased test items
3. Evaluators and educational practitioners who are engaged in the development of assessment tools should use logistic regression for bias correction
4. Measurement practitioners should make use of logistic regression for developing a valid , reliable gender fair test school type fair test with biased items revised or replaced
5. The subject curriculum should be made clear for teachers to be able to teach the concept effectively
6. Teachers should exposure learners to more than one textbooks

REFERENCES

Camilli G, Shepard LA (1994). *Methods for identifying biased test items.* Newbury Park, CA: Sage.

Embretson S, Reise SP (2000). *Item response theory for psychologists.* Mahwah New Jersey: Lawrence Erlbaum Associates Publishers.

Hambleton, R., & Rodgers, J. (1995).Item bias review.Practical Assessment, Research, and Evaluation, 4(6). Retrieved November 18, 2006, from http://PAREonline.net/getvn.asp?v=4&n=6

Hambleton R, Rodgers J (1995). Item bias review.Practical Assessment, Research, and Evaluation, 4(6).Retrieved March 18, 2009, from http://PAREonline.net/getvn.asp?v=4andn=6.

Hambleton RK, Swaminathan H, Rogers HJ (1991). *Fundamentals of item response theory.* Newbury Park, California: Sage Publications.

Hambleton RK, Swaminathan H (1995). Item response theory: Principle and application. Boston: Kluwer.

Kimball MM (1989). *A new perspective on women's math achievement.* Psychol. Bull. 105: 198-214.

Lord FM (1980). *Applications of item response theory to practical testing problems.* Mahwah, NJ: Erlbaum.

Perrone M (2006). Differential item functioning and item bias: Critical consideration in test fairness. *Applied Linguistics*, 6(2): 1-3.

Kunnan, A. J. (2000). Fairness and justice for all. In A. J. Kunnan (Ed.), Fairness and validation in language assessment: Selected papers from the 19th Language Testing

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/brownbag.pdf

Schumacker, R. (2005). Test bias and differential item functioning. Retrieved November 18, 2006, from http://www.appliedmeasurementassociates.com/White%20Papers/TEST%20BIAS%20AND%20DIFFERENTIAL%20ITEM%20FUNCTIONING.pdf

Shohamy, E. (2000). Fairness in language testing. In A. J. Kunnan (Ed.), Fairness and validation in language assessment: Selected papers from the 19th Language Testing

Williams, V. (1997). The "unbiased" anchor: Bridging the gap between DIF and item bias. *Applied Measurement in Education,* 10, 253-267.

Zumbo, B. D. (1999). *A handbook on the theory and methods of differential item functioning (DIF): Logistic regression odelling as a unitary framework for binary and Likert-like (ordinal) item scores.* Ottawa, Canada: Directorate of Human Resources Research and Evaluation.