

A COMPARATIVE STUDY OF DIFFERENTIAL ITEM FUNCTIONING FOR COMPROMISED AND SECURED ECONOMICS ITEMS AMONG SECONDARY SCHOOL STUDENTS IN OGUN STATE

**Yusuf Olayinka Shogbesan, Ayobode Patricia Aowo & Isaac Bamikole
Ogunsakin**

Department of Educational Foundations and Counselling, Faculty of Education,
Obafemi Awolowo University, Ile-Ife, Nigeria

Abstract

This study investigated the extent to which moderating variables such as sex and school type affect the occurrence of differential item functioning for compromised and secured economics items. The study adopted the quasi-experimental research design with a population which comprised all the secondary school students in Ogun State and a sample of 1500 economics students selected through multi-stage sampling. The instruments used for the study were “Economics Achievement Test (EAT)” -adapted from NECO 2015 SSCE economics objective test paper III and an OMR-type answer sheet. Data collected was subjected to DIF analysis which was calibrated in R language. The results showed that sex and school type of examinees had a moderating effect on the occurrence of differential item functioning (DIF) for 34 (56.7%) and 42 (70%) compromised and secured item conditions. In addition, the effect of compromised and secured items contributed more to the occurrence of DIF based on sex and school type of the examinees. It was recommended that differential item functioning analysis should be carried out on all items suspected to be compromised as a statistical evaluation by test experts and examination bodies.

Keywords: item compromise, differential item functioning (DIF), secured items

Introduction

Tests are meant to elicit information about the latent ability of an individual and provide evidence for taking educational decisions about the individual test-taker. Based on scores from tests, schools decide who is to be promoted, examining agencies decide who is to be certified, higher institutions decide who is to be admitted and for which course and organizations decide who is to be employed or recruited. Since these decisions provide information about examinees' knowledge and skills, the scores should reflect the most accurate estimates of skills and abilities. To adequately perform this function, tests must be free from measurement errors that can negatively influence reliability and validity. One factor that introduces measurement error is the occurrence of cheating during examinations.

Cheating in examinations has become a serious threat to the reliability and validity of examinations. It is regarded as any action that violates the rules for administering a test and gives an examinee an unfair advantage over others, or any action on the part of an examinee or test administrator that decreases the accuracy of the

intended inferences arising from the examinee's test score or performance (Cizek, 2001). Cheating can be caused by a breach in the security of examination questions. This has been acknowledged as a major problem that has impacted negatively on the integrity of public examinations in Nigeria (Ojerinde, 2015). The breach can be a result of leakages in the questions prior to the commencement of the examinations and are traceable to printers or persons connected with the custody of the question papers (Chinyere, 2014).

Cheating at the pre-testing stage is a form of fore-knowledge of test items and it results in item compromise. Adewale (2004) states that candidates offer incentives to individuals who have access to test items to leak the items to them before the examinations commence. They then study the materials and memorise answers to the questions raised in the examination papers. This scenario is termed 'item compromise'. Its implication is that the test takers become familiar with the test items which may positively affect their scores. When item compromise occurs, estimates of an examinee's performance may no longer be accurate.

Item compromise is a security breach which may undermine the psychometric properties of a test. It occurs when test contents have been distributed beyond defined valid usage boundaries (Zara & Pearson, 2006) or, specifically, when test-takers have access to the test questions on a test (Dragow, Nye, Guo, & Tay, 2009; Lievens & Burke, 2011). This gives a student undue advantage over other students in answering examination questions. Producers of educational assessments are now greatly concerned with these fraudulent testing behaviour resulting from benefiting from item compromise (Eckerly, 2017; Mcleod, Lewis & Thissen, 2003).

According to McLeod, Lewis and Thissen, (2003), the "normal" IRT model produces the probability of an item response for varying values of θ (ability). Item pre-knowledge modifies the IRT model because the probability of a correct response to an item depends on (a) the probability of answering the item correctly based on the test taker's pre-knowledge of the item and (b) on the test taker's underlying proficiency where the test taker did not have pre-knowledge of the specific item.

This compromise situation may also lead to the occurrence of test bias since the ability of tests to be fair, reliable, and valid primarily depends on whether the test items are secured. If they are not, it will be necessary to investigate or establish a measure that will lead to the detection of differential item functioning (DIF) or item bias. Zumbo (1999) and other researchers asserted that although DIF is necessary for declaring item bias, it is not a sufficient condition. Rather, a follow up item analysis such as content analysis and empirical evaluation will have to be carried out to declare an item biased (Zumbo, 1999). This study intends to investigate the occurrence of DIF for compromised and secured sets of items with sex and school type as moderating variables.

Differential item functioning (DIF) can also occur when examinees with the same ability from different groups, after controlling for their overall ability, have different probabilities of successfully answering an item (McCarty, Oshima and Raju, 2007; Rezaee & Shabani, 2010). DIF can be categorized into uniform and non-uniform. Uniform DIF refers to a situation which occurs when the probability of correctly

responding to an item is greater for one group than for the other across all the levels of proficiency. On the other hand, non-uniform DIF implies that the difference in the probability of a correct response is not the same at all levels of proficiency between the two groups being compared. This implies that the probability of correctly answering an item that exhibited non-uniform DIF is higher for one group at some given points on the scale than for the other. (Rezaee & Shabani, 2010).

Research has shown that DIF occurs when the responses of individuals having the same ability of interest show systematic differences simply based on their membership of certain groups. Amuche and Fan (2014) conducted a DIF study in relation to school type (private and public schools) and school location (urban and rural schools) using 60 items from National Examinations Council's Senior School Certificate Examinations Biology 2012 and a sample of four hundred and forty seven (447) candidates in Taraba State, Nigeria. Logistic regression DIF analysis showed that 10 items were biased in relation to school type and 8 items in relation to school location. Angoff (1993) reported that a DIF study was carried out to investigate the behaviour of items that had been compromised when a "performance contracting" organization engaged to develop and teach an educational programme in reading and mathematics in elementary schools had actually been teaching the children the correct answers to the items on the tests that were to be used for evaluating the programme. When compared with the control group, the group that benefitted from the compromised items recorded higher scores as they found the test items relatively easier. This implies that item compromise situations adversely affect item psychometric properties and raise fairness and validity concerns.

O' Leary (2013) compared candidates' performance on 80 scored and 20 non-scored items while assuming that only scored items were exposed and non-scored items were not compromised. 531 candidates (6.4%) were flagged for Differential Person Functioning (DPF). Many DIF studies have investigated only items considered secured which had not been made known to examinees prior to taking the examinations. It is imperative to investigate the degree of occurrence of differential item functioning when test items are compromised as compared to when they are secured. This study intends to provide answers to the following research questions:

Research Questions

1. Does sex have a moderating effect on the occurrence of differential item functioning (DIF) for the compromised and secured economics items?
2. Does school type have moderating effect on the occurrence of differential item functioning (DIF) for compromised and secured economics items?

Methods

The study adopted the quasi-experimental research design. From a population of 520,537 secondary school students in Ogun State, a sample of 1500 economics students in public and private secondary schools were selected using multi-stage sampling procedure.

Three local government areas (LGAs) were selected from each of the three senatorial districts in the state using simple random sampling technique. Four secondary schools (2 public and 2 private) were chosen from each selected local government area using stratified random sampling technique with school type as stratum. Proportional sampling technique was employed to select SSIII Economics students from each of the 36 secondary schools.. The research instrument was titled “Economics Achievement Test” (EAT). It comprised 60 multiple-choice test items each of which had a five-option format adapted from NECO SSCE 2015 Economics Paper III. The items were administered in two forms (A and B). Before administering the tests, 50% of the items were clustered as compromised (experimental items) while 50% were clustered as secured items (control items) and vice-versa. Less than an hour before the examination, items 1-30 were shown to the examinees under Form A while items 31-60 were not exposed to them. The situation was reversed for Form B to allow for comparison. Items 1-30 were not exposed to examinees while items 31-60 were shown to them before the examination. The examinees were informed that the compromised items would be part of the items they would attempt during the test. The data collected was subjected to DIF analysis using the multidimensional item response theory (MIRT) package in R-statistics software (version 3.6).

Results

Research Question One: Does Sex have moderating effect on the occurrence of differential item functioning (DIF) for the compromised and secured economics items?

Table 1: Effects of Sex on the Occurrence of Differential Item Functioning (DIF) for Compromised and Secured EAT Items

Item	Compromised Items			Secured Items			Decision	
	X ² statistic	p.value	Evaluation	X ² statistic	p.value	Evaluation		
*1	3.972	0.137	NO DIF	**1	6.269	0.044	NO DIF	No Effect
*2	-0.856	1.000	NO DIF	**2	7.322	0.026	DIF	Effect
*3	0.918	0.632	NO DIF	**3	14.054	0.001	DIF	Effect
*4	9.834	0.007	DIF	**4	-4.250	1.000	NO DIF	Effect
*5	10.885	0.004	DIF	**5	-1.471	1.000	NO DIF	Effect
*6	3.784	0.151	NO DIF	**6	18.684	0.000	DIF	Effect
*7	1.189	0.552	NO DIF	**7	-3.530	1.000	NO DIF	No Effect
*8	4.912	0.086	NO DIF	**8	2.144	0.342	NO DIF	No Effect
*9	6.630	0.036	NO DIF	**9	1.757	0.415	NO DIF	No Effect
*10	4.120	0.127	NO DIF	**10	-9.359	1.000	NO DIF	No Effect
*11	12.173	0.002	DIF	**11	8.429	0.015	DIF	No Effect
*12	13.439	0.001	DIF	**12	-4.484	1.000	NO DIF	Effect
*13	11.291	0.004	DIF	**13	21.638	0.000	DIF	No Effect
*14	8.998	0.011	DIF	**14	6.810	0.033	DIF	No Effect

*15	13.005	0.001	DIF	**15	-3.088	1.000	NO DIF	Effect
*16	9.462	0.009	DIF	**16	-1.738	1.000	NO DIF	Effect
*17	6.883	0.032	NO DIF	**17	9.951	0.007	DIF	Effect
*18	28.562	0.000	DIF	**18	6.551	0.038	NO DIF	Effect
*19	-0.237	1.000	NO DIF	**19	5.660	0.059	NO DIF	No Effect
*20	1.506	0.471	NO DIF	**20	-11.758	1.000	NO DIF	No Effect
*21	20.534	0.000	DIF	**21	-7.660	1.000	NO DIF	Effect
*22	1.075	0.584	NO DIF	**22	-6.663	1.000	NO DIF	No Effect
*23	5.181	0.075	NO DIF	**23	-5.902	1.000	NO DIF	No Effect
*24	5.383	0.068	NO DIF	**24	-3.680	1.000	NO DIF	No Effect
*25	1.169	0.557	NO DIF	**25	-5.108	1.000	NO DIF	No Effect
*26	8.393	0.015	DIF	**26	-6.648	1.000	NO DIF	Effect
*27	6.435	0.040	NO DIF	**27	1.148	0.477	NO DIF	No Effect
*28	9.981	0.007	DIF	**28	15.834	1.000	NO DIF	Effect
*29	10.780	0.005	DIF	**29	3.634	0.163	NO DIF	Effect
*30	2.213	0.331	NO DIF	**30	0.360	0.835	NO DIF	No Effect
**31	2.567	0.277	NO DIF	*31	43.672	0.000	DIF	Effect
**32	0.019	0.991	NO DIF	*32	2.154	0.341	NO DIF	No Effect
**33	2.595	0.273	NO DIF	*33	13.210	0.001	DIF	Effect
**34	0.830	0.660	NO DIF	*34	14.083	0.001	DIF	Effect
**35	1.949	0.377	NO DIF	*35	19.411	0.000	DIF	Effect
**36	1.229	0.541	NO DIF	*36	1.660	0.436	NO DIF	No Effect
**37	-2.950	1.000	NO DIF	*37	0.466	0.792	NO DIF	No Effect
**38	3.894	0.143	NO DIF	*38	1.670	0.434	NO DIF	No Effect
**39	-4.941	1.000	NO DIF	*39	1.783	0.410	NO DIF	No Effect
**40	2.196	0.333	NO DIF	*40	13.024	0.001	DIF	Effect
**41	1.039	0.595	NO DIF	*41	9.426	0.009	DIF	Effect
**42	9.285	0.010	DIF	*42	-0.372	1.000	NO DIF	Effect
**43	5.827	0.054	NO DIF	*43	7.980	0.019	DIF	Effect
**44	0.906	0.636	NO DIF	*44	0.211	0.900	NO DIF	No Effect
**45	2.725	0.256	NO DIF	*45	2.196	0.334	NO DIF	No Effect
**46	-5.827	1.000	NO DIF	*46	31.472	0.000	DIF	Effect
**47	3.114	0.211	NO DIF	*47	2.932	0.231	NO DIF	No Effect
**48	-1.385	1.000	NO DIF	*48	2.071	0.355	NO DIF	No Effect
**49	-5.523	1.000	NO DIF	*49	0.096	0.953	NO DIF	No Effect
**50	-2.713	1.000	NO DIF	*50	0.557	0.757	NO DIF	No Effect
**51	3.936	0.140	NO DIF	*51	10.458	0.005	DIF	Effect
**52	2.685	0.261	NO DIF	*52	2.506	0.286	NO DIF	No Effect

**53	1.602	0.449	NO DIF	*53	9.765	0.008	DIF	Effect
**54	-5.582	1.000	NO DIF	*54	11.383	0.003	DIF	Effect
**55	5.771	0.056	NO DIF	*55	21.150	0.000	DIF	Effect
**56	-6.299	1.000	NO DIF	*56	29.164	0.000	DIF	Effect
**57	0.068	0.967	NO DIF	*57	14.721	0.001	DIF	Effect
**58	-2.358	1.000	NO DIF	*58	1.146	0.564	NO DIF	No Effect
**59	-3.903	1.000	NO DIF	*59	18.432	0.000	DIF	Effect
**60	7.443	0.024	DIF	*60	0.527	0.768	NO DIF	Effect

Source: Aurthors' Analysis, 2020

***EAT Items (Form A)**

**** EAT Items (Form B)**

Table 1 shows the items that exhibited differential item functioning for the compromised and secured items subset of EAT with respect to the sex of the examinees. The table shows that 28 items (46.7%) of EAT Form A functioned differentially among male and female examinees while 9 items (15%) of EAT Form B functioned differentially among male and female examinees. Out of the 28 items of the EAT (Form A) that functioned differentially among male and female examinees, 13 (Items 4,5,11,12,13,14,15,16,18,21,26,28 and 29) are compromised items while 15 items (Items 31,33,34,35,40,41,43,46,51,53,54,55,56,57,59) are secured items. Also, among the 9 items of the EAT Form B that functioned differentially among male and female examinees, 2 items (Items 42 and 60) are compromised items while 7 (Items 2, 3, 6, 11, 13, 14 and 17) are secured items. The result implies that differential item functioning (DIF) occurred due to the sex of the examinees for the compromised and secured economics items.

Comparatively, the sex of the examinees have a moderating effect on the occurrence of differential item functioning (DIF) for 34 (56.7%) compromised and secured item conditions with 3 items (5%), Items 11, 13 and 14, exhibiting DIF for both the compromised and secured item conditions. 19 items (31.7%), Items 2, 3, 6, 17, 31, 33, 34, 35, 40, 41, 43, 46, 51, 53, 54, 55, 56, 57 and 59, exhibited DIF only when the items were secured as compared to when the items are compromised. 12 items (20%), Items 4, 5, 12, 15, 16, 18, 21, 26, 28, 29, 42 and 60) exhibited DIF only when the items were compromised as compared to when the items are secured. However, 26 items (43.3%) (Items 1, 7, 8, 9, 10, 19, 20, 22, 23, 24, 25, 27, 30, 32, 36, 37, 38, 39, 44, 45, 47, 48, 49, 50, 52 and 58) do not exhibit DIF for both the compromised and secured item conditions. The implication of these results is that although differential item functioning (DIF) occurs due to the sex of the examinees for the compromised and secured economics items, item compromise condition also contributed to the occurrence of DIF due to the sex of the examinees. This is evident in the number of items (12 items) that exhibited DIF only when the items were compromised. It can therefore be concluded that sex has a moderating effect on the occurrence of differential item functioning (DIF) for the

compromised and secured economics items with the compromised item relatively contributing more to the occurrence of DIF due to the sex of the examinees.

Research Question Two: Does school type have a moderating effect on the occurrence of differential item functioning (DIF) for the compromised and secured economics items?

Table 2: Summary statistics showing moderating effect of School Type on the occurrence of Differential Item Functioning (DIF) for the compromised and secured EAT items

Item	Compromised Items			Secured Items			Decision	
	X ² statistic	p.value	Evaluation	Item	X ² statistic	p.value		Evaluation
*1	5.665	0.059	NO DIF	**1	58.962	0.094	NO DIF	No Effect
*2	54.448	0.000	DIF	**2	39.822	0.089	NO DIF	Effect
*3	86.322	0.000	DIF	**3	71.696	0.063	NO DIF	Effect
*4	4.566	0.102	NO DIF	**4	10.060	0.007	DIF	Effect
*5	17.468	0.000	DIF	**5	2.841	0.509	NO DIF	Effect
*6	33.021	0.000	DIF	**6	18.395	0.062	NO DIF	Effect
*7	6.145	0.046	NO DIF	**7	8.481	0.014	NO DIF	No Effect
*8	80.163	0.031	NO DIF	**8	94.790	0.000	DIF	Effect
*9	3.999	0.135	NO DIF	**9	10.627	0.060	NO DIF	No Effect
*10	24.637	0.000	DIF	**10	10.011	0.079	NO DIF	Effect
*11	96.172	0.000	DIF	**11	81.546	0.014	NO DIF	Effect
*12	42.165	0.000	DIF	**12	27.539	0.206	NO DIF	Effect
*13	98.220	0.088	NO DIF	**13	112.847	0.000	DIF	Effect
*14	81.885	0.000	DIF	**14	67.259	1.000	NO DIF	Effect
*15	44.539	0.000	DIF	**15	29.912	0.580	NO DIF	Effect
*16	95.273	0.000	DIF	**16	80.647	0.314	NO DIF	Effect
*17	30.968	0.000	DIF	**17	16.341	1.000	NO DIF	Effect
*18	111.710	0.000	DIF	**18	97.084	0.038	NO DIF	Effect
*19	1.925	0.382	NO DIF	**19	12.701	1.000	NO DIF	No Effect
*20	110.779	0.000	DIF	**20	-96.152	0.820	NO DIF	Effect
*21	6.389	0.041	NO DIF	**21	8.237	0.070	NO DIF	No Effect
*22	8.671	0.013	DIF	**22	5.956	0.112	NO DIF	Effect
*23	119.438	0.000	DIF	**23	-104.811	0.479	NO DIF	Effect
*24	44.135	0.000	DIF	**24	-29.508	0.076	NO DIF	Effect
*25	7.495	0.024	NO DIF	**25	7.132	0.064	NO DIF	No Effect
*26	-34.997	0.664	NO DIF	**26	49.623	0.000	DIF	Effect
*27	0.828	0.661	NO DIF	**27	13.798	0.131	NO DIF	No Effect

*28	16.986	0.000	DIF	**28	-2.360	1.000	NO DIF	Effect
*29	11.904	0.003	DIF	**29	2.722	0.046	NO DIF	Effect
*30	13.730	0.001	DIF	**30	0.897	0.071	NO DIF	Effect
**31	28.894	0.000	DIF	*31	-14.267	1.000	NO DIF	Effect
**32	7.905	0.019	NO DIF	*32	6.721	0.054	NO DIF	No Effect
**33	4.393	0.111	NO DIF	*33	10.234	0.066	NO DIF	No Effect
**34	-104.066	0.733	NO DIF	*34	118.692	0.000	DIF	Effect
**35	-3.986	1.000	NO DIF	*35	18.612	0.945	NO DIF	No Effect
**36	2.532	0.282	NO DIF	*36	12.094	0.427	NO DIF	No Effect
**37	20.436	0.000	DIF	*37	-5.809	0.477	NO DIF	Effect
**38	5.220	0.074	NO DIF	*38	9.406	0.038	NO DIF	No Effect
**39	-15.959	0.627	NO DIF	*39	30.585	0.000	DIF	Effect
**40	19.622	0.000	DIF	*40	-4.996	0.064	NO DIF	Effect
**41	1.250	0.535	NO DIF	*41	13.376	0.709	NO DIF	No Effect
**42	6.831	0.033	NO DIF	*42	7.796	0.128	NO DIF	No Effect
**43	38.154	0.000	DIF	*43	-23.528	0.096	NO DIF	Effect
**44	137.308	0.000	DIF	*44	-122.682	0.350	NO DIF	Effect
**45	5.961	0.051	NO DIF	*45	8.665	0.097	NO DIF	No Effect
**46	5.460	0.065	NO DIF	*46	9.166	0.099	NO DIF	No Effect
**47	82.309	0.000	DIF	*47	-67.682	1.000	NO DIF	Effect
**48	46.004	0.000	DIF	*48	-31.378	0.046	NO DIF	Effect
**49	26.531	0.000	DIF	*49	-11.905	0.572	NO DIF	Effect
**50	-46.388	0.056	NO DIF	*50	61.015	0.000	DIF	Effect
**51	78.049	0.000	DIF	*51	-63.422	0.090	NO DIF	Effect
**52	81.122	0.000	DIF	*52	-66.495	0.063	NO DIF	Effect
**53	53.898	0.000	DIF	*53	-39.272	0.040	NO DIF	Effect
**54	2.203	0.065	NO DIF	*54	12.424	0.002	DIF	Effect
**55	28.814	0.000	DIF	*55	-14.188	0.056	NO DIF	Effect
**56	34.463	0.000	DIF	*56	-19.836	0.504	NO DIF	Effect
**57	6.823	0.033	NO DIF	*57	7.803	0.135	NO DIF	No Effect
**58	96.537	0.000	DIF	*58	-81.910	0.078	NO DIF	Effect
**59	0.554	0.758	NO DIF	*59	14.072	0.405	NO DIF	No Effect
**60	52.889	0.000	DIF	*60	-38.262	0.930	NO DIF	Effect

Source: Aarthors' Analysis, 2020

*EAT Items (Form A)

** EAT Items (Form B)

Table 2 shows the items that exhibited differential item functioning for the compromised and secured items subset of EAT with respect to school type of the examinees. The table shows that 23 items (38.3%) of the EAT Form A functioned differentially among

examinees in the public and private schools while 19 items (31.7%) of the EAT items in Form B functioned differentially among examinees in the public and private schools. Out of the 23 items of EAT Form A that functioned differentially among examinees in the public and private schools, 19 items (Items 2,3,5,6,10,11,12,14,15,16,17,18,20,22,23,24,28,29 and 30) are compromised items while 4 (Items 34,39,50,54) are secured items.. Also, among the 19 items of the EAT Form B that functioned differentially with examinees in the public and private schools, 15 (Items 31, 37,40,43,44,47,48,49,51,52,53,55,56,58, and 60) are compromised while 4 items (Items 4, 8, 13 and 26) are secured items respectively. The implication of the result is that differential item functioning (DIF) occurred due to school type of the examinees for the compromised and secured economics items.

Comparatively, school type of the examinees has a moderating effect on the occurrence of differential item functioning (DIF) for 42 (70%) compromised and secured item conditions with 34 items (56.7%)(Items 2, 3, 6, 17, 31, 33, 34, 35, 40, 41, 43, 46, 51, 53, 54, 55, 56, 57 and 59) exhibiting DIF only when the items were compromised as compared to when the items were secured. 8 items (13.3%)(Items 4, 5, 12, 15, 16, 18, 21, 26, 28, 29, 42 and 60) exhibited DIF only when the items were secured as compared to when the items were compromised. However, 18 items (30%)(1, 7, 9, 19, 21, 25, 27, 32, 33, 35, 36, 38, 41, 42, 45, 46, 57 and 59) did not exhibit DIF for both the compromised and secured item conditions. The implication of these results is that although differential item functioning (DIF) occurred due to school type of the examinees for the compromised and secured economics items, the item compromise condition contributed more to the occurrence of DIF due to school type of the examinee. This is evident in the number of items (34 items) that exhibited DIF only when the items were compromised. School type, therefore, has a moderating effect on the occurrence of differential item functioning (DIF) for the compromised and secured economics items with the compromised item relatively contributing more to the occurrence of DIF due to the school type of the examinees.

Discussion of Findings

The results of research question one showed that 28 items, comprising 13 compromised items and 15 secured items on EAT Form A functioned differentially among male and female examinees respectively. Similarly, 9 items comprising of 2 compromised items and 7 secured items on EAT items Form B functioned differentially among male and female examinees. Finally, the results of Research Question Two showed that 23 items, comprising 19 compromised items and 4 secured items on EAT Form A, functioned differentially among private and public schools respectively. Correspondingly, 19 items comprising of 15 compromised items and 4 secured items on the EAT (Form B) functioned differentially among private and public schools respectively. Specifically, the result revealed that sex and school type of individual examinees had a moderating effect

on the occurrence of differential item functioning (DIF) for compromised and secured item conditions.

The results obtained in research question five and six imply that although differential item functioning (DIF) occurred due to sex and school type of the examinees for the compromised and secured economics items, the item compromise condition also contributed to the occurrence of DIF due to sex and school type of the examinees. This is so because the investigation of differential item functioning (DIF) helps in assessing the extent to which candidates' item pre-knowledge impacts item performance. It determines the degree of item degradation, gathers information to drive examination maintenance and compares performance of DPF-flagged candidates to DPF non-flagged candidates (O' Leary, 2013).

The findings of this study on the occurrence of DIF due to sex and school type of examinees agree with Angoff (1993) that when compared with the control group, the responses on the compromised tests confirmed the suspicion that the tests have indeed been compromised. The occurrence of DIF due to gender also agrees with the findings of Adedoyin (2010), who in his study investigated gender biased items in public examinations and found that 5 items were gender biased. Ogbekor and Onuka (2013) who investigated differential item functioning method as an item bias indicator in Delta State and found out that out of the sixty (60) 2010 National Examinations Council (NECO) economics questions, 18 items functioned differentially based on school type and school location of the examinees. Similarly, Osadebe and Agbure (2018) investigated the occurrence of differential item functioning in social studies multiple-choice questions of Basic Education Certificate Examinations in 2014 and reported DIF among 9 items which functioned differentially for male and female students.

Moyo and Nenty (2017) in a DIF study revealed that 29 out of 40 items of 2013 BGCSE Agriculture Examination displayed DIF due to the gender of the examinees. Igomu and Akpan (2014) in a study which examined the occurrence of item bias and DIF due to school type and school location using the 2012 multiple-choice Biology questions -conducted by National Examination Council (NECO) found out that 10 out of the sixty items were biased in relation to school type while 8 items were biased with respect to school location. On the effect of item compromise of the occurrence of DIF, the findings agreed with those of O' Leary (2013) in a case study comparing candidates' performance on scored (80) and non-scored (20) items considered to be compromised and uncompromised items respectively. O'Leary found that 531 candidates (6.4%) were flagged for differential person functioning (DPF) given the item compromise condition of the item. Item compromise has, therefore, helped to compare performance of DPF-flagged candidates to DPF non-flagged candidates.

Conclusion

It can be concluded from the study that sex and school type had moderating effects on the occurrence of differential item functioning (DIF) for the compromised and secured

economics items with the compromised items contributing more to the occurrence of DIF due to sex and school type of the examinees.

Recommendations

Based on the findings of the study, the following recommendations are made.

1. Differential item functioning analysis should be carried out on all items suspected to be compromised as a form of statistical evaluation by test experts and examination bodies.
2. Examination agencies and test experts should ensure that test items are not compromised to improve the reliability, validity, usability, fairness and credibility of tests and decisions made from such tests.
3. Examination bodies and test developers should consider feasible approaches to eliminate cheating in any form and sanction cheats appropriately to serve as deterrent to others.
4. Given the need for test security, more items should be developed and stored in the item banks to avoid undue item rotation as it can cause item compromise.

References

- Adedoyin, O. O. (2010). Using IRT approach to detect gender biased items in public examinations: A case study from the Botswana Junior Certificate Examination in Mathematics. *Educational Research and Reviews Academic Journals*, 5(7), 385-399
- Adewale, G. (2004). Examination malpractice: A stigma on school effectiveness in Nigeria. *Nigerian Journal of Educational Research and Evaluation*. 5 (1). 1-10.
- Angoff, W. H. (1993). Perspectives on DIF methodology. In P. W. Holland, & H. Wainer (Eds.): *Differential Item Functioning*. (3-23): Hillsdale, Lawrence Erlbaum Associates Publishers.
- Cizek, G. (2001). An overview of issues concerning cheating on large-scale tests. Paper presented at the annual meeting of the National Council on Measurement in Education. Seattle
- Drasgow, F., Nye, C. D., Guo, J., & Tay, L. (2009). Cheating on proctored tests: The other side of the unproctored debate. *Industrial and Organizational Psychology*, 2, 46-48.
- Igomu, A. C. & Akpan, F. (2014). An Assessment of Item Bias using Differentially Item Functioning Technics in NECO Biology Conducted Examination in Taraba State, Nigeria. *American International Journal of Research in Humanities, Arts and Social Sciences*. <http://www.iasir.net>. 2328-3696
- Jurich, D. P., DeMars, C. E. & Goodman, J. T. (2012). Investigating the Impact of Compromised Anchor Items on IRT Equating Under the Non-equivalent Anchor Test Design. *Applied Psychological Measurement*. 36: 291-308, doi:10.1177/0146621612445575.

- Lievens, F., & Burke, E. (2011). Dealing with threats inherent in unproctored Internet testing of cognitive ability: Results from a large-scale operational test program. *Journal of Occupational and Organizational Psychology*, 84, 817-824.
- McCarty, F. A., Oshima, T. C., & Raju, N.S. (2007). Identifying Possible Sources of Differential Functioning Using Differential Bundle Functioning With Polytomously Scored Data. *Applied Measurement in Education*, 20(2), 205–225
- McLeod, L., Lewis, C., & Thissen, D. (2003). A Bayesian method for the detection of item pre knowledge in computerized adaptive testing. *Applied Psychological Measurement*, 27(2), 121–137.
- Moyo, S. E. & Nenty, H. J. (2017). Dimensionality analysis of students' performance in 2013 BGCSE agricultural examination: Implications for differential item functioning. *Research Journal of Educational Sciences*, 5(1), 1-10
- Ogbebor, U., & Onuka, A. (2013). Differential item functioning method as an item bias indicator. *Educational Review*, 4(4), 367-373.
- Ojerinde, D. (2015). Innovations in Assessment: JAMB Experience. *Nigerian Journal of Educational Research and Evaluation*. 14 (3), 1-9.
- O'Leary, S. L. (2013). *Extending the Role of Psychometrics in Integrated Security Plans*. Alpine Testing Solution
- Osadebe, P. U. and Agbure, B. (2018). Assessment of differential item functioning in social studies multiple choice questions in Basic Education Certificate Examination. *European Journal of Education studies* 4(9), 236-256
- Wollack, J. A., & Cizek, G. J., (2017). Security Issues in Professional Certification/licensure Testing. In S. Davis-Becker, & C. W. Burkendahl (2017): *Testing in the professions: credentialing policies and practices*. (New York): Routledge.
- Zara, A. & Pearson, V. (2006). Defining Item Compromise. Paper presented at the 2006 annual meeting of the National Council on Measurement in Education, San Francisco.
- Zumbo, B. D. (1999). *A handbook on the theory and methods of differential item functioning (DIF): Logistic regression modelling as a unitary framework for binary and likert-Type (ordinal) item scores*. Ottawa, Canada: Directorate of Human Resources Research and Evaluation, Department of National Defense.