International Review of Social Sciences and Humanities
Vol. 7, No. 2 (2014), pp. 63-82
www.irssh.com
ISSN 2248-9010 (Online), ISSN 2250-0715 (Print)

Identifying Location Biased Items in the 2010 Botswana Junior Certificate Examination Mathematics Paper One Using the Item Response Characteristics Curves

T. Mokobi
Department of Educational Planning and Research Services, Botswana
Ministry of Education, Gaborone, Botswana, Southern Africa
E-mail: tjapie@yahoo.com

O.O. Adedoyin
(Corresponding Author)
BA ISAGO University College, Gaborone, Botswana
Dean of Faculty of Education, Gaborone, Botswana, Southern Africa
Email: omobola_adedoyin@yahoo.com

(Received: 10-3-14 / Accepted: 22-4-14)

Abstract

This quantitative study was conducted to identify location biased items with respect to rural and urban schools in the 2010 Botswana Junior Certificate Examination Mathematics paper 1 using IRT Item Response Characteristics Curves. The study further identified rural/ urban location biased items with respect to gender of students. The 2010 Botswana Junior Certificate Examination Mathematics examination paper 1 consisted of forty (40) multiple choice test items. The sample for this study comprised of 4000 students randomly selected from a population of 36940 who sat for 2010 Botswana Junior Certificate Examination Mathematics paper 1. The sample of students randomly selected consisted of 2000 male students, of which 1000 were from rural schools and 1000 were from urban schools. The remaining 2000 students were females, 1000 from urban schools and 1000 from rural schools. 3PL (Multilog software) Item Response Theory (IRT) statistical analysis was used to generate the Item Characteristics Curves (ICCs) for the corresponding groups rural/urban, rural / urban with respect to gender. The ICCs for the corresponding groups were compared to identify rural/urban location biased items. The findings of the study revealed that from the 24 items that fitted the IRT (3PLM) model, six (6) items were rural /urban location biased items. The study further found out that three (3) items were rural /urban location biased with respect to males and six (6) items were rural /urban location biased with respect to females. It is recommended that test developers in Africa should always endeavour to create bias free items for testing and examination purposes and the connotations reflected in test or examination items should be relevant to the life experiences of examinees responding to the items.

Keywords: Item bias, Item Response Theory (IRT), Item Characteristics Curves (ICCs).

Introduction

Item bias occurs when examinees of one group are less likely to answer an item correctly than examinees of another group due to some characteristics of the test item or testing situation that is not relevant to the testing purpose. According to Pedrajita and Talisayon (2009), Hambleton and Rodgers (2005), item bias is the presence of some characteristics of an item that result in different performance for individuals of the same ability but from different ethnic, sex, cultured or religious group. Test item bias can also be explained as differences in test scores that can be attributed to demographic variables such as gender, ethnicity, socioeconomic background, or age and can also be attributed to testing content that has not been taught i.e. asking for information that disadvantaged students who have not had equal opportunity to learn. Camilli and Shepard(1994) cited by Pedrajita&Talisayon(2009), stressed that “Bias is a major factor for tests to be considered unfair, inconstant, and contaminated by extraneous factors” Test bias is a concern that has been there for a long time and according to Baghi, Heibatollah, Ferrara and Steven (1989), “the study of test items that function differently for subpopulations of examinees has been a concern for test developers”. They continued by highlighting that this concern is especially central in competency based testing, where graduation certification is contingent in passing one or more tests. Wiberg (2007) noted that, in order to draw valid conclusions from an achievement test it is essential that the test is a valid measurement of what it intends to measure. Zumbo (2009, p.76), stressed that item bias occurs when items function differently for certain groups of examinees or respondents, and the existence of item bias, violates the assumption of measurement invariance, which holds that measurement properties should not be affected by any context. That a test item is not biased is an important consideration in the selection and use of any psychological test, that is, it is essential that a test is fair to all applicants, and is not biased against a segment of population taking the test items. In many cases, test items are biased due to the fact that they contain sources of difficulty that are irrelevant or extraneous to the construct being measured, and these irrelevant factors affect performance of the examinees.

Bias of test items in examinations may lead to some group of students being disadvantaged and the society might believe that passing is a privilege for certain groups and a taboo for others when they can perform equally well. For those who construct the examinations it is important that they know if the examinations disadvantage certain groups in the society because examination test items that are biased will produce results from which invalid references will be made. For example, if students from the same subgroup have failed the examination because some items were biased towards their subgroup, it might not mean that the students did not know what was being tested but simply that the items had disadvantaged them. These undesirable biases disadvantage certain groups of people taking the test and some students end up performing badly due to the biased test items. It is very necessary that test items are always fair to all examinees, and that the test items are not biased against certain group of students or examinees. It is therefore important to address test item bias as Medie and Fetzer (2008) stated that test bias is a fundamentally important issue in testing as pervasive and systematic sources of error can lead to erroneous inferences regarding the interpretation and use of test scores. Bias can also result in systematic errors that distort the inferences made in any selection and classification. According to Camilli and Shepard (2007), bias in use-social consequences occur when treatment assigned on the basis of test result vary in quality. A test could be a valid predictor of an outcome but the use of the test might lead to undesirable consequences. A fair and unbiased use of test involves more than psychometric validity, it encompasses the consequences to the decision made on the basis of test scores.

The purpose of the study is to detect location biased test items from the 2010 Botswana Junior Certificate mathematics examination paper one .To achieve this Item Response Theory (3PL) model software will be used to produce the item response characteristic curves (ICC) which will be used to detect bias on Junior Certificate mathematics paper one examinations responses by Junior Secondary School students brought about by rural /urban location and rural/urban location with respect to gender of students.

The following research questions guided in investigating the Junior Certificate mathematics paper one examinations for bias amongst candidates in rural and urban schools.

- Are items of the mathematics JC paper 1 biased to candidates in relation to rural or urban schools?
- Are items of the mathematics JC paper 1 biased to male candidates in rural schools or urban schools?
- Are items of the mathematics JC paper 1 biased to female candidates in rural schools or urban schools?

Literature Review

The literature revealed that a lot of tests or examinations papers contain items that were biased towards a certain group, more especially in relation to gender. Studies conducted by Abedalaziz (2011) and Karami (2011) have indicated that some test items were biased and disadvantaged students who belonged to different gender groups. Research carried out by Maliki, Ngban and Ibu (2009) have also indicated that in terms of performance in mathematics, male students outperformed females, students from urban schools outperformed students from rural schools and students from private schools outperformed students from public schools. Various types of test bias exist and they consist of content bias, atmosphere bias, bias in use-prediction and bias in use-social consequences.

From studies conducted by Adedoyin (2010) on test item bias, there were indications to suggest that mathematics examinations in Botswana contain items that were biased in relation to gender. There are different kinds of bias that may be encountered in tests ranges widely and they include gender bias, religious bias, geographic bias, linguistic bias and racial ethnic heritage bias. Other types of test bias could easily be encountered in the process of testing and they include content bias, atmosphere bias, and bias in use-social consequences. Content bias occurs when the content of the test items gives a systematic advantage to a particular group of test takers, this type of bias reflects differences in the opportunities to learn the material tested. Test items may be biased and unfair to the members of any group if they have not had the opportunity to learn the material. However, if members of various groups have had equal opportunities to learn the test contents, any observed differential performance may not be persuasive evidence of content bias. Another type of bias, the atmosphere bias could arise as a result of the testing conditions on the examinees' performances. It could emanate from the type of motivation elicited, factors related to the examinees-testers interaction, and factors in the evaluation and scoring of responses. The goal in testing is to minimize any possible test condition effects and this is usually accomplished by using standard testing conditions.

Many methods for detecting test item bias in the measurement of ability exist and they include among others: item characteristic curve, regression method, chi-square method and transformed item difficulty method among others. For the purpose of this study in identifying location biased test items in 2010 Junior Certificate Mathematics paper 1, IRT item characteristics curves (ICC) was used to detect the biased test items.

Using IRT Item Characteristic Curves (ICC) to Detect Biased Test Items

Item characteristic curve approach of detecting test item bias, states that a test is unbiased if all the individuals having the same underlying ability have equal probability of getting the item correct regardless of subgroup membership (Pine,2006). This means that an item is said to be unbiased if the item characteristics curves for the item measured on two groups are identical. If the item characteristics curves are not identical, then the item is biased and the area between the group ICCs serve as a measure of item aberrance (Lord, 2002).

Using the ideas of Anastasia and Urbina (2006), in figure 1 below, item 1, according to the two ICCs, there is a significant difference between the two groups. But in item 2, it can be observed that the ICCs were identical or very close for the two groups, it can then be concluded that item 2 was not biased. But item1 was biased since the ICCs of the two groups were not identical or similar.

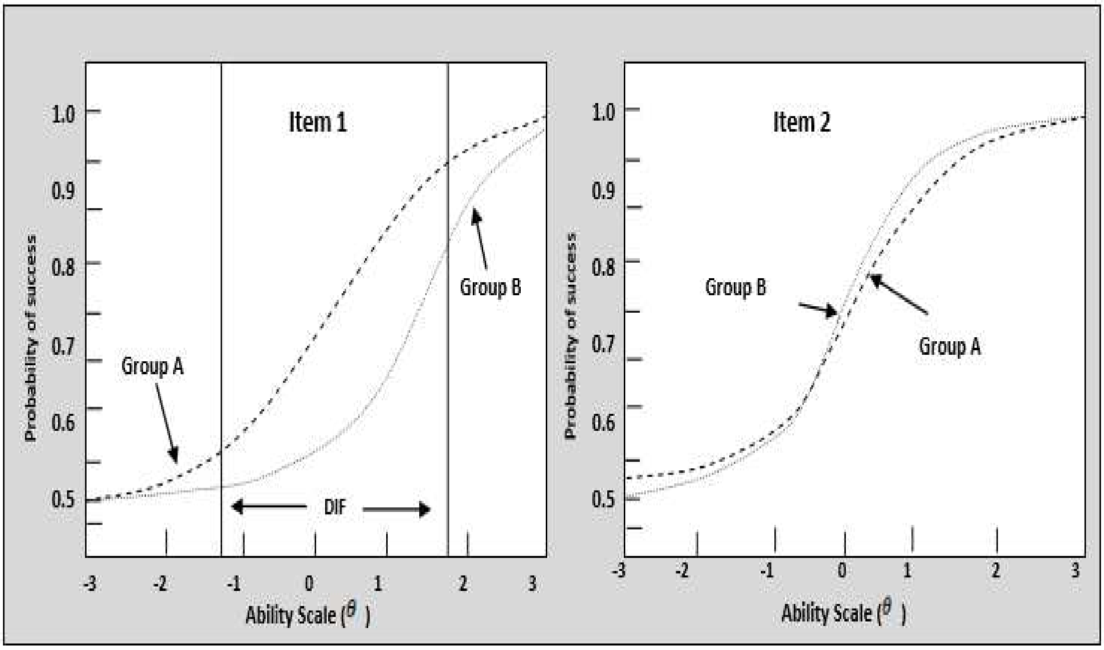


Fig. 1: Item Characteristic Curves (ICC) for two items, illustrating large and small amounts of differential items functioning (DIF). (Graph adapted from Anastasia & Urbina, 2006)

Methodology

The data for this study was obtained from Botswana Examinations Council (BEC). The population for this study was 36939 students. The researchers used stratified random sampling to divide the population into 18668 females and 18271males. Secondly, the researchers used cluster random sampling to divide the males and females into those who attend schools in rural and urban areas. The result of the sampling procedure was a population divided into four clusters which are; Males from schools in a rural location; Males from schools in an urban location; Females from schools in a rural location; and Females from schools in an urban location.

From each of the clusters simple random sampling was used to select 1250 participants to make up 5000 participants who were used as a sample for the purpose of checking if the data is uni-dimensional and the overall fit of the model to the data. It is of high importance to check for uni-dimensionality as “multidimensionality can be mistaken for item bias with latent trait models as a result of differences among ICCs” Milardo (2000).It is also important to check if the data fit the model as “the utility of the IRT model is dependent upon the extent to which the model accurately reflects the data” Adedoyin (2010).

From the sample that was selected to check for uni-dimensionality and the overall fit of the model to the data, the final sample of 4000 examinees was randomly selected for use to generate the ICC curves using IRT(3PLM) model software. Baghi et al (1989) also indicated that the three parameter model of the IRT requires a minimum of 1000 cases per group to estimate item parameters.

Instrumentation

The researchers collected the data for this study from BEC. The data were responses of students for the 2010 mathematics paper one JC examination. The examination was a multiple choice paper which consisted of 40 items. The examination was administered for one and a half hours. Out of the 40 test items, 24 test items of the 2010 mathematics paper 1 that fitted the IRT (3PL).

To address the research questions, IRTstatistical method using 3-PL (Multilog Software) was used to analyse the responses from the different sub-groups (those who attended school in rural areas and those who attended school in urban areas, male who attended school in rural areas and males who attended school in urban areas, females who attended school in rural areas and females who attended school in urban areas).The parameter estimates for each group were produced with the corresponding Item Characteristic Curves. The ICCs for the two corresponding groups in each research question were compared for items that were biased.

Results

Test for Unidimensionality

The method used to assess uni-dimensionality for this study was confirmatory factor analysis. This was performed to determine whether or not a dormant factor existed among all the items. The confirmatory factor analysis performed on the 40 items of the 2010 JC Mathematics paper one yielded nine eigen values greater than one. The first eigen value was 5.909 greater than the next eight eigen values of 1.492, 1.096, 1.088, 1.060, 1.029, 1.022, 1.017 and 1.010. The first factor explained 14.772% of the variance in the data set, while the second factor explained 3.73% of the remaining variance and the rest of the variance was explained by the remaining 38 factors and uni-dimensionality was confirmed.

Test for Model Fit

To determine whether the test item fitted the model, Chi-square goodness of fit statistics was performed. A Chi-square test was run on the data set using Bilog to establish whether the items fit the 1PL, 2PL and 3PL models. Table 1 showed the results of the chi-square statistics. The Chi-square goodness of fit analysis showed that only one item (item 10) fitted the 1PL model, eleven items (items 1, 3, 6, 7, 10, 14, 15, 23, 27, 33 and 39) fitted the 2PL model and 24 items (items 1, 3, 4, 6, 7, 10, 14, 15, 21, 23, 24, 25, 26, 27, 29, 30, 31, 32, 33, 34, 35, 38, 39and 40) fitted the 3PL model. The 3PLM was used for analysis and to generate the item characteristics curves (ICCs) for each item based on location of school (rural/urban).

Table 1: Results of the Chi-square statistics for the 1PL, 2PL and 3PL IRT models

Items	1PL			2PL			3PL		
	Chi-square	**p**	**df**	**Chi-square**	**P**	**df**	**Chi-square**	**P**	**df**
1	60.3	0.0000	9.0	15.4	0.0812**	9.0	16.4	0.0585**	9.0
2	94.3	0.0000	9.0	34.2	0.0001	9.0	27.0	0.0014	9.0
3	361.1	0.0000	9.0	15.5	0.0774**	9.0	13.7	0.1339**	9.0
4	247.6	0.0000	9.0	25.0	0.0016	8.0	13.2	0.1556**	9.0
5	154.4	0.0000	9.0	78.8	0.0000	8.0	67.0	0.0000	8.0
6	82.3	0.0000	9.0	13.5	0.1409**	9.0	9.7	0.3741**	9.0
7	28.3	0.0008	9.0	15.5	0.0788**	9.0	5.7	0.773**	9.0
8	240.4	0.0000	9.0	61.8	0.0000	9.0	44.7	0.0000	9.0
9	428.4	0.0000	8.0						
10	16.4	0.0593**	9.0	12.3	0.1967**	9.0	7.9	0.5469**	9.0
11	63.3	0.0000	9.0	56.4	0.0000	9.0	41.6	0.0000	9.0
12	112.0	0.0000	9.0	21.8	0.0094	9.0	19.9	0.0182	9.0
13	329.6	0.0000	9.0	53.2	0.0000	9.0	67.4	0.0000	9.0
14	38.7	0.0000	9.0	8.6	0.4726**	9.0	6.7	0.6718**	9.0
15	45.0	0.0000	9.0	11.6	0.2356**	9.0	16.4	0.0596**	9.0
16	85.9	0.0000	9.0	58.1	0.0000	9.0	42.9	0.0000	9.0
17	429.8	0.0000	9.0	69.6	0.0000	7.0	35.0	0.0000	8.0
18	76.6	0.0000	9.0	24.8	0.0032	9.0	19.1	0.0242	9.0
19	222.5	0.0000	9.0	29.3	0.0003	8.0	28.7	0.0007	9.0
20	387.5	0.0000	8.0	20.6	0.0083	8.0	20.6	0.0146	9.0
21	21.7	0.0099	9.0	21.1	0.0122	9.0	6.6	0.6812**	9.0
22	405.9	0.0000	8.0	58.6	0.0000	8.0	39.2	0.0000	8.0
23	57.8	0.0000	9.0	15.3	0.0840**	9.0	12.8	0.1736**	9.0
24	93.6	0.0000	9.0	60.9	0.0000	9.0	12.7	0.1747**	9.0
25	140.3	0.0000	9.0	40.3	0.0000	9.0	14.1	0.1196**	9.0
26	57.7	0.0000	9.0	64.9	0.0000	9.0	5.8	0.7557**	9.0
27	46.8	0.0000	9.0	14.4	0.1080**	9.0	15.6	0.0749**	9.0
28	252.5	0.0000	9.0	39.3	0.0000	9.0	27.7	0.0011	9.0
29	47.6	0.0000	9.0	18.2	0.0330	9.0	6.0	0.7404**	9.0
30	143.8	0.0000	9.0	31.5	0.0002	9.0	8.3	0.4999**	9.0
31	180.5	0.0000	9.0	20.9	0.0075	8.0	10.3	0.3234**	9.0
32	194.3	0.0000	9.0	28.8	0.0007	9.0	9.0	0.4404**	9.0
33	47.9	0.0000	9.0	10.3	0.3294**	9.0	4.6	0.8708**	9.0
34	128.6	0.0000	9.0	98.4	0.0000	9.0	12.8	0.1723**	9.0
35	171.7	0.0000	9.0	23.5	0.0028	8.0	14.3	0.1126**	9.0

36	103.9	0.0000	9.0	41.7	0.0000	9.0	40.8	0.0000	9.0
37	146.5	0.0000	9.0	29.6	0.0003	8.0	24.8	0.0032	9.0
38	68.7	0.0000	9.0	19.4	0.0222	9.0	14.1	0.1204**	9.0
39	97.8	0.0000	9.0	4.0	0.9111**	9.0	8.4	0.4980**	9.0
40	136.1	0.0000	9.0	17.0	0.0298	8.0	15.1	0.0893**	9.0

**The items with probability greater than the alpha level of 0.05 significant level.

Answering Research Questions

Are items of the mathematics JC paper 1 biased to candidates in relation to rural or urban schools?

From the 24 items (1, 3, 4, 6, 7, 10, 14, 15, 21, 23, 24, 25, 26, 27, 29, 30, 31, 32, 33, 34, 35, 38, 39, 40) that fitted the model, only 5 items (14, 26, 29, 33, 39) exhibited rural/urban location biased items towards a particular group (students who attended schools in rural areas or students who attended schools in urban areas).The Item Characteristic Curves for the five (5) identified test items were not the same for both the students who attended schools in rural areas and students who attended schools in urban areas, implying that the five (5) test items were location biased. Table 3 summarised the results of both groups and the item parameter estimates for the 24 items.

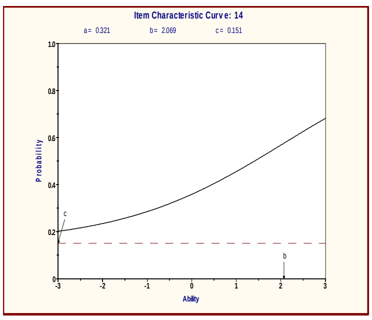
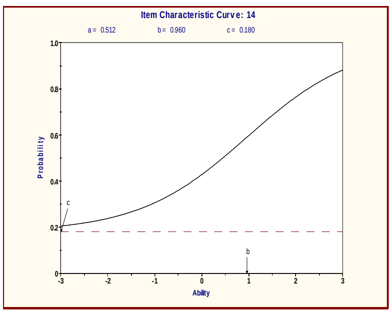
Table 3: Summary of the item parameter estimates for the Students who Attend Schools in Rural (SASRA) areas and Students who Attend Schools in Urban Areas (SASUA)

Items	Groups	Item discrimination (a)	Item difficulty (b)	Guessing Parameter (c)
Item 1	SASRA	0.604	-0.629	0.116
	SASUA	0.706	-0.530	0.214
Item 3	SASRA	-0.009	-7.990	0.116
	SASUA	-0.002	-3.000	0.036
Item 4	SASRA	1.248	0.757	0.154
	SASUA	1.239	0.348	0.125
Item 6	SASRA	0.821	1.165	0.125
	SASUA	0.738	0.455	0.080
Item 7	SASRA	0.747	1.203	0.180
	SASUA	0.756	0.789	0.140
Item 10	SASRA	0.923	1.350	0.357
	SASUA	0.750	1.046	0.316
Item 14**	SASRA	0.321	2.069	0.151
	SASUA	0.512	0.980	0.180
Item 15	SASRA	0.960	1.137	0.432
	SASUA	0.606	1.214	0.457
Item 21	SASRA	0.707	1.104	0.239
	SASUA	0.976	0.741	0.291
Item 23	SASRA	0.503	-0.029	0.000
	SASUA	0.617	-0.092	0.131
Item 24	SASRA	1.217	1.583	0.333
	SASUA	1.299	1.393	0.316
Item 25	SASRA	1.293	1.065	0.183
	SASUA	10146	0.581	0.161
Item 26**	SASRA	1.069	1.992	0.127

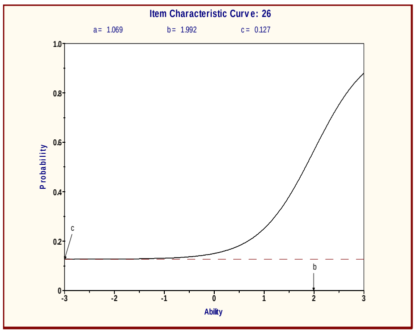
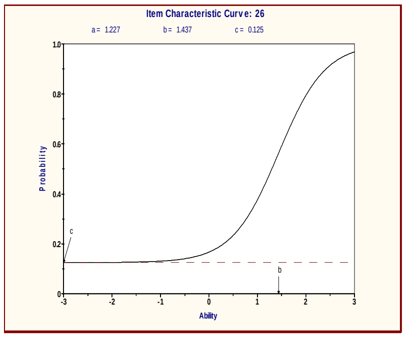
	SASUA	1.227	1.437	0.125
Item 27**	SASRA	0.543	-0.331	0.037
	SASUA	0.654	-0.116	0239
Item 29**	SASRA	0.741	1.774	0.394
	SASUA	0.718	1.099	0.353
Item 30	SASRA	1.012	2.333	0.314
	SASUA	0.933	1.788	0.289
Item 31	SASRA	1.149	0.947	0.106
	SASUA	1.230	0.673	0.143
Item 32	SASRA	1.378	2.448	0.263
	SASUA	1.220	2.265	0.257
Item 33**	SASRA	0.625	0.142	0.233
	SASUA	0.877	-0.103	0.373
Item 34	SASRA	2.033	1.986	0.186
	SASUA	2.101	1.579	0.188
Item 35	SASRA	0.887	0.735	0.134
	SASUA	1.035	0.308	0.129
Item 38	SASRA	1.062	1.175	0.144
	SASUA	0.794	0.761	0.124
Item 39**	SASRA	0.198	2.089	0.000
	SASUA	0.757	1.868	0.286
Item 40	SASRA	0.914	0.312	0.148
	SASUA	0.881	0.221	0.160

** Items exhibiting the most obvious location bias

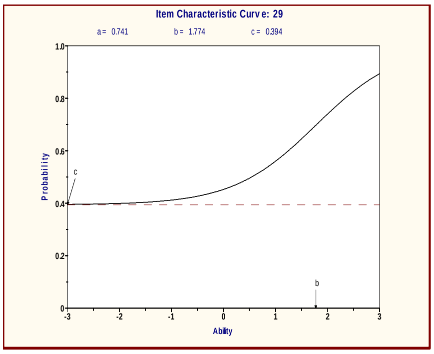
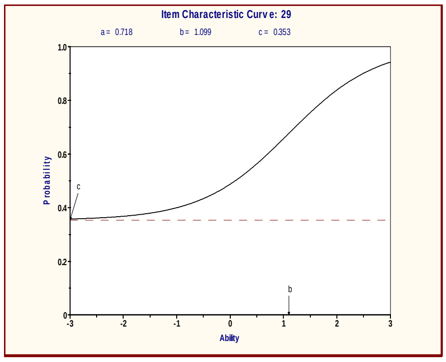
Item 14: Comparing between the two item characteristics curves for rural and urban schools,the two curves were not identical because the item characteristic curve for students who attended schools in rural areas was just about horizontal and shifted to the right indicating that the item could not discriminate well amongst this group and it is more difficult for students who attended schools in rural areas than those who attended schools in urban areas .It can then be concludedthat item 14 was biased towards students who attended schools in rural areas.

Item 14. Graph for students who attended schools in rural areas	**Graph for students who attended schools in urban areas**
	

Item 26: The ICC for students who attended schools in rural areas was shifted towards the right than the ICC for students who attended schools in urban areas indicating that the item was more challenging for students who attended schools in rural areas.

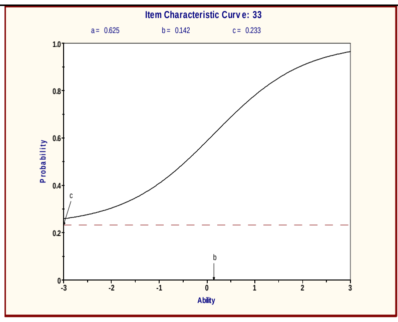
<table>
<tr><th>Item 26.Graph for students who attended schools in rural areas</th><th>Graph for students who attended schools in urban areas</th></tr>
<tr><td>
</td><td>
</td></tr>
</table>

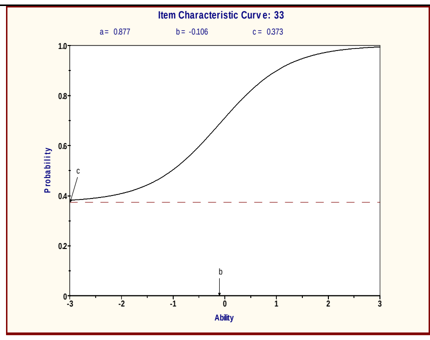
Item 29: From the ICC graphs, the item proved to be more difficult for students who attended schools in rural areas than those who attended schools in urban, since the ICC of students who attended schools in rural areas shifted more towards the right than students who attended schools in urban areas. This item was biased towards students who attended schools in rural areas.

<table>
<tr><th>Item 29. Graph for students who attended schools in rural areas</th><th>Graph for students who attended schools in urban areas</th></tr>
<tr><td>
</td><td>
</td></tr>
</table>

Item 33: The ICCs for this item indicated that guessing factor was more pronounced for students who attended schools in urban areas than those who attended schools in rural areas. Therefore the item was biased towards students who attended schools in rural areas.

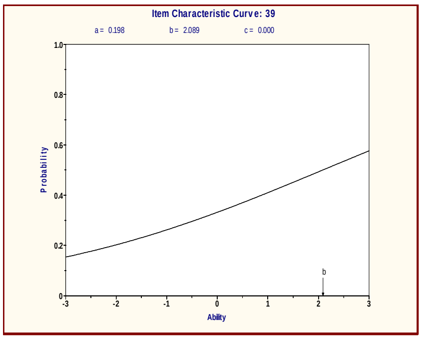
Item 33 Graph for students who attended schools in rural areas	Graph for students who attended schools in urban areas

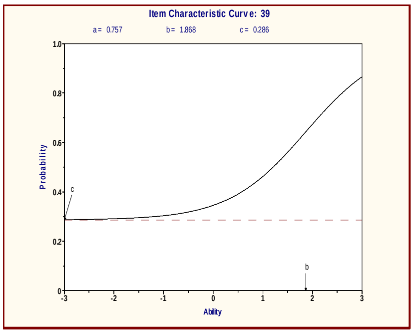




Item 39: The two Item Characteristic Curves were not identical at all. This was an indication that the item was biased towards one group. The item proved to be more difficult for students who attended schools in rural areas than those who attended schools in urban areas. This item was the most obvious test item that exhibited test item bias.

Item 39. Graph for students who attended schools in rural areas	**Graph for students who attended schools in urban areas**





Are items of the mathematics JC paper 1 biased to male candidates in rural schools or urban schools?

From the 24 items (1, 3, 4, 6, 7, 10, 14, 15, 21, 23, 24, 25, 26, 27, 29, 30, 31, 32, 33, 34, 35, 38, 39, 40) that fitted the model, three (3) items (14, 33, 39) exhibited the most obvious location biased items towards a particular group (male students who attended schools in rural areas or male students who attended schools in urban areas).The Item Characteristic Curves for the three (3) identified test items were not identical for both the male students who attended schools in rural areas and male students who attended schools in urban areas, implying that the three (3) test items were location biased. Table 4 summarised the results of both groups item parameter estimates for the 24 items.

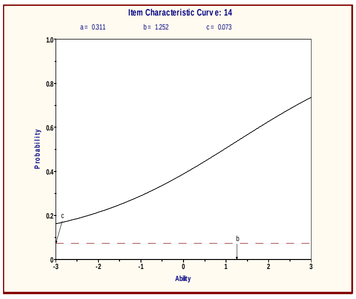
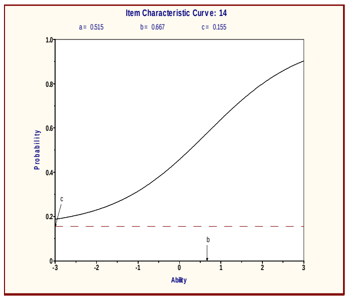
Table 4: Summary of the item parameter estimates for the Male Students who attended Schools in Rural (MSASRA) areas and Male Students who attended Schools in Urban Areas (MSASUA)

Items	Groups	Item discrimination (a)	Item difficulty (b)	Guessing Parameter (c)
Item 1	MSASRA	0.840	-0.078	0.244
	MSASUA	0.835	-0.263	0.226
Item 3	MSASRA	-0.061	-7.625	0.000
	MSASUA	-0.021	-28.170	0.049
Item 4	MSASRA	1.371	0.696	0.140
	MSASUA	1.433	0.441	0.168
Item 6	MSASRA	0.943	1.161	0.153
	MSASUA	0.898	0.236	0..057
Item 7	MSASRA	0.611	1.150	0.143
	MSASUA	0.795	0.816	0.139
Item 10	MSASRA	1.126	1.404	0.394
	MSASUA	0.789	0.979	0.343
Item 14**	MSASRA	0.311	1.252	0.013
	MSASUA	0.515	0.667	0.155
Item 15	MSASRA	1.053	1.152	0.375
	MSASUA	0.719	0.992	0.383
Item 21	MSASRA	0.767	1.274	0.234
	MSASUA	1.184	0.806	0.319
Item 23	MSASRA	0.629	0.125	0.062
	MSASUA	0.734	-0.098	0.135
Item 24	MSASRA	1.893	1.569	0.351
	MSASUA	1.387	1.323	0.300
Item 25	MSASRA	1.604	0.962	0.188
	MSASUA	1.145	0.479	0.165
Item 26	MSASRA	1.367	1.877	0.124
	MSASUA	1.119	1.618	0.140
Item 27	MSASRA	0.678	-0.029	0.144
	MSASUA	0.702	-0.200	0.256
Item 29	MSASRA	0.976	1.396	0.394
	MSASUA	1.025	0.800	0.381
Item 30	MSASRA	1.170	2.071	0.288
	MSASUA	1.417	1.607	0.292
Item 31	MSASRA	1.648	1.063	0.145
	MSASUA	1.595	0.823	0.174
Item 32	MSASRA	1.066	2.401	0.260
	MSASUA	1.928	2.067	0.275
Item 33**	MSASRA	0.834	0.165	0.302
	MSASUA	1.109	-0.071	0.425
Item 34	MSASRA	1.691	1.935	0.178
	MSASUA	2.108	1.545	0.169
Item 35	MSASRA	0.899	0.688	0.098
	MSASUA	1.288	0.379	0.168
Item 38	MSASRA	1.451	1.186	0.166
	MSASUA	0.891	0.879	0.141
Item 39**	MSASRA	0.274	1.241	0.000
	MSASUA	0.698	1.943	0.316
Item 40	MSASRA	1.068	0.458	0.193

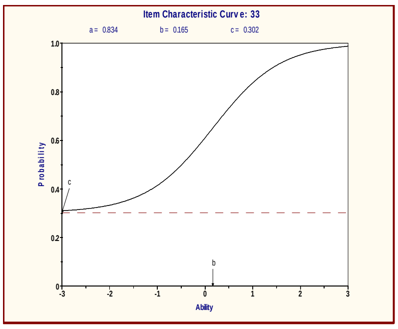
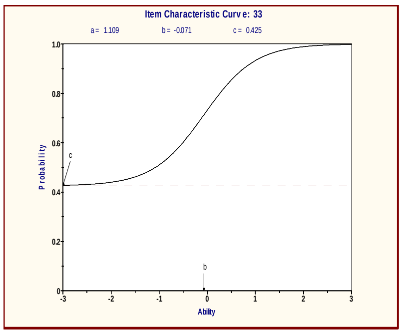
MSASUA	0.831	0.265	0.137

** Items exhibiting the most obvious location bias for male students

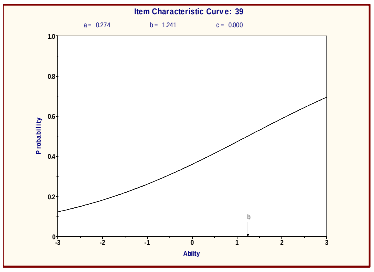
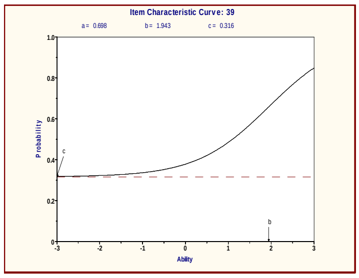
Item 14: The item proved to be more difficult for male students who attended schools in rural areas as indicated by the different ICCs. The item difficulty parameter for males who attended rural schools shifted towards the right more than the ICC of male students who attended schools in the urban areas. Also male students who attended schools in urban areas were more likely to get the correct answer by guessing than their counterparts as demonstrated in the ICC graphs. Therefore the item was biased towards male students who attended schools in rural areas.

Item 14. Graph for males who attended schools in rural areas	**Graph for males who attended schools in urban areas**
	

Item 33: For this item the probability of a low ability student getting the correct answer is more pronounced in students who attended school in urban areas as their ICC shifted up. That is male students who attended schools in urban areas were more likely to get the correct answer by guessing than those who attended schools in rural areas. Therefore the item was biased towards male students who attended schools in rural areas.

Item 33. Graph for males who attended schools in rural areas	**Graph for males who attended schools in urban areas**
	

Item 39: Comparing the two ICC curves, the graph for the males who attended schools in urban areas shifted up than the ICC graph for male who attended schools in rural areas. This indicated that within male students who sat for the examinations, this was the most obvious test item that exhibited test item bias. This item was biased to male students who attended schools in rural areas. The guessing factor (MSASRA c = 0.000 and MSASUA c = 0.316) was more pronounced for male students who attended schools in urban areas.

Item 39. Graph for males who attended schools in rural areas	**Graph for males who attended schools in urban areas**
	

Are items of the mathematics JC paper 1 biased to female candidates in rural schools or urban schools?

From the 24 items (1, 3, 4, 6, 7, 10, 14, 15, 21, 23, 24, 25, 26, 27, 29, 30, 31, 32, 33, 34, 35, 38, 39, 40) that fitted the, five (5) items (14, 26, 33, 35, 39) exhibited the most obvious location biased items towards a particular group (female students who attended schools in rural areas or female students who attended schools in urban areas).The Item Characteristic Curves for the five (5) identified test items were not identical for both the female students who attended schools in rural areas and female students who attended schools in urban areas, implying that the 5 test items were location biased. Table 5 summaries the results of both item parameter estimates for the 24 items.

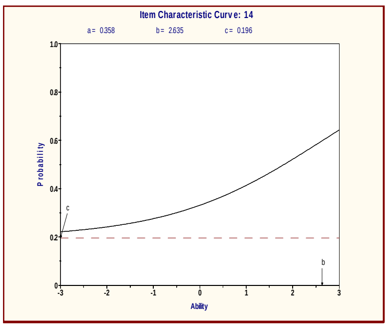
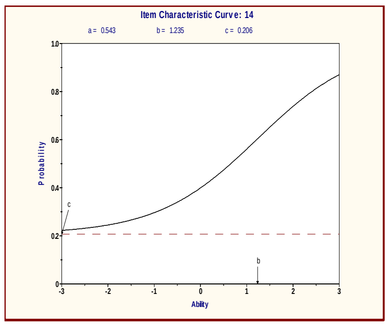
Table 5: Summary of the item parameter estimates for the Female Students who Attended Schools in Rural (FSASRA) areas and Female Students who Attended Schools in Urban Areas (FSASUA)

Items	Groups	Item discrimination (a)	Item difficulty (b)	Guessing Parameter (c)
Item 1	FSASRA	0.467	-1.280	0.000
	FSASUA	0.612	-0.686	0.275
Item 3	FSASRA	0.043	1.160	0.000
	FSASUA	0.026	2.540	0.73
Item 4	FSASRA	1.217	0.833	0.182
	FSASUA	1.100	0.232	0.072
Item 6	FSASRA	0.682	1.103	0.070
	FSASUA	0.651	0.762	0.116
Item 7	FSASRA	0.949	1.186	0.206
	FSASUA	0.724	0.789	0.160
Item 10	FSASRA	0.701	1.172	0.286
	FSASUA	0.660	1.071	0.271
Item 14**	FSASRA	0.358	2.635	0.196
	FSASUA	0.543	1.235	0.206
Item 15	FSASRA	1.220	1.262	0.532
	FSASUA	0.622	1.872	0.575
Item 21	FSASRA	0.742	1.032	0.282
	FSASUA	0.764	0.615	0.240
Item 23	FSASRA	0.444	-0.046	0.000
	FSASUA	0.502	-0.138	0.111
Item 24	FSASRA	0.898	1.508	0.301
	FSASUA	1.269	1.478	0.336
Item 25	FSASRA	1.050	1.161	0.176
	FSASUA	1.090	0.631	0.136
Item 26**	FSASRA	0.896	2.082	0.131
	FSASUA	1.453	1.270	0.113

Item 27	FSASRA	0.499	-0.467	0.000
	FSASUA	0.557	-0.316	0.123
Item 29	FSASRA	0.681	2.189	0.408
	FSASUA	0.414	1.435	0.280
Item 30	FSASRA	1.183	2.511	0.350
	FSASUA	0.693	2.045	0.292
Item 31	FSASRA	0.845	0.745	0.036
	FSASUA	0.854	0.354	0.034
Item 32	FSASRA	1.337	2.597	0.258
	FSASUA	0.786	2.544	0.229
Item 33**	FSASRA	0.470	-0.193	0.070
	FSASUA	0.683	-0.318	0.264
Item 34	FSASRA	2.898	2.000	0.194
	FSASUA	2.268	1.615	0.209
Item 35**	FSASRA	0.937	0.819	0.190
	FSASUA	0.763	0.093	0.027
Item 38	FSASRA	0.750	1.085	0.091
	FSASUA	0.691	0.604	0.093
Item 39**	FSASRA	0.141	3.440	0.000
	FSASUA	0.709	1.811	0.238
Item 40	FSASRA	0.788	0.090	0.074
	FSASUA	1.009	1.009	0.212

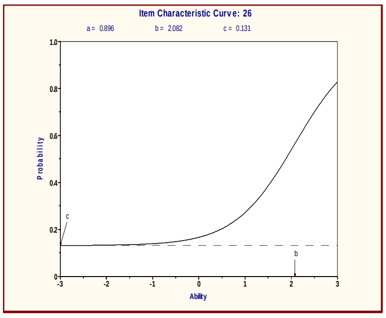
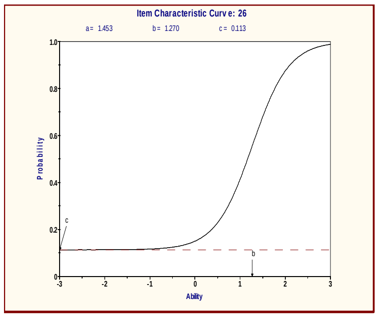
** Items exhibiting the most obvious location bias for female students

Item 14: The two ICCs were not the same, which is an indication of bias. The item was biased towards female students who attended schools in rural areas as it could not discriminate well amongst this group. The item difficulty parameter (FSASRA b = 2.635 and FSASUA b =1.235) as seen in the graphs indicated that the item was more difficult for female students who attended schools in rural areas than those who attended schools in urban areas as the ICC shifted towards the right.

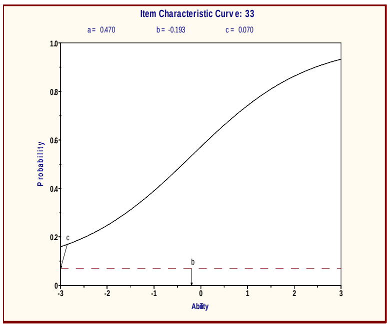
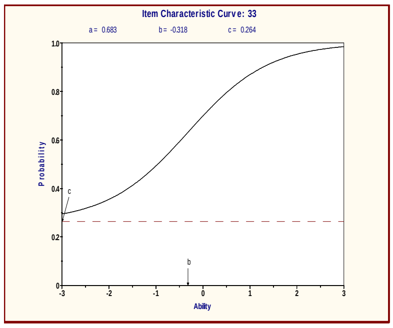
Item 14. Graph for females who attended schools in rural areas	**Graph for females who attended schools in urban areas**
	

Item 26: This item was biased towards female students who attended schools in the rural areas, because the ICC of these students shifted towards the right more than the ICC of female students who attended schools in urban areas. The item also proved to be more difficult for female students who attended schools in rural areas than those who attended schools in urban areas as shown by the item difficulty parameters (FSASRA b = 2.082 and

FSASUA b =1.270).

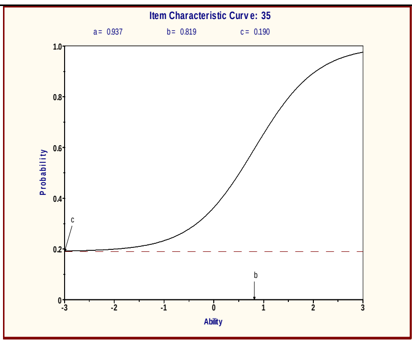
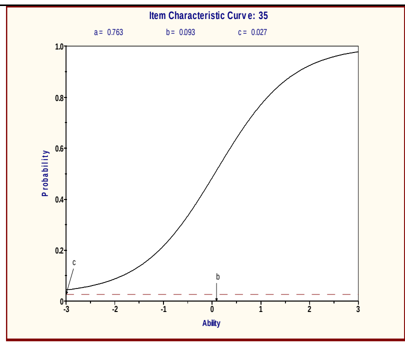
<table>
<tr><th>Item 26. Graph for females who attended schools in rural areas</th><th>Graph for females who attended schools in urban areas</th></tr>
<tr><td>
</td><td>
</td></tr>
</table>

Item 33: The two ICCs were different which indicated that the item was biased. The graph of female students who attended school in urban areas shifted up indicating that the item was easy for these students. The guessing factor (FSASRA c = 0.264 and FSASUA c = 0.070) was more pronounced for female students who attended schools in urban areas than those who attended schools in rural areas i.e. female students who attended schools in urban areas were more likely to get the correct answer by guessing than those who attended schools in rural areas. Therefore the item was biased towards female students who attended schools in rural areas.

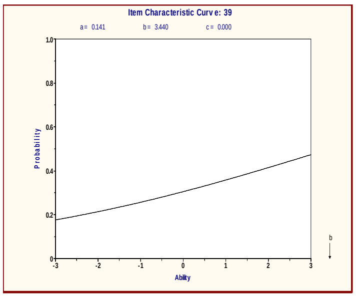
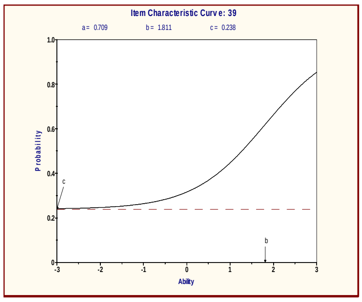
<table>
<tr><th>Item 33. Graph for females who attended schools in rural areas</th><th>Graph for females who attended schools in urban areas</th></tr>
<tr><td>
</td><td>
</td></tr>
</table>

Item 35: The two ICCs were different which indicated that the item was biased. Comparing the two ICC curves the guessing factor (FSASRA c = 0.190 and FSASUA c = 0.027) was more pronounced for female students who attended schools in rural areas than those who attended schools in urban areas. That is, female students who attended schools in rural areas were more likely to get the correct answer by guessing than those who attended schools in urban areas. Therefore the item was biased towards female students who attended schools in urban areas.

Item 35. Graph for females who attended schools in rural areas	Graph for females who attended schools in urban areas

Item 39: This item was the most obvious test item that exhibited test item bias. The ICC curves were very different, the ICC for females who attended schools in rural areas had a flat shape and it could not discriminate between the female examinees. This was a very poor item for females who attended schools in rural areas from the shape of the ICC. It can be concluded that this item was biased to female students who attended schools in rural areas.

Item 39. Graph for females who attended schools in rural areas	**Graph for females who attended schools in urban areas**
	

Discussions and Summary

Are items of the mathematics JC paper 1 biased to candidates in relation to whether the school they attend is in a rural or urban area?

The findings of the study revealed that six items which was 15%of all the items in the examination were location biased. All of these six items were location biased towards students who sat for their examinations in schools that are in rural areas. Campbell and Silver (1999) and also Barker (1985) have attributed the poor performance in rural areas to being largely due to a deficit in rural education including lack of funding, lack of a varied curriculum, lower scores on achievement tests and higher dropout rates. Researchers have attributed poor performance in rural areas to lower scores on achievement and not much research have been done to investigate the influence of the assessment tools on performance of these students. One could then speculate that the assessment tools have an influence on these students performance. This is because the current study's findings have indicated that 15% of the items in the examination were biased towards students in rural area. This gave students in urban areas an advantage over their counterparts in terms of performance. These items could have been biased towards students in rural areas due to exposure. Students in urban areas are exposed to a lot of things at an early age as they attend pre-primary which is not there or limited in rural areas. This allows them to have an early understanding and appreciation of the concepts. Attending pre-school also increases the level of motivation for students in urban areas. For example, item 29 could have been biased towards students in rural areas due to exposure, as the item was about naming a cross-sectional shape of an object given the object and naming a solid figure given the net diagram of the figure respectively. Student in urban areas could have had exposure to these shapes form an early age and for a long time giving them an advantage over students in rural areas. The students from urban areas could have the confidence and motivation to deal with the questions easily.

Item 14 was about time difference. It asked about the starting time in Botswana of a live broadcast of a game which started at 0815 hours in Australia which has a six hour time difference ahead. This question could be biased towards the rural child because in Botswana, most international life games are broadcast through television. The urban child in Botswana watches these games a lot while their counterparts do not as most families in rural areas do not have a television. Since the urban child watches these games a lot he/she would calculate the starting times of these games so that she/he will not miss the game. By doing so the urban child gets a lot of practice in calculating time differences.

Are items of the mathematics JC paper 1 biased to male candidates in rural schools or urban schools?

This question was basically included to find out those items that are location biased for male students only. Three items were location biased within males but all the items were also biased nationally. This meant that there were no questions that were unique to males only in terms of location bias.It emerged from the findings of the study that there were no items that were only location biased towards males and not biased nationally.

Are items of the mathematics JC paper 1 biased to female candidates in rural schools or urban schools?

This question was basically included to find out those items that were location biased for female students only. This would help understand the extent to which items were location biased within females. Item 33 and 35 (33% of location biased items within females) were the only location biased items within females. Item 33 was biased towards the female students from rural schools and item 35 was biased towards the female students from urban schools. Item 33 could have been biased towards the female students from rural schools because it was a technical question.

Item 33 required that the student identify the dimensions of a right-angled triangle without any picture which was why it was biased towards the female students from rural schools. Mostly, female students from urban schools have parents who understood the importance of education. These parents could have informed and encouraged their female child that technical careers were not only for male students but for females as well. Therefore, the female students from urban areas were advantaged in technical questions unlike female students from rural areas who were informed that there are jobs for males. Item 35according to the findings was biased towards the female students from urban areas mainly because of the guessing factor. That is, female students from rural school were able to get the correct answer by guessing than their counterparts.

The findings of the study present certain implications for educators, teachers, test developers, curriculum developers, policy developers and Botswana Examination council. All stakeholders need to be involved in developing assessment tools that are free of bias. Assessment bias reduces the validity of examinations as it leads to misguided decisions taking into consideration that examinations are used as gate keepers to higher institutions of learning and career path. The study has shown that location bias exists in mathematics examinations in Botswana. If mathematics examinations are location biased then stakeholders should be concerned of those other factors that may contribute to examinations being biased. Such factors as, to mention but a few: socio economic status, language, tribe, religion, gender etc. Creating an assessment tool that is free of bias is a difficult thing. Teachers and test developers ought to understand factors that contribute or would make tests/examinations biased in order to reduce bias in tests/examinations. In service workshops and training should be organized on issues of assessment bias so that teachers and test/examination developers would always have these issues in mind as they develop these tests and examinations. The findings of the study revealed that all location biased items were biased towards students who attended their schools in rural areas. The view that schools in rural areas perform poorly than schools in urban areas could be brought about by a faulty assessment tool. From the findings one would say test item bias should be reduced if not eliminated so that examinations are reliable and valid.

Conclusions

One important and immediate conclusion based on the findings of the study is location bias exist in mathematics examinations in Botswana. Further, the existence of location biasness is shifted towards the students who attend schools in rural areas. The same applies to males and females who attend schools in rural areas. Almost all location biased items were biased towards them. The impression that students who attend schools in urban areas outperform students who attend schools in rural areas might be brought about may be because students from schools in urban areas would have an upper hand even before the start of the examination due to test item bias. The current study has indicated that the assessment tool is an important factor to be considered when researchers look into factors that influence students' performance in mathematics. This is because if the assessment tool is biased, then obviously performance of the students will be influenced, be it in a positive or negative manner. Test developers should bear in mind that whenever one or more items in a test are biased, then the validity of the test or examination is at stake.

Recommendations

From the findings of the study the research has put forward the recommendations as follows:

- There should be more empirical studies on test item bias especially in Botswana where such studies are limited.
- Examination bodies should commit itself to eliminating or reducing gender and location biased items in national examinations.
- There is need for teachers and examiners to be trained by Examination Bodies and the Ministries of Education on item writing. This would enable them to produce valid, reliable and bias free assessment tools which will ultimately improve the quality of assessment.
- IRT framework should be incorporated into educational assessment. This would help provide for an objective assessment.
- The Item Characteristic Curves should be used to detect any kind of bias in test items. This would enable examiners to produce tests that have reduced bias or are free of bias.

References

[1] N. Abedalazi, Detecting DIF using item characteristic curve approaches, *The International Journal of Educational and Psychological Assessment*, 8(2) (2011), 1-15.

[2] O.O. Adedoyin, Using IRT approach to detect gender biased items in public examination: A case study from the Botswana junior certificate examination in mathematics, *Educational Research and Reviews*, 7(2) (2010), 385-399.

[3] O.O. Adedoyin, H.J. Nenty and B. Chilisa, Investigating the invariance of item difficulty parameter estimates based on CTT and IRT, *Educational Research and Reviews*, 3(2) (2008), 83-93, Retrieved August 15 (2013) from http://www. academicjournals.org/ERR.

[4] A. Anastasia and S. Urbina, *Psychological Testing*, (2006), New Delhi: Prentice Hall.

[5] Baghi, Heibatollah, Ferrara and Steven, A comparison of IRT, delta plot, Mantel-Haenszel techniques for detecting differential item functioning across subpopulations in the Maryland test of citizenship skills, *Paper Presented at the Annual Meeting of the American Educational Research Association*, (1989), Francisco, USA.

[6] B. Barker, Curricular offerings in small and large high schools: How broad is the disparity? *Research in Rural Education*, 3(1985), 35-38.

[7] G. Camilli and L.A. Shepard, Identifying biased test items by differential item functioning analysis using contingency table approaches: A comparative study, *Education Quarterly*, 67(1) (1994), 21-43.

[8] P.F. Campbell and E.A. Silver, *Teaching and Learning Mathematics in Poor Communities*, (1999), Reston, VA: National Council of Teachers of Mathematics.

[9] G. Camilli and L.A. Shepard, *Methods for Identifying Biased Test Items*, (1994), Newbury Park, CA: Sage.

[10] G. Camilli and L.A. Shepard, *Methods for Identifying Biased Test Items (Vol. 4)*, (2009), Thousand Oaks, CA: Sage Publications.

[11] R.K. Hamleton, H. Swaminathan and H.J. Rogers, *Fundamentals Odd Item Response Theory*, (1991), Newbury Park, Califonia: Sage Publications.

[12] R. Hambleton and J. Rodgers, Item bias review, *Practical Assessment, Research and Evaluation*, 4(6) (1995), Available: http://PAREonline.net/getvn.asp?v=4&n=6.

[13] H. Karami, Detecting gender bias in a language proficiency test, *International Journal of Language Studies*, 5(2) (2011), 27-38.

[14] F.M. Lord, *Applications of Item Response Theory to Practical Testing Problems*, (2002), Hillsdale, New Jessey: Erbaum Press.

[15] A.E. Maliki, N.N. Ngban and J.E. Ibu, Analysis of students' performance in junior secondary school mathematics examination in Bayelsa State of Nigeria, *Studies on Home and Community Science*, 3(2) (2009), 131-134.

[16] A.W. Meade and M. Fetzer, A new approach to assessing test bias, *Paper Presented at the 23th Annual Conference of the Society for Industrial and Organizational Psychology*, April (2008), San Francisco, CA.

[17] A. Milardo, An examination of the mantel statistical procedure's effectiveness inidentifying item bias in the Maryland state performance assessment program (MSPAP), (2000), University of Baltimore.

[18] R.A. Morales, Evaluation of mathematics achievement test: A comparison between CTT and IRT, *The International Journal of Educational and Psychological Assessment*, 1(1) (2009), 19-26.

[19] J.Q. Pedrajita and V.M. Talisayon, Identifying biased test items by differential item functioning analysis using contingency table approaches: A comparative study, *Education Quarterly*, 67(1) (2009), 21-43.

[20] S.M. Pine, Applications of item response theory to the problem of test bias, *Research Report of the Department of Psychology*, (2006), University of Minnesota.

[21] H. Swaminathan and H.J. Rogers, Detecting differential item functioning using logistic regression procedures, *Journal of Educational Measurement*, 2(7) (2005), 361-370.

[22] M. Wiberg, Measuring and detecting differential item functioning in criterion-referenced licensing test: A theoretic comparison of methods, Educational Measurement, *Technical Report No. 2*(2007).

[23] B.D. Zumbo, A handbook on the theory and methods of differential item functioning (DIF): Scores, Ottawa, ON: Directorate of Human Resources Research and Evaluation, (2009), Department of National Defence.