Journal of CUDIMAC (J-CUDIMAC)
ISSN 0794-4764 (Print) ISSN 2651-6063 (Online)
Vol. 8, No.1, September, 2020

http://cudimac.unn.edu.ng/volume-8/

Application of Rasch Model in Measuring Differential Item Functioning (DIF) of Students Attitude to Geography in South East Nigeria

[1]Ikwueze, C.C. and [2]Ani, M.I.

[1]Department of Science Education, Michael Okpara University of Agriculture, Umudike, Abia State, Nigeria
[2]Department of Educational Foundations, Alex Ekwueme Federal University, Ndufu Alike Ikwo Ebonyi State, Nigeria

Abstract

This study sought to detect Differential Item Functioning (DIF) in Geography Attitude Scale (GAS), using Rasch model procedure. The objective of the study was specifically to ascertain how the items of Geography attitude scale function with respect to gender and school type. The design of the study was instrumentation research design and the area of the study was south-east Nigeria. Two research questions guided the study. The sample was made up of 850 students from 78 secondary schools in three states of the south east (Abia, Anambra and Enugu State). Simple random sample and proportional stratified random sample were employed in sampling of the states, schools and subjects. An 80-item draft instrument, called Geography Attitude Scale (GAS) was developed by the researcher and trial tested. The responses were subjected to factor analysis after which 38 items were dropped and 42 items emerged. Conditional maximum likelihood estimation technique of the WINSTEPs 3:81software was used to analyze the data in order to answer the research question one and two. DIF effect discovered was approximately eight (8) items which were discarded, making the items of the GAS 34 in number. It was recommended among others that the instrument should be used to compare attitude level among sub groups of students.

keywords: Assessment, Attitude, Geography, Rasch model, and Differential Item Functioning

Introduction

According to Bloom (1956), assessment can be carried out in the three domains of learning, that is, cognitive domain, which deals with knowledge and intellectual development of skills; psychomotor domain, which included manipulative and motor skills and the affective domain focuses on attitude and value. In its simplest form, affective domain characterizes the emotional area of learning reflected by the beliefs, values, attitudes, and behaviors of learners (Kara, 2009). It deals with how learners feel while they are learning, as well as with how learning experiences are imbibed so they can direct the learner's attitudes, opinions, and behavior in the future (Miller, 2005). According to Federal Republic of Nigeria (FRN) (2013), affective behavior consists of attitudes,

interest, emotional adjustment, values, beliefs, social relations, habits, and overall lifestyle. Krathwohl, Bloom and Masia (1964) opine that affective educational outcomes can be arranged in a hierarchy, according to complexity. The order starts with an ability to listen to ideas, followed by responding in interaction with others and manifesting values or attitudes suitable to a particular situation. The highest levels involve displaying a commitment to principled practice on day-to-day basis, as well as a willingness to revise one's judgment and change one's behavior in light of new evidence (Shephard, 2008).

Assessment in the affective domain is important because, a close look at the provisions of the new policy on education in Nigeria (FRN) (2013) unveiled that the first two of the four educational aims and objectives in the country are in the affective domain. The first is the inculcation of the right type of values and attitudes for the survival of the individual and Nigerian society (Nigeria Consulate, 2012). In more distinct terms, the teachers are needed to acquire a more practical and suitable orientation that would awaken and preserve the interests of the pupils in various school subjects and programmes, and other similar conditions in the vast world outside the classroom (Nworgu, 2004). Their teachings are expected to facilitate the development in individuals of favorable attitudes toward issues of unity, social integration, patriotism, civic duties, and other socio-political ideals.

As teachers try to fulfill these objectives, they are expected to resolve often the level to which their instructional procedures are succeeding in supporting the desired affective behavioral changes in their students. However, assessment of the affective domain may at times be more significant than the cognitive because it can assist an instructor to intervene in students who seems to "give up on themselves" in the classroom. In assessing the affective domain, with feedbacks in hand, a teacher can change the lesson plan based on the student's desire. Therefore, the affective domain must be assessed periodically during and after teaching, so as to monitor changes in the students and modify the lesson plan. These affective features are showed in terms of students' regular attendance to class, punctuality to school, honesty, active participation in class activities, sociability, helping others, affection, carrying out assignments, attentiveness and so many others (Gano-Phillip and Friedman, 2009).

Teachers do not often carry-out formal assessment of affective learning, contrary to what they do in the cognitive domain where the child's intellect, knowledge and the ability to think are usually assessed using a type of assessment method or the other. This is because it is not always feasible to unveil a person's inner state of belief, motivation, or perception. Also, people are reluctant to produce answers observed as socially unwelcome, and therefore tend to report what they think their attitudes should be rather than what they know to be (Phillips, 2011). Affective traits are also easily impelled by a person's mood or feeling, which deteriorates from day to day or even hour to hour. Other challenges teachers may encounter while assessing affective learning are difficulties in promptly stating desired affective learning outcomes because they involve opinions, beliefs, and attitudes (Smith, 1991). The absence of clearly explained affective learning outcomes makes the assessment of this feedback more demanding.

Attitude as one of the attributes of affective domain has been defined by many authors from different areas of study. Nevertheless, the elements that comprise these definitions are alike and complementary. According to Ajzen, cited in Pulka, Kwentishe and Ibrahim (2014), attitude is

defined as the predisposition to respond in a generally favorable or unfavorable manner with respect to the object of the attitude. He further said that attitude is an emotional and mental readiness or a preliminary tendency based on experience, knowledge, emotion or motivation on any subject, social topic or event. Kind, Jones, and Barmby (2007) viewed attitude as having different components which include cognitive (knowledge, belief, and ideas); affective (feeling, like, dislike) and behavioral (tendency, towards an action).

The attitude that one has towards an object makes one to make judgment as to whether the object is good or bad, healthy or unhealthy, satisfying or unsatisfying, relevant or irrelevant (Crano and Prislin, 2006). Also, Ball, cited in Phillps (2011: 14), concluded by combining the common features from many definitions, that "attitude is an implicit cue- and drive- producing response to socially salient characteristics and possesses evaluative properties". He further explained that an attitude is within the individual, and cannot be seen, felt, touched, or observed.

A student's attitude regarding a school subject most of the times shows the effort he/she put in during class, his/her engagement in activities involved in the lesson, learning outcome related to the subject (Dibyajyoti, 2012). It also implies a positive or negative predisposition towards a school subject and every activity connecting to the subject in question.

Attitude towards geography is a separate affective conduct that has been learned over the years in association to geography achievement. Attitude towards geography can be defined as a positive or negative emotional disposition towards geography as a school subject. It can be an individual's attitude toward geography as the emotions a student associates with geography, his or her beliefs toward geography and how he or she act towards geography.

Geography is the study of the different environments, places, and spaces of the Earth's surface and their interactions. It seeks to respond to the question of why things are the way they are. According to Oformata (2008), the modern academic discipline of geography is rooted in ancient practices concerned with the characteristics of places, in particular their natural environments and people, as well as the inter relation between the two. Its specific identity was first framed and named some 2000 years ago by Greeks, where geo and graphein were merged to mean "earth writing" or "earth description". Geography is the scientific study of the location of people and activities around the earth, and reason for their distribution (Necati, 2010). Walker (2000:56) observed that, "it is a science in the sense that it is systematically studied and organized". That is to say, that it has the order of gaining knowledge, it has a specific procedure through which knowledge is increased and fact can be measured. Oformata (2008) defined geography as the study of the earth surface as the home of man or more precisely a science of spatial relationship, which focuses attention mainly on the interaction between man and his environment.

Geography is useful for both the students who are likely to proceed with the subject at higher level and those who will not continue. It endows students with a body of knowledge to make them operational and socially useful in the quick changing world. Geography is a distinct and dynamic science and or social science discipline that deals with the study of mankind and his physical environment (Akintade, 2012). It, therefore, aids individuals to comprehend the value of their environment and its extensive natural asset. As a subject, it covers vast areas, expressively, and it is intellectually inspiring. It indicates an association with all other school subjects. It inculcates in the students the need to treasure and create a sense of obligation towards their own community.

Geography is very vast but a fascinating subject which is connected to other subjects such as the Biology, mathematics, Social Studies, Economics, and Agriculture. Hence, entails a lot of hard work to learn and appreciate it at the senior secondary school level. But, with positive attitude and strategy, it is always a delight to study it. Nevertheless, there has been an observable quick decline in the number of students that register Geography at the senior secondary school level in South-East Nigeria. According to Aydin (2007), geography is generally seen as a difficult subject to learn at school, and many students consider geography as a compilation of statistical data and knowledge. It can be seen as the lowest valley in Africa, the deepest river in Nigeria and the largest city or the least region. Great numbers of students do not see the relevance of geography lessons. Teacher's attitude and association with students has a very important influence on the student's attitude toward a subject like geography.

The way he/she associates with the students and passes across his teachings goes a long way to influence the student's response and attitude towards the subject. Students' attitude may vary due to the change of condition of the student, change of teacher, change of teaching method, learning materials for the student, period of holding the lesson during the school periods and change of home situation. Other challenges in the teaching and learning of geography, as stated by Obika (2004), include change in content, changes in learning and instructional materials, and changes in evaluation/assessment. Main concern, as she stated, are students' poor attitude, interest, and unwillingness to attend geography lessons.

Another important challenge according to the present study is on how teacher can develop items that have the similar difficulty level for individuals of two groups with the same level of ability, that is reliable and valid and the suitable psychometric property he/she should use in approximating the child's mental skill in all subjects including geography. A better way of solving these problems is to develop a good instrument for measurement of students' attitude and for dictating Differential Item function (DIF). That is the ability of the teacher to develop an instrument set of items, in order to check if a Differential Item function (DIF) can be noticed.

Differential item function is a notable challenge in item response theory. It exists if the likelihood of a right response among equally able person differs in subgroups, for instance, if the difficulty of an item relies on the membership to a racial, ethnic or gender subgroup. Then the achievement of a group can be reduced because these items are associated to particular knowledge that is less present in this group. The implication is measurement bias and possibly discrimination.

Differential item function study is expected to display whether an instrument could be one of the causes of the gender disparity in test performance. Item facility for each unit are established as well as items with maximum and minimum DIF values to probe if an item format provides any advantage or disadvantage to any of the group of students. An attitude test is said to have DIF effect if the functions separately for a particular subgroup of test. Ogbabor (2012) states that DIF measure characteristics that not necessary or items that are irrelevant to the test. Frequently, examination/attitude items are considered biased because they contain source of difficulty that are not relevant to the construct being measured and these extraneous sources affects test-takers performance (Zumbo, 1999).

Ultimately, the probability of responding to an item positively should only rely on a testee's ability level on the trait being assessed including any useful item parameters such as difficulty or

discrimination. If the probability also depends on a construct-irrelevant subgroup factor, such as race, ethnicity, gender or socio-economic status, then DIF may be present hence reducing the validity of such items. Suitable testing exercise suggests that developers should recognize items with DIF and elude using them in operational test except they are judged to be fair and are required for valid measurement.

In general, there are three main DIF analysis methods: the generalized Lai Eton test (Lai Eton, 2002), logistic regression and item response theory (IRT). The Lai Eton DIF statistic was proposed as a method for detecting DIF by Pallant and Tennant (2007). It has been extensively used in educational measurement because of its easy application in testing programs. Nevertheless, it is often used to find uniform DIF for dichotomous items. The logistic regression procedure for DIF was introduced by Swaminwthan and Rogers (1990). It can discover both uniform and non-uniform DIFs and can likewise include exogenous variables in the models. Item response theory (IRT) DIF procedures have gained multiple attentions because they can model differences in item difficulty and discrimination parameters. Differences in difficulty of items between groups reflect uniform DIF while differences in item discrimination parameters show non uniform DIF.

Rasch model as one of models of IRT investigate DIF only in the threshold (location) parameters. This methodology has strict requirements to maintain the elegance of the Rasch model (e.g. Sum score sufficiency). Any item that differs in its ability to discriminate among respondents compared to other items in a measure is considered a misfiting item to the Rasch model (Smith, 1991). DIF takes place when examinees from varying groups show differing probabilities of success on (or endorsing) an item after matching on the underling construct that the item is intended to measure (Zumbo, 1999: 12). According to Lai Eton (2002), values of 0.5 logits DIF contrast would be vital for likert scale. Meanwhile, Wright and Panchalakesan in Pallant and Tennant (2007) argue that the size of gender DIF which is less than 0.5 logit is considered unimportant (DIF negligible) values, above 0.5 logits show that the difference is noticeable. Also, Bond and Fox (2007) suggest these DIF indication based on the studied groups which are (i) DIF contrast ± 0.5 (DIF contrast > + 0.5 – 0.5) and (ii) $p < 0.05$). This yardstick given by Bond and Fox was used to pinpoint differential items and eliminate from the scale in this study.

A large number of research studies have been conducted in this DIF. Obinne, Nworgu, and Umobong (2013) in a study “DIF of tests, carried out by the two major examination bodies in Nigeria using Two-parameter model of IRT”, the result from the study showed that the biology examination items by the two examinations bodies; West Africa Examination Council (WAEC) and National Examination Council (NECO) discriminated equally among urban and rural examinees. This implies that students’ scores in such examinations are determined mostly by the group to which an examinee belongs and not by the ability. 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. Studies have also shown that no significant difference existed between performance of males and females in mathematic in the urban areas (Oluwatayo, 2011). Akintade (2012) observed that difference in the attitude of senior secondary students towards offering geography is a significant factor that affects learners in their choice.

However, Unal (2012) in his study on high school students’ attitude toward geography courses showed that male students have more positive attitude towards geography than female students.

There is no significant difference between male and female students' attitude towards geography. According to Omare, as cited in Philips (2011), there is a major concern not only to identify attribute of attitude and interest but to have a proper measure in assessing the affective behavior of students, which attitude is one of them in school subject like geography. Also, it is relevant that the tests utilized by teachers be free of systematic demographic subgroup bias. The CTT which is frequently used psychometric properties in determining the quality of the instrument used for objective assessment of students' achievement does not provide an accurate assessment of student's affective feedbacks.

CTT-based method of assessing bias are fundamentally insufficient, especially strategies that based their assessment of bias on the presence of group mean differences on overall tests scores across demographic groups. IRT techniques give a dominating means of testing items for bias, using what is called DIF analysis. This is because the measurement of an examinees' ability from responses to test item or survey is not limited to a particular test. Rather, it can be measured by any collection of items that are considered to be measuring the same trait. Another desirable property of IRT is that it provides measures of precision of ability estimate at each ability level. The IRT models include; one-parameter (Rasch model), two-parameter, and three-parameter. The Rasch model of IRT is a mathematical formulation linking the likelihood of the outcome when an individual attempts a single item to characteristics of the person and the item. It is therefore one of the families of psychometrical models for the measurements of achievement, ability, attitude and personality tests, and is arguably the least complex member of this family. Rasch model was advanced by psychometricians as a new measurement system to work on the limitation of CT measurement. Not minding the limitations of CTT, researchers in Nigeria still use it in their studies and the Rasch model which is one of the best models of IRT is not utilized. The model does not receive sufficient recognition from Nigeria researchers. All these necessitated further investigations on the use of IRT (Rasch model) to detect geography attitude items with DIF. Hence, this study seeks to trail the presence of gender and school-type DIF in geography attitude items, using Rasch model.

Purpose of the Study

The main purpose of this study was to develop and apply Geography Altitude Scale (GAS) using Rasch model. Specifically, the study was designed to:

1. determine how items of Geography Attitude Scale function with respect to school type.
2. determine how the items of Geography Attitude Scale function with respect to Gender.

Research Questions

The following research questions guided this study:

1. To what extent do the items of Geography Attitude Scale function with respect to school type (single and mixed schools)?
2. To what extent do the items of Geography Attitude Scale function with respect to Gender (male and female)?

Research Methodology

In carrying out the research, an instrumentation design was employed. All the senior secondary schools two (SS2) students in South-East Nigeria who offered Geography within 2015/2016 academic session were used. Multi-stage sampling technique was used to determine the sample of the study. The sampling stages involved in this study were sampling of the states (using simple random sampling technique), schools and sampling of research subjects (using stratified proportionate random sampling technique). With a stratified proportionate random sampling technique, a sample of eight hundred and fifty (850) students were selected from a total of 24,921 from 78 schools selected from 1,104 SS2 students. in three of the five states in south east geo-political zones of Nigeria (Abia, Anambra and Enugu states). This sample size was 3.4% of the total population from each of the state selected. Geography Attitude Scale (GAS) was face validated by three experts in Measurement and Evaluation and Geography Education. GAS is a self-rating questionnaire consisting of 80 structured items with four options (of strongly agree, agree, disagree and strongly disagree) only. It contained attitude generating statement designed to find out the attitude of senior secondary two students (SS2) towards Geography. The drafted instrument was subjected to factor analysis for data reduction and structure detection. Forty-two (42) items that were loaded into six factors survived factor validation. Thus, 38 items were dropped because they did not meet up with the minimum acceptance value of .350.

The validated instrument was trial-tested using 100 students from Kogi state. The reliability of the instrument calculated using Cronbach Alpha was o.775 and Rasch rating model software WINSTEPS was used to obtain Rasch person and item internal consistency of 0.93 and 0.95 respectively. Two research questions were posed and the instrument and were analyzed using Rasch rating scale model software WINSTEP (Version 3: 81).

Results

Research question 1:

To what extent do the items of GAS function with respect to Gender (male and female)?

To answer this research question, DIF measures according to gender contrasts and probability levels were presented in Table 1

Table 1: Differential Item Function (DIF) Scores with Respect to Gender

Item Number	Male DIF measure	Female DIF measure	DIF contrast	Probability
1.	-.49	-.26	.23	.013
2.	.53	-.10	-.63	.000
3.	.04	-.21	-.26	.142
4.	-.33	-.19	.14	.063
5.	.57	.02	-.55	.001
6.	-.36	.04	.40	.060
7.	-.26	.01	.30	.054
8.	-.66	.01	**.67**	.026
9.	.79	.10	-.69	.039
10.	-.28	.03	.30	.426
11.	-.82	-.02	**.81**	.039
12.	-.23	-.02	.22	0.67
13.	.35	.07	-.28	.440
14.	.38	.06	-.32	.050
15.	-.27	-.03	.24	.078
16.	.40	.09	-.31	.523
17.	-.43	.02	.45	.359
18.	.54	.00	-.54	.004
19.	.39	.05	-.34	.258
20.	-.07	-.07	.00	.343
21.	-.11	.12	.22	0.75
22.	.54	.08	-.46	.035
23.	.49	.05	-.43	.456
24.	.11	.07	-.05	.601
25.	.22	.04	-.18	.084
26.	-.41	.05	.46	.173
27.	-.57	-.07	.50	.415
28.	-.21	-.09	.11	.207
29.	.02	.02	.00	1.000
30.	-.18	.09	.27	.057
31.	-.42	.05	.47	.063
32.	-.37	-.12	.26	.075
33.	.07	-.07	-.14	.130

34.	-.04	.07	.11	.209
35.	-.26	.13	.39	.081
36.	.63	.01	-.62	.004
37.	.42	.00	-.42	.163
38.	-.04	-.04	.00	1.000
39.	-.22	-.08	.14	.135
40.	-.12	-.07	.05	.564
41.	.08	.13	.05	.561
42.	.20	.8	-.02	.812

Table 1 shows how the items functions with respect to gender. The values of DIF contrast for the 42 items range from -.69 to 0.81. The maximum contrast is 0.81 logits, while the minimum DIF contrast is -0.69. The items with the maximum and minimum DIF contrast values are 9 and 11 respectively. According to Bond and Fox (2007) who carried out a study on group differences, suggested these DIF indicators based on the studied groups which are (i) DIF contrast ± 0.5 (DIF contrast ≥ + 0.5 ≤ - .5), and (ii) $p < 0.05$). Hence, the researcher detects DIF using DIF contrast greater than .5 logits and less than -.5) as showing noticeable and significant difference respectively. From the above table, noticeable gender DIF could be observed in 6 items whose gender DIF contrast were above .5 and less than -0.5 logits. These items numbers are 8, and 11 for items with positive DIF effect and items 2,5,18, and 36 for items with DIF contrast less than -.5. These items represent 14.2% of the items. Similarly, all of the items have their probability values less than .05 also suggesting significant DIF effects. The 2 items will be excluded from the scale. Two items have DIF contrast zero in the table. These items have no DIF effect between themselves.

Research question 2:
To what extent do the items of GAS function with respect to School type (mixed- sex and single-sex)?

To answer this research question, DIF measures according to school type contrasts and probability levels were presented in Table 2.

Table 2: Differential Item Functioning (DIF) Scores with Respect to School Type

Item Number	Mixed sex DIF measure	Single sex DIF measure	DIF contrast	Probability
1.	-.39	-.34	-.05	.602
2.	.29	.09	.20	.022
3.	-.05	-.14	.09	.319
4.	-.49	.00	-.48	.065
5.	.39	.15	.24	.085
6.	-.32	.05	-.37	.045
7.	-.21	-.04	-.17	.064
8.	-.34	-.23	-.11	.232
9.	.77	.00	**.77**	.014
10.	-.25	.05	-.30	.054
11.	-.59	-.13	-.46	.075
12.	-.25	.03	-.28	.092
13.	.17	.23	-.06	.488
14.	.20	.18	.02	.814
15.	-.23	-.03	-.20	.067
16.	.31	.14	.17	.048
17.	-.34	-.01	-.33	.068
18.	.30	.18	.12	.166
19.	.35	.04	.31	.056
20.	.02	-.18	.20	.076
21.	.10	-.08	.18	.047
22.	.52	.04	.47	.060
23.	.46	.02	.44	.058
24.	.04	.14	-.10	.278
25.	.16	.08	.08	.369
26.	-.20	-.10	-.10	.275
27.	-.33	-.25	-.08	.392
28.	-.23	-.05	-.18	.049
29.	-.01	.05	-.05	.549
30.	-.26	.22	-.48	.052
31.	-.32	.02	-.34	.041
32.	-.35	-.10	-.25	.066
33.	.06	-.07	.13	.151

34.	.02	.02	.00	1.000
35.	-.12	.04	-.17	.063
36.	.54	.01	**.53**	.000
37.	.29	.10	.19	.028
38.	.00	-.08	.08	.386
39.	-.17	-.12	-.05	.577
40.	-09.	-.09	.00	1.000
41.	.15	.06	.09	.316
42.	.14	.23	-.09	.307

Table 2 above shows the DIF value based on school type. The values obtained ranged from -0.48 to 0.77. Still using the range suggested by Bond and Fox (2007) i.e. DIF contrast greater than .5 logits and $p<0.05$ as showing noticeable and significant difference, school type DIF could be observed in 2 items e.g. items 9 and 36 with DIF contrasts .77, and .53 respectively. For these items, their logit values were above .5 and probability values equally less than 0.05 (.014 and .000) respectively. The 2 items represent 5% of the items. These 2 items will be removed from the GAS. Two items such as items 34 and 40 show DIF contrast of .00, indicating that the items function equally for the two groups.

Discussion

The research question 1 and 2 aimed at finding out the different items function with regard to gender and school types respectively. Differential item functioning (DIF) occurs when a test item favours or hinders a characteristic exhibited by group members or a test taking population. Bond and Fox (2007) suggest that DIF indicators based on the studied groups which are (i) DIF contrast ± 0.5 (DIF contrast ≥ + 0.5 ≤ .0.5) and (ii) $p<0.05$. The researcher identified DIF using DIF contrasts greater than .5 logit as showing noticeable and significant different respectively. As seen from the table, gender DIF could be observed in 6 items whose gender DIF contrast was ± 0.5 logits.

Also, the result of the study as reported showed that there were two items without differential effects, meaning that the DIF contrast was .00. In all, 86% of the GAS items (36 items) function equally among the groups. This result is in agreement with the study, carried out by Ariffin, Idris and Ishak (2010), whose finding detected items with DIF effect.

Likewise, to find out how the different items function with respect to school type (mixed and single sex school), the ranged suggested by Bond and Fox (2007) was also used i.e. DIF contrast greater than .5 logits and $p<0.05$ as showing significant difference school type DIF could be observed in 2 items e.g. items 9 and 36. For these items, there logit values were above .5 and P-values equally less than 0.05. The 2 items represent 4.7% of the items. Two items such as item 34 and 40 show DIF contrast of .00, 95% of the GAS items (40 items), function identically among the two groups (mixed sex and single sex schools). According to Zumbo (1999), for an item to be biased, DIF is required but not sufficient. Therefore, the 2 items with DIF effects will be kept aside for

further investigation concerning items bias. This is in line with the study by Xavier (2012) that detected two items that were potentially problematic, DIF wise and were consequently discarded.

Conclusion

This study sought to detect Differential Item Functioning (DIF) in Geography Attitude Scale (GAS), using Rasch model procedure. The objective of the study was specifically to ascertain how the items of Geography attitude scale function with respect to gender and school type. The findings of the study revealed that most items of the GAS have the difficulty value of equally high or equally low for male and female students. Nevertheless, gender DIF effects can be seen in 6 items. The other remaining items show contrast of .00 or less than meaning that the items function equally for male and female students. It was also revealed that school type DIF effects were also observe in a very small percentage of the items (50%) for mixed sex schools and single sex schools' students. Apart from 2 items that was affected, other items have equal or nearly equal strength for students in mixed sex and single sex schools. Therefore, the 8 items with DIF effects were removed from the scale, making items of the GAS 34 in number.

Recommendations

From the study, the following recommendations made based on the results of this study.

1. The instrument, GAS should be used to assess attitude of students towards geography in all secondary school setting, especially in senior secondary level; whether public, private or mission secondary schools. This is the may purpose of this research study; to develop an attitude scale for assessing attitude to Geography.

2. Giving the obvious advantages of IRT over other popular measurement framework, the government should encourage our examination bodies such as WAEC, NECO, NABTEB, etc to adopt this measurement framework. This will ultimately surmount the measurement problems we frequently encounter in Nigeria. Such measurement problem as test score equating has nearly gone into extinction in the foreign countries that have adopted the IRT measurement framework. IRT framework can also do the magic for us in Nigeria. This in line with the findings of research question 1 and 2, that discussed DIF of a test. According to these findings, it is essential that item is fair to all applicants in a population (no bias). Based on this, examination bodies in Nigeria are expected to place the examinees on the correct ability level because the IRT analysis is able to describe the test items and the abilities of the examinees.

3. Secondary schools' teachers in Nigeria should be oriented on the usage of IRT for psychometric analysis of their examinations. This way the quality of our test items in such school will get more refined and measurement problems associated with the presently used framework will get obliterated. This recommendation is in agreement with the findings of research questions 1,4 and 5 that dealt with misfit items and SEM of test items.

References

Akintade, B.O. (2012). Considering the determinants of selecting geography as a discipline: The case of senior secondary school students in Ilorin, Nigeria. *Ozean Journal of Social Sciences,* 5(1): 202.

Arber, N. (2003). Assessment for learning: *Teaching Geography*, 28(1): 42-6.

Ariffin, S.R., Idris, R. & Ishak, M.N. (2010). Differential item functioning in Malaysian Generic skills instrument (MGSI). *Journal Pendidikan Malaysia*, 35(1), 1-10.

Aydin, F. (2009). Secondary school students' opinions about geography course: A qualitative study. *Archives of Applied Science Research,* 4(2): 105-110.

Bloom, B.S. (1956). Taxonomy of educational objectives. New York: David McKay Company Inc.

Bond, T.G. & Fox, C.M. (2007). Applying the Rasch model: fundamental measurement in the human sciences. Second Edition. New Jersey: Lawrence Erlbaum Associates Publishers London.

Crano, W.D., Prislin, R. (2006). Attitude and persuasion. *Annual Review of Psychology*, 57: 345-374.

Dibyajyoti, M. (2012). Achievement in mathematics effect of gender and positive / negative attitude of students. *International Journal of Theoretical and Applied Sciences,* 4(2): 157-162.

Federal Republic of Nigeria (2013). National policy on education, Lagos. NERC Press.

Gano-Phillips, S. & Friedman, S.J. (2009). Affective education and experiential learning. Presentation as part of the "3+3+4" symposium on general education at the Chinese University of Hong Kong. http://www.cuhk. edu.hk/oge/rcge/ conference/090603.htm.

Krathwohl, D.R., Bloom, B.S & Masia, B.B. (1964). Taxonomy of educational objectives, the classification of educational goals. Handbook II: Affective domain. New York: David McKay Co, Inc.

Lia, J.S. & Eton, D.T. (2002). Clinically meaningful gaps. *Rasch Measurement Transactions*, 15 (4), 850.

Miller, M. (2005). Teaching and learning in affective domain. In M. Drey (Ed.). Emerging perspectives on learning, teaching and technology. Retrieved October 14, 2009, from http://projects.coe.uga.edu/ep/tt/.

Necati, T. (2010). High school attitude towards geography and the questions they wonder about scientific *Research and Essays,* 5(13): 1236-2733.

Nigeria Consulate, (2012). New York. http://www.nigeriahouse.com/ ducation.aspx

Nworgu, B.G. (2004). Educational research basic issues and methodology. Ibadan: Wisdom Publisher Limited.

Nworgu, B.G. (2011). Differential Item Functioning: A Critical Issue in regional quality assurance. Paper Presented in NAERA Conference.

Obika, G.A. (2004). Development and preliminary validation of interest of students in economics in secondary schools. Unpublished UNN Thesis.

Obinne, A.D.E, Nworgu, B.G & Umobong, M.E (2013). An investigation into differential item functioning of test conducted by the two major examination bodies in Nigeria Journal of Advance in Educational Research, 2(1): 1-8

Oformata, G.E.K. (2008). Geographic thought: 2nd edition. Enugu: Jamoe Enterprises.

Ogbebor, U.C. (2012). Differential Item Functioning. Economic Questions Paper of National Examination Coucil on Delta State Nigeria. Unpublished M. ed thesis. University of Ibadan.

Oluwatayo, J.A. (2011). Gender difference and performance of secondary school students in mathematics. *European Journal of Educational Studies,* 3(1): 201-2019.

Pallant, J.F. & Tennant, A. (2007). An introduction to Rasch measurement model: An example using the hospital anxiety and depression scale (HADS). The International of Educational and Psychological Assessment, Pp 1-11. 46, 1-18.

Pedrajita, J.Q. (2012). Using Logistic Regression to detect Biased Test Items. The international of education and psychological assessment journal.

Philips, J.M. (2011). Role of goal orientation, ability, need for achievement, locus of control in the self-efficacy and goal-setting process. *Journal of Applied Psychology,* 82: 792-802.

Pulka, B.M., Kwentishe, R. & Ibrahim, B. (2014). An evaluation of students' attitude towards entrepreneurship education in some selected universities in worth east Nigeria. *Global Journal of Management and Business Research,* 4(8): 11-133

Shepherd, K. (2008). Higher education for sustainability: Seeking affective learning outcomes. *International Journal of Sustainability in Higher Education,* 9(1): 87-98.

Smith, R.M. (1991). The distributional properties of Rach item fit statistics. *Educational and Psychological Measurement,* 51, 541-565.

Swaminathan, H., and Roggers, H.J. (1990). Detecting differential item functioning using logistic regression procedures. *Journal of Educational Measurement,* 27: 361-370

Unal, O. (2012). High school students' attitude towards geography courses (Karabuk sample-Turkey. *World Applied Science Journal,* 17(3): 340-346.

Walker, D.A. (2000). The IEA six subject survey: An empirical study of education in twenty-one centuries. Stockholm: Almquist and Wisksell.

Xavier, E.J. (2012). Measuring job satisfaction among Kentucky head principles using the Rasch rating scale model. Unpublished Ph.D Thesis, University of Kentucky.

Zumbo, B.D. (1999). A handbook on the theory and methods of differential item functioning (DIF): Logistic regression modeling as a unitary framework for binary and likerk-like (ordinal) item response scores. Ottawa, Canada. Directorate of human resources research and evaluation. http://www.wordengine.jp/research/Pdf/IRT.reliability-and-standard-error.Pdf