

Gender Based Comparative Analysis of Item Parameters of Geography Qualifying Examinations Using Item Response Theory in Senior Secondary Schools in South West, Nigeria

Original Research Article | Volume 1 | Issue 5 | 2026 | Article Number: 100

Accepted: 05 October 2026 | Published: 10 October 2026 | ISSN: 2979-8582 (Online)



Crossref : <https://doi.org/10.67693/BJCR-SU458VD3>



Dr. Kolawole Alaba Gbenga

Department of Educational Foundation, Faculty of Education, Federal University of Oye Ekiti, Nigeria

Correspondence: Dr. Kolawole Alaba Gbenga, alliswellgbengha@gmail.com

Abstract

The study investigated the gender based comparative analysis of item parameters of geography qualifying examinations using item response theory in senior secondary schools in south west, Nigeria. The study adopted an ex-post facto research design. It examined the responses of Senior Secondary School II students who sat for the 2021 Geography qualifying examinations in Southwest Nigeria. The sample comprised 1,200 SS II students' responses to 50 multiple-choice Geography items from the Ekiti State Qualifying Examination (ESQE) and the Oyo State Qualifying Examination (OSQE), and 60 multiple-choice items from the Ogun State Qualifying Examination (OGQE). A multistage sampling procedure was employed to select the sample. Students' responses to 50 multiple-choice Geography items from the ESQE and OSQE, and 60 items from the OGQE, were obtained from the respective State Ministries of Education in the form of Optical Mark Reader (OMR) sheets. Correct responses were scored 1 while incorrect responses were scored 0. The data were analyzed using inferential statistics in line with the stated hypotheses. Item parameters for the 1-Parameter, 2-Parameter, and 3-Parameter Logistic Models were estimated using the BILOG-MG software. The hypotheses were tested using correlation analysis, t-test, Analysis of Variance (ANOVA), Levene's test, Tukey's post hoc test, and Shapiro-Wilk's test at a 0.05 level of significance. The findings showed a significant gender difference in one-parameter indices in Oyo and Ogun States, while Ekiti State differed from both. In Oyo and Ekiti, male and female examinees perceived the items as having different difficulty levels, whereas items functioned similarly across genders in Ogun. Conversely, the 2-parameter results indicated similar difficulty perceptions for both genders in Oyo and Ogun, but not in Ekiti, while vulnerability to guessing was similar across genders in Oyo and Ekiti but differed in Ogun. It was therefore recommended that item analysis be conducted by experts before final examination administration, qualified subject teachers be employed to replace those who have left service to enhance performance, and teachers be trained on item analysis and test construction to

improve item quality.

Keywords: Gender-Based Analysis, Item Response Theory, Geography Examinations, Item Parameters, Senior Secondary Schools, Nigeria

INTRODUCTION

In Nigeria, education is the bedrock upon which economic, social, and political development rests and a critical driver of national growth. Providing quality education is therefore essential for building the manpower required to advance the economy. The Nigerian educational system is structured into three tiers: primary education as the foundational level, secondary education, and tertiary education. According to the Federal Government of Nigeria (FGN) (2013), three years of junior secondary school and six years of primary education make up the nine years of formal education that make up basic education. In Nigeria, it appears that just one-third of pupils who start basic education finish senior secondary school in either an academic or technical program. This study focuses on the secondary school level with the aim of ascertaining the quality of test items used in assessing students' abilities. A test is an instrument designed to elicit a sample of behaviour in order to measure a student's ability, assess acquired knowledge, and determine actual change in behaviour. Tests should be constructed to cover all areas taught and possess appropriate difficulty levels that adequately challenge even high-ability students who have mastered the content. Furthermore, a good test must be able to discriminate between brilliant and dull students and minimize vulnerability to guessing. Examination comprises a combination of two or more tests administered to examinees at the end of a cycle, term, or academic session to assess overall knowledge gained and abilities acquired. Its purpose is to determine whether students have attained the requisite knowledge and competencies to progress to the next educational level. A qualifying examination is administered to public senior secondary school students to identify those eligible for government sponsorship and those with the aptitude to excel in the Senior School Certificate Examination (SSCE). In Nigeria, this examination serves as an effective tool for assessing SS2 students' achievement, knowledge, interest, skills, and aptitude. The persistently low performance of students in external examinations such as WAEC and NECO underscores the need to critically evaluate the SS2 qualifying examination items used to determine candidates' readiness. At the Senior Secondary School (SSS) level, some integrated subjects are separated into distinct disciplines, one of which is Geography. A sound knowledge of Geography is essential for fostering positive attitudes toward the environment and developing practical skills for its preservation.

Researchers concur that formative and summative assessments are essential for improving geography education, as assessment provides valid, reliable, and timely data to stakeholders including teachers and policymakers to enhance learning (Lane & Bourke, 2017). Assessment is central to teaching and learning, informing identification, classification, placement, and monitoring of students (Ojerinde, 2011), and enables inferences about examinees' competencies to guide decision-making. Since assessment outcomes shape teachers' perceptions of students' abilities and instructional quality, there is a need for clarity on optimal approaches for capturing learning in geography. Although studies have examined specific assessment areas, a comprehensive synthesis of empirical evidence to inform effective practice remains lacking.

For Geography items to adequately measure students' quality and ability, item analysis must be conducted. Item analysis entails the statistical evaluation of students' responses to examination items and their interrelationships. It yields constructive feedback on item quality, improves examination effectiveness, and enhances validity and reliability. Decisions on retaining or discarding items should be largely informed by

item analysis outcomes, which can also guide improvements in instructional delivery. Item analysis is examined in terms of its importance, determinants, parameters, and applications, with key parameters including internal consistency, difficulty indices, discrimination indices, and distractor analysis (Rezigalla, 2021).

Item Response Theory (IRT) comprises measurement models that explain how observed item responses relate to an underlying latent trait. IRT models employ non-linear monotonic functions to describe the relationship between an individual's level on a latent variable and the probability of a specific item response (Cappelleri, Lundy & Hays, 2014). It seeks to model the link between examinees' psychological traits and their responses, as well as between test performance and the underlying trait. Based on response patterns, IRT predicts an examinee's ability and the likelihood of correctly answering an item (Ojerinde, 2013). Item analysis is commonly conducted on previously administered or pilot tests to enhance their utility for subsequent applications. Academic performance is a key indicator of students' achievement across subject areas. In formal educational settings, achievement is typically assessed through standardized test scores, graduation rates, and classroom evaluations. Academic performance reflects the learning outcomes attained upon completion of an educational programme. In educational research, performance is viewed as observable behaviour within a defined, often experimental context, thus representing the demonstrable dimensions of achievement at a particular point in time (Olanipekun & Aina, 2014).

STATEMENT OF THE PROBLEM

The persistent poor performance of senior secondary school students in external examinations in Southwest Nigeria, especially in Geography, remains a major concern for government and education stakeholders. Reports from WAEC and NECO and Chief Examiners' reports indicate consistently high failure rates, particularly in the 2016/2017 to 2018/2019 academic sessions. To address this, state governments introduced the SS2 Qualifying Examination to determine students eligible for sponsorship to SSCE. However, mass failure in WASSCE and NECO persists. While prior studies have linked poor performance to socio-economic factors, teaching methods, and environment, little attention has been given to the psychometric quality of qualifying examination items. Although WAEC and NECO conduct item analysis, there is no evidence that Geography items in the State Qualifying Examinations (SQE) in Southwest Nigeria undergo similar validation. Variations in item quality across states may contribute to disparities in SSCE outcomes. The inadequate standardization of SS2 Geography SQE items may thus undermine their predictive validity, necessitating empirical investigation.

The steady decline in enrolment for Geography in external examinations may be linked to repeated failures and the use of unstandardized SS2 qualifying examinations that inadequately prepare students for SSCE. The poor standardization of State West Qualifying Examination (SWQE) Geography items may limit their ability to predict external examination performance, thereby discouraging enrolment. Beyond Classical Test Theory (CTT), this study seeks to analyze SS2 SWQE Geography items within the Item Response Theory (IRT) framework, which evaluates students' abilities based on item-level performance.

Research Hypotheses

The following research hypotheses were formulated to guide the study:

1. There is no significant difference between male and female students in the one parameter of Geography items of the qualifying examination constructed for students in Senior Secondary School 2 (SSS 2) in each of the states in southwest.
2. There is no significant difference in the two parameters functioning power between male and female students taking the Geography qualifying examination in Senior Secondary School 2 (SSS

- 2) in each of the states in southwest.
3. There is no significant difference between male and female students in the three parameters of the Geography items of the qualifying examination for Senior Secondary School 2 students in each of the states in the southwest.

MATERIALS AND METHODS

The study adopted an ex-post-facto research design, as the responses of SS2 students who sat for the 2021 Geography Qualifying Examination in Southwest Nigeria were examined in their natural state without manipulation. The population comprised all public senior secondary school students who sat for the 2021 Geography Qualifying Examinations in Ekiti, Oyo, and Ogun States. A sample of 1,200 students' OMR sheets was selected through a multistage sampling procedure, covering responses to 50 multiple-choice Geography items for Ekiti and Oyo States and 60 items for Ogun State. The research instruments were the adopted 2021 Geography Qualifying Examination items for the three states. Students' responses were obtained from the respective State Ministries of Education. Correct responses were scored 1 and incorrect responses 0. Data were analysed using inferential statistics.

RESULTS AND DISCUSSIONS

Item parameters (1PL, 2PL, and 3PL) were estimated using the BILOG-MG program. The hypotheses were tested at the 0.05 level of significance using t-test and Levene's Test.

Hypothesis 1: There is no significant difference between male and female students in the one parameter of Geography items of the qualifying examination constructed for students in Senior Secondary School 2 (SSS 2) in each State in the Southwest.

Table 1: t-test statistics of one-parameter indices for qualifying geography tests in each State of Southwest based on gender.

			Oyo		Ekiti		Ogun	
			EVA	EVNA	EVA	EVNA	EVA	EVNA
Levene's Test for Equality of Variances	F		17.26		0.16		4.48	
	Sig.		0.14		0.69		0.63	
t-test for Equality of Means	t		4.45	4.45	-1.02	-1.02	8.96	8.96
	df		98.00	69.07	98.00	97.89	118.00	115.27
	Significance	Two-Sided p	0.00	0.00	0.31	0.31	0.00	0.00
	Mean Difference		1.13	1.13	-0.21	-0.21	1.59	1.59
	Std. Error Difference		0.25	0.25	0.21	0.21	0.18	0.18

	95% Confidence Interval of the Difference	Lower	0.63	0.62	-0.63	-0.63	1.24	1.24
		Upper	1.63	1.64	0.20	0.20	1.95	1.95

* Equal variances assumed=EVA; Equal variances not assumed=EVNA

The result in Table 1 reveals that the mean difference is statistically significant ($t = 4.45$, $df = 98$, $p = 0.00$ (i.e $p < 0.05$)) between male and female one parameter indices for qualifying geography test in Oyo state. This implies that the test items' difficulty varies between male and female students in Oyo State. Also, a similar result is observed in Ogun State ($t = 8.96$, $df = 118$, $p = 0.00$ (i.e $p < 0.05$)), while one-parameter values in Ekiti differs from other States ($t = -1.02$, $df = 98$, $p = 0.31$ (i.e $p > 0.05$)). The result in Ekiti showed a non-significant difference in the mean score of male and female students in the qualifying geography test.

Hypotheses 2: There is no significant difference in the two parameters functioning power between male and female students taking the Geography qualifying examination in Senior Secondary School 2 (SSS 2) in each State of the Southwest

Table 2: t-test statistics of two-parameter indices for qualifying geography test in each State of Southwest based on gender.

			OYO				EKITI				OGUN			
			b-param		a-param		b-param		a-param		b-param		a-param	
			EVA	EVNA	EVA	EVNA	EVA	EVNA	EVA	EVNA	EVA	EVNA	EVA	EVNA
Levene's Test for Equality of Variances	F		0.54		1.82	4.66			1.01		5.55		0.01	
	Sig.		0.47		0.18	0.03			0.32		0.02		0.94	
t-test for Equality of Means	T		-3.39	-3.39	-1.90	-4.77	-4.77		-0.97	-0.97	-1.49	-1.49	-1.34	-1.34
	Df		98	96.46	98	98	84.95		98	97.97	118	104.16	118	117.40
	Significance	Two-Sided p	0.00	0.00	0.06	0.06	0.00	0.00	0.17	0.17	0.14	0.14	0.18	0.18
	Mean Difference		-0.97	-0.97	-0.40	-0.40	-0.91	-0.91	-0.32	-0.32	-0.29	-0.29	-0.31	-0.31
	Std. Error Difference		0.29	0.29	0.21	0.21	0.19	0.19	0.33	0.33	0.19	0.19	0.23	0.23
	95% Confidence Interval of the Difference	Lower	-1.54	-1.54	-0.82	-0.82	-1.28	-1.28	-0.97	-0.97	-0.67	-0.67	-0.76	-0.76
		Upper	-0.40	-0.40	0.02	0.02	-0.53	-0.53	0.33	0.33	0.09	0.09	0.15	0.15

* Equal variances assumed=EVA; Equal variances not assumed=EVNA

Table 2 revealed that independent t-test statistics for Oyo state remarked mean difference that is statistically significant for the difficulty indices between male and female students on the qualifying geography test ($t = -3.39$, $df = 98$, $p = 0.00$ (i.e., $p < 0.05$)) while the discriminating indices between male and female

showed a non-significant difference ($t = -1.90$, $df = 98$, $p = 0.06$ (i.e., $p > 0.05$)). Also, similar trend is observed in Ekiti state for difficulty and discriminating indices in the test items (b-parameter- $t = -4.77$, $df = 98$, $p = 0.00$ (i.e., $p < 0.05$) and a-parameter- $t = -0.97$, $df = 98$, $p = 0.17$ (i.e., $p > 0.05$)). However, the mean statistics in Ogun state showed a non-significant difference for difficulty and discriminating indices in the test items (b-parameter- $t = -1.49$, $df = 118$, $p = 0.14$ (i.e., $p < 0.05$) and a-parameter- $t = -1.34$, $df = 118$, $p = 0.18$ (i.e., $p > 0.05$)). Consequently, the null hypothesis, which stated that there is no significant difference between male and female students in the two-parameter of Geography items of the qualifying examination constructed for Senior Secondary School 2 students in each state, was not rejected. This implies that in Oyo and Ekiti states both male and female examinees see the test items as having the different difficulty level. Also, the test distinguishes between those that know the material from those that did not the same way between the male and female students. More so, in Ogun state the test functions the same way between male and female students, that is the students have similar abilities to respond to the test items equally, and discriminate well between the low and high achievers

Hypotheses 3: There is no significant difference between male and female students in the three parameters of the geography items of the qualifying examination for Senior Secondary School 2 students in each of the states in the southwest.

Table 3: t-test statistics of three-parameter indices for qualifying geography test in each State of Southwest based on gender.

States				b-param		a-param		c-param	
OYO				EVA	EVNA	EVA	EVNA	EVA	EVNA
	Levene's Test for Equality of Variances	F		1.23		0.01		15.35	
		Sig.		0.27		0.94		0.00	
	t-test for Equality of Means	t		-1.22	-1.22	-0.22	-0.22	0.38	0.38
		df		98	95.35	98	97.75	98	57.85
		Significance	Two-Sided p	0.23	0.23	0.83	0.83	0.70	0.70
		Mean Difference		-0.28	-0.28	-0.06	-0.06	0.04	0.04
		Std. Error Difference		0.23	0.23	0.28	0.28	0.11	0.11
		95% Confidence Interval of the Difference	Lower	-0.75	-0.75	-0.62	-0.62	-0.18	-0.18
			Upper	0.18	0.18	0.50	0.50	0.26	0.26
				a_param		b_param		c_param	

EKITI			EVA	EVNA	EVA	EVNA	EVA	EVNA
	Levene's Test for Equality of Variances	F	0.57		0.03		1.96	
		Sig.	0.45		0.87		0.16	
	t-test for Equality of Means	t	2.30	2.30	3.04	3.04	1.45	1.45
		df	98	97.32	98	71.87	98	96.42
		Significance	Two-Sided p	0.02	0.02	0.00	0.00	0.15
		Mean Difference		0.87	0.87	0.88	0.88	0.05
		Std. Error Difference		0.38	0.38	0.29	0.29	0.03
		95% Confidence Interval of the Difference	Lower	0.12	0.12	0.31	0.30	-0.02
			Upper	1.63	1.63	1.46	1.46	0.11
			a_param		b_param		c_param	
			EVA	EVNA	EVA	EVNA	EVA	EVNA
	Levene's Test for Equality of Variances	F	0.01		8.85		55.70	
OGUN		Sig.	0.94		0.00		<.001	
	t-test for Equality of Means	t	0.39	0.39	-1.57	-1.57	2.04	2.04
		df	118	117.77	118	111.47	118	61.74
		Significance	Two -Sided p	0.70	0.70	0.12	0.12	0.04
		Mean Difference		0.09	0.09	-0.33	-0.33	0.34
		Std. Error Difference		0.23	0.23	0.21	0.21	0.17
		95% Confidence Interval of the Difference	Lower	-0.37	-0.37	-0.74	-0.74	0.01
			Upper	0.55	0.55	0.09	0.09	0.67

* Equal variances assumed=EVA; Equal variances not assumed=EVNA

Table 3 revealed that independent t-test statistics for Oyo state remarked mean difference that is not statistically significant for the difficulty, discriminating and guessing indices between male and female students on the qualifying geography test (b-param-t = -1.22, df = 98, p = 0.23 (i.e., p > 0.05); a-param-t = -0.22, df = 98, p = 0.83 (i.e., p > 0.05) and c-param-t = -1.22, df = 98, p = 0.23 (i.e., p > 0.05)). In Ekiti state, difficulty and discriminating indices in the test items showed statistically significant difference between male and female students (b-parameter-t = 3.04, df = 98, p = 0.00 (i.e., p < 0.05) and a-parameter-t = 2.30, df = 98, p = 0.02 (i.e., p < 0.05)) while the guessing showed a non-significant difference (t = 1.45, df = 98, p = 0.15 (i.e., p > 0.05)).

The mean statistics in Ogun state showed a non-significant difference for difficulty and discriminating indices in the test items (b-parameter-t = -1.57, df = 118, p = 0.12 (i.e., p > 0.05) and a-parameter-t = 0.39, df = 118, p = 0.70 (i.e., p > 0.05)) while the guessing showed a significant difference (t = 2.04, df = 118, p = 0.04 (i.e., p < 0.05)). This implies that in Oyo and Ogun states both male and female examinees see the test items as having the same difficulty level, though differ in Ekiti state. Also, in Oyo and Ogun state, the test distinguishes between those that know the material from those that did not the same way between the male and female students, though it differs in Ekiti state. More so, vulnerability to guessing between male and female students are the same in Oyo and Ekiti states, but different in Ogun state.

CONCLUSION AND RECOMMENDATIONS

The study concluded that students' test-taking abilities across the three states are unrelated, and that poorly functioning items should be eliminated with examinations rescored using only valid items before results are published. The findings affirm that systematic item analysis enhances the quality of multiple-choice questions and overall assessment standards; therefore, it should be prioritized during item selection prior to final administration by assessment experts and Ministry of Education test developers, while government should ensure that vacancies created by transfer, retirement, or resignation of Geography teachers are filled with qualified subject specialists rather than teachers from related fields or PTA staff to improve students' performance.

References

- Cappelleri, J.C., Lundy, J.J., & Hays, R. D. (2014). Overview of classical test theory and item response theory for quantitative assessment of items in developing patient-reported outcome measures. Doi: 10.1016/j.clinthera.2014.04.006
- Federal Government of Nigeria, (2013). National Policy on Education, (6th ed.). Lagos: NERDC Press, Retrieved June, 15, 2018, <https://educateokad.files.wordpress.com/2016/22/national-policy-on-education-polie-2013-rpdf>
- Lane, R., & Bourke, T. (2017). Assessment in Geography education: A systematic review. International Research in Geographical and Environmental Education. <https://doi.org/10.1080/10382046.2017.1385348>
- Ojerinde, D. (2011). *Public examination in Nigeria: A collection of papers*. India: Melose Books and Publishing Limited.
- Ojerinde, D. (2013). *Classical test theory (CTT) VS item response theory (IRT): An evaluation of the comparability of item analysis results. Joint Admissions and Matriculation Board (JAMB), Abuja, Nigeria*. Lecture Presentation at the Institute of Education, University of Ibadan.
- Olanipekun, S. S., & Aina, J. K. (2014). Improving students' academic performance in Nigerian schools:

The role of teachers. *International Journal of Research in Humanities and Social Studies*, 1(2), 1-6.

Rezigalla, A.A. (2021) : Item Analysis: Concept and Application DOI:
<http://dx.doi.org/10.5772/intechopen.100138>



© 2026 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY 4.0) licence (<https://creativecommons.org/licenses/by/4.0/>)