

EVALUATION OF THE PSYCHOMETRIC PROPERTIES OF EDO STATE SENIOR SECONDARY MULTIPLE-CHOICE QUESTIONS (MCQS) ECONOMICS UNIFIED PROMOTION EXAMINATION USING ITEM RESPONSE THEORY

Dr. Osemeikhian Edowaye

Dept of Guidance & Counselling

Faculty of Education

Ambrose Alli University, Ekpoma, Edo State.

osemeikhianedowaye@aauekpoma.edu.ng

+2347038884972

Abstract

The study evaluated the psychometric properties of Edo State senior secondary MCQs Economics unified promotion examination using item response theory. The study was guided by three research questions/objectives. The study adopted the ex-post facto design. The study population was 7376 SSII Economics students in Edo State while sample was 4056 SSII students who sat for July 2024/2025 Economics Unified Promotion Examinations (UPE) in Edo State using multistage sampling procedure. The study extracted students' responses from the marked optical mark recognition (OMR) answer sheets of the sample from the study population (Economics students). The examinations were multiple choice types. Data collected were calibrated with Multidimensional parameter logistics item response theory (3-PL IRT) using item analysis statistics built in Supplementary Item Response Theory (sirt package in R software). The result revealed that the 2024/25 UPE for SSII Economics multiple choice questions have high difficulty, discrimination and guessing indices. The study recommended that Economics teachers and other teachers alike should be encouraged to use good psychometric properties in their test items in order to improve students learning.

Keywords: Economics, Item Discrimination, Item Difficulty, Item Response Theory; Psychometric

Introduction

Tests must meet predetermined standards of quality to ensure that the scores obtained from them provide an appropriate basis for making valid and reliable decisions. Consequently, psychometricians and other experts in test development are required to systematically assess and characterise the psychometric properties of tests and their constituent items. Understanding these properties provides evidence that the data generated from a test can serve as a sound basis for interpreting examinees' performance and making informed decisions. However, examiners and test takers are generally more concerned with the results obtained from a test administration than with the technical characteristics underlying the test and its constituent items. Psychometric properties refer to the statistical characteristics that determine the effectiveness, accuracy, consistency, and appropriateness of a test in measuring the construct for which it was designed.

Similarly, the psychometric properties of test items are the measurable characteristics of individual items that provide a basis for evaluating their quality and functioning (Ohiri, 2023). These properties can be assessed using several indicators, including item difficulty, item discrimination, item mean, guessing parameter, and distractor indices. However, the present study focuses specifically on item difficulty, item discrimination, and the guessing parameter.

Examinations are used in education as a means of assessment to gauge students' mastery of subject matter and growth in their skills. Of the examinations given in Nigeria, the internal school examination which prepares students for external examinations are particularly significant because students performance on it will have a significant impact on their future prospects on examinations such as the WAEC, NECO and career development (Ndung'u, 2015). Therefore, psychometric properties of test items are important as the examination itself. This is because the psychometric properties of test items are crucial for ensuring that assessments are accurate, objective, and meaningful especially for internal examination where psychometric properties of examination are not emphasized.

Concerns about the standardization of internal examination and improved academic performance of students some state governments, like Edo State, have developed the Unified Promotion Examination (UPE) program in an effort to assist and enhance public tests such as the Senior Secondary Certificate Examination (SSCE). It is anticipated that the UPE would provide consistency in student evaluations and teachers standard (teaching and assessment) throughout the state-owned secondary schools by implementing conference marking, identical topic work plans, and uniform promotion and scoring standards (Osun Education Factsheet, 2019). Nonetheless, the psychometric properties of these UPE questions provide difficulties for this commendable program. This is so because most teachers or test constructors lack the basic knowledge on item test quality like item difficulty, discrimination and guessing, thus hindering their ability to construct standardized test items.

According to Osadebe (2015), poorly written achievement tests have contributed to a portion of senior secondary school students' poor performance in the Senior School Certificate Examination (SSCE) in Economics. As a result, high-quality achievement tests in Economics are necessary to allow senior secondary economics teachers to evaluate their students' performance in the subject before external examinations. Eleje, Onah, and Abanobi (2018) noted that strong performance in senior secondary economics can benefit students in external examinations as well as in university education while studying economics, banking, finance, accounting, and other relevant disciplines. However, this strong academic performance and the dream of studying the course of their choice may be hindered by poor psychometric qualities of test items. Therefore, in order to avoid this hindrance and prepare Economics students for their future dream course, this study finds it necessary to evaluate the psychometric qualities of the Edo State Economics 2024/2025 July

Unified Promotion Examination (UPE) using item response theory. Using the 1, 2, and 3 parameter logistic models, Shogbesan (2024) investigated the IRT parameter estimations of Economics multiple-choice items and revealed that the discriminating index estimated using the 2 Parameter Logistic (PL) and 3PL models showed that, respectively, 28 and 25 items are bad items while 32 and 35 items are good items. Additionally, the difficulty index calculated using the 1, 2, and 3 PL models reveals that, although 2 items were classified as tough, the remaining 35, 33, and 23 items were classified as moderately difficult, and the remaining 23 items, 25 items, and 35 items, respectively, were simple items. Additionally, the 3PL model's results indicated that just 9 items were thought to be susceptible to guesswork, whereas 51 items were not.

The difficulty and discrimination indices of NABTEB Economics essay test items between 2013 and 2015 were determined by Okoye and Ndubuzor (2021). For the study, two research questions were raised and answered. The research methodology used for the study was a descriptive survey. The study's population included 582 Senior Secondary 3 Economics students at state government technical institutions in Anambra State in 2017–2018, as well as all of the Economics essay test questions administered by NABTEB (May/June) from 2013–2015. Using basic random sampling, a sample of 307 S.S.3 pupils was acquired for the research. The entire set of May/June 2013–2015 NABTEB Economics essay assessments served as the study instrument. By calculating difficulty indices (p) and discrimination indices (d), the collected data were analyzed. The results of the NABTEB Economics essay test indicated that seven out of twelve items in each of the years 2013 and 2014 and eight out of twelve items in the year 2015 came within acceptable ranges of difficulty indices. Additionally, the results demonstrated that the NABTEB Economics essay exam items fell within acceptable limits of discrimination indices for seven out of twelve items in each of the years 2013 through 2014 and five out of twelve items in 2015.

Psychometric qualities of multiple-choice problems linked to graphs in WASSCE Economics were investigated by Maxwell (2019). Out of the 851 Senior Secondary (SSIII) students in the three local government areas (Isi-uzo, Enugu-east, and Enugu-south) that comprise the zone that provides Economics, 150 students were chosen using a multi-stage sampling procedure. A 50-item WAEC Economics Graph Related Multiple-Choice Questionnaire (WEGRMCQ) was used as the data collecting tool. From WAEC previous questions from the years 2012 to 2017, the objective questions were chosen at random. The sampled students were given the questions with assistance from the Economics teachers at the corresponding sampling schools. Among other things, the outcome demonstrated that a higher proportion of all the pupils who were evaluated had fair item discrimination and high difficulty indices.

The purpose of this study is to evaluate the psychometric properties of Edo State Economics Unified Promotion Examination (EUPE) using Item Response Theory (IRT). Specifically, the objectives are:

- i. To determine the item difficulty indices of SS2 test items of the 2024/2025 July Economics multiple choice questions promotion examination.
- ii. To determine the item discrimination indices of SS2 test items of the 2024/2025 July Economics multiple choice questions promotion examination
- iii. To determine guessing parameters of SS2 test items of the 2024/2025 July Economics multiple choice questions promotion examination

Research Questions

The following research questions guided the study.

- i. What is the item difficulty indices of SS2 test items of the 2024/2025 July economics multiple choice questions promotion examination?
- ii. What is the item discrimination indices of SS2 test items of the 2024/2025 July economics multiple choice questions promotion examination?
- iii. What is the guessing parameters of SS2 test items of the 2024/2025 July economics multiple choice questions promotion examination?

Methodology

The study adopted the ex-post facto design. The population consisted of 7376 SSII students in Economics who sat for July 2024/2025 Unified Promotion Examinations (SUPE) in Edo State. Using multistage sampling procedure, first, Edo State was stratified into three Senatorial Districts. Secondly, from each senatorial district, 3 local government areas were selected, thus making a total of 9 Local Government Areas (LGAs) selected using simple random sampling technique. Thirdly, all the 120 secondary schools and 4056 students who sat for the July 2024/2025 Economics Unified Promotion Examinations (SUPE) in Edo State were selected to participate in the study based on the existence of SSII students in the schools. The 4056 students made up the sample for the study

Table 1: Distribution of LGAs, number of secondary schools and Economics Students in SS2 for 2024/25 session

	Oredo	Egor	Ovia N.E	Iguebe n	Esan central	Esan west	Owan East	Etsako West	Etsako Central	Total
No of School	14	12	15	8	12	15	15	21	8	120
Students	805	618	421	137	283	341	620	674	157	4056

Source: *Ministry of Education, Edo State*

The study extracted students' responses from the marked optical mark recognition (OMR) answer sheets of the sample from the study population (Economics students). The examinations were multiple choice types. The responses of the

sample population were scored in binary format by assigning 1 (one) to item that was correctly answered and 0 (zero) to an item that was not correctly answered. Data collected were calibrated and analysed with Multidimensional parameter logistics item response theory (3-PL IRT) using item analysis statistics built in Supplementary Item Response Theory (sirt package in R software).

Result

Table 1: Item Parameter Estimates for 20- SUPE Items.

<i>Items</i>	<i>Model</i>	<i>a</i>	<i>b</i>	<i>c</i>
SUPE1	L3	2.89	** 2.31	*** 0.07
SUPE2	L3	* 0.17	** 2.74	*** 0.19
SUPE3	L3	* 0.32	** 2.73	*** 0.08
SUPE4	L3	* 0.09	10.31	*** 0.06
SUPE5	L3	* 0.27	** 1.07	*** 0.09
SUPE6	L3	* 0.07	8.81	*** 0.07
SUPE7	L3	* 0.10	** 3.03	*** 0.10
SUPE8	L3	* 0.56	** 2.10	*** 0.10
SUPE9	L3	* 0.16	8.32	0.45
SUPE10	L3	* 0.08	** 2.81	*** 0.09
SUPE11	L3	* 0.01	0.08	*** 0.20
SUPE12	L3	* 0.07	** 0.68	*** 0.06
SUPE13	L3	* 0.72	** 1.84	*** 0.12
SUPE14	L3	3.78	** 1.00	*** 0.18
SUPE15	L3	* 0.07	** 1.03	*** 0.10
SUPE16	L3	* 0.16	5.42	0.38
SUPE17	L3	* 0.18	** 2.46	*** 0.11
SUPE18	L3	* 0.67	** 1.89	*** 0.04
SUPE19	L3	* 1.75	** 2.04	*** 0.19
SUPE20	L3	* 0.22	7.85	*** 0.14

NB: * - discriminating “a”, ** - difficulty “b” and *** guessing – “c”

It could be observed from table 1 that the result of item calibration fitted into 3-PL model of item response theory gave parameter estimates, that is, difficulty ‘b’, discriminating ‘a’ and guessing ‘c’ for all the items of UPE for SSII 2024/2025 Economics. It was observed that for the item discriminator ‘a’, the value for the UPE for SSII 2024/2025 Economics ranges between 0.01 and 3.78. The typical range for discrimination is between 0 and 2. For the UPE test items, 18 items were within the typical range. Items with discrimination below 0.65 are considered to have low discrimination, item between 0.65 and 1.34 are moderate, 1.35 and 1.69 are considered high while 1.70 and above are very high (Baker, 2001). Thus, 2 items in UPE were outside the typical range of 0 to 2.

Table 2: Discrimination Percentage

s/n	Discrimination level	Range	Frequency	Percentage
1	low	below 0.65	15	75
2	Moderate	0.65 to 1.34	3	15
3	High	1.35 to 1.69	0	0
4	Very high	1.70 and above	0	0
5	Out of range	Outside 0 to 2	2	10

Source: researcher's computation 2026.

The result of table 2 revealed that 18 item representing 90% are within discrimination range while 2 representing 10% are outside the range, hence the 2024/25 SUPE for SSII Economics items were within discrimination range. Although, majority of the items have poor discrimination index. A breakdown showed that 75% of the items have poor discrimination, 15% has moderate discrimination index and nil percentage for high and very high discrimination.

Table 3: Difficulty Percentage

s/n	Difficulty range	Range	Frequency	Percentage
1	Good items	-3 to +3	14	70
2	Bad items	Outside -3 and +3	6	30

Source: researcher's computation 2026.

The difficulty parameter 'b' for 2024/25 UPE for SSII Economics examination items ranges between 0.08 and 10.31. Item with lower 'b' value are considered to be easier than item with higher 'b' value. However, the typical range for difficulty is between -3 and +3. In table 3 majority of the 2024/25 SUPE for SSII Economics test items (14 items) falls within this range of difficulty while 6 items fall outside the range. This means that 70% of the 2024/25 SUPE for SSII Economics examination items are within difficulty index, while 30% are regarded as bad items or need to be raised since their values indicate that they are extremely difficult.

Table 3: Guessing Percentage

s/n	Difficulty range	Range	Frequency	Percentage
1	Acceptable items	between 0 and 0.35	18	90
2	Unacceptable items	Outside 0 to 0.35	2	10

Source: researcher's computation 2026.

For the guessing parameter 'c', for 2024/25 SUPE for SSII Economics examination ranges between 0.04 and 0.45. Typical, the range of guessing is between 0 and 0.35 is considered acceptable and if otherwise unacceptable (Barker, 2001). UPE for SSII Economics (18 items) fall within this range of guessing while 2 items fall outside the range. This reveals that 90% of the 2024/25 UPE for SSII Economics

examination falls within the guessing range, while 10% are out of the guessing range.

Discussion of Findings

The finding of this study has showed that the psychometric properties of Edo State Senior Secondary MCQs Economics Unified Promotion Examination in 2024/25 are good and acceptable. Majority of the items fall within the range of discrimination, difficulty and guessing indices. The finding of this study is in consonance with Okoye and Ndubuzor (2021) who investigated the difficulty and discrimination indices of the Economics essay test items created by NABTEB between 2013 and 2015 and revealed that NABTEB Economics essay test indicated that seven out of twelve items in each of the years 2013 and 2014 and eight out of twelve items in the year 2015 came within acceptable ranges of difficulty indices. In same vein, the finding of the study is in line with Maxwell (2019) who investigated the psychometric qualities of multiple-choice problems linked to graphs in WASSCE Economics and found among others that a greater percentage of the entire students assessed recorded high difficulty indices as well as fair item discrimination.

Conclusion

Item Response Theory (IRT) and related models are important to the practice of educational and psychological measurement because they provide a framework for considering issues and addressing technical problems. Using IRT, the study concluded that most of the items in 2024/25 UPE for SSII Economics multiple choice questions were suitable because their psychometric properties were within the acceptable and good ranges. The few items that failed to meet the acceptable level of difficulty, discrimination and guessing indices should be eliminated or reviewed before their inclusion in future examinations.

Recommendations

- i. There should be no prejudice against Unified Promotion Examinations since the Edo Unified Promotion Examinations showed appreciable level of validity/reliability in the construction of their items especially in the area of discrimination, difficulty and guessing indices in Economics in 2024/25 UPE for SSII.
- ii. Economics teachers and other teachers alike should be encouraged to use good psychometric properties in their test items in order to improve students learning
- iii. Lastly, Economics teachers and other teachers alike should analyze students' responses to test items for psychometric properties before administering them to the examinees

References

- Adekunle, F. T. & Faleye B. A. (2019). Path-analytic study of the factors influencing students' performance in senior secondary school in Osun State, Nigeria. *FUDMA Journal of Educational Foundations*, 2(2), 1-14.
- Baker, F. B. (2001). *The basic of item response theory. test calibration*. Eric clearing House on Assessment and Evaluation. University of Maryland, College Park, MD. 136- 330.
- Eleje, L. I., Onah, F. E. & Abanobi, C. C. (2018). Comparative study of classical test theory and item response theory using diagnostic quantitative economics skill test item analysis results. *European Journal of Educational and Social Sciences*, 3(1), 57-75.
- Eniayewu, J. (2015). Effect of instructional materials on teaching of economics in secondary schools in Akoko North-East Local Government Area of Ondo State. *Ikere Journal of Education*, 7, 117-120.
- Maxwell E. O.(2019). Assessment of the psychometric properties of WASSCE economics graph related multiple choice questions. *International Journal of Economics Education Research* 2(1)104-116
- Edo State Ministry of Education (2025). *Students' statistics*. Edo State secretariat , Sapele Road, Benin City, Edo State.
- Ndung'u, B. C. (2015). *Reliability, item difficulty and discrimination indices of science test item in commercial test papers and their correlation to students KCPE performance in science in Limuru Subcounty*, A research project submitted to the school of education, University of Nairobi
- Ohiri, S.C. (2023). Psychometric analysis at item level of WAEC May/June Mathematics multiple choice questions using the Item Response Theory. *Unizik Journal of Educational Research and Policy Studies* 16 (5), 66-76.
- Okoye, R.O & Ndubuzor, J.N. (2021). Empirical analysis of item difficulty and discrimination indices of National Business and Technical Examination Board (NABTEB) Economics essay test items from 2013-2015. *International Journal of Education and Social Science Research* 4(2)23-28
- Osadebe, P.U. & Agbure, B. (2019). Assessment of differential item function in social studies multiple choice questions in Basic Education Certificate Examination. *European Journal of Education Studies*, 6(8), 312-344

Osun Education Factsheet, (2018). Impact of an 8-year focus on education reform, assessing the performance of students in the West African Senior School Certificate Examination (WASSCE).

Shogbesan, Y. O. (2024). Item characteristics of national examination council's economics multiple-choice items: An item response theory exploration. *Journal of Psychometric Research*, 2(1), 8-18.