

Diagnostic Quantitative Economics Skill Test for Secondary Schools: Development and Validation Using Item Response Theory

Nkechi P.M. Esomonu
Department of Educational Foundations, Faculty of Education,
Nnamdi Azikiwe University, Awka Nigeria

Lydia I. Eleje*
Department of Educational Foundations, Faculty of Education,
Nnamdi Azikiwe University, Awka, Nigeria

Abstract
The study developed and validated a diagnostic test to measure economics quantitative skill of secondary school students. The test consisted of 50 multiple choice test items which were constructed based on economics curriculum. Six research questions guided the study. Preliminary validation was done by two experienced teachers in secondary economics and two experts in test construction. The pilot testing was conducted to ensure grammatical checking and compatibility of the test items while test try-out on 517 students yielded the data for item analysis. Analysis were attained through Bilog MG using Item Response Theory (IRT) three-parameter logistic model (3PL) to establish item difficulty, item discrimination, and the guessing value. The empirical reliability of the test was 0.87. The difficulty and discrimination level of the items indicated that 39 items be retained while 11 items be replaced or modified. The test was found to be of good quality, valid and highly reliable. The test is ready to assess examinees quantitative skill in economics in Nigeria and other countries.
Keywords: diagnostic test, quantitative economics skill, IRT (Item response theory), unidimensionality, reliability.

1. Introduction

Economics in secondary school is aimed at bringing about desirable behavioral changes in students. Such behavioral changes which are the products of the objectives of the teaching/learning situations can only be realized if the students who are taught economics understand the basic concepts. Some elements of mathematics incorporated into the economics syllabus has been posing serious problem for the students in Senior Secondary (SS) classes partly as a result of the carry over effects of the negative attitudes which they have towards mathematics (Adu, Ojelabi & Hammed, 2009). The problem posed by the introduction of mathematics in economics syllabus is evident in students' Senior School Certificate Examination (SSCE) results in economics. In 2010 specifically, the West Africa Examination Council (WAEC) Chief examiners' report indicated that an area of weakness was the inability of economics candidates to carry out simple calculations. Also in 2012, the WAEC Chief examiners' report shows that most of the candidates who did not score good marks in the economics questions lacked simple quantitative skills. The Chief examiners' suggested that candidates should be encouraged to develop their skills in quantitative economics before the examinations.

Consequently, educators and teachers need to seek a coherent instructional evaluation instrument (test) that can diagnose students' quantitative ability in economics. The essence of using tests and other evaluation instruments during the instructional process according to Alonge (2004) and Kolawole (2010) is to guide, direct and monitor students' learning and progress towards attainment of course objectives. Thus, the use of diagnostic quantitative economics skill test in teaching and learning process should be adopted to locate and identify students' quantitative weakness in the subject and proffer solutions adequately.

The quality of any test, for example diagnostic test, and the information the test generates is determined through item analysis of students' responses at any examinations. Item analysis according to Adedoyin and Mokobi (2013) is a process which examines students' responses to individual test items in order to assess the quality of those items and of the test as a whole. Classical Test Theory (CTT) and Item Response Theory (IRT) are the two approaches to test item analysis. One limitation with CTT approach to determine test quality through item analysis, is that students with the same number of correct answers may have different response patterns, that is, correct answers on different items and, thus, may not have the same level of proficiency (ability) measured by the test. Also the indexes of item difficulty and item discrimination depend on the group of examinees being tested and, therefore, do not adequately reflect the measurement quality of the test items. The limitations of CTT approach to determine test quality through item analysis are successfully addressed in the framework of item response theory (IRT).

With regards to the use of IRT in test development process, Ani (2014) and Ceniza and Cereno (2012) developed and validated test items using IRT. Ani used IRT in the development and validation of multiple choice test in economics while Ceniza and Cereno developed and validated mathematics diagnostic test for DORSHS second year high school students using IRT. It is obvious that educational researchers have paid little attention to the use of IRT in test development, also little attention have been paid to the use of IRT in development of diagnostic tests in economics. The only known diagnostic test in economics to the best of our knowledge is that of Eleje et.al. (2016). Eleje et.al, developed and validated diagnostic test that was based on selected content areas in secondary school economics using CTT (traditional approach to item analysis).

Thus, there is lack of a diagnostic test on specific skills in economics. In essence, the development and validation of skill based diagnostic economics test has not been researched on, thereby constituting an educational gap. The researchers hence deems it necessary to develop a diagnostic test on quantitative economics skill using item response theory (IRT). The paucity of diagnostic quantitative economics skill test to identify students' specific strengths and weaknesses in quantitative skill of economics necessitates this study.

The objective of this study therefore, is the development and validation of a diagnostic quantitative economics skill test (DQEST) for secondary schools using IRT. Based on the objective stated above, the study sought answers to the following research questions:

1. Is the test unidimensional?
2. Which of the IRT model represents a better fit for the DQEST data?
3. What are the item threshold levels of DQEST?
4. What are the item slope values of DQEST?
5. What are the guessing values of DQEST items?
6. How reliable is the entire test according to IRT model?

2.0 Literature Review

The literature was reviewed based on the following sub-headings.

2.1 Item Response Theory (IRT)

Item response theory (IRT) is a collection of measurement models that attempt to explain the connection between observed item responses on a scale and an underlying construct (Thorpe & Favia, 2012). Specifically, IRT models are mathematical equations describing the association between subjects' levels on a latent variable and the probability of a particular response to an item, using a non-linear monotonic function (Hays, Bjorner, Revicki, Spritzer & Cella, 2009). As in classical test theory, IRT requires that each item should be distinct from the others, yet should be similar and consistent with them in reflecting all important respects of the underlying attribute or construct. IRT makes it possible to scale test items for difficulty, to design parallel forms of tests, and to provide for adaptive computerized testing. The test item analysis of any examination is based on item parameters (discrimination, item difficulty and the guessing). Item parameters in IRT are estimated directly using logistic models instead of proportions (difficulty or threshold) and item-scale correlations (discrimination) (Adedoyin & Mokobi, 2013). There are a number of IRT models depending on the number of parameters (one, two and three-parameter models) and whether they handle dichotomous only or polytomous items more generally. The study made use of three parameter model.

2.2 The a parameter (Item discrimination)

Item discrimination (represented by *a* parameter) indicates how well an item can separates among respondents with abilities below (to the left of) the item location from those with abilities above (to the right of) the item location (Thorpe & Favia, 2012). The parameter is a measure that can be graphically expressed by the steepness of the item characteristic curve (ICC). A positive discrimination of a test item means that higher ability students have a high probability of answering an item correctly and lower ability students have a low probability of answering the item correctly. While a test item with negative discrimination implies that high ability students' have a low probability of answering an item correctly and low ability students have a higher probability of answering an item correctly. High discrimination level indicates that the item discriminates well between low and high skilled individuals.

Table 1: Interpretation of Values for Discrimination (a); from Baker (2001, p. 34)

0.01 – 0.34	Low
0.35 – 1.34	Moderate
1.35 – 2.00	High
2.01 and above	Very high

2.3 The b parameter (Item difficulty)

The difficulty of an item, known as the *b* parameter, is the point where the S-shaped curve has the steepest slope. An easy item functions among the low-ability examinees; a difficult item functions among the high-ability examinees (Baker, 2001, Thorpe & Favia, 2012). Thus difficulty is a location index along the *x*-axis, i.e. how far to the right or left the curve is displaced. The index of an item's location is the point on the *x*-axis at which the curve crosses the 0.5 probability value on the *y*- axis. The values of *b* greater than 1 indicate a very difficult item and items with low *b* values below -1 indicate easy items. When the values of *b* are between -0.5 and 0.5, then the test items with such difficulty indexes have medium difficulty level.

Table 2: Interpretation of Values for Difficulty (b); from Baker (2001, p. 34)

Less than -2	Very Easy
-0.50 to -2.00	Easy
-0.49 to 0.49	Average
0.50 to 2.00	Difficult
Greater than 2.00	Very Difficult

Some IRT models include a pseudo-guessing parameter, the *c* parameter which expresses the likelihood that an examinee with very low ability can be able to guess the correct response to an item and therefore has a greater-than-zero probability of answering correctly (Thorpe & Favia, 2012). The item guessing parameter c, is the lowest value that an ICC curve attains. For example, an examinee who randomly selects responses to items that have four response choices can answer these items correctly about 1 out of 4 times, meaning that the probability of guessing correctly is about 0.25.

2.4 Item response theory models

In the one-parameter logistic (1PL) model items differ only in difficulty; the slopes of the curves are equal (are held constant). Items are only described by a single parameter in terms of location or difficulty (b). The results in one-parameter models have the property of specific objectivity, meaning that the rank of the item difficulty is the same for all respondents independent of ability, and that the rank of the person ability is the same for items independently of difficulty. The equation for one parameter model is given by the following:

$$P(\theta) = \frac{1}{-1 + e^{-1(\theta - b)}}$$

Where b is the difficulty parameter and è is the ability level.

The two-parameter logistic (2PL) model assumes that the data have no guessing, but that items can vary in terms of location (b) and slope (a) (i.e., difficulty and discrimination parameters). The equation for the two-parameter model is given below:

$$P(\theta) = \frac{1}{1 + e^{-L}} = \frac{1}{1 + e^{-a(\theta - b)}}$$

Where: e is the constant, b is the difficulty parameter, and a, is the discrimination parameter. L =a (è - b) is the logistic deviate (logit) and è is an ability level.

An important feature of the two-parameter model is that the distance between an individual's trait level and an item's severity has a greater impact on the probability of endorsing highly discriminating items than on less discriminating items. In particular, more discriminating items provide more information (than do less discriminating items) and even more so when a respondent's level on the latent attributes is closer to an item's location of severity.

The three parameter logistic (3PL) model is named so because it employs three item parameters. Such as item difficulty, discrimination and guessing parameter. The equation for the three-parameter model is:

$$P(\theta) = c + (1 - c) = \frac{1}{1 + e^{-a(\theta - b)}}$$

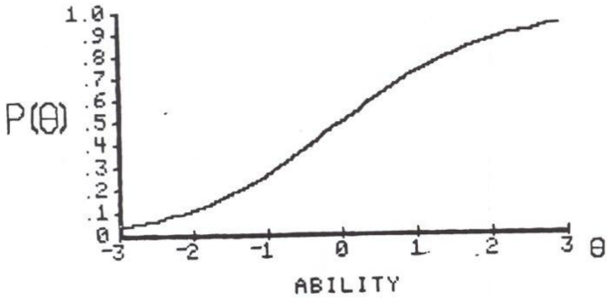
Where:
b is the difficulty parameter
a is the discrimination parameter
c is the guessing parameter and
è is the ability level

2.5 *Item characteristic curve*

The item characteristic curve (ICC) is the fundamental unit in IRT and can be understood as the probability of endorsing an item (for a dichotomous response) or responding to a particular category of an item (for a polytomous response) for individuals with a given level of the attribute. In the latter case, the ICC is sometimes referred to as a category response curve. Depending on the IRT model used, these curves indicate which items (or questions) are harder or more difficult and which items are better discriminators of the attribute. The various IRT models, which are variations of logistic (i.e., non-linear) models, are simply different mathematical functions for describing ICCs as the relationship of a person's level on the attribute and an item's characteristics (e.g., difficulty, discrimination) with the probability of a specific response on that item measuring the same attribute (Cappelleri, Lundy & Hays, 2014).

The parameter logistic equations when graphed produce plots that are called item characteristic curves (ICCs) (Fig. 1). When ICCs are plotted the ability of the examinee is denoted by theta (θ) on the *x*-axis, while the probability of an examinee correctly answering the question is denoted by P(θ) on the y-axis. ICCs typically take the shape of an S – shaped curve called o give (θ).

Fig-1. Example of Item Characteristics Curve (ICC)



The probability of the correct response is closer to zero at the lowest levels of the trait and it increases to the highest levels of the traits where the probability of correct response approaches 1(Adedoyin & Mokobi, 2013). To describe the ICC, two technical properties are used, the values of item difficulty and item discrimination.

2.6 Standard error of measurement

To Chatterji (2003), standard error of measurement is a statistical estimate of the amount of random error in the assessment of results or scores. In a students' test score, there is the possibility that the observed score might be less or more than the true score. That is, students' true score = student observed score + student error score. The scores obtained in an educational measurement are only estimates and may be considerably different from individuals' true scores. Thus, Meredith, Joyce and Walter (2007) pointed out that standard error of measurement allows one to determine the probable range within which the individual's true score fall.

2.7 Unidimensionality assessment

It is to check for the presence of a dominant component or factor that influences test performance. This dominant component or factor is referred to as the ability measured by the test (Hambleton, Swaminathan & Rogers, 1991). Any violation of this assumption would result in inadequacy of the model in describing the data and hence unreliable estimation of the examinee's ability. Therefore, the correct specification of the number of the latent dimensions is directly tied to the construct validity of the test (Rijn, Sinharay, Haberman & Johnson, 2016).

2.8 Analysis of fit

The analysis of statistical fit is a check on internal validity (Obinne, 2013). Within the latent trait test model, the internal validity of a test is assessed in terms of the statistical fit of each item to the model. According to Korashy (1995), if the fit statistic of an item is acceptable, then the item is valid. The IRT has three models: one-parameter, two-parameter and three parameter models. If a given set of items fits the model, this is the evidence that the items refer to uni-dimensional ability. Also, fit to the model, also, implies that item discriminations are uniform and substantial, and that there are no errors in item scoring. In BILOG, -2Log likelihood value is commonly used to check the goodness of model fit. Comparing the values from different models can indicate which model represents a better fit. However, the smallest -2Log likelihood value is the best. (Thorpe & Favia, 2012).

2.9 Diagnostic Test

In education, diagnostic test is an assessment tool that provides information about the strengths and weaknesses of the students in a given subject (Fushs, Fushs, Hosp, & Hamlett, 2003). Diagnostic test are scored using true test score criteria, this means that they are not averaged or normed (Dixon, 2009). Izard (2005) pointed out that scores in a criterion referenced test are interpreted as an individual performance of each student in the group on what he can do or not do rather than comparing the results with other groups of students.

In reference to diagnostic tests use, different studies have found that the effective use of diagnostic tests improves teaching and learning, and students' performance (Betts, Hahn & Zau, 2011; Kato, 2009; Patel, 2012; Richards, 2008). Thus, the development and validation of diagnostic quantitative economics skill test will make such test available to teachers in secondary schools, which if used, will improve teaching and learning, as well as students' achievement in economics. The paucity of diagnostic quantitative economics skill test to identify students' specific strengths and weaknesses in quantitative/calculative skill in economics necessitates this research.

2.10 Quantitative Economics Skills

A quantitative skill is any skill that involves using or manipulating numbers. It is the ability to reason using numbers (Riley, 2015). The quantitative skills are essential in these areas- statistics, economics, and algebra, but quantitative tools pop up in every discipline.

The quantitative skills in economics as listed by Riley (2015) include the ability of economics students to

A) Calculate, use and understand ratios and fractions
B) Calculate, use and understand percentages and percentage changes
C) Understand and use the terms mean, median and relevant quartiles
D) Construct and interpret a range of standard graphical forms
E) Calculate and interpret index numbers
F) Calculate cost, revenue and profit (marginal, average, totals)
G) Make calculations to convert from money to real terms
H) Make calculations of elasticity and interpret the result
I) Interpret, apply and analyze information in written, graphical and numerical forms.

3.0 The Test Development Process

The development of diagnostic tests involves these stages/processes: (1) Planning the test (2) Constructing the test items (3)Initial validation of the test (4) Pilot testing (5) Trial testing (6) Item analysis and (7) Assembling of final test (Alderson, 2005; Ceniza & Cereno, 2012). The detail explanation of each of the stages/processes are given below.

3.1 *Planning the Test:* Planning of the test entails the description of the population, sample and sample technique, and the content area of test. The study population is 917 senior secondary three (SS3) economics students in the 49 public secondary schools in Nigeria for the 2016/2017 academic year. The test try-out sample is 517 students randomly selected from 14 Nigerian' schools in different education zones. The selection of schools and students were done through disproportionate stratified random sampling and through proportionate stratified random sampling respectively. The content area of DQEST focused specifically on quantitative/calculative sub-skills in economics as given by Riley (2015). Table 1 shows the distribution of the quantitative skills in the test.

Table 3: Distribution of the quantitative sub-skills in economics

Quantitative sub-skills in Economics	Number of items
A- Calculate, use and understand ratios and fractions	21, 23, 24, 27
B- Calculate, use and understand percentages and percentage changes	3, 9,10, 28, 29
C- Understand and use the terms mean, median and relevant quartiles	4, 5, 6, 7, 8, 39, 40, and 41
D- Construct and interpret a range of standard graphical forms	38, 43, 44, 45, 46, and 47
E- Calculate and interpret index numbers	12, 13, 14, and 15
F- Calculate cost, revenue and profit (marginal, average, totals)	16, 17, 18, 19, and 20
G- Make calculations to convert from money to real terms	2, 34, and 35
H- Make calculations of elasticity and interpret the result	30, 31, 42, 48, 49, and 50
I- Interpret, apply and analyze information in written, graphical and numerical forms.	1, 11, 22, 25, 26, 32, 33, 36, and 37

3.2 Constructing the Test: The researchers constructed DQEST items in the form of objective test with a multiple choice format because of its versatility in content coverage (Winarni, 2002). The 50 items constructed were in line with the economics quantitative sub-skills as stipulated by Riley (2015) (see Table 1) and guidelines given by Alderson (2005) and Winarni, (2002).

3.3 Initial Validation of the Test: The 50 items constructed on its initial stage were given to two experts in educational measurement and evaluation, and two experienced O-level economics teachers for face and content validation. Their expert observations, comments and suggestions were included in the modifications of the test.

3.4 Pilot Testing: The 50 items in its preliminary stage were administered to few SS3 economics students to ensure grammatical checking and compatibility of the test items.

3.5 The Test Try-out: The researchers, with the help of the economics teachers of the participating schools, administered the DQEST test items to the students. The test try-out occurred when subject teachers and the students have completed the teaching and learning of quantitative economics sub-skills. The test try-out was for the purpose of item analysis.

3.6 Item Analysis: Item analysis was conducted on the test items mainly to ensure the quality of the items. The fifty (50) multiple choice items of DQEST were assessed first to determine the dominance of the first factor (unidimensionality), secondly for model fit and thirdly items were subjected to IRT psychometric analysis to examine the quality of the test items, in terms of item difficulty, item discrimination and guessing parameter estimates using the examinees' responses.
3.6.1 Test for Uni-dimensionality- The method used to assess unidimensionality in this study was exploratory factor analysis. It was performed to determine whether or not a dormant factor (ability measured by the test) existed among all items of the DQEST. This factor would represent the construct underlining the quantitative skills measured by the examination.
3.6.2 Test of model fit- The validity of the IRT model is dependent upon the extent to which the given responses reflect this model. Because of the dichotomous response format of the test items in this study, either the one-parameter logistic IRT (Rasch) model, the two parameter logistic (2PL) IRT model or the three parameter logistic (3PL) IRT model may be appropriate for the data. Goodness of fit statistics can be used to test for the amount of improvement in model fit to the data. In Bilog MG, -2Log likelihood value is commonly used to check the goodness of model fit, where higher values indicate a poorer fit of the data to the model. Comparing the values from different models can indicate which model represents a better fit (Rijn, Sinharay, Haberman & Johnson, 2016; Thorpe & Favia, 2012).
3.6.3 The fifty (50) items of DQEST were subjected to IRT psychometric analysis using Bilog MG software to examine the quality of the test items, in terms of item difficulty (threshold=b), item discrimination (slope=a) and guessing (asymptote=c) parameter estimates using the examinees' responses. An item is considered good for selection if it had b-value within -2 to +2, a-value within 0.35 to 1.69 and c value of 0.01 to 0.25.

3.7 Assembly/Selection of Good Items: The selection of items for inclusion in the final output of the test was determined through the verbal interpretation of each item. This is shown in Table 4 below.

Table 4. Decision table of difficulty and discrimination indices

Difficulty Level	Discrimination Level	Decision
Easy	Low	Revise
	Moderate	Retain
	High	Retain
Average	Low	Revise
	Moderate	Retain
	High	Retain
Difficult	Low	Revise
	Moderate	Retain
	High	Retain

To preserve the validity of the entire test and to maintain the total number of items in order not to omit any learning point, all items that were rejected was revised or replaced (Ceniza & Cereno 2012; Izard 2005).

3.8 Test Reliability

Reliability refers to the consistency of a measuring instrument. If the same result is approximately gotten repeatedly, the test should be considered reliable (Cherry, 2009). In IRT, there is reliability. For the set of parameters associated with each item in IRT model, Bilog MG computed a separation reliability index. This reliability was an index of the equality of the parameters. In the case of dichotomous data like the test conducted, the empirical reliability given by Bilog MG is equal to KR-20 (Ceniza & Cereno 2012). A reliability index within the range of 0.81to 1.0 indicates high reliability; 0.61 to 0.80 shows a moderate reliability; 0.41 to 0.60 means fair reliability; 0.10 to 0.40 means slight reliability and less than 0.10 means virtually no reliability.

4.0 Results

4.1 Research question 1

Is the test unidimensional?

Table 5: Total Variance Explained by the result of factor analysis

	Initial Eigenvalues		
Component	Total	% of Variance	Cumulative %
1	6.623	13.246	13.246
2	1.976	3.951	17.197
3	1.759	3.518	20.716
4	1.679	3.359	24.075
5	1.615	3.230	27.304
6	1.589	3.178	30.482
7	1.511	3.023	33.505
8	1.420	2.840	36.345
9	1.372	2.743	39.088
10	1.289	2.579	41.667
11	1.270	2.541	44.208
12	1.213	2.426	46.634
13	1.154	2.308	48.942
14	1.118	2.236	51.178
15	1.082	2.163	53.341
16	1.055	2.111	55.452
17	1.029	2.059	57.511
18	1.007	2.014	59.525
19	.966	1.932	61.457
20	.926	1.852	63.309
21	.911	1.823	65.131
22	.888	1.775	66.907
23	.887	1.773	68.680
24	.857	1.715	70.395
25	.821	1.642	72.037
26	.796	1.593	73.630
27	.786	1.571	75.201
28	.756	1.511	76.712
29	.729	1.458	78.170
30	.713	1.425	79.595
31	.690	1.380	80.975
32	.670	1.340	82.315
33	.659	1.319	83.634
34	.622	1.244	84.877
35	.589	1.177	86.055
36	.584	1.169	87.224
37	.563	1.126	88.349
38	.555	1.110	89.459
39	.538	1.076	90.535
40	.515	1.029	91.565
41	.496	.991	92.556
42	.490	.981	93.537
43	.476	.951	94.488
44	.463	.926	95.413
45	.446	.891	96.305
46	.409	.818	97.123
47	.406	.812	97.935
48	.360	.720	98.655
49	.344	.689	99.344
50	.328	.656	100.000

Extraction Method: Principal Component Analysis.

Table 5 shows that diagnostic quantitative economics skill test (DQEST) had eighteen eigen values greater than one. The first eigen value was 6.623 greater than the next seventeen eigen values (1.976, 1.759, 1.679, 1.615, 1.589, 1.511, 1.420, 1.372, 1.289, 1.270, 1.213, 1.154, 1.118, 1.082, 1.055, 1.029 and 1.007). The rest of the variance was explained by the other 49 factors.

4.2 Research question 2

Which of the IRT model represents a better fit for the DQEST data?

Table 6: Model fit information

-2Log likelihood$_{1\text{-}PL}$ = 30901.938
-2Log likelihood$_{2\text{-}PL}$ = 30332.641
-2Log likelihood$_{2\text{-}PL}$ =30314.198

4.3 Research questions 3, 4 and 5.

What are the item threshold, slope and guessing values/levels of DQEST items based on three parameter logistic (3PL) model?

Table 7: Item threshold (difficulty estimates), slope (discrimination), and asymptote (guessing) values of DQEST items.

Item	Threshold (Difficulty)	Difficulty Level	Slope (Discrimination)	Discrimination Level	Asymptote (Guessing)	Remark
1	-0.76	Easy	0.78	Moderate	0.28	Retain
2	-0.83	Easy	0.58	Moderate	0.20	Retain
3	-0.51	Easy	0.95	Moderate	0.16	Retain
4	-1.19	Easy	0.85	Moderate	0.18	Retain
5	-0.82	Easy	1.04	Moderate	0.17	Retain
6	2.05	Very Difficult	1.59	High	0.26	Reject
7	1.94	Difficult	1.21	Moderate	0.16	Retain
8	).19	Average	0.81	Moderate	0.19	Retain
9	).01	Average	0.71	Moderate	0.15	Retain
10	).70	Easy	0.93	Moderate	0.22	Retain
11	0.55	Difficult	1.28	Moderate	0.30	Retain
12	0.42	Average	1.43	High	0.31	Retain
13	1.41	Difficult	0.71	Moderate	0.19	Retain
14	0.26	Average	0.88	Moderate	0.23	Retain
15	2.35	Very Difficult	1.27	Moderate	0.17	Reject
16	0.73	Difficult	1.20	Moderate	0.11	Retain
17	0.13	Average	1.40	High	0.14	Retain
18	0.78	Difficult	1.94	High	0.18	Retain
19	1.40	Difficult	0.64	Moderate	0.17	Retain
20	0.36	Average	0.69	Moderate	0.18	Retain
21	0.26	Average	1.11	Moderate	0.18	Retain
22	0.46	Average	0.81	Moderate	0.13	Retain
23	2.14	Very Difficult	0.95	Moderate	0.28	Reject
24	2.46	Very Difficult	1.60	High	0.18	Reject
25	0.35	Average	1.28	Moderate	0.12	Retain
26	0.51	Difficult	0.73	Moderate	0.13	Retain
27	0.88	Difficult	1.06	Moderate	0.26	Retain
28	0.59	Difficult	1.20	Moderate	0.18	Retain
29	1.17	Difficult	1.10	Moderate	0.24	Retain
30	0.86	Difficult	1.63	High	0.21	Retain
31	0.69	Difficult	1.45	High	0.31	Retain
32	2.06	Very Difficult	1.08	Moderate	0.24	Reject
33	2.61	Very Difficult	0.77	Moderate	0.33	Reject
34	1.57	Difficult	1.36	Moderate	0.19	Retain
35	1.17	Difficult	2.22	High	0.27	Reject
36	2.63	Very Difficult	0.54	Moderate	0.16	Reject
37	1.05	Difficult	1.36	High	0.20	Retain
38	0.57	Difficult	0.74	Moderate	0.16	Retain
39	2.51	Very Difficult	0.20	Low	0.17	Reject
40	2.66	Very Difficult	1.29	Moderate	0.26	Reject
41	1.20	Difficult	1.21	Moderate	0.16	Retain
42	0.87	Difficult	0.60	Moderate	0.21	Retain
43	0.43	Average	0.56	Moderate	0.13	Retain
44	1.11	Difficult	0.68	Moderate	0.29	Retain
45	0.80	Difficult	0.80	Moderate	0.20	Retain
46	0.77	Difficult	1.60	High	0.24	Retain
47	1.73	Difficult	1.15	Moderate	0.31	Retain
48	3.80	Very Difficult	0.56	Moderate	0.30	Reject
49	1.62	Difficult	0.68	Moderate	0.15	Retain
50	0.86	Difficult	0.83	Moderate	0.12	Retain

As seen in Table 6. The item threshold (difficulty) and item slope (discrimination) values indicated that thirty nine (39) items be retained while eleven (11) items (6, 15, 23, 24, 32, 33, 35, 36, 39, 40, 48) be rejected. The rejected items were either revised or replaced. The guessing value were within the range of 0.00 to 0.33.

4.4 Research question 6.
How reliable is the entire test according to IRT model?

Table 8: Summary statistics from item analysis results

N	517
Mean	-0.0042
Standard Deviation	0.9350
Variance	0.8742
Empirical Reliability	0.8659

The empirical reliability index as given by IRT model (Bilog MG) is 0.87. This is the Kuder-Richardson formula 20 (KR-20) reliability index (Wu et. al., 2007).

5.0 Discussion of Findings
The discussion of findings of this study was done based on the following sub-headings: (1) dimensionality assessment of DQEST, (2) analysis of DQEST model fit, (3) item threshold values (difficulty levels of DQEST), (4) item slope values (discrimination levels) of DQEST, guessing values of DQEST items, and reliability of DQEST.

5.1 Dimensionality Assessment of DQEST
Table 5 shows that the exploratory factor analysis yielded eighteen eigen values greater than one. The first eigen value was 6.623 greater than the next seventeen eigen values (1.976, 1.759, 1.679, 1.615, 1.589, 1.511, 1.420, 1.372, 1.289, 1.270, 1.213, 1.154, 1.118, 1.082, 1.055, 1.029 and 1.007). The first factor explained 13.246% of the variance in the data set. The rest of the variance was explained by the other 49 factors with 6 factors each having a percentage of variance between 3 and 4, 11 factors each having a percentage of variance between 2 and 3, 22 factors each having a percentage of variance between 1 and 2, and 10 factors each having a percentage of variance between 0 and 1. This implies that there exists a dominant component or factor referred to as the ability measured by the test (Hambleton, Swaminathan & Rogers, 1991). Thus, diagnostic quantitative economics skill tests developed had quantitative skill as a dominant factor.

The exploratory factor analysis done on the 50 items of the diagnostic quantitative economics skill test (DQEST) was to check for the presence of a dominant component or factor that influences test performance. This was necessary because any violation of dimensionality assumption would result in inadequacy of the model in describing the data and hence unreliable estimation of the examinee's ability and the correct specification of the number of the latent dimensions is directly tied to the construct validity of the test (Rijn, Sinharay, Haberman & Johnson, 2016).

5.2 Analysis of DQEST Model Fit
In BILOG, -2Log likelihood value is commonly used to check the goodness of model fit (Klose, 2014). According to Thorpe and Favia (2012), the smallest -2Log likelihood value indicates that the model represents a better fit or is the best. Also, if the fit statistic of an item is acceptable, then the item is valid (Korashy, 1995). As observed in Table 6, the model fit information indicates that -2log likelihood for one parameter (1PL), two parameter (1PL) and three parameter (1PL) logistic models were 30901.938, 30332.641, and 30314.198 respectively.

Therefore, the 3PL model fits better than the 1PL and 2PL model. Thus, three parameter logistic (3PL) model was used in this study to estimate the item parameters and to generate item characteristics curves (ICC).

5.3 Item Threshold (Difficulty) Values of DQEST
In Table 7, it was revealed that items numbers 7, 11, 12, 13, 14, 16, 17, 18, 19, 20, 21, 22, 25, 26, 27, 28, 29, 30, 31, 34, 35, 37, 38, 41, 42, 43, 44, 45, 46, 47, 49, and 50, that is, thirty two (32) items or sixty six percent (66%) of the DQEST within the b-value range of -2 to +2 had positive difficult estimates while eight (8) items that is (34%) items, of these numbers 1, 2, 3, 4, 5, 8, 9, and 10 within b-value range of -2 to +2 had negative difficulty estimates. The positive values show that 32 items are difficult (hard) while 8 negative items are easy. Ten (10) items were outside the -2 to +2 range and thus were revised.

The study finding was in line with the study of Ani (2014), who pointed out that difficulty parameter or the threshold parameter value tells us how easy or how difficult an item is. The selection of items based on the b-value range of -2 to +2 corresponds with the criteria stipulated by (Baker, 2001).

5.4 Item Slope (Discrimination) Values of DQEST
As seen in Table 6, thirty nine (39) items had Moderate discriminating values, ten (10) items indicated High discriminating values while one item showed Very High discriminating values. The discriminating parameter reveals how well an item discriminate between respondents below and above the item threshold parameter, as

indicated by the slope of the item characteristics curves (Reeve & Fayers, 2005). The selection of items based on the a-value range of 0.01 - .34 as Low, 0.35 - 1.34 as moderate, 1.35 - 2.00 as High, 2.01 and above as Very high was in agreement with the criteria description of Baker (2001).

It was also observed from Table 7 that the item threshold (difficulty level) and item slope (discrimination level) values indicated that thirty nine (39) items be retained while eleven (11) items (6, 155, 23, 24, 32, 33, 35, 36, 39, 40, 48) be rejected. The rejected items were either revised or replaced and re-tested.

5.4Asymptote (Guessing) Values of DQEST Items

The findings of the study as seen in Table 7, revealed that guessing values (c-values) of the items were ranged from 0.00 to 0.33. Table 7 also indicates that items within the range of 0.00 to 0.25 were thirty seven (37) items (that is items number 2, 3, 4, 5, 7, 8, 9, 10, 13, …………….22, 24, 25, 26, 28, 29, 30, 32, 34, 36, 37, 38, 39, 41, 42, 43, 45, 46, 49 and 50) while the remaining thirteen (13) items fall within the c-value range of 0.26 to 0.33. The 37 items within the c-value range of 0.00 to 0.25 were desirable and the probability of getting the answer correctly by mere guessing was low as seen in a study by Ani (2014), and Adedoyin and Mokobi (2013). The 13 thirteen items that fall within c-value range of 0.26 to 0.33 were not very good and the probability of getting an answer correctly by mere guessing is high. A previous study by Harris (2005) came to a similar conclusion. Harris concluded that the items with 0.30 or greater c-values were regarded as not very good, rather c-values of 0.25 or lower are desirable. Similarly, Akindele (2003) pointed out that items do not have perfect c-values since examinees do not guess randomly.

5.6The Reliability of DQEST

The summary results in Table 8 shows that the empirical reliability of the DQEST was 0.8659 (approximately 0.87). This implied that there was 87% certainty of the consistency of the test items in yielding approximately same result repeatedly (Cherry, 2005). This implied that the test was very reliable. The finding was supported by Ceniza and Cereno (2012). According to Ceniza and Cereno, the reliability coefficient within the range of 0.81 to 1.0 indicated high reliability, 0.61 to 0.80 signified a moderate reliability, 0.41 to 0.60 signified fair reliability, 0.10 to 0.40 signified slight reliability, and less than 0.10 signified no reliability. Thus, DQEST reliability was high.

Since the output of this study was the first diagnostic quantitative economics skill test in Africa, to the best of the researchers' knowledge it will provide a valuable instrument in the field of economics education. However, a limitation is attributed to this study. IRT is a new concept in education measurement in Nigeria, hence obtaining the software packages for Item Response Theory (IRT) analysis is difficult, costly and rarely available.

6.0 Conclusions

These study findings indicated that the diagnostic quantitative economics skill test being developed is of good quality. In the dimensionality assessment done on the DQEST items through exploratory factor analysis, the first factor explained 13.246% of the variance in the data. This implied that DQEST had quantitative skill as dominant factor/ability being measured by the test. It was also noted that the 3PL model is the best model fit for the DQEST data, since it had the lowest -2log likelihood value. This finding according to Obinne (2013) and Korashy (1995) is evidence of the internal validity of the test. The item threshold and item slope values of the test showed that 39 (78%) of the items that met the criteria stipulated by Baker (2001) be retained while eleven (11) items be revised or replaced. Also, the (37) items that met the Baker's criteria were regarded as desirable, that is, the probability of getting the answer correctly by mere guessing is low while thirteen (13) probability of getting the answer correctly by mere guessing is high. As noted in the findings of this study and according to Ceniza and Cereno (2012), the empirical reliability of the overall test was 0.87. This implied that DQEST is of high reliability and there was 87% certainty of the consistency of the test items in yielding approximately same result repeatedly (Cherry, 2009).

Therefore, it could be concluded that the diagnostic test developed in this study is unidimensional, valid, highly reliable and fair. Consequently, DQEST is of good quality, validated, reliable and can now be used in assessing examinees' quantitative economics skill in secondary economics. It is an instrument that can measure quantitative economics skill of secondary school economics in Nigeria, and other countries.

7.0 Recommendations

Based on the findings and conclusions of this study, the researchers' recommend that the developed diagnostic quantitative economics skill test (DQEST) should be used by the examinees to assess their quantitative skill in economics prior to or from any internal and external economics achievement examination. Economics teachers are also encouraged to use DQEST in identifying quantitative ability of their students and to recommend as well as effect remediation promptly. School authorities should use the DQEST to assess their economics teachers' mastery of the subject as well as their coverage of syllabus. The researchers finally recommend the instrument for use by any possible researches in other African countries including other countries of the world on development of diagnostic tests.

8.0 Acknowledgement
We acknowledge Michael Akinsola Metibemu PhD for his assistance in DQEST data analysis using IRT soft-way.

References

Adedoyin, O.O. & Mokobi, T. (2013). Using IRT psychometric analysis in examining the quality of junior certificate mathematics multiple choice examination test items. *International Journal of Asian Social Science, 3(4), 992-1011.* Retrieved from http://www.aessweb.com/journal etail.php?id=5007

Adu, E.O., Ojelabi, S.A., & Hammed, A. (2009). Quantitative ability as correlates of students' academic achievement in secondary school economics. *An international multi-disciplinary journal Ethiopia, 3*(2), 322-333.

Akindele, B. P. (2003). *The development of an item bank for selection tests into Nigerian universities: an exploratory study.* (Unpublished Doctoral dissertation). University of Ibadan, Nigeria.

Alderson, J.C. (2005). *Diagnosing foreign language proficiency: The interface between learning and assessment.* London: Continuum.

Alonge, M.F. (2004). *Measurement and evaluation in education and psychology.* Ado – Ekiti: Adebayo Printing Nig. Ltd.

Ani, E.N. (2014). *Application of item response theory I the development and validation of multiple choice test in economics.* (Master's thesis). University of Nigeria, Nsukka.

Betts, J.R., Hahn, Y., & Zau, A.C. (2011). *Does diagnostic math testing improve student learning?* California: Public policy Institute (PPIC).

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

Ceniza, J.C., & Cereno, D.C. (2012). *Development of mathematic diagnostic test for DORSHS.* Retrieved from http://www.doscst.edu.ph/index.ph[/academics/graduateschool/publication/category/5-volum-1-issue-1-2012?

Chatterji, M. (2003). Designing and using tools for educational assessment. *Journals of education:* Retrieved from http://www.columbia.edu/~mb1434/EdAssess.htm

Cherry, K. (2009). Reliability – what is reliability? Retrieved from http://psychology.about.com/od/researchmethods/

Dixon, J.J. (2009). *The diagnostic prescriptive assessment.* Retrieved from http://www.eddps.com/dpa.html.

Eleje, L.I., Esomonu, N.P.M., Agu, N.N., Okoye, R.O., Obasi, E., & Onah, F.E., (2016). Development and validation of diagnostic economics test for secondary schools. *World Journal of Education, 6(3), 90.*

Fuchs, L.S., Fuchs, D., Hosp, M.K., & Hamlett, C.U. (2003). The potential for diagnostic analysis within curriculum-based measurement. *Assessment for Effective Intervention,* 28 (3&4), 13-22.

Hambleton, R.K., Swaminathan, H. & Rogers, H.J. (1991). Fundamentals of item response theory. Newbury Park, CA: Sage Publications.

Harris, D. (2005). Educational measurement issues and practice: comparison of 1-, 2-, and 3- parameter IRT models. DOI: 10.1111/j.1745-3992.1989.tb00313.x.

Hays, R. D., Bjorner, J. B., Revicki, D. A., Spritzer, K. L., Cella, D. (2009). Development of physical and mental health summary scores from the patient-reported outcomes measurement information system (PROMIS) global items. *Quality of Life Research, (18), 873–80.*

Izard, J. (2005). *Quantitative research methods in educational planning. Overview of test construction.* Paris, France: International Institute for educational planning/UNESCO.

Kato, K. (2009). *Improving efficiency of cognitive diagnosis by using diagnostic items and adaptive testing.* (Doctoral dissertation). Retrieved from http://conservancy.umn.edu/bitstream/57169/1/Kato_umn_0130E_10759.pdf

Kolawole, E.B. (2010). *Principles of test construction and administration (Revised Edition).* Lagos: Bolabay Publications.

Korashy, A.F. (1995). Applying the Rash model to the Selection of items for mental ability test. *Educational and Psychological Measurement, 55(5) 753-763.*

Meredith, D. G., Joyce, P. G., & Walter, R B., (2007). *Educational research: an introduction* (8th ed.). United State of America: Pearson Press.

Obinne, A.D.E. (2013). Test item validity: item response theory (IRT) perspective for Nigeria. *Research Journal in Organizational Psychology & Educational Studies* 2(1). Retrieved from www.emergingresource.org

Patel, R.N. (2012). Construction and try-out of science diagnostic test for the trainees of primary teachers college. *International indexed and referred research Journal,* 3(31), 18-19.

Reeve, B. B. & Fayers, P. (2005). Applying item response theory modeling for evaluating questionnaire items and scale properties. In P. Fayers and R.D. Hays (Eds.), *Assessing quantity of life in clinical trials: method of practice.* (2nd ed.). USA: Oxford university press. Retrieved from http://cancer. Unic.edu/research/faculty/display member-plone.asp?ID-694.

Richards, B.J. (2008). Formative assessment in teacher education: The development of a diagnostic language test for trainee teachers of German. *British Journal of Educational Studies,* 56(2), 184-204. Retrieved from http://dx.doi.org/10.1111/j.1467-8527.00403.x

Rijn, R.W.V., Sinharay, S., Haberman, S.J. & Johnson, M.S. (2016). *Assessment of fit of item response theory models used in large-scale educational survey assessments.* DOI: 10.1186/s40536-016-0025-3

Riley, J. (2015). *Teaching the new a level economics: mastering quantitative methods.* Retrieved from http/beta.tutor2u.net/economics/…/teaching-the-new-a-level-economics-mastering

Thorpe, G. L. & Favia, A. (2012). Data analysis using item response theory methodology: an introduction to selected programs and applications. Retrieved from http://digitalcommons.lidrary.umaine.edu/psy_facpub/20

West African Examination Council (WAEC), (2012). Chief examiners' report. Retrieved from Weaconline.org.ng/e-learning-/Economics/econs22/Nw.html

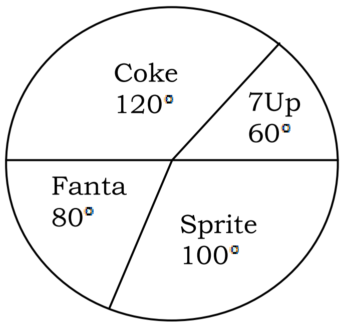
West African Examination Council (WAEC), (2010). Chief examiners' report. Retrieved from https://lifemagnanimous.files.wordpress.com/2013/06/chief-examiners-report-report-for-the-dec-2010-lshsce.pdf

Winarni, I. (2002). *The development of a diagnostic reading test of English for the students of medical faculty.* Malang: Brawijaya University.

Appendix A

Diagnostic Quantitative Economics Skill Test (DQEST)

Use the diagram below to answer questions 1-3. The total number of drinks sold by Bello in a year is 120, and this is represented by the pie chart below.



1. The quantity of coke sold by Bello is
   (a) 39.99 (b) 26.99
   (c)19.99 (d) 33.99

2. If the nominal GDP value in 2015 is N60m with a price index of 100. What is the real GDP value?
   (a)N60m (b)N30m (c)N40m (d)N50m

3. The percentage of 7up sold by Bello is
   (a) 16.67% (b) 17.67%
   (c) 18.67% (d) 16.70%

Use the following distribution to answer questions 4 to 7
2, 4, 5, 5, 4, 8, 6, 7, 6, 2, 4, 5, 6, 6, 7, 10, 12, 10, 5, 6

4. In the distribution above, the mean is (a) 6 (b) 4 (c) 5 (d) 10
5. The median is (a) 6 (b) 4 (c) 5 (d) 10

Use the table below to calculate the median quantity

Class interval	Frequency
25-29	2
30-34	3
35-39	7
40-44	10
45-49	15
50-54	2

6. The median quantity is (a) 32.5 (b) 35.2 (c) 43.25 (d) 45.25
7. The mean of the distribution above is
(a) 42 (b) 40 (c) 50 (d) 52
Use the folllowing set of distribution to answer question 8.
6, 6, 6, 8, 16, 8, 12, 22, 6, 20, 10, 16, 8, 22, 16, 16, 24, 14, 24, 10, 20, 24, 22, 22, 14, 16, 20, 14, 14, 14.
8. The frequency of 6 is (a) 6 (b) 4 (c) 3 (d) 5
9. If there are 10million people in the working class age groups (18-60 years) and 2million of them are unemployed.
The rate of unemployment is (a) 10% (b) 20% (c) 30% (d) 5%
10. If a 10 percent rate of tax is imposed on a commodity worth ₦1,000.00, the tax payable is
(a)100 (b)200 (c)10 (d)20
11. Mr. Obi's monthly salary is N50,000.00 and he pay's N250.00 as an income tax. Mr Obi's tax rate is
(a)0.5 (b)0.1 (c)0.25 (d)0.55
12. A country's export price index is N80.00 while that of import is N120.00.What is the terms of trade?
(a)67% (b)76% (c) 66.67% (d)65.76%
13. The terms of trade of the above country can be interpreted as
(a) favourable (b) unfavourable (c) balanced (d) unbalanced
14. Assuming the price of a tin of milk was 12k in 1987 but rose to 15k in 1958, then the index number will be
(a)100 (b)150 (c)125 (d)175
15. The answer in question 14 indicates that the value of money (a) fall by 25% (b) fall by 15% (c) rise by 25% (d) rise by 15%
Given that fixed cost is N500.00, variable cost is N1,500.00 and output is 50units.
16. Find the cost of producing one unit
(a)N2.00 (b)N60.00 (c)N50.00 (d)N40.00

17. Calculate for TC (a)₦500.00 (b)₦2000.00
(c)₦5000.00 (d)₦200.00

Use the table below to answer question 17 and 18

Output	Total Revenue (₦)
1	40
2	49
3	56
4	61
5	65

18. The marginal revenue when output is 4 units is
(a)₦5.00 (b)₦8.00 (c)₦56.00 (d)₦61.00
19. If total national income increases from ₦450m to ₦570m and total national consumption increases from ₦46m to ₦106m. The marginal propensity to consume (MPC) is
(a) 0.5 (b) 5 (c) 0.55 (d) 5.5
20. If a producer sells 1kg of rice for ₦2.00 and his marginal product is 100kg, what is the highest wage he can pay the marginal laborers?
(a) ₦20.00 (b) ₦30.00 (c) ₦200.00 (d) ₦100.00
Given that supply per period is a function of price and that the relationship is expressed as S= 60-1/3 p
21. Compute S when p is ₦210.00 (a) 10 (b) 20 (c) -10 (d) -20
The population of Mubi Local Government Area at the beginning of year 2013 was 6million. The births were 30,000. The deaths were 230,000. In the same year immigrants were 60,000, while emigrants were

30,000.

22. The council's population figure at the end of 2013 is
(a) 6,000,000 (b) 5,300,000 (c) 5,110,000 (d) 5,830,000.

Age Distribution Table

0-15	16-35	36-60	61 and above
10,000	3,000	5,000	2000

The above table shows the age distribution of a town in Nigeria.

23. What is the dependency ratio of the town?
(a) 10:3 (b)2:1 (c)3:2 (d)4:3

24. The dependency ratio shows that the town has ____________ dependency ratio.
(a) low (b) high (c) equal(d) constant

The demand and supply functions of a commodity are given as follows:

Qd= 20-2p

Qs = 6p-12

Where P= price in naira, Qd= Quantity demanded and QS= Quantity supplied.

25. The equilibrium price is (a) 2Naira (b)4 Naira (c)6Naira (d)8Naira

26. The equilibrium quantity is (a)20 (b)12 (c)6 (d)2

27. At what price will a buyer be ready to buy 6 oranges using the equation below: P= 1/2q + 2
(a)N3.00 (b) N4.00 (c) N5.00 (d)N6.00

Use the following demand schedule to answer questions 28-31

Price	Quantity demanded
15	60
13	70

28. Find the percentage change in demand
(a) 12.7% (b) 15.7% (c)16.7% (d)19.7%

29. Solve for percentage change in price
(a)13.3% (b)14.3% (c)15.3% (d)16.3%

30. Using the result in questions 28 and 29, derive co-efficient of price elasticity of demand(a)0.3 (b)1.3(c) -1.3 (d) -0.3

31. The co-efficient of price elasticity above can be interpreted as (a)inelastic demand (b)elastic demand
(c)zero elastic (d)unitary elastic demand

If investment is ₦60, government expenditure ₦50, consumption ₦100 + 0.7y, exports, ₦80 and imports ₦50.

32. What is the equilibrium level of national income?
(a) ₦900 (b) ₦800 (c) ₦700 (d)₦80

33. If marginal propensity to consume (MPC) is 0.60. What is the multiplier (a)2 (b)2.5 (c)5 (d)2.8

Use the table below to answer question 34 and 35.

Year	Nominal GDP value in Naira	Price Index	Real GDP value in Naira
1960	543.3	19.0	2859.5
1965	743.7	20.3	A
1970	1075.9	24.8	B
1975	1688.9	34.1	4952.8
1980	2862.5	48.3	C

34. Determine the value of A. (a) ₦3255.5m (b) ₦3663.5m (c) ₦3666.5m (d) ₦3950.5m

35. The values of B and C is (a) ₦4215.3m & ₦5328.5m (b) ₦4251.3m & ₦5988.5m (c) ₦4338.3m & ₦5926.5m (d) ₦4325.3m & ₦5996.5m

Journal of Education and Practice
ISSN 2222-1735 (Paper) ISSN 2222-288X (Online)
Vol.8, No.22, 2017

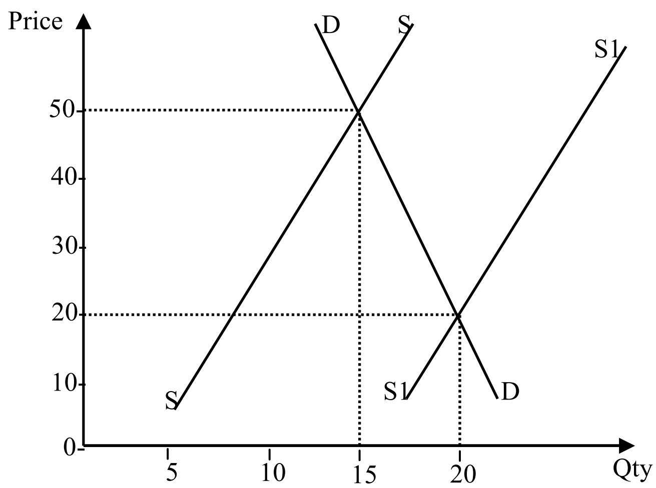
Use the table below to answer question 36 and 37

Item	₦ million
Final consumption expenditure	500
Gross private investment	150
Government purchases	300
Depreciation	25
Net factor income	-10

36. The gross domestic product GDP is equal to
(a) ₦825m (b)₦925m (c)₦940m (d)₦950m
37. The net capital formation is
(a) ₦115m (b)₦125m (c)₦915m (d)₦925m
38. The graph of the function X=a+by is
a)linear (b)quadratic (c)cubical (d)exponential
Use the data below to answer questions 39 to 41
1, 8, 2, 0, 12, 9, 13, 4, 3, 6, 7
39. The lower quartile (Q1) is (a)2 (b)3 (c)4 (d)1
40. The upper quartile (Q3) is (a)8 (b)9 (c)12 (d)7
41. The quartile range (Q2) is (a)7 (b)9 (c)2 (d)8
The price of a commodity Y increasef from 30k to 35k per unit and the quantity of another commodity X bought increased from 120 to 250.
42. What type of elasticity of demand is commodity X?
(a)Elastic demand (b)Inelastic demand
(c)Unitary elastic demand (d)Zero elastic demand
Use the market schedule below to answer questions 43 and 44.

Price per tuber of Yam	Quantity supplied
N5	600
N4	500
N3	400
N2	300
N1	250
N0.5	150

43. Thecurve of the market supply schedule above slopes
(a)upwards (b)downwards (c)bilateral (d)concave
44. The supply curve has _____________ slope
(a)positive (b)negative (c)higher (d)lower



The diagram above illustrates the demand for and supply of tomatoes. Use the information in the diagram above to answer questions 45 to 50.
45. The total revenue of the farmer at the initial equilibrium is (a)₦ 750 (b) ₦ 400 (c) ₦ 650 (d) ₦ 700
46. The total revenue of the farmer if the supply curve shifts to S_1S_1 is (a)₦ 750 (b) ₦ 400 (c)₦ 650 (d)₦ 700

47. The shift of the supply cureve from SS to S_1S_1indicates
(a) an increase in supply (b) a decrease in supply (c) an increase in price (d) a decrease in sale
48. The price elasticity of the farmer is
(a) -1.65 (b) -0.55 (c) 0.55 (d) 1.55
49. The co-efficient of price elasticity abve can be interpreted as (a)elastic demand (b) perfectly inelastic demand (c) inelastic demand (d) perfectly inelastic demand
50. If 20% rise in the price of milo leads to a 30% increase in quantity demanded of Burnvita, the cross elasticity of demand is (a) 3.0 (b) 2.5 (c) 2.3 (d) 1.5

Appendix B

Diagnostic Quantitative Economic Skill Test Answer (key) for Fifty (50) Multiple Choice Test Items

ITEM	KEY	ITEM	KEY
1	A	26	B
2	B	27	C
3	A	28	C
4	A	29	A
5	A	30	B
6	C	31	B
7	A	32	B
8	B	33	B
9	B	34	B
10	A	35	C
11	A	36	D
12	C	37	C
13	B	38	A
14	C	39	A
15	A	40	B
16	D	41	A
17	B	42	A
18	A	43	A
19	A	44	A
20	C	45	A
21	C	46	B
22	D	47	A
23	C	48	C
24	A	49	C
25	B	50	D