A comparative study of Achievement Scores of Students in Economics after Three Objective Test Formats

Abstract

This is a comparative study of achievement scores of students in economics after three objective test formats. Two research questions and three hypotheses guided the study. A 2x3 mixed ANOVA design was used to carry out the study. The sample for the study comprised of 533 Senior Secondary II (SSII) students offering Economics drawn using multi-stage sampling procedure from public secondary schools in Anambra State. Economics Achievement Test in short-answer, true-false and multiple-choice formats were utilized in data collection. The reliability indices of the instruments were 0.76, 0.78 and 0.83 respectively for short-answer, multiple-choice and true-false item formats. Kuder-Richardson Formular 21 was employed to measure the internal consistency reliability of the test scores for the three objective tests. Each student took the three objective tests. The collected data were analyzed using mean, standard deviation and two-way ANOVA. The findings of the study showed that students performed better with multiple-choice format relative to other objective test formats. While the tested hypotheses revealed that a significant difference existed between the achievement scores of male and female SSII students exposed to short-answer test and multiple-choice formats, no significant difference existed between achievement scores of male and female SSII students exposed to true/false test format in economics. It was further revealed that there was no interaction effect between objective test formats and gender of secondary school students studying economics. It was recommended that examination bodies should keep relying on multiple-choice since it leads to higher achievement scores than others.

Keywords: Achievement test, achievement scores, economics, objective test, students.

Introduction

There may be doubts expressed by examiners regarding comparability of scores obtained using the various objective test formats. Such examiners may not have empirical evidence to support their position (Esomonu et al., 2020). Objective tests are tests in which students are provided with a specific problem and given limited number of choices from which they could make choice of the correct answers or to supply answers limited to not more than three words, or at most a phrase. Objective test formats consist of true-false, short-answer and multiple choices, among others (Eleje, et al., 2017).

True-False (TF) format is a type of objective test format that allows only two options to a question. Its items have the advantage of having more content sampling and wider material coverage than any other item format ((Eleje et.al, 2016; Okoye, 2015). According to Chandratilake et.al (2011), the chance of guessing in TF item is 50%. Hence, there is strong possibility that students can score higher marks by chance in TF format than other objective test formats.

UNDER PEER REVIEW

Short-answer item format is a type of objective test format in which students are asked to provide a short answer to a question or finish an incomplete statement by filling in a blank with the correct word, number, symbol, or short phrase (Thawabieh, 2016). Short-answer item formats are deemed more appropriate than multiple choice and true-false items for measuring high inference mental skills or abilities where the examiner is expected to construct an answer (Nuri & Ceylan, 2016).

Multiple-choice format is one of the objective test formats that give room for easy scoring. It consists of a stem which provides the student with a problem and then a list of possible answers. There is little doubt that multiple-choice test format is cost-saving and capable of providing reliable and valid inferences about some kinds of conceptual knowledge (National Research Council, 2012). However, multiple-choice test items cannot be used to assess the ability of the student to organize and summarize ideas (Okoye, 2015). Other disadvantages of this format include that it takes time to construct and gives room for guessing (Eleje & Esomomu, 2018). These disadvantages are capable of influencing students' academic achievement.

Out of the three objective test formats, the one to choose seems dependable on the examiner's preference. Consequently, researchers like Lawan, Ashiru and Magaji (2017) did a comparison of medical and dental students' performance using multiple-choice questions (MCQs) and short answer questions (SAQs). The analysis of data was done through the adoption of independent sample t-test to compare the mean scores of students in SAQs and MCQs. The findings of the study pointed to the fact that the grades of students was in favour of multiple choice. Furthermore, there was the existence of a statistically significant interaction between question format and grades. Also, Pradeep, Ghildiyal, Ravindra and Sanjeev (2017) comparatively analyzed first year medical students' performance in physiology subject assessed with short-answer type questions and multiple choice questions in India. Survey research design constituted the research design for the study. Conventional type short-answer questions and multiple-choice questions in Physiology constituted the instruments for data collection scores were subjected to analysis using frequencies, percentages and Chi-square. The revelation of the findings of the study is that the performance of students, on exposure to short-answer type questions was significantly higher than that of students assessed by totally multiple-choice questions. Likewise, Javid (2014) investigated the comparison between Multiple-choice (MC) and True-false Test formats in Iranian Intermediate EFL learners' vocabulary learning using survey research design. The instruments used in this study were 50 multiple choice grammar and vocabulary items used as a pre-test and standardized multiple choice tests used as the post-test. Data collected were compared by

UNDER PEER REVIEW

using Mann Whitney-U test. The results of the statistical analysis revealed that there was a significant difference in the performance of students in favour of true-false. Additionally, researchers have equally showed that gender impacts student's academic achievement (Eleje, et al., 2020; Thawebieh, 2016). In this study, the possible effect of gender on students' Economics achievement score in various objective test formats was also investigated.

The objective test format to choose depends on the examiners' preference. Their preference or otherwise may be based on sundry reasons which include: guessing factor, coverage of syllabus, uniformity in scoring and interpretation of scores. Could it be that the achievement scores of students are linked to the type of objective test format to which they are exposed? The comparability of achievement scores of students exposed to the three formats of objective test is ascertainable by empirical investigation. Hence, the needs for this present study.

Statement of the Problem

Some examiners may not be comfortable with true-false because it could give rise to high marks as a result of the guessing factor. This, they suspect could lead to having undeserved scores that do not match the students' ability. Examiners' lack of preference for true-false could be hinged on the fact that it narrows the students' options to two.

Some others may be uncomfortable with short-answer for some other reasons ranging from issues of uniformity in scoring and interpretation of scores. There may be doubts expressed by examiners regarding comparability of the scores obtained using the three formats. Such examiners may not have empirical evidence to support their position. It therefore becomes necessary to compare scores from the three formats to see if they are comparable.

Research Objectives

Specifically, the study sought to:

1. compare the achievement scores of secondary school students exposed to all three formats in Economics.
2. compare the achievement scores of male and female secondary school student in Economics.

UNDER PEER REVIEW

3. ascertain if there is correlation between the effect of gender and achievement scores of students in objective test format in Economics.

Research questions

1. How comparable are the mean achievement scores of secondary school students exposed to short-answer, multiple-choice, and true false item formats in Economics subject?
2. How comparable are the mean achievement scores of male and female student in Economics?
3. What is the correlation between the effect of gender and achievement scores of students in objective test format in Economics?

Hypotheses

Three hypotheses were formulated to further guide the study and were tested at 0.05 level of significance.

1. There is no significant difference in the mean achievement scores of secondary school students exposed to short-answer, multiple-choice and true false item formats in Economics subject.
2. There is no significant difference between mean achievement scores of male and female students in Economics.
3. There is no significant interaction effect between gender and test format on academic achievement of students Economics.

Methodology

The study adopted a 2x3 mixed ANOVA design. The population of the study comprised 5793 (2684 males and 3109 females) SSII students offering Economics in public secondary schools in Anambra State. The choice of SSII students was underlined by the fact that this was the stage where students were given the opportunity of selecting their subjects of choice in advance preparation for Senior School Certificate Examination. The sample of the study comprised 533 Senior Secondary II (SSII) students for the 2018/2019 academic session obtained through multi-stage sampling procedure. Again, intact classes of all SS II students offering Economics in nine schools obtained during the sampling were used for the study. Three instruments were used for data collection. Each of them was

UNDER PEER REVIEW

an achievement test titled "Economics Achievement Test" covering 10 units of 2018/2019 SSII First Term Economics syllabus, and addressing exactly the same contents and problems, except for the fact that they were constructed using different test formats viz: short-answer, true-false and multiple-choice. Each of the tests contained 25 items. Copies of the three different achievement test formats titled Economics Achievement Test (EAT) (one constructed with short-answer item format, the second with true-false item format and the third with multiple-choice item format) were given along with the purpose of the study, research questions and hypotheses, to three experts, one from the Department of Educational Foundations, another from the Department of Science Education, and the third from the Department of Economics, all in Nnamdi Azikiwe University, Awka. They were requested to assess the three instruments with regard to their content validity, appropriateness of the framing of items, correctness of the keys and their relatedness to the research questions and purpose of the study. The reliabilities were determined using Kuder-Richardson Formula 21, after administering the instruments to 30 students from public secondary schools randomly selected in Imo State. In estimating the reliability coefficient, each of the items was scored dichotomously, that is for each item, a mark was awarded for a correct answer and zero for a wrong answer. Reliability indices of 0.76, 0.78 and 0.83 were respectively found for short-answer item format, multiple-choice item format and true-false item format, and so were considered high enough to confirm the instruments as reliable. Three different groups were used for the study each participant was exposed to the three test formats on three different days to avoid testing effect. In scoring the instruments, one mark was awarded for each correct response. Mean and standard deviation were used to answer the research questions. The null hypotheses were tested at 0.05 alpha level using two-way ANOVA for repeated measures. The computation was done using a Statistical Package for Social Science (SPSS), Version 23.

Results

The results from data analysis were presented based on research question and hypothesis.

Research question 1

What are the mean achievement scores of secondary school students exposed to short-answer, multiple-choice and true false item formats in Economics subject?

Table 1.

Mean and Standard Deviation Scores of Secondary School Students Exposed to Short-Answer, Multiple Choice and True/False Item Formats (N=533).

UNDER PEER REVIEW

Sources of variance	N	Mean	SD
Short answer	533	17.09	3.14
Multiple choice	533	21.34	2.80
True false	533	18.58	2.74

Results presented in Table 1 indicate the means and standard deviations of secondary school students exposed to short answer, multiple-choice and true false item formats. The mean scores of 17.09, 21.34 and 18.58 reported in the table show that students' achievement was best under multiple-choice item format and least when exposed to short answer.

Research question 2.

What are the mean achievement scores of male and female students in Economics?

Table 2.

Mean Scores of Male and Female Secondary School Students Exposed to Short-Answer, Multiple-Choice and True/False Item Formats (N=533).

Sources of variance	N	Mean	SD
Male	188	19.38	2.92
Female	345	18.80	2.87

Results presented in Table 2 indicate the mean scores of male and female secondary school students exposed to short answer, multiple choice and true/false item formats in Economics subject. The mean scores were 19.38, and 18.80 mean. The result showed that the male students had higher mean score the female students.

Hypothesis 1

There is no significant difference in the mean achievement scores of secondary school students exposed to short-answer, multiple-choice and true false item formats in Economics subject.

Table 3.

Repeated Measures Analysis of Variance on Mean Achievement Scores of Students Exposed to Short-Answer, Multiple-Choice and True-False Item Formats in Economics

UNDER PEER REVIEW

Source	Sum of Squares	Df	Mean Square	*F*	*P*	Decision
Item Formats	4552.391	1.905	2389.868	735.906	<.001	Significant
Item formats*Gender	5.584	1.905	2.931	.903	.402	Not Significant
Error	3284.821	1011.487	3.248			

PEER REVIEW

The results presented in Table 3 show that there is a significant difference in students' academic achievement as a result of the item format they were exposed to, $F(1.905,1011.487) = 735.906$, $p < 0.05$. Therefore, the null hypothesis which postulated no significant difference is rejected. Having found out that there is a significant difference, it is necessary to make pairwise comparison to found out where the difference lies, the result of pairwise comparison is shown in Table 4.

Table 4.

Pairwise Comparisons

Measure: MEASURE_1

(I) Item Formats	(J) Item Formats	Mean Difference (I-J)	Std. Error	Sig.b	95% Confidence Interval for Differenceb	
					Lower Bound	Upper Bound
1	2	-4.242*	.119	<.005	-4.528	-3.956
	3	-1.486*	.101	<.005	-1.730	-1.242
2	1	4.242*	.119	<.005	3.956	4.528
	3	2.756*	.102	<.005	2.512	3.000
3	1	1.486*	.101	<.005	1.242	1.730
	2	-2.756*	.102	<.005	-3.000	-2.512

From the table it can be seen that there is a significant difference between Multiple choice and True-false, between True-false and Short-answer and between Short-answer and Multiple-choice.

From the values of means in Table 1, it then means that the mean of Multiple-choice is significantly higher than that of True-false and that of True-false is significantly higher than Short-answer.

Hypothesis 2.

There is no significant difference between mean achievement scores of male and female students in Economics.

Table 5.

Comparison of Achievement Scores of Male and Female Students

Source	Sum of Squares	Df	Mean Square	*F*	*p*	Decision
Intercept	177405.617	1	177405.617	28260.313	.000	
Gender	41.495	1	41.495	6.610	.010	Significant
Error	3333.381	531	6.278			

The results shown in Table 5 indicate that there is a significant gender main effect on academic achievement of students in Economics, F (1,531) = 6.610, p <0.05, hence the null hypothesis was rejected. As shown in Table 2, the difference is in favour of male students.

Hypothesis 3.

There is no significant interaction effect between gender and test format on academic achievement of students in Economics.

As shown in Table 3, there is no significant interaction between gender and item format on students' academic achievement in Economics, F (1.905, 1011.487) = .903, p>0.05. The null hypothesis that postulated no significant interaction effect of gender and item format on students' academic achievement is therefore not rejected.

Discussion of Findings

The result of the study revealed that there was a significant difference in the mean achievement scores of secondary school students exposed to short answer, multiple-choice and true/false item formats. The finding of the present study agrees with the position of Javid (2014) that

UNDER PEER REVIEW

the test formats through which the participants were assessed had a significant effect on the scores they receive in the final exam. Javid further posited that there was a significant difference in the performance of students with different test formats.

The findings of this study indicated that students recorded best performance with multiple-choice format and least performance with short answer format on exposure to three objective test formats. This could be traced to the fact that students tend to be most favourably disposed to multiple-choice format relative to other objective test formats. The mean for short answer format was least probably because the test called for a more demanding task of recalling while multiple-choice called for recognition. The findings of the current study contradict the position of Pradeep, Ghildiyal, Ravindra and Sanjeev (2017) that the performance of students exposed to short answer type questions was significantly higher than those who were assessed using multiple-choice questions. This contradiction may be traced to the peculiarity of the students in the area of study. Additionally, it could be linked to the fact that secondary school students are more intellectually mature than those used by Predeep et al, and more used to multiple choice formats than true false.

The findings of this study also indicated that the mean for multiple-choice is higher than that of true-false even though the guessing factor for true-false is greater than that of multiple-choice, but then the test called for a more demanding task of guessing. The finding of the present study agrees with the position of Simbak, et al. (2014) which revealed that both academic year results pointed out that the students got higher marks in multiple-choice than true-false.

The findings of this study further indicated that male students consistently did better in academic achievement in Economics across the three item formats. This could be traced to the fact that the mean for male students across the three item formats was higher than their female counterparts. The findings of the current study agree with the position of Thawebieh (2016) that there was a significant difference in the performance of male and female students but disagree with the findings of Eleje et al., (2020).

Conclusion

Based on the findings of the study it was concluded that students perform best with multiple-choice format relative to other objective test formats, and that male students do better than their female counterpart in Economics across the three test formats.

UNDER PEER REVIEW

Consent

As per international standard or university standard, Participants' written consent has been collected and preserved by the author(s).

Recommendations

In line with the findings of the study, the following recommendations were arrived at:

1. Teachers should pay more attention to the female students in order to bridge the gap in their achievement between them and their male counterparts.
2. Test constructors should rely on the use of True-false since it does not lead to unnecessary high scores due to guessing factor, as people might think.
3. Examination bodies such as WAEC, JAMB should keep relying on Multiple-choice since it leads to higher scores than other test formats.

UNDER PEER REVIEW

REFERENCES

Chandratilake, M., Davis, M., & Ponnamperuma, G. (2011). Assessment of medical knowledge: The pros and cons of using true/false multiple choice questions. *Indian National Medical Journal, 24*(4), 225-228. Retrieved from http://www.nmji.in/archives/Volume-24/Issue-4/Medical-Education.pdf.

Eleje, L.I., Abanobi, C.C., & Obasi, E. (2017). Development and validation of economics achievement test for secondary schools. *Asian Journal of Education and Training*, 3(1), 6-17. Doi: 10.20448/journal.522/2017.3.1/522.1.6.17

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), 90-116. doi:10.5430/wje.v6n3p90.

Eleje, L.I., & Esomonu, N.P.M. (2018). Test of achievement in quantitative economics for secondary schools: Construction and validation using item response theory. *Asian Journal of Education and Training*, 4(8), 18-22. Doi: 10.20448/journal.522.2018.41.18.28

Eleje, L.I., Esomonu, N.P.M., Okoye, R.O., Agu, N.N., Ugorji, C.O., Okoi, O.A., & Abanobi, C.C. (2020). Students' academic achievement in secondary-school quantitative economics: Effect of feedback with remediation. *The Asian Institute of Research Education Quarterly Reviews*, 3(4), 479-488. DOI: 10.31014/aior.1993.03.04.155

Esomonu, N.P.M., Esomonu, M.N., & Eleje, L.I. (2020). Assessment big data in Nigeria: Identification, generation and processing in the opinion of the experts. *International Journal of Evaluation and Research in Education*, 9(2), 357-363. DOI: 10.11591/ijere.v9i2.20339

Javid, L. (2014). The comparison between multiple-choice (MC) and multiple true-false (MTF) test formats in Iranian intermediate EFL learners' vocabulary learning. *Procedia - Social and Behavioural Sciences, 98* (1), 784 – 788.

Lawan, A. H., Ashiru, H. Y., & Magaji, G. T. (2017). Differences in multiple choice and short answer questions performance among medical and dental trainees as reflected in their anatomy

UNDER PEER REVIEW

course. *Journal of Contemporary Medical Education, 1* (1), 1-8. Doi 10.5455/jeme.201707706060134.

National Research Council (NRC), (2012). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. Washington, DC: The National Academies Press.

Nuri, D., & Ceylan, G. (2016). The impact of item format on student's score and item's statistics. *International Journal of Research in Engineering and Social Sciences, 6* (11), 18-25.

Okoye, R. O. (2015). *Educational and psychological measurement and evaluation (2nd Ed.)*. Awka: Erudition Publishers.

Pradeep K., Ghildiyal J P., Ravindra, S. R., & Sanjeev K. (2017). Comparison of performance in physiology subject assessed by short essay type questions and multiple choice questions of first year medical students. *Asian Journal of Medical Sciences, 8* (1), 82–84. DOI: 10.3126/ajms.v8i1.15977.

Simbak, N., Aung, M., Ismail, S., Joush, N., Ali, T., Yassein, W., Reuben, H. (2014). Comparative study of different formats of MCQs: Multiple true-false and single best answer test formats in a new medical school of Malaysia. *International Medical Journal, 21* (6), 562-566.

Thawabieh, A. M. (2016). A comparison between two test item formats: Multiple-choice items and completion items. *British Journal of Education, 4*(8), 32-43.

UNDER PEER REVIEW