WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Effects of Mathematics Test Items Arranged in Top-bottom and by Order of Topics, on the Performance of Senior Secondary Schools Students in Rivers-South Senatorial District.

Obilor, Esezi Isaac (PhD)[1] and Miwari, Goodluck Uwaks (PhD)[2]

Department of Educational Foundations,
Faculty of Education, Rivers State University, Port Harcourt, Nigeria
Phone: 08037273533 Email: esezirukanimgboh@gmail.com/goodmiwaks@gmail.com

ABSTRACT

The study examine the effect of mathematics test items arranged in Top-bottom (easy-to-hard items) and arrangement based on order of topics presentation on examinees' performance. With quasi-experimental research design adopted, 400 final year students of 8 intact classes from different schools, consisting of 50 students each were used for the study. Using test blueprint to guarantee sampling adequacy, the researcher constructed a Mathematics Achievement Test which was used for data collection. The reliability indices of the Achievement Test established using Cohen's Kappa (K) procedure and Kuder Richardson formula 20 (KR_{20}), were respectively K = 0.70 and r = 0.73. Mean achievement differences were used to address the research questions while hypotheses were tested at the 0.05 level of significance using t-test statistic. Results indicated that test item arrangement based on top-bottom (easy-to-hard) format has significant positive effect on examinees' performance, and Examinees show outstanding performance when items are arranged on bases of earlier topics treated, compared to when items are arranged beginning with last topics presented. Consequently, it was recommended that stakeholder in the business of test construction, administration and users alike to ensure that test items are arranged in a top-bottom format where by testee first encounter the less difficult items to gain courage before proceeding to the more difficult item. Arrangement must marry with topical presentation. In other words, while considering the easiness of the item, the earlier topics presented during instruction should be considered first.

Keywords: Test, Testees' Performance, Top-Bottom, First-Topic-First, Last-Topic-First

INTRODUCTION

Measurement, as a systematic process of quantifying the characteristics of a phenomenon, can be achieved by the use of a measuring tool and for the behavioural sciences and education, this tool is often called test. According to Blair and Simpson (2013), a test is made up of items or tasks that are designed to elicit the quality or efficiency of an individual's behaviour in a situation for which he has or has not been specifically trained. Orluwene (2012), advanced that a test is an instrument that can be utilized in detecting some qualities, traits characteristics, attributes and the likes possessed by a person, an object or a thing; and can be used to quantify behaviour; it is an instrument, a device or a system that proposes a sequence of tasks that a testee/examinee is to respond and the results are used as measurement of specified trait.

A test involves a task to be performed; a testee unconsciously reveals the traits or behaviour in him by responding to the task which could be a sequence of items. It is on this note that Anastasi and Urbina (2006) describiing psychological test as an objective and standardized measure of a sample of behaviour, expounded that traditionally the function of a psychological test has been to measure differences between individuals and between the reactions of the same individual under different circumstances. According to these scholars, all tests: cognitive or non-cognitive, psychological or physiological are alike insofar as observations are made on a small but carefully chosen sample of an individual behaviour.

A test therefore is a device or any instrument by which the characteristics or behaviour of a phenomenon or an individual, whether observable or latent can be inferred or elicited. The device or instrument could constitute a set of tasks, a challenge situation to be attended to by an individual or phenomenon, and by responding to the task or challenge situation, the hidden behaviour or trait in the individual or phenomenon will be revealed. Test items are simply the individual tasks that constitute a test.

Every test consists of several independent but related tasks. Going by Ukwuije (2009) definition of test as a series of questions given to the testees or examinees to be answered in order to measure performance or knowledge, test items would constitute the series of questions given to the testee. Whether a series of questions to be answered or several independent but related tasks to be performed, they must be arranged in a particular fashion. Test item arrangement is the positioning of items of a test in a particular order or format. There are different types of test items arrangements, but in this work the Top-bottom format and arrangement by topics are considered.

Top-Bottom (or easy-to-hard) Arrangement of test items begins with easiest items through the easier ones to the difficult or hard ones. Top-Bottom format is analogous to the structure of a tree where the leaves which are light are on top, followed by the stem or middle then the base or roots which is heaviest. A test which is arranged in this format has the easiest (lightest) items first, followed by the easier (lighter) ones then the difficult (heavy) items. It could be viewed as ascending order or easy-to hard (ETH), all depicting that the simplest ones are encountered before the more difficult ones. Orluwene (2012) cited Onunkwo (2002), Hopkins and Antes (1978) and Martuza (1977) to have emphasized and recommended this arrangement. They enlisted the merits of this arrangement to include ensuring that test contains only related materials, motivates and minimizes anxieties and frustration.

Arrangement could be by topics as presented during instruction. Under this arrangement, it is either first topic first (FTF) or last topic first (LTF or first topic last-FTL). With the first topic first (FTF), items from the earlier topics introduced in the class during instruction appear first during test item arrangement, and the subsequent topics follow in that order. On the other hand, the reverse is the case where the last topics treated in the class appear first in the test item arrangement. Proponents of this format argue that the last topic treated is still fresh in the memory of testees and therefore arouses interest, reduces anxieties, less time is spent on tacking these items and encourages memory recall (Colman 2014).

Thorndike (1982), Popham (1975) & Mehren and Lehman (1992) cited in Uwah (2016) identified that one major importance of item arrangement is to curb examination malpractices. In their view, if items are arranged in such a way that adjacent students are not opportune to attempt the same item number at a given time, cheating will be frustrated. It has been argued that besides examination malpractices, item arrangement enhances academic performance.

Academic performance according to Umude (2010) is a measure of the ability of learners to recall facts and information at a given time. It is however premised that the degree of this ability to recall learned experiences is a function of several environmental factors. Academic performance of student has to do with the extent an individual perform in the course of carrying out academic activities or program. It is a varying quantity, not consistent depending on circumstances, situation and at time on the individual student. It refers to how students deals with their studies and how they cope with or accomplish different tasks given to them by their teachers.

Academic performance is important because it is strongly linked to valued positive outcomes. Howard (2009) observed that in the past, academic performance was often measured by "ear" than today. He stressed that teachers' observation formed the bulk of assessment but today summation or numerical method of determining how well a student is performing is a fairly recent invention (cited in Uwah 2016). Performance in school can be measured for regular grading by students demonstrating their knowledge by taking written or oral test, performing presentation, turning in homework and participating in class activities and discussion. This was the position of Obinna (1998) in Uwah (2016), when he noted that students are being evaluated for academic performance through home-work, assignments, tests and examinations. On one of the functions of academic performance, Ezewu (1983) decried that teachers may be motivated or demotivated by the academic performance of his students. To him, teachers find it difficult to relate with underperforming learners; students who perform better are liked by their teachers who in most cases call them to answer questions.

Statement of the Problem

In recent times there have been unpleasant reports on the performances of students both in internal and external examinations. Tei-Firstman (2011) cited Ali (2009) as saying that majority of students find it difficult to pass subjects like mathematics at first sitting while Dike (2012) decried that mathematics as a subject scares many students, some develop negative attitude towards the subject and the teacher and this has attendant negative effect on their performances in both internal and external examinations.

Experience has shown that blames concerning students' poor performances are usually levelled against the child himself, the peers, the parents, the standard of the items, students' poor test taking skills and sometimes the teacher; no mention is ever made concerning the arrangement of the test items and whether or not such arrangement has effect or influence on the students' performance. However, authorities have also maintained that such a sensitive aspect of testing cannot be overlooked while discussing factors affecting students' performance.

Appealing as these opinions and views may be, there is the need for an empirical evidence in support of these views. What positive or negative effect does a particular arrangement have on the performance of examinees and what can be done to improve the situation. This is the missing-link and it is against this backdrop, that this study seeks to examine the effect of test items arranged in Top-bottom and order of topics presentation, on examinees performance.

Purpose of the Study

The purpose of the study was to examine the effect of mathematics test items arranged in Top-bottom and order of topics presentation, on examinees performance. Specifically, the study:

1. Examine the effect of top-bottom arrangement of test items on examinees' performance in mathematics.
2. Investigate the effect of item arrangement based on order of topics presentation on examinees' performance in mathematics.

Research Questions

To guide the study, the following research questions were raised:

1. What effect does test item arrangement based on top-bottom have on examinees' performance in mathematics?
2. What is the effect of test item arrangement based on order of topic presentation on examinees' performance in mathematics?

Hypotheses

The following null hypotheses tested at 0.05 significance level guided the study:

1. There is no significant effect of top-bottom test item arrangement on examinees' performance in mathematics in senior secondary schools.
2. Test items arranged based on order of topics presentation has no significant effect on examinees' performance in mathematics in senior secondary schools.

Methodology

The study adopted the quasi-experimental research design which allows for the waiver of confounding variables that are impossible to eliminate or control; when the researcher is interested in independent variables which cannot be randomly assigned, and in such cases when experimental and control groups are such naturally assembled groups as intact classes (Kpolovie, 2010 & Kuol, 2018). Intact classes in classroom situation were used for the study and so could not permit random assignment of participants. They were partitioned into three experimental groups and one control group. Group A were assigned Top-Bottom arrangement, Group B – First Topics First (FTF), Group C - Last Topics First (LTF), and the control group D – which is the traditional or random arrangement.

The study was carried out in Rivers-South Senatorial District of Rivers State, Nigeria. A senatorial district that is naturally partitioned into riverine and upland communities and is sometimes referred to as the Orashi region as many of the communities lie along the Orashi river and its flood plains. Rivers State Post Primary Schools Board shows a record of 97 senior secondary schools in the senatorial district, from where a sample of final year students who are believed to be capable of providing the relevant data for the study were chosen for the study. The sample size for the study was 400 final year students, chosen through multistage and random sampling techniques. Initially, a school was randomly drawn from each of the 8 LGAs in the senatorial district, thereafter an intact SS3 class of 50 students was randomly drawn from each of the 8 chosen schools, giving a total of 400 students.

The instrument for data collection was a validated Mathematics Achievement Test (MAT) constructed by the researcher. The arrangement of the items was varied according to formats for various experimental groups. That is, Top-Bottom for experimental group A, while arrangement by topics was First Topic First (FTF) and Last Topic First (LTF) for experimental groups B and C respectively. The last arrangement is the random or traditional type for group D – the control group.

Cohen's Kappa inter-rater reliability procedure and Kuder-Richardson formula 20 (KR_{20}) were applied to establish the reliability of the instrument which yielded indices (K) = 0.70 and r = 0.73. Administration was by direct delivery method within the specified time limit stated in the instrument. This was done with the help of research assistants who are experienced classroom teachers.

Differences in mean achievement scores and standard deviation were used to answer research questions while t-test statistic was used to test hypotheses at 0.05 level of significance. Mean differences greater than zero were interpreted positive effect, otherwise negative. Null hypothesis was retained if observed t-value was less than its critical equivalent, otherwise rejected.

Results

Research Question 1: What effect does test item arrangement based on top-bottom format have on examinees' performance in mathematics

Table 1: Descriptive Statistics on the Effect of Top-bottom Test Items Arrangement on Examinees Performance

Group	**N**	$\overline{X}$	**SD**	$\overline{X}_A$-$\overline{X}_D$	**Remarks**
Group A	100	36.30	6.31		
				4.36	Positive Effect
Group D	100	31.94	8.67		($\overline{X}_A$- $\overline{X}_D$>0)

Table 1 shows a mean difference of 4.36 incurred by experimental group A over the control group (i.e. $\overline{X}_A$ - $\overline{X}_D$ = 4.36) and this means that experimental group A who were tested with the items arranged in top-bottom format, out-performed those of the control group who were tested with items arranged in traditional format. This suggests that items arranged in top-bottom format have positive effect on the academic performance of students in Mathematics.

Research Question 2: What is the effect of test items arrangement based on earlier topics presented, on examinees' performance in mathematics?

Table 2 (a):Descriptive Statistics on the Effect of Test Items Arrangement based on earlier Topics Presented during Instruction, on Examinees' Performance in mathematics

Group	N	$\overline{X}$	SD	$\overline{X}_B$-$\overline{X}_D$	Decision
Group B (FTF)	100	37.86	8.89		
				5.92	Positive effect
Group D (Control)	100	31.94	8.67		$\overline{X}_B$- $\overline{X}_D$>0

From table 2 (a), a comparison of the two groups (B and D), yielded a mean difference of 5.92 in favour of Group B. This suggests that test item arrangement beginning with the first topics presented has positive effect on the performance of senior secondary schools in mathematics.

Table 2 (b):Descriptive Statistics on the Effect of Test Items Arrangement Based on Last Topics Presented During Instruction, on Examinees Performance

Group	N	$\overline{X}$	SD	$\overline{X}_C$-$\overline{X}_D$	Decision
Group C (LTF)	100	31.90	10.18		
				-0.04	Negative though negligible effect
Group D (Control)	100	31.94	8.67		($\overline{X}_C$- $\overline{X}_D$< 0)

A mean score and standard deviation of 31.90 and 10.18 were respectively observed for Group C whereas 31.94 mean score and 8.67 standard deviation were observed for Group D (control group). This reveals a marginal mean difference of -0.04 which although is negative, its absolute value is negligible; implying that items arranged based on last topics presented during instruction have negative effect on examinees' performance in Mathematics.

The effect of arrangement based on a first topics presented was compared with the effect of arrangement based on last topics treated and the details of analysis is presented in table 2 (c)

Table 2 (c):Analysis of Mean difference of Effect of Test Items Arrangement Based on First Topics or Last Topic Presented during Instruction

Group	N	$\overline{X}$	SD	$\overline{X}_B$-$\overline{X}_C$	Decision
Group B (FTF)	100	37.86	8.89		
				5.96	Positive Effect
Group C (LTF)	100	31.90	10.18		($\overline{X}_B$-$\overline{X}_C$>0)

Table 2 (c) reveals a positive mean difference of 5.96 between items arrangement based on first topics presented and arrangement based on last topics presented, in favour of the former. This simply implies that when test items are arranged starting with the earlier topics taught and progresses to the latest topics, the performance of testees will be more outstanding than when items are arranged starting with the topic treated last. Thus, items arrangement based on earlier topics presented during instruction, enhances examinees' performance than when items are arranged based on last topics presented.

Hypothesis 1: There is no significant effect of top-bottom test items arrangement on examinees' performance in mathematics in senior secondary schools in Rivers State.

Table 3 t-test Analysis on the Effect of Test Item Arrangement Based on Top-bottom on Examinees Performance in Mathematics

Group	N	$\overline{X}$	SD	t_{cal}	Df	A	t_{crit}	Decision
Group A (Top-Bottom)	100	36.30	6.31					
				4.07	198	0.05	1.96	Reject Ho
Group D (Control)	100	31.94	8.67					

The t-test analysis from table 3 revealed an observed t-value of 4.07, and a critical value of 1.96 with 198 degrees of freedom at the 0.05 level of significance. Since the computed value of t is higher than the critical value, (ie. $t_{cal} > t_{crit}$), the null hypothesis that "there is no significant effect of test item arrangement based on top-bottom on examinee performance in Mathematics" is rejected, in other words, the computed t-value is too significant to be attributed to chance. The implication is that test item arrangement based on top-bottom (ascending order) format has significant effect on examinees performance in Mathematics.

Hypotheses 2: Test items arranged based on order of topics presentation has no significant effect on examinees' performance in mathematics in senior secondary schools in Rivers State.

Table 4 (a): t-test Analysis on Effect of Test Items Arranged Based on First Topics Presented during Instruction on Examinees Performance in Mathematics

Group	N	$\overline{X}$	SD	t_{cal}	df	A	t_{crit}	Decision
Group B (FTF)	100	37.86	8.89					
				4.77	198	0.05	1.96	Reject Ho
Group D (Control)	100	31.94	8.67					

[

The t-test analysis from table 4 (a) reveals a computed t-value of 4.77 which is significant when compared to the critical value of 1.96 obtained with 198 degrees of freedom at the 0.05 level of significance. Consequently, the stated null hypothesis was rejected, meaning that test items arranged beginning with first topics treated during instruction has a significant effect on examinees performance in Mathematics.

Conversely, the items were arranged beginning with topics treated last during instruction to test whether there is a significant effect on examinees scores in mathematics. The t-test analysis is presented in the table below.

Table 4(b): t-test Analysis on Effect of Test Items Arranged Beginning with Last Topics Treated on Examinees Performances.

Group	N	$\overline{X}$	SD	t_{cal}	df	A	t_{crit}	Decision
Group C (LTF)	100	31.90	10.18					
				-0.03	198	0.05	1.96	Accepted Ho
Group D (Control)	100	31.94	8.67					

From the above table, the computed t-value was -0.03 while the critical value was 1.96, and since the computed value was less than the critical, it means the observed value was insignificant at 0.05 level of significance and this led to the retainance of the stated null hypotheses. This also confirms the negligible marginal negative difference earlier observed between Group C and Group D. A comparative analysis was done to investigate the differential effect of items arranged beginning with first topics treated and last topics treated and the result is as presented in table 4(c).

Table 4(c): t-test Analysis on Differential Effect of Test Items Arranged Beginning with First or Last Topics Treated During Instruction.

Group	N	$\overline{X}$	SD	t_{cal}	Df	α	t_{crit}	Decision
Group B (FTF)	100	37.86	8.89					
				4.41	198	0.05	1.96	Reject Ho
Group C (LTF)	100	31.90	10.18					

Table 4(c) reveals a significant differential effect of items arranged beginning with first topics treated over items arranged beginning with last topics treated during instruction. The computed t-value was found to be 4.41 while the critical value was 1.96 at a level of significance of 0.05 with 198 degrees of freedom; and this computed value is too significant to be attributed to sampling error. This suggests that examinees performed better when test items are arranged beginning with the first topics treated during instruction, in other words, examinees would not do as better when items are arranged beginning with last topics treated or in random order or traditional format.

Discussion of Findings

The study revealed a significant positive effect on examinees performances when test items arrangement is on top-bottom format (easy-to-hard). The mean scores of examinees were found to be higher compared to those tested with items arranged in the traditional or random format. Also, the null hypothesis connecting the mean scores difference between the two groups could not be retained as the observed t-value was significantly higher than the critical, confirming that item arrangement based on top-bottom format has significant positive effect on the performances of examinees. This result suggests a corollary to the opinion of Mehren and Lehman cited in Tei-Firstman (2011) that test items should be arranged from easiest to the more difficult (top-bottom) to counter the sequence effect which occurs in a test taking environment whereby an examinee motivational moral may be altered by negative effect resulting from failure on a particular item.

A test taker therefore according to Mehren and Lehman in Tei-Firstman (2011) would be frustrated if faced from the beginning of a test with items he is unable to tackle and this could result to overall poor performance. Abbasian and Zadar (2019) however, found a contrary result to the current finding upon examining the relationship between test item arrangement and testees performance and test usefulness criteria in Iran, while Omay *et al.,* (2015) result, that different item ordering lead students at the same ability level display different performance on the same items and that a test form of sequential east-to-hard questions brings more advantages than that of a hard-to-easy sequence or a random version corroborate the cuurent finding. Other results in agreement with the current result were those of Uwah (2016), Tei-Firstman (2011) and Ollennu and Etsey (2015).

Secondly, the study revealed that item arrangement based on order of topic presentation has significant effect on examinees performances with items arranged beginning with first topics presented during instruction outperforming those arranged beginning with last topics presented during instruction. This result is in consonance with the result of Uwah (2016), when he investigated the effect of item arrangement on performance in mathematics among

secondary school students in Obio/Akpor Local Government Area and found that test item arrangement based on order of topics presentation has positive as well as significant effect on student's performance in mathematics, and he went further to interpret this result as students understanding first thing first. This interpretation conforms with the findings in the current study that arranging test items beginning with the first topics presented enhances the positive performances of examinees.

Implications of the Findings

The findings of the study have some educational implications, especially for teachers, test developers, test users, administrators, examination bodies, examination divisions of ministries of education and students. The result of the study revealed that item arrangement based on top-bottom (easy-to-hard) format has significant positive effect on examinees performance. This therefore implies teachers whose one major assignment is to design, develop and administer test to their students for proper assessment, will be more informed, careful and cautious to present students with test items beginning with the easier ones first, in order to captivate their interest and spur their motivation towards progressively attempting more questions with precision. A negative implication however, is a situation where teachers, examination and related bodies may in an attempt to capture examinees interest, in the test, present testees with substandard item which will not discriminate adequately.

Conclusion and Recommendations

Arrangement of test items in whatever format has influence on the performance of examinees. This confirms the sensitivity of this key aspect of testing as held by some scholars. Consequently, the study recommends that:

1. All stakeholder in the business of test construction, administration and users alike to ensure that test items are arranged in a top-bottom format where by testee first encounter the less difficult items to gain courage before proceeding to the more difficult item.
2. Beside the top-bottom format recommended above, the arrangement must marry with topical presentation. In other words, while considering the easiness of the earlier item, the earlier topics presented during instruction should be considered first (FTF).

REFERENCES

Abbasian, G. & Zadsar, H. (2019). Relationship between test item arrangement and Testee's performance and test usefulness criteria. *International Journal of English Language Teaching*, 6(2), 76-84. https://iject.sciedupress.com. Retrieved Dec. 15, 2019.

Adebayo, T. D. (2010). An appraisal of self-assessment of secondary schools' performance in Ondo State. *Journal of Educational Research and Evaluation*. 19(1), 5-11.

Ajayi, V. O. (2018). Difference between assessment, measurement and evaluation in Science Education. https://en.m.wikipedia.ng. Retrieved Oct. 4, 2019.[

Anastasi, A. & Urbina, S. (2006). *Psychological testing (7th Edition).* India: Dorling Kindersley.

Barbara, (2015). Item arrangement and knowledge of arrangement of test scores. *In Journal of Experimental Education,* 49(1). https://dei.org/10.1080/00220973. Retrieved Jan. 2, 2020.

Blair, G. M., Jones, R.S & Sampson, R. H. (2001). *Educational psychology (4th Edition).* New York: Macmillan Publishers Co., Inc.

Heick, T. (2019). An Efficient Classroom: Six factors of academic achievement. *Teach thought*, 14, 321-337. www.teachthrough, http://www.sciencedirect.com. Retrieved Sept. 22, 2019

Koul, L. (2018). *Methodology of educational research (4th ed.).* India: Vikas Publishing House PVT Ltd.

Onunkwo, G.I.N. (2002). *Fundamentals of educational measurement and evaluation*. Owerri: Cape Publishers International.

Orluwene, G.W. (2012). *Introduction to test theory and development process.* Port Harcourt: Chris-Ron Integrated Services.

Tei-Firstman, R. I. (2011). *Test Item Arrangement on Students Test Scores*. Unpublished Thesis, University of Port Harcourt.

Umude, L. E. (2018). *Test and Measurement and its relevance to educational management*. Unpublished Thesis, University of Ibadan

Uwah, I. V. (2016). *Effect of test items arrangement on performance in mathematics among junior secondary school students in Obio/Akpor local government area of Rivers state.* Unpublished M. Ed dissertation, University of Port Harcourt, Nigeria.