EFFECT OF ITEM-POSITION CHANGE IN MATHEMATICS MULTIPLE-CHOICE AMONG SENIOR SECONDARY SCHOOL STUDENTS IN OGUN STATE, NIGERIA

TAIWO, O.
Department of Educational Foundations and Counselling Psychology
Faculty of Education, Lagos State University, Ojo
olutaiwo22@gmail.com

Abstract
This study determined the effect of item position in Mathematics multiple-choice among senior secondary school students in Ogun state, Nigeria. The study adopted the quasi-experimental research design. A sample of four hundred and fifty (450) Senior Secondary School Students drawn from the population using the simple random and stratified sampling techniques was used for data collection. The sample included two experimental group labelled 'A-B" and the control group labelled "C" This was selected from a population of 1,350 SS II students from nine public schools in the area. A test instrument tagged Mathematics multiple-choice tests consisting of 40 items adapted from WASCE/SSCE questions. Type A was arranged in Easy to hard. Type B was arranged in Hard to easy order while type C was arranged in random order all based on table of specification of the topic selected. Validity of the instruments was determined using table of specification while a general reliability index of 0.96 was determined using Kuder Richardson formular 20 (KR20). Analysis of variance was used to test the hypotheses. Results showed that the change in item order has significant effect on the performance of students, the F-value (F=39.297 & P< 0.05). The study concluded that item position change affects students' performance but random order performance is consistent. It was recommended among others that test constructors and professional examination bodies should endeavour to arrange items randomly or from easy to hard in order to boost students' morale.

Keywords: Test, item position, item arrangement and rearrangement, mathematics performance.

Introduction

A test is best defined as a procedure for sampling and describing behaviour through the use of scores generated from the test. This includes varieties of tools such as checklists, rating scales and observation schemes. The essential features of a standardized test are: standardized procedure, i.e. the procedure is administered uniformly over a group of persons and also focussed behavioural sample, i.e. the test is focused at a well-defined behavioural domain. Examples of domains in educational measurement are achievement in arithmetic, or language performance. Ferrier (2011)

Test as noted by Opara (2014) is an instrument or procedures which is designed to measure the knowledge, intelligence, ability, traits, skills, aptitude, interest, attitude which an individual or thing exhibits. It is a systematic procedure for observing an individual's behaviour as well as describing such behaviour or performance by numerical scale or category. Mathematics education holds the potency of making individuals to relate mathematics knowledge to everyday problem being encountered and hence develop the individuals to a level that they are intellectually and economically stable. Right from the historic days of early human societies to the present age, mathematics has played a fundamental role in the economic development of many countries of the world (Fennema & Sherman in Gegbe, Sundai & Sheriff, 2015). The knowledge of mathematics is paramount in the success of every man in his numerous day-to-day activities in life. It is observed by the researchers that majority of students find it difficult to pass mathematics test effectively. This implies that many students have develop test phobia for the subject and as such attach negative attitude toward the subject which at times are extended toward the teachers handling them.

Shohamy (2005) rightly stated, a test is considered as a good one if the method has little effect on the trait. To put it in another way if students' performance on a test is the result of the trait being measured rather than the testing method, that test is considered to be a good testing tool. Under testing conditions where examinees are close together and the answer sheet of examinees are easy to read, test supervisors are concerned about copying. The method of minimizing the likelihood of copying is to space examinees enough apart so that visual copying requires either superb visual acuity or indiscreet posturing of the body on the point of the examinee who intends to read someone else answer sheet. Unfortunately, the option of using direct and simple methods is not always available to test administrator. Nelson, 2010.).

One approach is to alter the layout of the answer sheet so that every other examinee marks his/her responses by successive items in a left to right manner on a horizontal grid, while the remaining examinees mark their responses in a top to bottom fashion on a vertical grid. Alternation of the placement of items in the test booklet is another procedure that is employed. Sometimes, the alteration of item placement is cosmetic. The same item may appear in different pages in two forms of the test items, but the item order is unchanged. In other situations, the actual ordering of the items is changed.; Pashler and Mozer, 2013. The phrase “within test context effects” refers to changes in performance on a particular type of test item which result from the context of questions that appear earlier in the same test (Stewart, 2002). It has traditionally been recommended that test items within sections should be arranged in an order of increasing difficulty. In other words, the easier items should be presented first, following by progressively more difficult items (Morsell, 2003; One explanation for this recommendation has been that examinees for which a test is a speeded could be disadvantaged as a result of spending time on hard items only in the test that they could be more profitably have spent on easy items near the end. A second explanation has been based on the belief that the anxiety aroused by the inability to answer a difficult test item affects performance in succeeding items. (Johansen and Palmeri, 2002). Item rearrangement as a possible solution to copying may produce undesirable context effects, including differential anxiety arousal. If the items cannot be re-arranged within a test item an alternative procedure to deal with copying is to re-arrange an items' options. Item positioning can also influence test performance which is ascribable to fact that examinees spend different time to respond to different questions based on the difficulty level of the item. Examinees spend much time on the difficult ones and at times run short of time. One of the problem is that the difficulty level of items are not the same and also item rearranging either from easy-to-hard, hard-to-easy and random can affect test performance. Mathy and Feldman, 2009

In an attempt to circumvent the negative effects imposed by social implications of testing, which include impedance to academic progress, forfeiture of professional advancement and promotion, and the stigma of being labelled a non-achiever. Pettijohn and Sacco (2001) reported that to prevent cheating on examinations, many professors will mix up the order of multiple-choice test questions from examination to examination without thought of the consequences the change of order may have on student examination performance and perceptions. However, they noted that questions have been raised regarding whether the order of test items influences student performance. Position effects are generally ignored when test items are presented to examinees in a conventional fashion with all items being presented to each examinee in the same order. In principle, there are two possibilities on how position of items can influence item difficulty in a rather general way.

However, position effects are inseparably connected with item difficulty parameters and can thus neither be confirmed nor quantified. Whereas position effects may not be a problem in conventional testing, they are certainly not acceptable in adaptive testing, where items are presented to the examinees in varying orders. Items presented at the beginning or towards the end of examination no longer have the same difficulty and 'fair' comparisons between examinees are no longer possible. Research on item context effects (whether the performance of test items change when the content, difficulty, or order of previous items is altered) has a long and diverse history in educational measurement. Debeer & Janssen 2013). In an extensive review, they concluded that there is some evidence of item context effects, the importance of these effects is not well understood. Since that review was published, the importance of item context effects on item parameter invariance and score equating is clear. These effects must be accounted for in the measurement process. However, the consequence of item context effects for measurement theory and practice in other areas is less apparent. Hence, there is the need to examine the effect of item rearrangement on test performance (Pui Chi Chiu 2012)

Testing plays a key role in promoting good performance and for decision making. In spite of their highly advantageous use, tests have varied social implications and limitations. Consequently, the use of tests in obtaining facts and data about people has received its fair share of criticisms. In an attempt to prevent cheating on examinations, many professors will mix up the order of multiple-choice test questions from examination to examination without the thought of the consequences the change of order or rearrangement may have on student examination performance and perceptions. However, they noted that questions have been raised regarding whether the order of test items influences student performance. In the light of the above the problem is, "can item position change in mathematics affects students' performance in secondary schools? In line with the study problem, the study aims at determining the effect of item position change in mathematics multiple-choice among senior secondary school in Ogun.

Research Hypotheses

The following null hypotheses tested at 0.05 alpha level were postulated to guide the study;

1. There is no significant effect of item position change on test performance in mathematics among senior secondary school students in Ogun
2. There is no significant interaction effect of demographic characteristics (gender, school type and location) and item rearrangement on test performance in Mathematics among students in Ogun.

Methodology

The study adopted the quasi-experimental research design. A sample of one hundred (450) senior secondary school students (SSII) was drawn from the population of one thousand, three hundred and fifty (1,350) using simple random and non-proportionate sampling techniques. The Mathematics adapted Test was used as instrument for data collection. This is a test instrument which has two types (Type A or B and C). Type A or B of Multiple choice questions contains 40 items arranged in Easy to Hard and Hard to Easy order Type C equally contained 40 items arranged in random order. Content validity of the instrument was determined using table of specification. Equally the reliability of the instrument was determined using Kuder Richardson formular 20(KR20). The reliability coefficient obtained was 0.96. The three types of instrument were administered on separate basis to fifty different groups of students from nine schools. Type A and B was administered to 50 students in the first school. Type C was administered to the control group. In other words, the test instrument was administered to one experimental group and one control group through the help of a research assistant and the classroom teachers. The data collected was separated to one-way analysis of variance (ANOVA) and Bonferroni's multiple comparisons post hoc test was used to test the hypothesis at .05 level of significance. The independent variables are item order with two levels as Random order (RDM), Easy-to-Hard (ETH) order and Hard-to-Easy order (HTE). The dependent variable was scores of the tests

Results

Hypothesis One: There is no significant effect of item position change on test performance in mathematics among senior secondary school students in Ogun

The results for the analysis of scores for item position change condition are given in Tables 1 and 2. Table 1 gives the results of the ANOVA and descriptive statistics for the three levels of item order, while Table 2 shows the results of the Bonferroni's multiple comparisons post hoc test indicated that at the .05 level of significance

Table 1: Analysis of Variance showing the effect of item rearrangement on students' power test performance in Mathematics

	Sum of Squares	df	Mean Square	F	Sig.	Effect size
Between Groups	1782.337	2	891.169	39.297	.000	0.276
Within Groups	4671.586	206	22.678			
Total	6453.923	208				
Mean score (SD) Easy-to-Hard (n=85) – 22.32±5.81 Hard-to-Easy (n=26) – 14.42±3.81 Random (n=98) – 23.69±3.90						

Table 1 showed the change in item order has significant effect ($F=39.297$; $p<0.001$) on the students' Mathematics test performance in the study area. Students exposed to random item order (RDM) performed better (23.69±3.90) than the ones exposed to easy-to-hard (ETH) (22.32±5.81) and hard-to-easy (HTE) (14.42±3.81) items order when subjected to power testing condition. The effect model reveals that item order accounts for 27.6% of the variation in students' performance in Mathematics.

Table 2: Bonferroni post hoc test of multiple comparisons of mean performance under power testing condition

	Mean difference		
	Easy-to-Hard	Hard-to-Easy	Random
Easy-to-Hard	-	7.895*	-1.376
Hard-to-Easy	-7.895*	-	-9.271*
Random	1.376	9.271*	-

*Mean difference is significant at * p<0.05*

Table 2 explains the post hoc test multiple comparisons of students' average performance using Bonferroni's approach. This however, explains that difference in test performance of students in Mathematics is significant only between hard-to-easy and easy-to-hard as well as hard-to-easy and random orders. That is, students' performance in random and easy-to-hard order of items significantly differs from that of hard-to-easy order; whereas, performance in random order does not significantly differ from that of easy-to-hard item order.

Hypothesis Two: There is no significant interaction effect of demographic characteristics (gender, school type and location) and item rearrangement on test performance in Mathematics among students in Ogun.

To test this hypothesis, students score in various schools- type, location and as well as gender were compiled and analyzed using multivariate analysis of variance and the interaction of both. The result is as presented in Tables 3

Table 3: Interaction effects of gender, school type and location among students' performance in Ogun by item position change

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Effect size
Corrected Model	2408.182[a]	5	481.636	24.167	.000	.373
Intercept	49387.351	1	49387.351	2478.070	.000	.924
Item order	712.991	2	356.496	17.888	.000	.150
Gender	3.146	1	3.146	.158	.692	.001
School type	269.002	1	269.002	13.497	.000	.062
Location	5 13.783	1	513.783	25.780	.000	.113
Error	4045.742	203	19.930			
Total	107434.000	209				
Corrected Total	6453.923	208				

R Squared = .373 (Adjusted R Squared = .358)

Table 3 shows the interaction effect of gender, school type and location on students' test performance in Mathematics by item position change. According to the table, the school type ($F_{1,203, 0.05}$=13.497; p<0.001), Location ($F_{1,203, 0.05}$=25.780; p<0.001) and items order ($F_{2,203, 0.05}$=17.888; p<0.001) have significant effect on students' test performance in Mathematics. However, gender has no significant effect on students' performance ($F_{1,203, 0.05}$=0.158; p=0.692). The effect size of the model indicates that change in items order accounts for 15.0% while location of the students accounts for 11.3% of the variation in students' test performance in the study area. Hence, use of change in item order to improve students' performance depends whether the students' school is private or public and whether the student is located in particular area i.e. urban or rural

Discussion of Findings

The effect of item position change on students' test performance reflected some variables that have been put into consideration such as easy-to-hard, hard-to-easy and random order. This construct or variables are important in decision making in construction of items. This study focused on the effect of item position change on test scores in mathematics. Overall, the findings revealed the impact of item position change on test scores, in the hypotheses, showed the change in item order has significant effect. That is, students exposed to random item performed better than those who are exposed to easy-to-hard and hard-to-easy, this was in line with chidomere (1989) who found higher scores with the random order compared against the hard-to-easy. Leary and Dorans (1985) argue that scores of item arranged in easy-to-hard (ETH) order declined because the difficulty level built up to a point where the items in the latter part of the test became too difficult and had to be guessed at or answered haphazardly. The results of this study however agree with those of educational measurement experts like Shepard(1977), MacNicol (1956), Skinner (1999) and Ahuman and Clock (1971) who asserted that changes in test format (or arrangement) can make a large difference in students' performance. The findings of this investigation also agree with that of a study by the Research Division of WAEC, Lagos (1993). They found significant differences in the performance of students in various subjects at the secondary school level when items were re-ordered. This finding corroborates work of Cronbach (1995) who found that scores on the easy-to-difficult item order were on the average significantly higher than scores on the difficult-to-easy order.

The study also revealed that under speeded condition there is significant difference only between easy-to-hard and random orders. This means students' performance in easy-to-hard order of items significantly differs from that of random order, whereas performance in easy-to-hard does not significantly differ from that of hard-to-easy order. This is in line with the work of Hambleton and Traub (1974) studied 11th graders' performance on an Algebra II Mathematics Test. They discovered the average number of correct answers for test questions arranged from easy-to-difficult was significantly higher than the test questions arranged from difficult-to-easy.

Recommendations

Based on the findings, the researchers recommend that;

1. appropriate adjustments to neutralize the impact of the difference in performance to facilitate the use of re-ordering of multiple-choice tests to arrest examination malpractices.
2. Teachers, professional bodies should train and re-train subject officers in the application of how to rearrange items.
3. From the findings random order is found to have positive effect, such effect being significant implies that test constructors can also follow such format in positioning items especial in a large stakes.
4. Classroom teachers, test constructors and other stakeholders involved in test construction should ensure that test items are arranged randomly because

based on the findings, it has positive effect or from easy to hard in order to give students the motivation to perform.

5. Though hard to easy order of test item arrangement is found to have a negative effect, such effect being insignificant implies that test constructors should avoid this format at every point in time. Also, teachers as well as professional educational bodies should always follow the order of topic in the curriculum in setting test or examination questions.

Conclusion

From the findings, it can be concluded that the effect of item position change its general sense, has an effect on students' performance in mathematics. But more specifically, each item order i.e. Random (RDM), easy-to-hard (ETH) and hard-to-easy (HTE) has different significant effect on students' performance. Findings from the study revealed that Random (RDM) order has a higher significant effect unlike the easy-to-hard and hard-to-easy order. It is good arrangement that enhanced better performance and such pattern should be encouraged among classroom teachers, evaluators as well as professional examination bodies.

References

Ahuman, S.W., & Clock, N.D. (1971). Item difficulty level and sequence effects in multiple-choice achievement tests. Journal of Educational Measurement, 9 (Summer), 105, 11.

Chidomere, R.C. (1989). Test item arrangement and student performance in principles of marketing examinations: A replication study. Journal of Marketing Education, (fall), 36-40. Classes. http://cran.r-project.org/web/packages/Ime4.

Cronbach, L.J. (1995). Essentials of psychological testing (4th ed.). New York: Harpers and Row.

Debeer, D., & Lawrence, I.M. (2013). Modelling item-position effects within an IRT framework. Journal of Educational Measurement, 50, 164-185. Dissertation, University of Lancaster.

De Boeck, P., & Wilson, M. (2004). A framework for item response models. In P. De Boeck & M. Wilson (Eds.), Explanatory item response models (pp.3-42). New York, NY: Springer.

Flaughter, D.A., Melton, R.S. and Myers, C.W. (1986). Item arrangement under typical test conditions. Educational and Psychological Measurements, 28, 813-824.formats. Language Testing, 26, 219-244 fundamentals and applications and contemporary guidance.

Gegbe, B. Sundai. A and Sheriff. V.K. (2015). Factors contributing to students' poor performance in Mathematics at West African senior school certification examination (A case study: Kenema City, Eastern Province Sierra Leone). International Journal of Engineering Research and General Science 3, (2),

2015.
Hambleton, R.K. & Traub, R.E. (1974). the effects of item order on test performance and stress. The journal of Experimental Education, 43(1), 40-46.
Leary, L., and Dorans, N. (1985). Implication for altering the context in which test items appear: A historical perspective on an immediate concern. Review of Educational Research, 55(3), pages 387.
MacNichol, K. (1956). Effects of varying order of difficulty on an unspeeded verbal test. An unpulished manuscript. Princeton, New Jersey: Educational Testing Services.
Opara, I. M. (2014). Psychological Testing, Principles and Techniques. Owerri Career publishers.
Pettijohn II, T.F., & Sacco, M.F. (2001). Multiple- choice exam order influence on student performance, completion time and perceptions. Journal of instructional Psychology, 34 (3), 142-149.
Pui Chi Chiu (2012): Effect of item position on mathematics assessment. A thesis from the university of Kansas, AERA 2012.
Sax, G. and Cromach, T. (1996). The effects of various forms of item arrangement on test performance. Journal of Educational Measurement, 3, 309-311.
Shohamy, E. (2005). Does the testing method make a difference? The case of reading.