TEST-TAKING SKILLS AS PREDICTORS OF PERFORMANCE IN MULTIPLE CHOICE MATHEMATICS TEST AMONG SENIOR SECONDARY STUDENTS IN LAGOS STATE, NIGERIA.

Dixon, O. O. & Erinosho, S. Y.

Abstract

This study investigated the predictive ability of students' test-taking skills on performance in multiple choice mathematics tests. Non-experimental, cross-sectional correlation research design was adopted in the study. Multi-stage sampling procedure was used to select eighteen (18) schools in three educational districts in Lagos State out of which a total of 1733 students (1006 SSII and 727 SSIII) were drawn. Two instruments were used to collect data: Multiple choice Mathematics Test (MMT, r = 0.75) developed and validated by the expert and an adapted Test-taking Skill (TSS) scale covering reported testing activities of students before-, during- and after-test (α = 0.78). The findings revealed a significant positive relationship between scores on TSS and MMT. Before-test skill contributed the highest to the variance in the prediction of performance on MMT. In addition, significant differences in the reported testing skills were found between the high, average and low scorers. The study recommended the need for mathematics teachers to integrate the teaching of testing skills as part of classroom instructional activities in order to support the students, especially weak ones, through constant practice and regular engagement with learning materials.

Keywords: Classroom Assessment, Multiple Choice Test, Test-taking Skill, and Mathematics Performance

Introduction

Tests are commonly used as assessment tools to measure cognitive learning and proficiency of students. A test is an instrument designed for measuring and quantifying the behaviour of students in all the domains of learning (Fook & Sidhu, 2010; Amajuoyi, Joseph & Udoh, 2013). When used appropriately, it enables the teacher to provide a realistic picture of learners' ability, conceptual difficulties, and learning needs. Two types of achievement tests are typically used to measure cognitive learning: the open-ended (essay) and open-ended (objective) types. Multiple Choice (MC), also known as close-ended, tests had gained popularity in its use in many contexts over the years. They are an attractive option for classroom assessment because they can be quickly administered; are easy to score; have greater objectivity and reliability in scoring (Scully, 2017); flexible to measure complex outcomes at all levels of learning (Ajayi & Omirin, 2013; Scully, 2017); and have a higher sampling of content per unit time (Tarrant, Knierim, Hayes & Ware,2006). Multiple Choice tests have also received criticisms for their inability to tap higher order thinking skills and susceptibility to guessing the answer (Popham, 2010; Haladyna, 2012). Haladyna condemned their use in medical education since multiple choice items are incapable of assessing cognitive process beyond recall or recognition of knowledge. Scully (2017), however, argued that multiple-choice items have the capacity to assess certain higher-order skills if the test developer is skilful in test construction.

Students must learn specific testing skills that would enable them to maximize the cognitive knowledge gained during the instructional process to allow the test to adequately serve as a measure of their achievement. Non-cognitive skills that enable students to perform at optimal level in a testing situation are referred to as test taking skills (Rupp, Ferne & Choi, 2006; Bicak, 2013; Dodeen, Abdelfattah & Alshumrani, 2014). Tested individuals apply them before, during and after a testing situation. Test-taking skills help testees in test preparation, time management, control of test anxiety, and feedback mechanisms for correcting weaknesses in future testing situations. It has been shown that test-taking skills improve students' motivation to learn and their attitude towards learning (Dodeen, 2009); reduces examinee's tension and anxiety over their ability to communicate what they know in a test situation (Dooden et.al., 2014); locus of control (Hong, Sas, & Sas, 2006); and academic achievement (Bicak, 2013). On the contrary, poor test-taking skill is likely to negatively affect student performance test wisdom (Biçak, 2013; Lewandowski, Berger, Lovett, & Gordon, 2016; Bensley *et.al,* 2016). Test-taking strategies differ among low and high achievers (Hong, Sas & Sas, 2006; Stenlund, Eklöf, & Lyrén, 2017) with high achievers reporting more active use of test-taking skills.

The study draws on Dekeyser Skill Acquisition Theory (1989) propounded to explain cognitive and behavioural stages in skill construction and development. Like Piaget's notion of stages in cognitive development, Deskeyer describes a framework to explain how skill and thinking development progress in stages from a novice stage to a state of being proficient (DeKeyser, 2007). The theory classifies skills acquisition into three stages: declarative, procedural and automatic. Declarative knowledge could be acquired through direct instruction and observation usually, during teaching and learning exercises. Procedural learning relies on declarative knowledge to apply the rules pertaining to a task at a given time. The automatic stage is displayed with complete fluency and spontaneity, rarely showing any errors.

In most situations, practice of procedural knowledge is needed to achieve automatic skill fluency. The learner must constantly engage with the skill knowledge through large amount of practice or constant engagement with same or closely related tasks (DeKeyser, 2007). Practice with a given task is marked by decrease reaction time and error rate and interference from other tasks. Practice should be engaging, purposeful, and distributed (Rohrer, 2009). Skill Acquisition Theory (SAT) is important in mathematics since the subject involves structured and logical operations, defined tasks and explicitly defined processes. Working through problems therefore requires that students develop skills to deal with the operations, logic and rules that are required to develop an answer. Such skills can be acquired through practice of similar problems as part of test preparation and developing fluency in managing testing situations.

In Nigeria, MC test-type is widely used in public examinations like the Senior School Certificate Examinations, and Unified Tertiary Matriculation Examinations. The performance in mathematics has been worrisome due to consistent poor results over the years. This could be due to insufficient knowledge or poor testing behaviour. Even if classroom instruction were effective, students' ability to display their knowledge in examinations could be affected by their test-taking behaviour. WAEC Chief Examiners' reports for 2015, 2016 and 2017 revealed that some of the challenges encountered by the students are inability to interpret word problems, deficiency in writing answers to the required degree of accuracy and inability to reason logically. These observations are indicators of poor testing behaviour among students; they suggest that intervention is needed.

While several studies have attempted to address the problems associated with quality learning and pedagogical strategies in mathematics, there is a dearth of research on the testing behaviour of students as it affects performance in mathematics in Nigerian secondary schools. A student might have the knowledge but be unable to display it in a testing situation because of deficiencies in test-taking strategies. It is therefore necessary to examine the inter-relatedness of students' mathematics performance and test-taking strategies. In addition, different test types require different test-taking skills because of the variation in the tasks and structure of answering mode. This study therefore focused on multiple choice mathematics test to examine the connection between test taking skills and performance among senior secondary school students in Nigeria.

The Research Questions

(i) What is the predictive ability of test-taking skills on performance among senior secondary school students' that were tested on multiple choice mathematics items?

(ii) How do the reported test-taking skills vary with the performance levels of the students?

Methods

The study adopted a non-experimental, cross-sectional descriptive research design to investigate the inter-relatedness of mathematics performance and test taking skills among Senior Secondary (SS) II and III students in public schools in Lagos State, Nigeria. Multi-stage sampling technique was used to select eighteen (18) out of the three hundred and eight (308) senior secondary schools in Lagos State. In the first stage, a purposive sampling technique was used in selecting three districts that each has at least 50 secondary schools from the six districts in the state. Proportional sampling was adopted in the second stage to select 10% of the total senior secondary schools in each district. Eighteen schools were then drawn through random sampling. In the last stage, random sampling was used to select seventy (70) SSS II and fifty (50) SSS III students from each of the eighteen schools. In all, one thousand, seven hundred and thirty-three (1733) students participated in the study, comprising one thousand and six (1006) SSS II and seven hundred and twenty-seven (727) SSS III students.

Two instruments were used to collect data -Test-Taking Skill Scale (TSS) and Mathematics Multiple choice Test (MMT). TSS was adapted from the Test-taking Skill Scale developed by Dodeen (2008) and Bicak (2013). The TSS is a 3-point Likert scale, consisting of 28 items distributed into four sub-scales: before-test strategies (6 items); during-test strategies with three sub-scales of structural organisation (4 items), time management (4 items), test-wisdom (6 items); and after-test strategies (8 items). To determine the reliability of TSS, Principal Component Analysis (PCA) was employed. The Kaiser-Meyer Olkin measure of sampling adequacy suggested that the item sample was factorable (KMO=.657). In addition, 23 of the 28 items correlated at least 0.3 with at least one other item, suggesting reasonable factorability. The Cronbach alpha reliability of the sub-scales are: before-test strategies (α=0.60); during-test strategies (α=0.85), time management (α=0.85), test-wisdom (α=0.83) and after-test strategies (α=0.78) and TSS (α=0.78), indicatingacceptable internal consistency. MMT is a twenty-five (25) item multiple choice test with five (5) response options covering five topics that had been taught in all the schools. To determine the reliability of MMT, it was administered in a classroom situation with the assistance of the class teachers in two schools that were not part of the sample, and then a reshuffled version was re-administered after two weeks. The test-retest correlation coefficient, (r=0.75), shows that the test has high reliability. The approval of the school principals, heads of departments, mathematics teachers as well as the consent of the students were sought before administering the instruments. MMT and TSS were administered on different days in each of the sampled schools with the assistance of the mathematics teachers and two research assistants.

Results

Research Question 1: What is the predictive ability of test-taking skills on performance in mathematics among senior secondary school students' that were tested on multiple choice items?

Correlation and multiple regression analyses were undertaken to examine the relationship between scores on MMT and TSS. As shown in Table 1, the scores on the TSS subscales have positive and significant correlation coefficients with performance on MMT. By implication, students with high scores in before-test, during-test and after-test skills tend to score high on MMT. Before-test skill has a stronger and significant relationship with mathematics performance for both SSS II and SSS III (r=0.43 and r=0.38) compared with during-test and after-test skills.

In addition, the multiple regression model indicates that the three predictor variables produced positive variance ($R^2 = 0.26$, F= 109.89, $p < 0.05$ for SSS II and $R^2 = 0.17$, F= 47.62, $p < 0.05$ for SSS III). The coefficient of determination of 0.26 and 0.17 implies that 26% and 17% of performance on MOT is caused by changes in the independent variables of test-taking skills. This was considered sufficient enough to determine the statistical significance of the coefficient of determination judging by the sample size. As can be seen in Table 1, before-test skills (β=1.98), during test skill (β =0.30) and after test skill (β =1.40) have a significant positive regression weights. This implies that for SSS II, students with higher scores on these scales wereexpected to have higher performance on MMT, after controlling for the other variables in the model. By implication, before-test skills had a stronger weight than other test-taking skills when students were tested on multiple choice items. Similarly, for SSS III, before-test skills (β =1.25), during test skills (β =0.37) and after test skill (β =1.20) had positive and significant regression weights. Before-test skill also had a stronger regression weight than during and after-test skills. In sum, it is evident that test-taking skills contributed significantly to mathematics performance on multiple test format and test-preparation skills contribute the highest to the variance in mathematics performance.

Table 1: Multiple Regression Showing Predictive Ability of TSS on MOT

	SSS II			**SSS III**		
Predictors	**B**	**β**	**R**	**B**	**β**	**R**
Test preparation skill	1.98**	0.31	0.43	1.25**	0.21	0.38
During test skill	0.30**	0.10	0.32	0.37**	0.11	0.29
After test skill	1.40**	0.25	0.41	1.20**	0.23	0.35
R^2	0.26			0.17		
F	109.89**			47.62**		

**$p<0.05$

The regression equation derived from the table are as follow:

For SSS II y = 1.98 (test preparation) +0.30 (during test) + 1.40 (after test) +10.46

For SSS III y = 1.25 (test preparation) +0.37 (during test) + 1.20 (after test) +11.13

y = performance on MMT

Research Question 2: How does the reported test-taking skills vary with the achievement levels of the students?

The mean scores on MOT were classed as high (> 70), average (< 70 > 40) and low (< 40) according to the grading convention in schools. Table 2 shows the variation in the reported mean scores on TSS according to the achievement level of the students. The significant F value ($p < .05$) for the mean score difference on TSS subscales indicate that there is a significant difference in the reported use of test taking skills according to achievement level.

For SS II, the high scorers reported significantly higher before-test skills ($F = 37.40$, $p < 0.05$), during-test skills ($F = 38.05$, $p < 0.05$), after test skills ($F = 29.32$, $p < 0.05$) and overall test-taking skills ($F = 37.40$, $p < 0.05$) compared with average and low scorers. Moreover, for SS III, high scorers reported significantly higher before-test skills ($F = 13.54$, $p =< 0.05$), during test skills ($F = 20.48$, $p < 0.05$), after-test skills ($F = 41.90$, $p =< 0.05$) and overall test-taking skills ($F = 291.91$, $p =< 0.05$) compared with average and low scorers. All of the performance sub-groups reported least use of after-test skills at both class levels. In addition, Table 3 provides the multiple comparisons of the significance of the mean differences in scores on MOT according to the achievement classifications. As observed, the highest difference in mean scores was obtained by the high scorers and least by the low scorers in both SSII and SSIII. By implication, the high scorers reportedly used the test skills more than the other subgroups.

Table 2: Variation in Test -taking Skills according to Achievement Level on MMT

			N	Mean	Stdev		SS	df	MS	F	Sig.p
SSS II	Test preparation	Low	137	12.28	1.22	Between Groups	169.29	2	84.64	37.50	**
		Average	826	13.16	1.57	Within Groups	2263.97	1003	2.26		
		High	43	14.42	0.98	Total	2433.26	1005			
		Total	1006	13.10	1.56						
	During test	Low	137	28.10	2.05	Between Groups	703.79	2	351.89	38.05	**
		Average	826	30.15	3.25	Within Groups	9276.46	1003	9.25		
		High	43	32.12	0.66	Total	9980.25	1005			
		Total	1006	29.96	3.15						
	After test	Low	137	14.56	1.91	Between Groups	169.03	2	84.51	29.32	**
		Average	826	15.42	1.70	Within Groups	2890.74	1003	2.88		
		High	43	16.09	0.29	Total	3059.77	1005			
		Total	1006	14.74	1.74						
	Test taking skills	Low	137	55.80	3.80	Between Groups	1553.34	2	776.67	36.29	**
		Average	826	57.88	4.85	Within Groups	21467.88	1003	21.40		
		High	43	62.63	1.76	Total	23021.23	1005			
		Total	1006	57.80	4.79						
SSS III	Test preparation	Low	85	12.49	1.27	Between Groups	66.80	2	33.40	13.54	**
		Average	622	13.33	1.63	Within Groups	1786.14	724	2.47		
		High	20	14.10	0.45	Total	1852.94	726			
		Total	727	13.25	1.60						
	During test	Low	85	28.29	2.07	Between Groups	318.44	2	159.22	20.48	**
		Average	622	30.06	2.91	Within Groups	5629.63	724	7.78		
		High	20	31.95	0.22	Total	5948.07	726			
		Total	727	29.90	2.86						
	After test	Low	85	15.46	1.82	Between Groups	83.80	2	41.90	14.42	**
		Average	622	14.62	1.71	Within Groups	2103.07	724	2.90		
		High	20	16.00	0.00	Total	2186.87	726			
		Total	727	14.76	1.74						
	Test taking skills	Low	85	56.25	3.92	Between Groups	583.82	2	291.91	14.95	**
		Average	622	58.01	4.55	Within Groups	14135.72	724	19.52		
		High	20	62.05	0.22	Total	14719.54	726			
		Total	727	57.91	4.50						

** p Significance at 5%

Table 3: Multiple Comparisons Analysis of the TSS between High, Average and Low Achievers

Class Achiever (J)	Test-taking skills	Achiever(I)		Mean Difference (I-J)	Sig.p
SSS II	Test preparation	Low	Average	-.644	**
		Low	High	-.746	**
		Average	High	-.442	**
	During test	Low	Average	-.522	**
		Low	High	-.765	**
		Average	High	-.329	**
	After test	Low	Average	-.229	**
		Low	High	-.306	**
SSS III	Test preparation	Low	Average	-.715	**
		Low	High	-.751	**
		Average	High	-.283	**
	During test	Low	Average	-.410	**
		Low	High	-.672	**
		Average	High	-.377	**
	After test	Low	Average	-.650	**
		Low	High	.691	**
		Average	High	-.393	**

** p significance at 5%.

Discussion and Conclusion

The results indicated that the three categories of test-taking skills contributed significant positive variance to the prediction of performance in multiple choice mathematics test. This has confirmed that there is interaction between test-taking skills and performance of students in mathematics. That is, a student who possesses adequate test-taking skills is likely to do well if the content is well learned. Rupp *et al*; (2006) reported a direct interaction between test performance and use of certain test-taking strategies. Experience shows that large numbers of students in schools are deficient in the skills they needed in mathematics tests and minimally engaged in practising problem solving. As intoned by Riccomini, Smith, Hughes & Fries. (2015), “the best way to learn and master mathematics concepts is through practice and repetition”. Consequently, practice of mathematics rules through procedural knowledge is key to building fluency in test taking skills. In the same vein, students' practice of worked examples could help them in developing skills to know what procedure might be most appropriate in a given situation.

Also, the findings of the study established that before-test skill contributed more than during and after test skills in predicting mathematics performance in multiple-choice tests. The practice before a test allows for effective preparation and promotes confidence building. Self-confidence in mathematics problem solving that is developed through practice of related problems is found to play a significant role in mathematics achievement (Mohd, Mahmood & Ismail**,** 2011; Smith & Smith, 2002). Doodeen (2005) and Dodeen (2014) confirmed that appropriate use of test-preparation skills is important for getting the knowledge or information to answer related test questions.

In addition, the results showed that differences in performance in mathematics are explained by the variance in the reported test-taking scores of students. The high-scorers reported higher use of before-test, during-test, and after-test skills than the average and low scorers. Other studies (Gbafournia, 2013; Stenlund *et al*, 2017; Hong *et al*, 2006) also reported that high achievers use more of test-taking strategies than the low achievers. High achievers are more likely to develop the motivation to engage more regularly with mathematics tasks; motivation to continue with engagement on a task is related to achievement level. Therefore, the implication of the outcomes of this study is that teachers must teach testing skills as part of classroom instructional activities in mathematics to support the students, especially weak ones, through constant practice and regular engagement with learning materials. This will help students to develop the confidence to express their knowledge in testing situations and to achieve acceptable levels of performance in mathematics.

Conclusion

The outcomes of this study provided further evidence to show that students' performance in mathematics is affected by their test-taking behaviour. The reported before-test and during-test skills contributed significant variance in prediction of mathematics performance on the test type at both SSII and SSIII levels. These outcomes have implication for improving the instructional process in mathematics in secondary schools. The consistent worrisome performance in mathematics examination may not be because students have not learned the subject content well enough, rather it may be that students lack the procedural knowledge (skills) for answering the questions.

5.3 Recommendations

The following recommendations are based on the findings of this study:

1. Mathematics teachers should integrate teaching of appropriate testing skills as they teach the subject content to enable students, especially weak ones, effectively express their knowledge in classroom.

2. Teachers need to learn to use low stake tests and practice exercises more regularly to enable students have more interaction with the subject content and help them develop good study habits.

3. Examination bodies should take special interest in helping teachers to develop the appropriate assessment skills that will make them support their students to overcome conceptual difficulties and, in turn, improve learning.

References

Ajayi, B. K. & Omirin, M. S. (2013). The effect of number right and corrected scoring methods on multiple choice agricultural science test scores. *International Journal of Educational Sciences,* 5(3), 345-348.

Amajuoyi, I. J. Joseph, E. U. & Udoh, N. A. (2013). Content validity of May/June West African Senior School Certificate Examinations (WASSCE) questions in chemistry. *Journal of Education and Practice,* 4(7), 15-21.

Bensley, D. A. Rainey, C. Murtagh, M. P. Flinn, J. A. Maschiocchi, C. Bernhardt, P. C. & Kuehne, S. (2016). Closing the assessment loop on critical thinking: The challenges of multidimensional testing and low test-taking motivation. *Thinking Skills and Creativity,* 21, 158-168.

Biçak, B. (2013). Scale for test preparation and test taking strategies. *Educational Sciences: Theory and Practice,* 13(1), 279-289.

DeKeyser, R. M. (2007). Introduction: Situating the concept of practice. In R. M. DeKeyser (Ed.), *Practice in a Second Language. Perspectives from Applied linguistics and Cognitive psychology* (pp. 1-18). Cambridge, England: Cambridge University Press..

Dodeen, H. (2008). Assessing test-taking strategies of university students: Developing a scale and estimating its psychometric indices. *Assessment & Evaluation in Higher Education,* 33(4), 409-419

Dodeen, H. M. Abdelfattah, F. & Alshumrani, S. (2014). Test-taking skills of secondary students: The relationship with motivation, attitudes, anxiety and attitudes towards tests. *South African Journal of Education,* 34(2), 01-18.

Dodeen, H. (2009). Test-related characteristics of UAEU students: Test-anxiety, test-taking skills, guessing, attitudes toward tests, and cheating. *Journal of Faculty of Education,* 26(1), 31-66.

Dodeen, H. (2014). Teaching test-taking strategies: Importance and techniques. *Psychology Research,* 5(2), 108-113.

Fook, C. Y. & Sidhu, G. K. (2010). Authentic assessment and pedagogical strategies in higher education. *Journal of Social Sciences,* 6(2), 153-161.

Gbafournia, N. (2013). The relationship between using multiple-choice test-taking strategies and general language proficiency levels. *Procedia - Social and Behavioural Sciences,* 70, 90 – 94

Haladyna, T. M. (2012). *Developing and Validating Multiple-choice test items*. NY: Routledge.

Hong, E. Sas, M., & Sas, J. C. (2006). Test-taking strategies of high and low mathematics achievers. *The Journal of Educational Research,* 99(3), 144-155.

Lewandowski, L. J. Berger, C. Lovett, B. J. & Gordon, M. (2016). Test-taking skills of high school students with and without learning disabilities. *Journal of Psychoeducational Assessment,* 34(6), 566-576.

Mohd, N. Mahmood, T. F. P. T. & Ismail, M. N. (2011). Factors that influence students in mathematics achievement. *International Journal of Academic Research,* 3(3), 49-54.

Popham, W. J. (2010). *Everything School Leaders Need to Know about Assessment.* CA, USA: Corwin Press.

Riccomini, P. J. Smith, G. W. Hughes, E. M. & Fries, K. M. (2015). The language of mathematics: The importance of teaching and learning mathematical vocabulary. *Reading & Writing Quarterly,* 31(3), 235-252.

Rohrer, D. (2009). The effects of spacing and mixing practice problems. *Journal for Research in Mathematics Education,* 40, 4-17.

Rupp, A. A. Ferne, T. & Choi, H. (2006). How assessing reading comprehension with multiple-choice questions shapes the construct: A cognitive processing perspective. *Language Testing,* 23(4), 441-474.

Scully, D. (2017). Constructing multiple-choice items to measure higher-order thinking. *Practical Assessment, Research & Evaluation, 22, Article 4*. Available at https:///schorlarworks.umass.edu/pare/vol22/iss1/4

Smith, L. F. & Smith, J. K. (2002). Relation of test-specific motivation and anxiety to test performance. *Psychological Reports,* 91(3), 1011-1021.

Stenlund, T. Eklöf, H. & Lyrén, P. E. (2017). Group differences in test-taking behaviour: An example from a high-stakes testing programme. *Assessment in Education: Principles, Policy & Practice,* 24(1), 4-20.

Tarrant, M. Knierim, A. Hayes, S. K. & Ware, J. (2006). The frequency of item writing flaws in multiple-choice questions used in high stakes nursing assessments. *Nurse Education Today*, 26(8), 662-671.