EFFECT OF FORMATIVE TESTING ON STUDENTS ACHIEVEMENT IN JUNIOR SECONDARY SCHOOL MATHEMATICS

Ajogbeje Oke James, PhD

Department of Mathematics, College of Education, Ikere – Ekiti, Ekiti State Nigeria

Abstract

The concern about improving the outcome of mathematics instruction requires that concerted efforts be made to find out an instructional strategy that can be used and operated upon in order to improve the quality of secondary school graduates in mathematics. Hence, this study investigated the effect of formative testing on students' achievement in junior secondary school mathematics. The research design is quasi – experimental design. The sample consisted of 312 JSS II students assigned to three experimental groups and one control group. Four instruments including three formative tests and Mathematics Achievement Test (MAT) were constructed, validated, and used for the collection of all revelant data. The data collected were analyzed using t – test and Analysis of Covariance (ANCOVA). The results of the study revealed that there was a significant difference in the level of achievement of members of the four treatment groups in their posttest scores after correcting for initial group differences. The formative test with feedback and remediation group performed better than the other three groups. The results of the study also revealed that male students did not perform better than female students in the posttest score.

Keywords: Formative Test, Remediation, Feedback, Gender, Achievement in Mathematics, Junior School Mathematics

1. Introduction

The consistent mass failures of most secondary school leavers in May/June examinations conducted by West African Examination Council (WAEC), National Examination Council (NECO) and National Business and Technical Examination Board (NABTEB) prompted the Federal Government to set up a panel to investigate the reported mass failure of students in the 2009 Senior Secondary Certificate Examination (SSCE). Of all

the candidates who sat for the 2009 May/June SSCE, 84% failed (Information on Nigeria Education, 2009). Theimport of this on the candidates' future and the nation's manpower development should be a cause for concern for the country's leaders, stakeholders in the education industry and the nation as a whole. Different researchers have identified different factors adduced as being responsible for the consistent poor performance of students in mathematics. These include among others lack of proper digestion and utilization of research findings by mathematics teachers, sex – stereotyping, transfer of poor attitudes of older students to the younger ones, and poor self-concept towards mathematics, instructional/classroom characteristics, societal factors and school factors (Nwoji, 1999);teachers' characteristics (Onocha and Okpala, 1985); anxiety, motivation, reasoning ability, problem solving skills and instructional strategy (Udousoro, 2000).

From the foregoing, certain key factors emerge which seem to contribute more to the problem of poor achievement in mathematics. The essence of using tests and other evaluation instruments during the instructional process is to guide, direct and monitor students' learning and progress towards attainment of course objectives (Alonge, 2004; Kolawole, 2010). The utilization of formative testing in the teaching – learning process involve breaking up the subject matter content or course into smaller hierarchical units for instruction; specifying objectives for each units; designing and administration of validated formative test; offering a group based remediation in areas where students are deficient before moving to another units and then administration of summative test on completion of all units.The breaking up of subject or course into small units makes for adequate preparation for the test by the students. Moreover, such frequent testing enables the student to get more involved and committed to the teaching –learning process thereby enhancing their performance (Bandura, 1982). Bloom, Hastings and Madaus (1971) opined that formative evaluation is useful to both the students (as a way of diagnosing students' learning difficulties and the prescription of alternative remedial measures) and to the teacher (as means of locating the specific difficulties that the students are experiencing within subject matter content and forecast summative evaluation result).

According to Gronlund and Linn (1990), formative evaluation serves three specific uses namely: (i) to plan corrective action for overcoming learning deficiencies; (ii) to aid in motivating learners and (iii) to increase retention and transfer of learning. According to them, students' responses to a formative test could be analyzed to reveal group and individual errors needing correction. Hence, formative testing is a strategy designed to identify learners' learning difficulties with a view to providing remediation measures to enhance the

performance of majority of students. The operations of the school are legally defined and structured around inflexible units of time as a result of the examination system that is operated. Kok – Aun, Toh and Brian (1990) corroborated this position and opined that the busy routine of teachers in the discharge of their duties does not permit the luxury of the ideal one – to – one observation of the students.

Some researchers have used strategies that can be seen as components of mastery learning like the use of feedback and remedial instructions (Burrows and Okey, 1979; Afemikhe, 1985;Erinosho, 1988; Saido, 1989; Ughamadu, 1990; Odulaja, 1993 and Ajogbeje, 2012). Okey (1977) and Godson and Okey (1978) from different studies found that the utilization of diagnostic tests with remediation in appraising learning weaknesses enhances the acquisition and retention learning tasks among students. Pizzini, Treagust and Cody (1982) also established in their study aimed at determining whether or not formative evaluation can be effective or could facilitate goal attainment in a biochemistry course, that the use of formative evaluation can be effective in producing desired learning outcomes to facilitate goal attainment.

Similarly, Alonge (1986) had reported that the result of investigation into the extent to which cognitive entry characteristics and formative evaluation measured students' academic performance among University undergraduates show that formative evaluation has the highest predictive strength to academic achievement out of all variables, that is, certificate worth and Joint Admission and Matriculation Board (JAMB) results considered. In a similar study carried out among Polytechnic students,Ajogbeje (1998) reported that cognitive entry characteristics [West African School Certificate (WASC) and Polytechnics and Colleges Entrance Examination (PCEE)] are not significantly related to academic achievement of Polytechnic students in mathematics and that most of the students with good grades in WASC and PCEE examinations often times rely too much on these results which, in turn, affect their academic achievement. However, the study revealed that semester results [i.e. continuous assessment scores] are the best predictors of academic achievement in mathematics.

Ughamadu (1990) in his study on the interactive effect of formative testing and cognitive style on students' learning outcomes in secondary school chemistry found that analytical students exposed to formative testing with remediation performed significantly higher in composite concept attainment at classification and formal level than global students.However, in a study carried out on continuous assessment as predictors of students' grades SSCE Chemistry, Oluwatayo (2007) reported that formative test (continuous assessment scores) are weak predictors of excellent grades in SSCE Chemistry. One would

wonder why continuous assessment scores do not predict senior school certificate if formative test is effectively conducted. Common to all these studies is the fact that formative testing allows for a diagnosis of the learners' learning difficulties. However, there are variations in the efficacy of the strategies adopted in the studies. In addition, some of the strategies are less applicable because of some obstacles inherent in mastery learning. The present study therefore is an attempt to investigate the effect formative testing and gender on students' achievement in junior secondary school mathematics.

3. Research Hypotheses

In order to address the above stated problem, the research study was designed to test the validity or otherwise of the following propositions:

HO1: There is no significant difference in the academic achievement of students in the experimental and control groups in their posttest scores in mathematics.

HO2: There is no significant difference in the students' achievement from different treatment groups in their posttest scores in mathematics.

H03: There is no significant difference in the academic achievement of male and female students
in the experimental and control groups in their posttest scores in mathematics.

4. Research Method

The study employed quasi – experimental design. The sample for the study consisted of 312 students [138 males and 174 females] drawn from four junior secondary schools in Akure South Local Government Area of Ondo State, Nigeria. These four schools which are co – educational and operating the same mathematics syllabus were selected using purposive sampling technique. Each school acted as a treatment group. The three experimental groups namely formative test with feedback and remediation group, formative test with feedback group and formative test only group were exposed to expository class teaching followed by a formative class test with feedback and remediation, expository class teaching followed by formative class test and feedback and expository class teaching and formative class test respectively. The fourth group (non – formative test or control group) was also exposed to the expository class teaching without class test, feedback and remediation after each unit.

Four instruments namely Formative Test I, II and III(which were administered on the respondents after the coverage of each selected topic during treatment) and Mathematics

Achievement Test (MAT) which served as pretest and posttest to the respondents on the topics covered during treatment were used to collect all the revelant data for the study. The MAT, the Formative Test I, II and IIIwere reviewed and vetted for face and content validities by two experienced junior secondary school mathematics teachers and two test experts in the area of test construction with bias in mathematics. Kuder Richardson formula 21 (KR_{21}) was used to establish a reliability coefficient estimate of 0.72 for MAT, 0.82, 0.78 and 0.75 for the formative tests I, II and III respectively. The data collected were subjected to t – test and Analysis of Covariance (ANCOVA) to test the rejection or otherwise of the stated hypotheses at 0.05 level of significance. Multiple Classification Analysis (MCA) test was used on significant variables to find out the magnitude of differences among the groups while Scheffe's Post Hoc analysis was used where a null hypothesis was rejected.

5. **Results**

The results of the data analysis carried out are presented below.

Hypothesis one was aimed at determining whether formative evaluation would result in significantly higher achievement or not. The mean scores and standard deviations of the posttest scores are shown in table 1.

Table 1: Mean and Standard Deviation of PosttestScores for the TreatmentGroups

Groups	Formative Test With Feedback & Remediation		Formative Test With Feedback		FormativeTest Only		Non- Formative Test	
	Mean	S.D	Mean	S.D	Mean	S.D	Mean	S.D
Pretest	17.27	3.79	15.26	4.80	15.10	4.72	13.50	3.44
Posttest	27.45	3.38	21.20	4.56	17.09	3.96	14.43	3.34

Table 1 revealed that on posttest scores, the formative test with feedback and remediation group had a mean score of 27.45, formative test with feedback group had 21.20 while formative test group only obtained 17.09 and non – formative test group (control group) had 14.43. The mean scores of the experimental group (that is, those groups whose students undertook treatments which were based on the various components of formative evaluation) and the control group were then compared using t – test.

Table 2: Mean Score of Experimental and Control Groups PosttestScores in Mathematics.

Groups	N	Mean	S.D	df	t_{cal}	t_{tab}
Experimental (Feedback and Remediation + Feedback without Remediation + Formative Test Only)	238	22.20	5.82	310	10.94*	1.96
Non – Formative Test (Control)	74	14.43	3.34			

From table 2 the $t_{cal}>t_{tab}$ showing that the observed difference in the performance of students is in favour of the experimental group. The result in table 2 further sho that the treatment given was significantly effective to the observed higher achievement of the experimental group on achievement in mathematics. The hypothesis was therefore rejected and it may be concluded that formative evaluation significantly results in higher achievement in mathematics. This finding becomes more meaningful when the performances of students in each experimental subgroup were compared with that of the control group on measure of achievement in mathematics. It was therefore possible to assess the significance of evaluation component to the result obtained.The results in rows two and three in table 3 showed that the $t_{cal}> t_{tab}$. The observed difference in the mean score of students on the measure of achievement in mathematics was in the favour of the formative test with feedback and remediation group. It may therefore be concluded that expository teaching with formative test, feedback and remediation treatment was significantly effective in the observed higher achievement of the experimental group on measure of achievement in mathematics.

Table 3: Comparison ofMean Score of each of the Experimental and Control Groups on Posttest Scores in Mathematics

Subgroups	N	Mean	S.D	df	t_{cal}	t_{tab}
Feedback and Remediation	85	27.45	3.38	157	24.39*	1.96
Control	74	14.43	3.34			
Feedback without Remediation	82	21.20	4.56	154	10.48*	1.96
Control	74	14.43	3.34			
Formative Test Only	71	17.08	3.96	143	4.37*	1.96
Control	74	14.43	3.34			

A further investigation between the performance of the students who received expository teaching with formative test and feedback and the control group as shown in rows four and five of table 3 revealed that the $t_{cal}> t_{tab}$ and the observed difference in the mean score on the measure of academic achievement in mathematics was in favour of the formative test with feedback group. This indicates that the expository teaching with formative test and

feedback treatment was significantly effective to the observed higher achievement of the experimental group on measure of achievement in mathematics. Columns six and seven in table 3 also revealed that the $t_{cal} > t_{tab}$ and the observed difference in the mean score on the measure of academic achievement in mathematics was in favour of the formative test only group. This indicates that the expository teaching with formative test only treatment was significantly effective to the observed higher achievement of the formative test only group on measure of achievement in mathematics. It can therefore be concluded that the administration of treatment (expository teaching with the three components of formative evaluation) was more effective than instruction only in enhancing the performance of students in mathematics.

The focus of hypothesis two was to determine the effect which the different treatments (formative test with feedback and remediation, formative test with feedback and formative test only) had on the student's achievement in the posttest.

Table 4: ANCOVA of the Posttest Scores According to Treatment Groups

Source	Sum of Squares	Df	Mean Square	F – cal.	Sig.
Corrected Model	7736.157	4	1934.039	131.955*	.000
Intercepts	7516.848	1	7516.848	512.857*	.000
Pretest	47.907	1	47.907	3.269	.072
Treatment	6672.181	3	2224.060	151.743*	.000
Error	4499.638	307	14.657		
Corrected Total	12235.795	311			

The summary of ANCOVA in table 4 shows that the main effect of treatment on achievement in mathematics was significant [F (3, 307) = 151.743, P < 0.05]. The F – calculated of 151.743 was significant. The data was further subjected to multiple classification analysis (MCA) in order to determine the magnitude and direction of the effect as presented in table 5. The control group has an adjusted mean of 15.21 while the formative test with feedback and remediation group had 26.64.The formative test with feedback group had 21.24 and the formative test only group had 17.19. The multiple R^2 in table 5 reveals that only 8.7% of the variation of the posttest scores is accounted for by the different treatment strategies.

Table 5: MCA of Posttest ScoresAccording toTreatments

Variable + Category	N	Unadjusted Deviation	Eta	Adjusted for Independent + Covariate	Beta	Adjusted Mean
Treatment						
Feedback with Remediation	85	7.09		6.28		26.64
Feedback without Remediation	82	0.84		0.88		21.24
Formative Test Only	71	-3.28		-3.18		17.19

Control	74	-5.93	0.28	-5.15	0.30	15.21
Multiple R^2						0.087
Multiple R						0.295

Grand Mean = 20.36

In order to determine the treatment condition that caused the rejection of the null hypothesis, Scheffe's Post Hoc analysis (as shown in table 6) was carried out on the adjusted mean scores of the four groups. The result showed that the formative test with feedback and remediation group achievement was significantly higher than that of the formative test with feedback group, formative test only group and non – formative test group. The formative test with feedback group also achieved significantly better than the formative test only group and non – formative test group. While the formative test only group equally achieved significantly better than the non – formative test group. The non – formative test group has the least effect over other groups.

Table 6: Scheffe's Post Hoc Analysisfor PosttestScores ofTreatment Groups

Groups	Mean Score	Feedback&Remediation	Feedback no Remediation	Formative Test Only	No Formative Test
Feedback & Remediation	27.45				
Feedback no Remediation	21.20	*			
Formative Test Only	17.08	*	*		
Non - Formative Test	14.43	*	*	*	

Therefore, the null hypothesis which stated that there was no significant difference in the achievement of students from different treatment groups in their posttest scores in mathematics was rejected since significant differences existed between the groups.

Hypothesis three intends to find out the effect which gender had on the student's achievement in the posttest. To test this hypothesis, ANCOVA was computed to correct for differences that might exist at pretest level among the subjects.

Table 6:ANCOVA of the Posttest Scores According to Gender

Source	Sum of Squares	df	Mean Square	F – cal.	Sig.
Corrected Model	1064.169	2	532.085	14.717*	.000
Intercepts	4556.451	1	4556.451	126.029*	.000
Pretest	1062.819	1	1062.819	29.397*	.000
Gender	0.194	1	0.194	0.005	.942
Error	11171.626	309	36.154		
Corrected Total	12235.795	311			

The summary of ANCOVA presented in table 6 showed that the effect of gender on achievement in mathematics was not significant F (1, 309) = 0.005, P > 0.05. The obtained F – calculated of 0.005 was not significant. Therefore, the null hypothesis which stated that there was no significant difference in the academic achievement of male and female students in the experimental and control was not rejected since there is no significant differences between the groups.

6. Discussion

The result of the study showed that formative evaluation (viz, formative test, feedback and remediation) enhanced the performance of students. This supported previous findings which established the effectiveness of formative evaluation in improving performance (Afemikhe, 1985; Erinosho, 1988, Ajogbeje, 2012). This outcome could be explained in terms of the feedback and remediation which the students received. The feedback and remediation treatment must have helped all the students who benefited from formative evaluation treatments. The poor performance which was recorded in respect of the non – formative test group may be due to the fact that they did not have opportunities to explore their problems. Forthe significant difference in academic achievement between the four treatment groups, the outcome for formative test with feedback and remediation was not unexpected. Remediation was expected to help in correcting the mistakes made.

This result was in agreement with studies conducted by Okey (1977), Godson and Okey (1978), Burrow and Okey (1979), Afemikhe (1985), Erinosho (1988) and Ajogbeje (2012) which found that students exposed to formative testing with remediation achieved higher than students exposed to formative testing with feedback only and the students exposed to only instruction without formative testing in mathematics. Similarly, the outcome of formative test with feedback also agreed with the findings of Bridgeman (1974), Bardwell (1981) and Ajogbeje (2012) that feedback from tests motivates students intrinsically. Thus, a person who is informed of his successful performance on a test would begin to develop interest in the area and explore means by which he will continue to do well in subsequent tasks.

All the same, the finding contrasted the finding of Erinosho (1988) that formative test with feedback treatment was not more effective than formative test treatment only. The outcome of formative test treatment only group also agreed with the findings of Saudargas, Madsen and Scott (1977), Pizzini, Treagust and Cody (1982), and Ajogbeje (2012) but contrasted that of Erinosho (1988). The findings of Saudargas, Madsen and Scott (1977) and Ajogbeje (2012) showed that the use of formative testing led to effective outcome in that

students developed more consistent study habits when they were tested daily than when they were tested weekly or within three weeks intervals. The non – significant result obtained when gender was considered agreed with the findings of Afemikhe (1985), Oladunni (1995) and Ajogbeje (2012) which found no gender differences in the junior high school. However, the result does not agreed with those studies carried out by Campbell and Beaudry (1998) and U. S. Department of Education (2000, 2001) which found sex – related differences in mathematics achievement.

Conclusion

Based on the findings of this study, it could be concluded that when formative tests are used for diagnostic purposes, the cognitive results obtained are usually better than when given as a series of summative tests. This was the case when the results of the formative tests served as a basis for finding out the sources of difficulties. In this way, the teacher is able to give necessary remediation and correctives. It was recommended that school administrators should emphasize to their teachers on regular basis that the teaching of mathematics in junior secondary schools, should be carried out by providing regular diagnostic tests and adequate feedback and remediation for the learners. They should allow and provide incentives for teachers to attend seminars, workshops, conferences and in – service trainings to enhance their performances and acquire necessary skills for constructing formative tests, and how to blend formative evaluation with classroom instruction procedures. Curriculum designers should take into cognizance while designing the learning tasks for learners that learning in mathematics is not solely a cognitive affair. Hence, mathematics curriculum should be designed to include the use of methods/strategies and material/media which would make the learning of mathematics very active, investigative and adventurous.

References:

Afemikhe, O. A. (1985). The effect of formative testing on students'achievementin secondary school mathematics, Unpublished Ph. D.Thesis, University ofIbadan, Ibadan.

Ajogbeje, O. J. (1998). Cognitive entry characteristics and formative evaluationasa predictor ofacademic achievement in mathematics among polytechnicstudents, Unpublished M. Ed.Thesis, University of Ado – Ekiti, Nigeria.

Ajogbeje, O. J. (2012). Path – analytic model and the effect of some teaching strategies on variables affecting achievement in junior secondary school mathematics in Ondo state. Unpublished Ph. D. Thesis, Ekiti State University, Ado – Ekiti, Nigeria.

Alonge, M. F. (1986). Cognitive entry characteristics and formative evaluationasmeasures ofacademic performance among university undergraduates,*African Journal of EducationalResearch*, 1(1), 103 – 107.

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

Bandura, A. (1982). Self efficacy mechanism in human agency, *AmericanPsychologist,* 77(2), 122 – 147.

Bardwell, R. (1981). Feedback: How does it function? *Journal of ExperimentalEducation,* 50(1), 87 – 95.

Bloom, B. S.; Hastings, J. T. & Madaus, G. F. (1971). *Handbook on formativeand summativeevaluation of students learning.* New York: McGraw – Hill.

Bridgeman, B.(1974). Effects of test performance feedback on immediatelysubsequent test performance, *Journal of Educational Psychology*, 6(1), 62–66.

Burrows, C. & Okey, J. R. (1979). The effects of a mastery learning strategy onachievement,*Journal of Research in Science Teaching*, 16(1), 33 – 37.

Campbell, J. R. & Beaudry, J. S. (1998). Gender gap linked to differentialsocialization for high– achieving senior mathematics students. *Journal ofEducational Research*, 91, 140 – 148.

Erinosho, S. Y. (1988). The effect of formative evaluation on the performanceofstudents in physics, Unpublished Ph. D. Thesis, University of Ibadan,Ibadan.

Gronlund, N. E. & Linn, R. L. (1990). *Measurement and evaluation in teaching(6th Edition),* New York: Macmillan.

Information on Nigeria Education (2009). Presidency probes mass failure inSSSCE, NECO exams, Nigerian News Nov 3, 2009, retrieved on December3, 2010. http://www.informationnigeria.org/2009/11/presidency-probes-mass-fa...

Kolawole, E. B.(2010). *Principles of test construction and administration(RevisedEdition).*

Lagos: Bolabay Publications

Kok – Aun, Toh & Brian E. W. (1990). Measuring students performance,*Educational Research,*32(1), 59 – 65.

Nwoji, Q. J. R. (1999). Evaluating the use of learning resources for primaryscience education: Implication for the learner.*STAN Annual ConferenceProceedings,* 245 – 249.

Okey, J. R. (1977). Consequences of training teachers to use mastery learningstrategy, *Journal of Teacher Education*, 28, 57 – 62.

Okpala, P. N.& Onocha, C.O. (1988). *Measurement and Evaluation inEducation,*Jattu–Uzairue:Sterling – Harden Publishers.

Oladunni, M.O. (1995). Effects of mathematics language and problem solvingstrategies on achievement of students in mathematics, *Journal ofEducational Research and Evaluation,*1(1), 153 – 163.

Oluwatayo, J.O. (2007). Continuous assessment scores as predictors of studentsgrade in senior school certificate chemistry examination, *Journal ofResearchin Education,* InternationalResearch and Development Institute,4(20), 81 – 84.

Pizzini, E. L.; Treagust, D. F. & Cody, J. (1982). Utilizing formative evaluationtoenhance the understanding of chemistry and the methods and proceduresofscience, *Journal of Researchin Science Teaching*, 19(9), 769 – 774.

Saudarges, R. N.; Madsen C. H. & Scott J. W. (1977). Differential effects offixedand variable –time feedback on production rates of elementaryschoolchildren, *Journal of AppliedBehaviour Analysis,* 10(1), 673 – 678.

Udousoro, U. J. (2000). The relative effects of computer and text – assistedprogrammed instruction on students' learning outcomes in mathematics,Unpublished Ph. D. Thesis,University of Ibadan, Ibadan.

Ughamadu, K. A. (1990). Interactive effect of formative testing and cognitivestyleon students learning outcomes in senior secondary school chemistry,Unpublished Ph. D. Thesis,University of Ibadan, Ibadan.

U. S. Department of Education (2000). *The condition of education 2000(NCES2000 – 062)*. Washington, DC: National Center for Education Statistics,Office of EducationalResearch and Improvement.

U. S. Department of Education (2001). *The nation's report card: Mathematicshighlights 2000 (NCES 2001 – 518)*. Washington, DC: National Center forEducation Statistics, Office ofEducational Research and Improvement.