International Journal of Education, Learning and Development
Vol. 9, No.10, pp.35-42, 2021
Print ISSN: 2054-6297(Print)
Online ISSN: 2054-6300 (Online)

THE EFFECT OF TIME OF THE DAY AND DURATION ON STUDENTS ACHIEVEMENT IN MATHEMATICS

Denis Achung Uyanah Ph.D.
Department of Curriculum and Instructional Technology, University of Cross River State, Calabar

Henry Ojating, Ph.D
Department of Educational Foundations and Childhood Education University of Cross River State, Calabar

ABSTRACT: *The purpose of this study was to investigate the effect of time of the day and duration of a test on secondary school students' mathematic achievement. A 3 x 4 factorial balanced design was applied. A 50-item multiple choice mathematics achievement test and two equivalent forms of it were administered on a sample of 488 students, randomly assigned to 12 treatment combinations of time of the day (morning, afternoon and evening) and duration (45 min, 60 mins., 75mins, 90mins). The parallel form reliabilities were* $r_{AB} = .823, r_{AC} = .891$ *and* $r_{BC} = .856$*, where form A was administered in the morning, form B in the afternoon and form C in the evening. The students mathematics achievement scores were recorded as number right. The resulting data were analyzed using two-way ANOVA while F-ratio and LSD tests were used to test for significance of effect. The results showed that test duration, time of the day and their interaction effects on students' mathematics achievement, were significant, with those who took the test for 90 mins and in the afternoon, obtaining the highest mean score (*$\bar{x} = 57.974$*), while those who took the test for 45mins and in the evening obtaining the least mean mathematics achievement score (*$\bar{x} = 19.535$*). The implications of those results to test designers and administrators are discussed.*

KEYWORDS: Test length, duration of test, mathematics achievement, unbanced, completely randomized, factorial designs.

INTRODUCTION

One of the issues that has been of consistent concern to all child education stakeholders, is how best to explain the performance of students in examinations, internal or external. Some explained students achievement in terms of students' characteristics, some use teachers' characteristics, some home variables, some school environment, some examination body related variables, some test related variables like test difficulty, validity, reliability, clarity of test items etc.

Still, some others try to explain students' performance using the influence of some external factors on some variables internal to the students that tend to trigger internal reactions that either inhibit or enhance achievement behaviour. Interestingly, before any examination, students, parents and

@ECRTD-UK https://www.eajournals.org/
https://doi.org/10.37745/ijeld.2013

teachers are seen engaged in all kinds of activities, in an attempt to ensure students' success. Every factor perceived to have positive influence on achievement is exploited to the fullest. Efforts are made to eliminate those factors whose influences are negative. Even with all these, performance, especially in mathematics, continue to decline (Uyanah, 2019). At the end of examination students are heard most of the times, complaining that the invigilator did not allow them enough time to answer all the questions they knew the correct answers. Some complain about the time allowed by the examiner, saying that given the difficulty level of the items, more time ought to have been given. They also complain of the time of the day, saying mathematics examinations should be either in the morning or evening hours, when the atmospheric temperatures are low, believing low temperature enhance cognitive processes. These claims need to be investigated, to provide empirical evidence to substantiate the claims. This is what, this study seeks to do.

Educational administrators and examination time-table planners and some experts in educational and psychological testing, demand that tests that are highly cognitive like mathematics should be administered in the morning when the sun is not high or in the evening when the sun is low and temperatures are low. They argue that in the afternoon, the temperature is generally high, and that such high temperatures do not allow students to concentrate. These argument seem to suggest students' maths achievement is a function of atmospheric temperature typified by time-of-the-day (morning, afternoon or evening).

The implication is that students mathematics achievement depends on the duration of the test, the time of the day and the interaction between these two factors. If this holds, then examiners are wrong to label students whose achievement scores are low, as failures, when in actual fact, the examiners are responsible for their low achievement scores. The effect of length, measured in terms of time allocated by the test designer or examiner has long been a subject of research. Majority of the studies related test length to test reliability (Odo, 2016). Few have looked at the actual score obtained by the students. A study by O'kwu and Anyagh (2010) investigated the effect of timing in mathematics tests on students' achievement. Two multiple choice Mathematics tests (MAT1 and MAT2) based on SS2 scheme of work were administered on 250 students drawn from five secondary schools randomly selected from 30 secondary schools in Makurdi metropolis. Two research hypotheses were raised for the study and tested at .05 level of significance. The student's t-statistic was used to test the hypotheses. One of the findings showed that students who were given a little extra time by invigilators (for MAT1) performed significantly better than the other students.

In an experimental study carried out by Odo (2016), utilizing a 3 x 4 balanced factorial design, with complete randomization, time-of-the-day and duration were factors. A 40-item mathematics achievement test was the data collection instrument. Balancing was done by randomly deleting some group members from the analysis. The results show significant main and interaction effects of time-of-the-day and duration. This study had some short-comings. The same test was given: morning, afternoon and evening. The possibility of those who took the test in the afternoon and

@ECRTD-UK https://www.eajournals.org/
https://doi.org/10.37745/ijeld.2013

evening have pre-knowledge of the test from those who took the test in the morning, cannot be ruled out. This effect will show-up in their score. All the students were in one hall, with duration only indicated boldly on their question paper. This made the hall congested and cheating cannot be ruled out. If these issues were resolved, it is possible that the results would have been different.

Similar results were obtained from a study by Jensen, Berry and Kummer (2013) who examined the effects of exam length on student performance and cognitive fatigue in an undergraduate Biology classroom. The examination tested higher order thinking skills. The findings showed that lengthier exams led to better performance on assessment items.

METHODOLOGY

The research design adopted for this study was completely randomized factorial design. The treatment combinations, A_1, B_1,A_1, B_2….A_3, B_4 were written on 516 (the population of SS_{one} and SS_{two} students in the school) equal size papers folded into small balls. Using a table of random numbers, the students were made to select one ball each. The balls were well shuffled after each selection. Thus, there were 160 for the morning session, 156 afternoon and 172 evening with 28 absentees during testing.

The 488 consisted of 227 boys and 261 girls, age 15 – 18yrs. The instrument for data collection was a 50-item multiple choice achievement test, with four (4) approximately equally plausible options. The test was previously designed based on SS_{one} second term scheme of work and validated by Uwanede (1996) with an internal consistency of .893. Two equivalent (paralled) forms were developed from this test so that one form was administered in the morning, another in the afternoon and another in the evening. The paralled form reliability estimates were predetermined in a trial test as $r_{AB} = .823, r_{AC} = .891$ and $r_{BC} = .856$. With these psychometric properties, the tests were considered good enough for the study.

Having obtained approval from the school principal, 12 science teachers (Biology, Physics, Chemistry, Integrated Science and Mathematics) were entisted to serve a invigilators in 12 different classrooms in all (four morning, four afternoon and four evening). With $1.2m^2$ spacing. It was assumed that the two mathematics teachers were equally effective, covered the entire scheme of work and equally motivated. The two sets of students had all been taught the content of the second term SS_{one} scheme of work. The students were told that the test was one step in process of choosing the most suitable time for future mathematics examinations. The sneaky behaviour of the students resulted in 40 per group in the morning, 39 afternoon and 43 evening. This variation was considered large enough to alter the designed. The scores were recorded as “number right” and two (2) points awarded for each correct answer, then totaled per testee.

@ECRTD-UK https://www.eajournals.org/
https://doi.org/10.37745/ijeld.2013

RESULTS

Two-way analysis of variance (ANOVA) was used in analysis the data while F-ratio and LSD tests were used to test for significance. the descriptive statistics, time-of-the-day by duration are presented in Table 1.

Table 1
Descriptive statistics of students' mathematics achievement: time-of-the-day by duration

Time-of-the-day	Duration	N	Mean	Std.Dev.	Std error
Morning	45 min	40	28 525	7 190	
	60 mins	39	26 775	8 882	
	75 mins	41	46 700	11 960	
	90 mins	40	31 050	8 370	
	Total	**160**	**33 263**	**12 132**	
Afternoon	45 min	39	22 949	10 846	
	60 mins	39	30 436	8 466	
	75 mins	38	49 897	10 799	
	90 mins	40	57 974	13 927	
	Total	**156**	**40 314**	**18 016**	
Evening	45 min	44	19 535	10 096	
	60 mins	42	42 605	11 054	
	75 mins	43	43 977	11 704	
	90 mins	43	47 674	13 652	
	Total	**172**	**38 448**	**16 058**	
Total	45 min	122	23 574	10 149	
	60 mins	122	33 525	11 738	
	75 mins	122	46 762	11 670	
	90 mins	122	45 516	16 403	
	Total	**488**	**37 344**	**15 831**	

The results in Table 1 show that for the morning session, those who took the test for 75 mins had the highest mean score ($\bar{x} = 46.700$) followed by 90 mins group ($\bar{x} = 31.050$) and the least were the 45 mins group ($\bar{x} = 28.525$). In the after session, the 90 mins group had the highest mean score ($\bar{x} = 57.974$) followed by the 75 mins group ($\bar{x} = 49.897$) and the least were the 45 mins groups ($\bar{x} = 22.949$). For the evening session, the pattern of the afternoon session was maintained. With respect to time-of-the-day, those two took the test in the afternoon had the highest mean score ($\bar{x} = 40.314$) followed by the evening group ($\bar{x} = 38.448$) and the least were the morning group ($\bar{x} = 33.263$). With respect to duration those who took the test for 75 mins had the highest mean score ($\bar{x} = 46.762$) followed by the 90 mins group ($\bar{x} = 46.516$)and the least were the 45 mins group ($\bar{x} = 23.574$).

@ECRTD-UK https://www.eajournals.org/
https://doi.org/10.37745/ijeld.2013

The results of the two-way ANOVA of the students mathematics achievement by time-of-the-day and duration, are presented in Table 2.

Table 2
Two-way ANOVA of students' mathematics achievement: time-of-the-day by duration

Source of variation	Sum of squares	df	Mean square	F-value	p-value
Corrected model	66661.467	11	6060.133	52.076*	.000
Intercept	679295.012	1	679295.012	5837.312*	.000
Time-of-the-day	4251.051	2	2125.525	18.265*	.000
Duration	44051.500	3	14683.833	126.181*	.000
Time-of-day x duration	18527.089	6	3087.848	26.534*	.000
Error	55392.697	476	116.371		
Total	802616.000	488			
Corrected total	122054.164	487			

***significant at .05 level P < .05**

The results (Table 2) show that the p-value (.000) associated with the computed F-values (52.076, 5837.312, 18.265, 126.181 & 26.534) for the corrected model, intercept, time-of-the-day, duration and the interaction between time-of-the-day and duration, are all less than .05. Consequently, the null hypotheses concerning the effect of these factors, were rejected. Put together, these results showed that the effects of time-of-the-day, duration and their interaction on secondary school students' mathematics achievement are all significant. To locate the pair of mean mathematics achievement responsible for the observed significant results, least significant difference (LSD) test was applied. The results for time-of-the-day are presented in Table 3.

Table 3
LSD multiple (pairwise) comparison of students' mathematics achievement by time-of-the-day

Time-of-the-day	Morning	Afternoon	Evening
Morning	33.263**	7.052*	5.185*
Afternoon	.000	40.316	1.866
Evening	.000	.118	38.448

***significant at .05 level. P < .05**
****values along main diagonal are group means, above it are mean differences (MD) and below it are corresponding p-values**

From Table 3, only the difference between afternoon and evening means scores was not significant (MD=1.866, P=.118> .05). All other paired comparisons were significant (p<.05).

@ECRTD-UK https://www.eajournals.org/
https://doi.org/10.37745/ijeld.2013

The LSD test results for duration are presented in Table 4

Table 4
LSD multiple (pairwise) comparison of students mathematic achievement by duration of test

Duration	45mins	60mins	75mins	90mins
45mins	23.574**	9.951*	23.189*	21.943*
60mins	.000	33.525	13.238*	11.992*
75mins	.000	.000	46.762	1.246
90mins	.000	.000	.367	45.516

***significant at .05 level. P <.05**
****values along main diagonal are group means, above it are mean differences (MD) and below it are corresponding p-values**

The results showed (Table 4) that only the difference between 75mins mean score and 90mins mean score, was not significant (MD=1.246, p=.367>.05). All other paired comparisons were significant ($99.951 \leq MD \leq 23.189, p = .000 < .05$).

The model parameters were estimated for the main effects of time-of-the-day and duration as well as their interactions were estimated. The results are presented in Table 5.

Table 5
Estimates of model parameters

Time-of-the-dayβ	Duration $\propto_1$ 45mins	$\propto_2$ 60mins	$\propto_3$75mins	90mins $\propto_4$	β_i
Morning β_1	25.615*	.795	19.348*	0	-16.64*
Afternoon β_2	-6.886*	-22.469*	-4.379	0	10.300*
Evening β_3	0	0	0	0	0
$\propto_i$	-28.140*	-5.070*	-.3698	0	

Intercept=47.674*; *significant at .05 level. P <.052. 0411 $\leq t \leq$ 28. 980

The sum of parameter estimates across factor levels for the two factors, are not all zero, while these sums in the population are zero. This means the model is truly a random effect model. This means these parameters, particularly the significant ones can be used in predicting a students' mathematics achievement, given the levels of the factors.

DISCUSSION

The results that the duration of the test has a significant main effect on students' mathematics achievement was expected. These results agree with those of Jenson, Berry and Kummer(2013), Odo (2016) and the position taken by O'kwu and Anyagh (2010). The non-linearity of the relationship between duration and mathematics achievement should be noted. It means that

@ECRTD-UK https://www.eajournals.org/
https://doi.org/10.37745/ijeld.2013

increasing the duration of a test towards infinity does not increase mathematics achievement continuously. Truly, increasing duration of a test does not bring the testee new abilities required to answer the questions correctly. It may rather give way to interperson communications in the examination hall, promotive cheating behaviour. Once what a testee knows is exhausted, there is no way he can improve, unless he cheats.

The relationship rather looks parabolic in nature, with low scores for short duration, approaching maximum when the duration is approximately what the test designer had specified, the decreasing from here as duration increases. This strengthen the results of Uwanede (1996) , Uyanah ( ), Odo (2016).The significant effect of time-of-the-day was quite substantial and agreed with the results observed by Odo (2016) though Odo's (2016) results were not substantial. This limitation in the level of agreement was not anticipated. However, it seems to be so because there has been no agreement both in theory and research results, as to the direction of the effect. Some measurement experts have said that whatever the effect that should have been observed is frequently absorbed by the level of motivation of the testee and the relative importance. The expected effect may been jeopardized by the fact that Calabar weather is generally cool, such that there is virtually no difference in temperature between morning, afternoon and evening and the study was carried out during the rainy season.

The significance of the interaction effect of duration and time-of-the-day was expected but not the way it came out to be. This effect has suggested by theory. The results are specifically instructive to both test designers and administrators, who fixed test duration and time-of-the-day arbitrarily. Test administrators particularly in the habit of cutting short the duration of test. Such action may go down will with the low achievers but not with high achievers. The point further strengthen by the redundancy of the parameters involving evening and 90mins for a 60mins test. This means that these extremes should not be advocated at all. There are very scanty literature, both theoretical and empirical that suggests this interaction effect. For this reason alone, a replication is needed, with location and ability level of the students built into the design.

References

Jenson, J. L., Berry, D.A., and Kummer, T. A. (2013). Investigating the effects of exam length on performance and cognitive fatigue. PLoS ONE 8(8): e70270.

Uwanede, A. (1996). Effect of number of response options and test length on achievement at mathematics examination. Being a HND thesis submitted to the Department of Statistics, the Polytechnic, Calabar.

Odo, E. O. (2016). Analysis of students mathematics achievement by duration and time-of-the-day. A B.Ed. thesis submitted to Department of Educational Foundation and Administration, Cross River University of Technology, Calabar.

O'kwu, E. I and Anyagh, P. I. (2010). Effect of timing in mathematics tests on students' achievement in Makurdi metropolis. *Multidisciplinary journal of research development.* 14(1), 110-114.

@ECRTD-UK https://www.eajournals.org/
https://doi.org/10.37745/ijeld.2013

Rindler, S. E. (1980). Effect skipping over more difficult items on time-limited tests: implication for test validity. *Journal of Educational and Psychological Measurement,* 40,1140 -1148.

Uyanah, D. A. (2019). *Application of SQC models in analysing students performance in WAEC – conducted SSCE examinations* (1982 - 2016). A Ph.D thesis submitted to the Graduate School, University of Calabar, Calabar, Nigeria.

Uyanah, D. A. (1996). *Correlates of the reliability of affective and cognitive measures.* A masters thesis submitted to the graduate school, University of Calabar, Calabar, Nigeria.

Uyanah, D. A. (2003). Achievement test score as a function of number of response options and test length. *Nigerian Education Journal*, 5(1), 20-27.