Journal of Education and Practice www.iiste.org
ISSN 2222-1735 (Paper) ISSN 2222-288X (Online)
Vol.9, No.23, 2018

Pupils' Academic Achievement in Genetics After Hands-on Teaching: Multiple-Choice Testing Versus Performance-Based Testing

Muma Elias[1*] Prof. Zulu J.N[2]
1.Mukuba University, School of Natural Sciences, P.O. Box 20382, Kitwe, Zambia
2.University of Zambia, School of Natural Sciences, P.O. Box 32379, Lusaka, Zambia

Abstract
Pupils' academic achievement in genetics can be assessed using different assessment techniques. Effective assessment procedures are important for most biology teachers who usually face challenges to identify appropriate assessment tools to improve the teaching and learning of genetics in schools. However, literature regarding choice of a suitable assessment tool with respect to other tools available in measuring pupils' academic achievement in genetics is unavailable. It is against this background that this study sought to determine the correlation between multiple-choice testing (MCT) and performance-based testing (PBT) on pupils' academic achievement in genetics in the Copperbelt province of Zambia after hands-on teaching. The study showed a statistically significant difference in the effects of multiple-choice testing on pupils' academic achievement in genetics from those of performance-based testing: Pupils performed better in multiple choice tests than in the performance-based test. The study also demonstrated that there was a weak correlation between multiple-choice testing and performance-based testing regarding pupils' academic achievements in genetics for pupils taught genetics using hands-on method and those taught the subject using the conventional-based methods. The study further showed a small effect size (coefficient of determination, R^2 value) between MCT and PBT for all the four groups under study with the highest being 0.112 for the pre-test experimental group, E1. Based on the results, the study recommended that biology teachers should be employing MCT alongside PBT in assessing pupils' academic achievement in genetics.
Keywords: Academic achievement, Multiple-choice testing, Performance-based testing, Hands-on teaching.

1. Introduction
Assessment is considered as one of the challenging areas in genetics education. Effective assessment procedure is one that can be used to improve student's learning. In view of this, biology teachers are expected to use appropriate assessment tools such as multiple choice testing and performance-based testing which would help them measure the actual learning-teaching outcomes of genetics. The rest of this section, describes the background to the study, statement of the problem, purpose of the study and the hypotheses the study investigated.

1.1 Background
Assessment of pupils' academic achievement in genetics is generally recognized as an integral part of the teaching and learning of genetics in schools. According to Ahmad-Fuad (2005), assessment drives learning and learning drives practice. Assessment of pupils' achievement in genetics is therefore an important element of genetics education. There are two different assessment strategies that are used by teachers to measure pupils' achievements and progress in genetics education; namely, traditional assessment and alternative assessment. Examples of traditional assessment include multiple-choice testing (MCT). As for alternative assessments, they include open-ended questions, hands-on execution of practical work and portfolios (Dietel et *al*., 1991). Performance-based assessment is described by two concepts; namely, performance and authentic. On the other hand, traditional assessments are generally regarded as indirect and inauthentic. Traditional assessment tools such as multiple-choice test (MCT) often assesses lower-order thinking skills while alternative assessment tools such as PBT often assess high-order thinking skills of the learner (Mauch, 2005).

1.2 Statement of the Problem
There have been calls by biology educators and biology teachers to employ student assessment strategies that would be consistent with the nature of science. Accordingly, the Zambia's third education policy document called 'educating our future' (Ministry of Education, 1996) stresses the importance of using an improved pupil assessment method that goes beyond multiple choice testing, which has generally been criticized for providing results that are often inaccurate and inconsistent measures of pupil achievement. Considering this, educators have called for the use of performance-based testing that can provide more accurate information about what students know and what they can actively do to demonstrate their understanding of the content knowledge of genetics in real life contexts. There is, however, little empirical evidence to support or refute expert opinion that

multiple-choice testing is significantly different from performance-based testing when hands-on teaching methods are used.

1.3 Purpose of the Study

The study sought to determine the effects of multiple-choice testing and performance-based testing on pupils' academic achievement in genetics in the Copperbelt province of Zambia using the hands-on teaching method.

1.4 Hypotheses

The following null hypotheses guided the study and were tested at 0.05 confidence level of significance.

Ho1: There is no significant difference in the academic pupil achievement between multiple-choice testing and performance-based testing in genetics for pupils taught using hands-on methods and those taught using traditional-based methods

Ho2: There is no significant correlation between multiple-choice testing and performance-based testing regarding pupils' academic achievement in genetics after hands-on and traditional-based methods of teaching.

2. Literature Review

There is a paucity of knowledge regarding the information involving the assessment of pupils' achievement in genetics when multiple-choice test and performance-based test are used after using hands-on teaching of genetics. A study by Akpan *et al.* (2014) on the effect of test types on students' achievement revealed that pupils' achievement was affected by the type of test used. Specifically, the study showed that pupils did much better in multiple choice tests compared to performance-based tests. The study claims multiple choice test offers pupils the opportunity to guess correct responses compared to alternative forms of testing. This trend of thought is supported by Hickson & Reed (2009) who established that pupils had higher test scores in multiple choice test than in performance-based test

On the other hand, some experts in educational testing argue that there is little or no difference between multiple-choice testing and performance-based testing regarding the type of knowledge which can be assessed. For example, Wainer & Thissen (1993), established that there was a strong correlation between multiple-choice testing and constructed response items. Similarly, Bridgeman & Rock (1993) found a high correlation between multiple-choice test score and essay test score. Walstad & Becker' s (1994) study also yielded similar results, which claimed a high coefficient of correlation between multiple choice test score and performance-based test score.

3. Theoretical Framework

The study adopted the input-process-output model as theoretical framework. This model is based on Ludwig's systems theory. Input generally refers to what goes into the system, and the process is what causes the change or transformation. Then what comes out of the system is generally referred to as the output (Kyoshaba, 2009). Figure 1 below depicts the conceptual framework employed in the study.

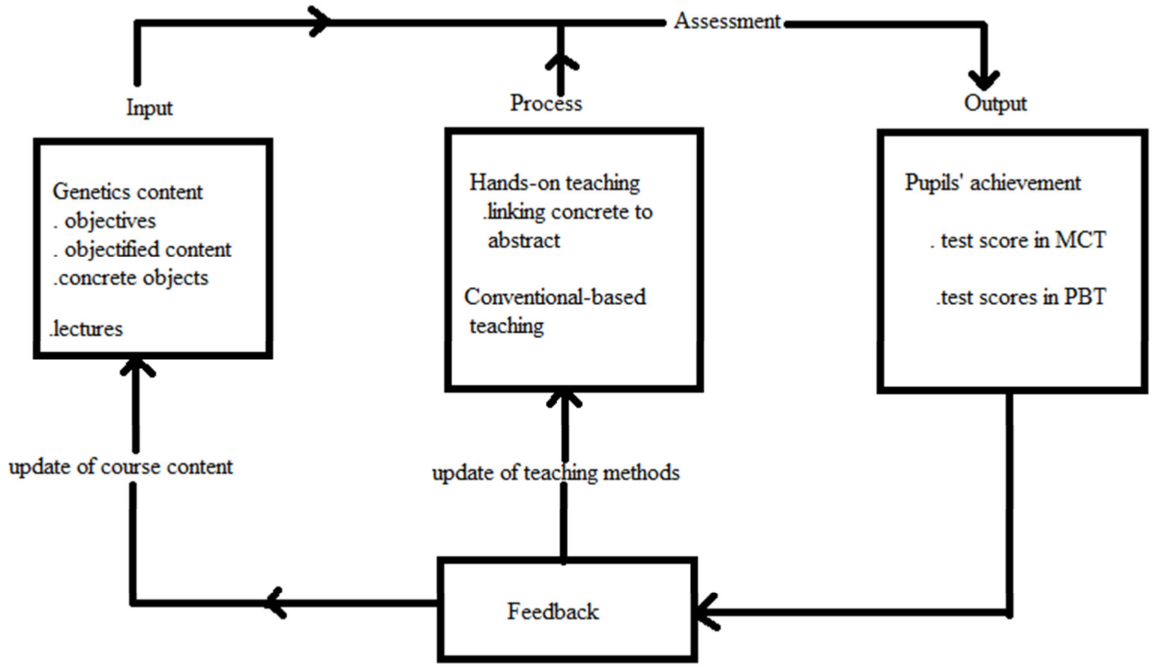

Fig1: Conceptual framework of pupils' academic achievement in genetics after hands-on: MCT versus PBT
Source: Adapted from Knoontz & Weihrich (1988)

4. Methodology and Procedure

This section focuses on the methodology of the study. Specifically, it looks at the plan and methods that were used in the study

4.1 Research design

This was a quantitative study that employed Solomon Four, Non–Equivalent Group Control Design. This design was chosen for the study because the subjects were sampled from intact biology classes. Intact biology classes were used because school authorities in the Copperbelt province could not allow 'disruption' of normal classes for the purpose of conducting a study (Fraenkel &Wallen, 2000). Secondly, Solomon Four, Non–Equivalent Group Control Design was believed to be able to address the major threats to internal validity such as a testing threat of the study (Kane & Trochim, 2006). Table 1 illustrates the notation for Solomon Four, Non –Equivalent Group Design:

Table 1: Notation of Solomon Four Non-Equivalent Control Group Design

Test	Pre-test	Treatment	Post-test
Pre-test Experimental Group (E1)	O1	X	O2
Pre-test Control Group (C1)	O3	-	O4
Unpre-test Experimental Group (E2)	-	X	O5
Unpre-test Control Group (C2)	-	-	O6

Source: Adapted from Fraenkel &Wallen (2000)

Legend:

E1 and E2 represent sampled schools which were taught using hands-on method of teaching
C1 and C2 represent sampled schools which were taught using conventional methods of teaching
X represents the treatment administered to the treatment group means no treatment
O1 and O3 are pre-test
O2, O4, O5 and O6 are post-test

Further, the study employed a number of hands-on activities which were essentially made out of low-cost physical materials. Specifically, concrete models, annotated drawings, photographs and strips of paper depicting hereditary features such as chromosomes and genes were used in designing hands-on activities which were used in teaching the experimental groups. The teacher's role in hands-on teaching was to help pupils link concrete objects to abstract concepts.

As for the control groups, they were taught genetics mainly by 'chalk and talk' method. In this method, learner participation was limited to listening from the teacher. That is, in this mode of teaching pupils were not accorded an opportunity to manipulate concrete objects.

At the end of three weeks (a period of interventions), pupils were subjected to the post-tests: First, pupils wrote MCT followed by PBT later on the same day. These tests were written under the same strict examinations conditions.

4.2 Population of the Study

The target population for the study was grade twelve (12) pupils from mixed-sex senior public secondary schools in the Copperbelt province of Zambia. This population was targeted for the study because genetics, as a topic of study, was covered in term three of grade 12. Public senior secondary schools were also selected because they were comparable in terms of class size, age of the pupils, and the learning facilities to what generally obtains in the country.

4.3 Study Sample

The sample size for the study was 155 pupils of which 84 were boys and 71 girls. The number of pupils in each group was as indicated in the table 2 below.

Table 2: Number of pupil participants in each group of the study

Groups	Number of pupils	Percent of the sample
E1	38	24.5
E2	36	23.2
C1	40	25.8
C2	41	26.5
Total	**155**	**100**

Since the number of pupils in each group was more than 30, this sample had met the minimum number of participants required for the Solomon Four Non-Equivalent Control Group design (Mugenda & Mugenda, 1999).

4.4 Sampling Technique

Prior to the sampling exercise, Copperbelt province was divided into four zones to ensure that the included schools in the sample were reasonably far apart so as to reduce interaction amongst themselves from those selected zones. This reduced the exchange of information between the experimental and control groups. As a result, contamination of the control group from the experimental group was excluded. Selection of the groups which participated in the study was done using simple random sampling which was achieved by a lottery method.

4.5 Data Collection instruments

The data for the study were collected using Genetics Achievement Tests in form of Multiple-Choice Test (MCT) and Performance-Based Test (PBT). The MCT comprised 30 multiple choice items. Each item had five alternative options to choose from. In order to reduce on guessing, a minus mark was awarded for a wrong answer. The PBT, was composed of six open-ended questions. Each question had three sub-question items. In order to increase content validity of the test items, expert validation of the instruments was undertaken. Besides, teachers made comments and observations which further assured the validity of the instruments. Further, these instruments were piloted in four mixed-sex public schools in the Copperbelt that were not sampled for the main study.

4.6 Data Analysis

Prior to data analysis, the data gathered were tested for normality. Accordingly, the pre-test scores obtained using MCT, which were not normally distributed, were analyzed using the Mann-Whitney U test whereas the pre-test scores obtained using PBT, which were normally distributed, were analyzed using an independent-sample t-test. The pretest scores were based on the ‘pretest experimental group, E1 and the pretest control group, C1 whereas the posttests covered all the four study groups.

5. Results

The results of the study are presented under the following sub-headings: the pre-test results, notably, the pupils’ background knowledge in genetics prior to the intervention and the post-test results

5.1 The pre-test results

The pre-test results are broken into two sections, namely, the pre-test results obtained using MCT and the pre-test results obtained using PBT.

5.1.1 The pretest results obtained using MCT

The specific non-parametric test employed to analyze the pretest results of MCT was the Mann-Whitney U test. Table 3 displays the results of the Mann-Whitney U test on pre-test scores of C1 and E1.

Table 3: Results of the Test Statistics of Mann-Whitney U test for pre-test scores of C1 and E1

	Test scores
Mann-Whitney U	789
Wilcoxon W	1650
Z	-0.296
Asymp. Sig. (2-tailed)	0.767

The Mann-Whitney test not significant at .05 level of significance

Table 3 indicates that the Mann-Whitney U test was not significant because the test yielded a p value of 0.767, which was greater than 0.05, (p = 0.767 > 0.05). This result means that there was no significant difference between the control group, C1 and the experimental group, E1, in terms of their background knowledge of genetics.

5.1.2 The pretest results obtained using PBT

The table below (Table 4) presents results for the independent-samples t-test on pretest scores of PBT.

Table 4: Results of the independent sample t-test scores on pre-test obtained using PBT (C1 and E1)

Variable	Group	N	Mean	SD	df	t-calculated	t-critical	p-value
PBT	C1	41	10.8	5.7	79	0.975	1.994	0.332
	E1	40	11.9	4				

The mean difference is not significant at .05 level of significance

According to table 4, the result of the independent sample t-test was not significant as the probability values were greater than 0.05, (t (79) = 0 .975, p = 0.332 > 0.05). Thus, there was no significant difference in the means of the pre-test experimental group, E1 and the pre-test control group, C1.

5.2 The posttest results

A one sample t-test was run on the posttest scores obtained in MCT and PBT for each of the four groups that participated in the study. Table 5 displays the results obtained.

Table 5: Results of a one sample t-test on posttest results of MCT and PBT for each group of the study

Group	Test Type	N	Mean	Std Dev.	t	Df	p -value (2 tailed)	95 % confidence interval of the difference	
								Lower	Upper
E1	MCT	38	83.47	8.63	59.65	37	0.000	80.64	86.31
	PBT	38	31.42	7.70	25.17	37	0.000	28.89	33.95
E2	MCT	36	67.11	7.16	56.28	35	0.000	64.69	69.53
	PBT	36	35.67	7.96	26.88	35	0.000	33.00	38.39
C1	MCT	40	53.98	13.04	26.18	39	0.000	49.81	58.14
	PBT	40	25.13	3.90	40.70	38	0.000	23.88	26.37
C2	MCT	41	64.24	8.78	46.87	40	0.000	61.47	67.01
	PBT	41	16.44	6.41	16.42	40	0.000	14.42	18.46

The means differed significantly between MCT and PBT for all groups under study at .05 level

Table 5 shows that the means between MCT and PBT differed significantly (p value = 0.000 < 0.05 for each group). This result, therefore, discounted the hypothesis stating that the mean of MCT scores in genetics did not differ significantly from that of the PBT score after using hands-on teaching and conventional-based teaching methods. As a result, the above stated hypothesis was rejected.

To determine whether there was no significant correlation between multiple choice testing and performance-based testing regarding pupils' academic achievement in genetics, a Pearson's correlation was run on the posttest scores for the four groups. Table 6 shows the results of the Pearson's correlation coefficients for each group of the study.

Table 6: Pearson's correlation between Multiple Choice Test and Performance-Based Test items

Study	Number of pupils	Variables	Correlation coefficient (r)	P value (2-tailed)	Effect size (Coefficient of determination, R^2)
E1	38	MCT and PBT	0.334	0.040	0.111556
E2	36	MCT and PBT	-0.002	0.991	0.000004
C1	40	MCT and PBT	0.088	0.590	0.007744
C2	41	MCT and PBT	0.086	0.593	0.007396

Correlation was weak for all groups under study and insignificant for E2, C1 and C2, but significant for E1at 0.05 level

Table 6 shows that the Pearson's correlation coefficient between MCT and PBT was weak for all the groups under study, since 0.1< | *r* | < 0.3. Further, the table shows that the correlations were not significant for the groups which participated in the study except for group E1. Additionally, table 6 shows that there was a narrow range of effect size of correlation coefficient, r^2 amongst the study groups.

6. Discussion

This section focuses on the discussion of the findings. The discussion is presented in line with the hypotheses whose outcomes are highlighted.

One of the hypotheses was that there was no significant difference in pupils' academic achievement in genetics when pupils were assessed using MCT and PBT after using hands-on and conventional-based methods of teaching genetics. This study found that there was a significant difference in pupils' academic achievement in genetics in favour of MCT. That is, pupils had higher mean test scores in multiple choice test than in performance-based test. This means that pupils did significantly better in MCT than in PBT. This result of the study agrees with the finding of Akpan *et al.* (2015) and that of Peuker *et al.* (n.d) who showed that pupils significantly achieved better in multiple choice test than in performance-based test.

There are many possible reasons why pupils performed significantly better in MCT than PBT. First, the difference in pupils' achievement in these tests may have been due to the point that multiple choice items are believed to reduce student fear and test anxiety (Birenbaum & Pinku (1997). This could have applied to the groups under study so much that their confidence was boosted whilst preparing for the MCT. This may be partly due to the faict that unlike PBT items, pupils writing MCT always have to select answers from a list of options rather than generating or constructing the responses by themselves (Weimer, 2018).

Secondly, it could be that pupils were influenced differently by test items in MCT and PBT. For example, MCT items accorded pupils the opportunity to check their answers after working through the problem and probably got some clues from the same options presented. Thirdly, it could be that pupils were in classes that were taught skills associated with multiple–choice testing such as risk-taking behavior and memorizing facts in genetics. In this way, multiple-choice test items could have fostered recall of previously learnt material in genetics. Thus, pupils can be expected to do better in MCT than PBT (Johnston & Becker, 1999). Similarly, a pupil trained in skills of answering constructed responses such as interpreting graphs, demonstrating a scientific phenomenon and logical argument is expected to achieve high test scores in PBT.

In this study, further tests were done regarding the second hypothesis which sought to establish whether there was no significant correlation between multiple-choice testing and performance-based testing concerning pupils' academic achievement in genetics after using hands-on method of teaching genetics. The results revealed that correlation between MCT and PBT was significant for the pre-test experimental group, E1 and not significant for the other experimental group, E2. These results suggest the presence of a relationship between MCT and PBT for the pre-test experimental group, E1 and absence of a relationship between the said variables for the second experimental group, E2. Further, the results revealed that the strength of the relationship between these two variables, (as indicated by the correlation coefficient, r values, in table 6 of this report), for the pre-test experimental group, E1 was weak and small. Similarly, the strength of the relationship between MCT and PBT (table 6) for the other experimental group, E2, was equally small. With respect to the correlation between MCT and PBT after using conventional-based methods of teaching, the relationship between the above stated two variables was found not to be significant and almost of negligible relationship. This result, therefore, suggests the absence of a relationship between MCT and PBT when genetics was taught using conventional-based methods.

Further, the lack of statistically significant correlation between MCT and PBT between C1 and C2 in terms of pupils' academic achievement could as well suggest that MCT and PBT have nothing or little in common. Put simply, the aspects of knowledge in genetics assessed by MCT could not be assessed by PBT. This result was at variance with those obtained by Wainer &Thissen, (1993), and Walstad & Becker (1994) who reported that there was strong relationship between multiple choice test and performance-based test. A strong relationship between MCT and PBT seems to suggest that MCT can be used in place of PBT. The converse is true. That is, MCT measures whatever concepts can be assessed by PBT. This notion is supported by Bridgeman & Rock (1993) who assert that MCT measures the same dimension of pupils' knowledge in genetics as much as PBT. These findings are in contradiction with previous results obtained by Hickson & Reed, 2009 and Woijas, n.d, which yielded a weak correlation between MCT and PBT. Hickson & Reed, (2009) and Woijas, (n.d) showed that pupils significantly performed better in PBT than in MCT.

One result which was in contradiction with other results of the study concerning the verification of the second hypothesis was un expectedly significant Pearson's correlation coefficient between MCT and PBT for group E1. Apart from this slight discrepancy, the correlation coefficients for all the four groups under study was weak. A weak correlation suggests that there was a lower likelihood of there being a relationship between MCT and PBT. The above account seems to suggest that MCT might not be used in place of PBT in every aspect. It is worth noting that, while MCT items can be used to measure attainment of both lower and higher order thinking skills in pupils as much as can be ascertained using PBT (Haldyna, 1997) MCT, may not measure the full range of complex thought represented in performance-based testing (Messick, 1995).

In terms of practical sense, the study noted that all the four groups under study yielded a weak relationship between MCT and PBT. This was due to the fact that each group of the study had a very small effect size (coefficient of determination, R^2 value), the largest being 0.1115 (table 6) for the pre-test experimental group, E1. According to Cornell & Berger (1987), an R^2 value of 0.1115, which as a percentage is equal to 11.15 percent accounts for only 11.15 percent of variance in MCT, while 88.85 percent of the variation occurs for reasons that cannot explain the relationship of MCT to PBT. This means that some pupils' achievement in genetics regarding MCT scores can be explained by a difference in PBT, the second variable. It is, however, worth noting that since a correlation only indicates the presence or absence of a relationship, and not the nature of a relationship, the 88.85 percent of the variation in mean test scores of MCT cannot be explained entirely by differences of mean test scores in PBT due to the point that there is always a possibility that a third variable influenced the results (Muhammad, 2012).

One interesting pattern noted in the results was that pupils taught genetics using hands-on methods had higher means in both the MCT and PBT than their counterparts taught using traditional-based methods. This result seems to suggest that the hands-on activities which were used in the hands-on teaching method were a better educational input than the unidirectional lecture method used in the traditional-based teaching. In view of this, the hands-on activities can then be assumed to be a more crucial educational input than lectures regarding pupils' academic achievement in genetics. Consequently, hands-on activities in genetics should be developed by teachers for quality output in schools. It is however worth noting that despite the experimental groups having been exposed to the hands-on method of teaching, one experimental group (E1) marginally outperformed the other experimental group(E2). This is possibly due to the fact that the two groups were composed of pupils of different pupil-characteristics such as study habits. Additionally, how their respective teachers handled the teaching process of genetics using hands-on activities might have been a factor in determining pupils' achievement (Hanushet, *et al*., 2006). This opinion is consistent with Chinelo (2012) who claims that transformation of educational inputs into the 'desired' output largely depends on the teacher.

7. Conclusion

The study showed that MCT and PBT significantly differed from each other in terms of their effects on pupil academic achievement in genetics: Pupils performed significantly better in multiple choice test than in the performance-based test. Consideration should, therefore, be given to assess pupils' learning outcomes in genetics using MCT alongside PBT (in a complementary role). Relying on MCT as the only assessment tool to provide information about the academic achievement of learners in genetics may only reflect a part of pupils' achievement. The study also demonstrated that there was a weak correlation between multiple choice testing and performance-based testing regarding pupil academic achievements in genetics for both groups: Those taught genetics using hands-on method and those taught the subject using the conventional-based methods. This finding was consistent with the earlier result in which the effects of MCT differed significantly from those of the PBT. Overall, the study revealed a weak linear relationship between these two forms of testing under study.

8. Recommendation

Given that the relationship between MCT and PBT was weak, biology teachers should use both forms of testing in assessing pupil's conceptual understanding of genetics. This implies MCT should not be used in place of PBT and vice versa. Further, the study recommends that a study regarding the use of MCT and PBT to measure pupils' misconceptions, misunderstanding and errors they might have concerning the learning of genetics be conducted to so that teachers prove more meaningful feedback to the learners.

References

Ahmad Fuad A.R., Rosniza, AA, Nik-Mohd Rizal F.M, Rogayah, J. & Salam A. (2005). Assessment of the students: tools used in UniversitiSains Malaysia. *Bangladesh Medical Journal*, 34:11-13

Akpan, A. G., Anongo, C.M., Ekong, J.N., & Okrikara, E. (2015). Influence of Selected Variables on Students' Academic Performance in Genetics and their Implications for Effective Application of Stem Education. *Journal of Emerging Trends in Educational Research and Policy Studies,* 6(4), 2141-6990.

Birenbaum, M. & Pinku, P. (1997). 'Effects of test anxiety, information organization, and testing situation on performance on tow test formats'. *Contemporary Educational Psychology*, 22(1), 23-38

Bridgeman, B. & Rock A.D. (1993). Relationships Among Multiple–Choice and Open–Ended Analytical Questions. Available at https://doi.org/10.1111/j.1745-3984.1993.tb00429.x

Brown, G., Bull, J. & Pendlebury, M. (19997). *Assessing student learning in higher education*, London, UK: Routlege.

Cornell, A J. & Berger, D. R. (1987). Factors that influence the value of the coefficient of determination in simple linear and nonlinear regression models. *Phytopathology* 77:63-70.

Deming, J.C. (1992). Assessing Student Achievement: Multiple-choice testing versus Performance assessment.

Available at https://digital.library.txstate.edu/bitstream/handle/10877/3454/fulltext.pdf?sequence
Dietel, J.R., Herman, J. L., & Knuth, A.R., (1991). What does research say about assessment? NCREL, Oak Brook. Available online: http://www.ncrel.org/sdrs/areas/stw_esys/4assess.htm
Dubey, S., Hirpara, N., Jain, S. & Gupta, A. (2015). Interpreting Research Findings with Confidence Interval. *Journal of Orthodontics & Endodontics*, 1(1), 8
Fraenkel, J. R. & Wallen, E.N., (1993). *How to design and evaluate research in education* (2nd ed.). New York.: McGrawHil Inc.
Haldyna, M.T. (1997). Writing Test Items to Evaluate Higher Order Thinking Skills. Allyn Bacon, Boston, MA.
Haney, W. & Madaus, G. (1989). *Searching for alternatives to standardized tests: whys, whats, and whithers*. *Phi Delta Kappan,* 70, 683-687.
Hickson, S. & Reed, R.W., (2009*). Do Constructed-Response and Multiple-Choice Questions Measure the Same Thing? (May 18, 2009). Available at SSRN: https://ssrn.com/abstract=1408040 or http://dx.doi.org/10.2139/ssrn.1408040*
Knoontz, H. & Weihric, H. (1988) *Management.* New York: McGraw-Hill, INC.
Kyoshaba, M. (2009). *Factors affecting academic performance of undergraduate students at Uganda Christian University* (Master's dissertation). Makerere University.
Kane, M. & Trochim, M.W., (2006). Hybrid Experimental Designs*. Retrieved July 23, 2016 from https://www.socialresearchmethods.net/kb/exphybrd.php*
Law, B. & Eckes, M. (1995). Assessment and ESL. Peguis publishers: Manitoba, Canada.
Mauch, J. (2005). Performance-Based Assessment versus Traditional Assessment in the Science Classroom. Sec 690.
Muhammad I. U. (2012). Coefficient of Determination: A model Selection Criteria. Retrievedon 28th July, 2018 at *https://en.wikipedia.org/wiki/Coefficient_of_determination*
Mugenda, A.G. & Mugenda, O.M. (1999). *Research methods: Quantitative and qualitative approaches.* Nairobi: ACTS Press.
Ministry of Education (1996) *Educating Our Future: National Policy on Education*. Government Printers, Lusaka
Messick, S. (1995). Validity of psychological assessment. American Psychologist, 50, 741-749.
Nasab, F.G., (2015). Alternative versus Traditional Assessment *Journal of Applied Linguistics and Language Research*, 2(6),165-178. Available online at www.jallr.ir
Neill, M. (February 26, 1996). Assessment reform at a crossroads. Education Week
Patrick A.R (2014).What is a statistically significant relationship between two variables? Available at https://simplyeducate.me/2014/05/29/what-is-a-statistically-significant-relationship-between-two-variables/
Peuker, J.M., Brook, J &Peuker, S. (n.d). Effect of Multiple-Choice Testing on Student Performance in an Introductory Engineering Course. Available at https://peer.asee.org/effect-of-multiple-choice-testing-on-student-performance-in-an-i..
Trochim, W.K. (2006). Hybrid Experimental Designs. Retrieved on March 12, 2018 from https//socialresearchmethods.net/kb/exphybrd.php
Wainer, H. &Thissen, D. (1993). *Combining multiple-choice and constructed response test scores: towards a Marxist theory of test construction*. Applied Measurement in Education*, 6, 103-118*
Walstad, W. B., & Becker, W. E. (1994). Achievement differences on multiple-choice and essay tests in economics. American Economic Review*, 84, 193-196.*
Winking, D. (1997). Critical issue: Ensuring equity with alternative assessments [online document]. NCREL (North Central Regional Educational Laboratory), Oak Brook: IL. Available online: http://www.ncrel.org/sdrs/areas/methods/assment/as800.htm
Weimer, M. (2018). Multiple-Choics Tests: Revisiting the Pro and Cons. Retrieved on July 26, 2018 from https://www.facultyfocus.com/articles/teaching-professor-blog/multiple-choice-tests-pros-cons/
Wojtas, K. (n.d.). *Multiple Choice vs. Performance Based Tests in High School Physics Classes*. Retrieved August 16, 2016 from http://www.pcs.cnu.edu/~David.Gore/Capstone/files/WojtasK.pdf