Canadian Journal of Educational and Social Studies
Vol. 2(6), 2022, pp. 35-48

Performance Assessment: Effective Tool Closing Academic Achievement Gap Between High and Low Ability Students

Patience Langee[1] & Abraham Gyamfi[2] & Abraham Yeboah[3] & Isaac Aboagye Adu[4] & Kingsley Kwasi Erzoah[5]

[1]St. John Bosco College of Education, Ghana
[2]Wesley College of Education, Ghana
[3]Akenten Appiah Menka University of Skills Training and Entrepreneurial Development, Ghana
[4]Ghana Education Service, Dunkwai-Offin, Ghana
[5]Ghana Education Service, Ellembelle, Ghana
Correspondence: Abraham Gyamfi, Wesley College of Education, Ghana
Email: abrahamgyamfi84@gmail.com

DOI: 10.53103/cjess.v2i6.78

Abstract

Globally, bridging the achievement gap between low and high ability students in examinations have been the goal of educators. Academic achievement disparities among students have been one of the major reasons bedeviling the attainment of target by many schools in external examinations. This study aimed at using performance assessment to narrow the gap between high and low ability students' achievement in tests. The descriptive survey was used to gather data from a 234 JHS students in Ahanta West Municipality obtained from a multi-stage sampling procedure. Teacher-made achievement tests were used to collect the data and were analysed with descriptive statistics and multivariate analysis of variance. Results of the analyses revealed that both high and low ability JHS students' achievements in tests were improved when assessed with performance assessment. It was concluded that educators at JHS should be in the known that performance assessment is the tool for closing gaps between high and low ability students' achievements in examinations. Therefore, assessors at the JHS level should adapt and integrate high percentage of performance assessment tasks in assessing students in order to close the gap between the low and high ability students' achievement in the tests and examinations.

Keywords: Performance Assessment, Tasks, High Ability Students, Low Ability Students, Students' Achievement

Introduction

As more and more educators use assessment, they increasingly recognize that the process has the power to transform instruction and the society. Lambert and Lines (2000) noted that assessment is the process of gathering, interpreting, recording and using information about students' responses to educational tasks in making relevant decisions. This means assessment is important to most stakeholders, especially the teacher as well as

the students. It makes teachers focus on individual learners, the learning community, or the educational industry as a whole and helps the student understand himself or herself better, based on the functionality of the decisions made using assessment results (Tamakloe, Atta, & Amedahe, 2005).

According to Wiggins (1989), the aim of assessing is primarily to educate and improve students' achievement. This aim can be achieved through a planned educational process based on some important advantages such as how to assess the students' achievement in advance (Demirel, 2007). Students' achievement as perceived by Mellroth (2014), is assessed either by comparing students with one another (norm-referenced) or according to goals in the curriculum (curriculum or criterion referenced). To Mellroth, the norm-referenced form makes it possible to study students' relative achievements, rank achievements and identify a top percentage population as the high ability students whereas the low ability students is defined by bottom populace. Gagné (2005) noted that, high ability students tend to maintain their position through their formal schooling. This implies, the high ability students are persistent to inert characteristics towards continuous achievement. The characteristics subsume self-motivation, self-efficacy and low level of test anxiety (Jaglois & Kitchel, 2014). These characteristics create the disparity between the high and low ability students in the classroom.

McCoach and Siegle (2001) noted that, low ability students are those who appear capable of succeeding in school but are nonetheless struggling with academic work. It is further stated that low ability students also, attend school without books or homework, and appear to choose not to study for exams (McCoach & Siegle). They are students who seem unphased by parents and teachers' pleas that their grades now will affect the rest of their professional lives. Dzulkifli and Alias (2012) also said low ability students are easily distracted emotionally, especially with the onset of anxiety and this result in pregnable performances. However, Dzulkifli and Alias argued that low ability students are more assertive. To them, it depicts they have strong personality; tend to be independent and dominant; have the capability to stand up for their rights and possess attributes needed for attainment of their achievements. The foregoing discussions bespeak students who perform poorly academically may still have strong and positive personal characteristics that can help them to succeed.

Researches have shown that, low ability students can be in a positive direction of achievement with their counterparts; high ability students, if an appropriate assessment procedure is employed for them (Maier, Adams, Burns, Kaul, Saunders & Thompson, 2020; Taylor et al.., 2016; Yaduvanshi & Singh, 2019). Evidence in Arhin's (2015) and Woodward, Monroe and Baxter's (2001) studies depict performance assessment has the capacity of giving the true achievement of students with different abilities. Performance assessment allows students to actively develop their approaches to demonstrate what they know about tasks or skills under defined conditions, knowing that their work will be

evaluated according to agreed-upon standards (Project Appleseed, the National Campaign for Public School Improvement, 2018).

Madaus and Dwyer (1999) stated that, "performance assessment requires examinees to construct or supply answers, perform or produce something for evaluation" (p. 690). In this form of assessment students are engaged in an action to accomplish a task. Nitko (2001) noted that in the assessment, students use acquired knowledge and skills from diverse fields in accomplishing the task which is related to a learning target. Thus, the requirement of performing an activity is rooted on knowledge and skills learnt in the classroom. Arhin (2015) in his view said students show their mastery of skills and competencies by doing activities they are capable of taking on their own. This implies that, in performance assessment, students develop their own solutions to tasks posed to them. The demonstration of proficiency occurs in solving complex problems such as a contextual mathematical problem which ensures that, both the process used in reaching a product and the product itself are assessed (Nitko).

According to Birenaum and Feldman (1998), performance assessment procedures permit the students to be assessed on complex learning outcomes which most at times involve hands-on activity to produce a real-life product. The procedures include conducting a science investigation, constructing an original product, providing a response as in writing an explanation of one's solution to a mathematics problem or writing a persuasive essay. This implies that, "performance tests measure skill or ability… and scoring often requires subjective judgment" (Frey & Schmitt, 2010. p 109). Performance assessment procedures measure students' cognitive thinking and reasoning skills and their ability to apply knowledge to solve realistic and meaningful problems. In other words, they are designed to closely reflect the performance of interest, allow students to construct or perform an original response, of which predetermined criteria is used to evaluate the student's work by an assessor. The performance assessment procedures are, generally, seen by educators as more valid indicator of students with different abilities achievements. However, performance assessment procedures are (a) harder to evaluate; (b) time consuming, especially during the administration stage; (c) labour intensive during development of tasks; (d) susceptible to unfairness due to subjectivity in scoring; and (e) less economical to be used on large participants (Mueller, 2016).

Despite the fact that students' academic achievements have been a matter of concern to stakeholders there is a paucity of studies making attempt to create the possibility of employing an appropriate assessment procedure to purportedly enhance the low ability students' achievement in examinations and bridge the gap with their counterparts, high ability students. Majority of studies conducted were outside Ghana (Maier et al.., 2020; Taylor et al.., 2016; Yaduvanshi, & Singh, 2019) and their findings are inapplicable in the Ghanaian context. None of the studies considered if performance assessment could make an attempt to close the achievement gap between the low and high ability students in

schools. Furthermore, all studies were conducted beyond the basic level of education. Based on these, this study is taken to fill the gap in knowledge and practicality of assessment at the basic school level.

Theoretical Framework: Constructivist Learning Theory

Constructivist learning theory posits that new knowledge generation is built from the fundamentals of one's own previous comprehension and experience (Stevens-Fulbrook, 2019). According to Stevens-Fulbrook, understanding of a new concept is based on reflections of prior knowledge and experiences acquired from theories and interactions in the environment. This means, past knowledge and experiences lie at the center of constructivism. As noted by Vygotsky (cited by Willson & Peterson, 2006), individuals are competent to develop a concept because their minds are not empty vessels.

There is an implication of this constructivist view of learning for assessment in cooperative education. It is likely that students' starting points are all different; each having differing levels of prior work and life experiences. Thus, assessment needs to acknowledge that each student learns different things from his or her perspective, each being of potential value and merit. For this reason, both high and low ability students were assessed using performance assessment tasks. This is because, performance assessment allows this free will of students using individual prior knowledge and experiences in providing responses to tasks.

Performance Assessment

Leon and Elias (1998) believed performance assessment originated from the Chinese proverb "I do I understand" because the assessment tasks require application of learnt content in realistic situations. The authors believed the assessment tasks encourage demonstration of learnt concepts in using activities that are thought provoking and require dedication and responsibilities from students. According to Lane (2010), performance assessment is an essential component and process in education yet with limited recognition and practice in many nations. Osterlind (2006) noted this assertion by saying that performance assessment got its recognition when traditional assessment of knowledge using multiple-choice could not give better account of individuals' knowledge in demonstrations, performance tasks, proficiency in writing skills, creation of products and even group work in tests. Khalanyane and Hala-hala (2014) argued further to say that performance assessment emphasizes on procedures employed to solve unique problems in society. This is an indication that tasks of performance assessment require students to act by performing or constructing a product other than recalling and selecting existing answers to a task. However, Khattri, Kane and Reeve (1995) believed performance assessment is unknowingly practiced by assessors in diverse procedures. Regardless of the fact that

performance assessment was informally known to assessors, the authors believed it is an old tool that is used to measure students' cognitive, affective and psychomotor skills.

Performance assessment is characterised by assessing multiple learning targets especially with tasks that require the students to create objects, produce a report or to put up a demonstration of an activity or event (Lane, 2010; Nitko, 2001). The assessment procedures require execution of more than one learning outcomes in accomplishing a task. An instance is when a student is asked to demonstrate a hands-on activity. In this process, the student will initially picture the whole activity at a glance, outline principles involved, gather and arrange relevant materials if required before putting up the demonstration just to ensure that the activity incorporate all necessary principles and assumptions. This confirmed what Herrera et al.., (as cited by Mussawey, 2009) said that, performance assessment tasks assess all aspects of student learning through determining means by which students assimilate information, store and apply information in novel ways. However, Nitko argued that, every learning target cannot be assessed using performance assessment. Such learning targets include declarative statements that need to be assessed with traditional tasks because they require recalling of facts.

Performance assessment has multiple names based on the diverse perception of authors. Brown and Hudson (as cited in Koné, 2015), said the assessment is also known as authentic assessment or performance–based assessment because the assessment procedures allow students to produce responses that are similar or exactly to the real thing in life by using their own productive skills. Brooks (1999) and Oosterhof (2001), on other hand, said performance assessment subsumes authentic and alternative assessments because all of them require the employment of higher-order thinking skills.

Darling-Hammond (as cited in Koné, 2015) posited that performance assessment judge's students on laid down criteria essential for the precise performance of the activity similar to the work place. Similarly, the Standards for Educational and Psychological Testing, indicating that performance assessments "emulate the context or conditions in which the intended knowledge or skills are actually applied" (American Educational Research Association [AERA], American Psychological Association [APA], & National Council on Measurement in Education [NCME], 1999, p.137). To the proponents, performance assessment has a broad range of assessment types subsuming authentic, alternative and performance-based assessments and these are interchangeable based on evidence of similar characteristics existing among them (Brooks; Herman, Aschbacher, & Winters, 1992; Frey & Schmitt, 2010; Meyer, 1992).

Research Hypothesis

H_O 1: There is no statistical difference in the performance of high and low ability students in Mathematics and Integrated Science when assessed using performance assessments.

H_A 1: There is a statistical difference in the performance of high and low ability students in Mathematics and Integrated Science when assessed using performance assessments.

Methodology

The study employed the descriptive design with the cross-sectional survey method. The choice of the design was with the notion that it allows to collect data from selected individuals in a single time period concerning the current state of JHS students' achievements when assessed with performance assessment in order to determine the assessment type capability of bridging the achievement gap between the low and high ability students.

Population

The population comprised of all JHS students in the Western Region of Ghana. The targeted population was the 58 public JHS students with an enrolment of 7527 in the Ahanta West Municipality in the 2019/2020 academic year. For the purpose of this study the accessible population was the second year (Form 2) JHS students in the Municipality with a total number of 2499 in five (5) educational Circuits as at December 18, 2019 (EMIS Report, 2019).

Sampling Procedures

The cluster sampling procedure with a multistage technique was employed to select the representative sample. The Municipality has five educational Circuits and these were used as clusters. Three out of the five cluster of schools (educational Circuits) were selected using the lottery method of simple random sampling. The simple random sampling by the lottery method technique was employed again to select the representative of four single stream schools in each selected cluster, totaling 12 schools with 488 participants in all intact classes in the selected schools.

Research Instrument

The instruments used for the study were teacher-made performance achievement tests. The researchers opted for achievement tests based on the assertion that it is the tool that can be designed to measure cognitive processes on mastery and proficiency in diverse areas of taught contents (Ary, Jacob & Sorensen, 2010; Cohen, Manion & Morrison, 2007). That is, the tools were suitable to measure the variable of interest, which is achievement of students. Furthermore, as opined by Ary et al. (2007) "the advantage of a researcher-made

test is that it can be tailored to be content specific; that is, it will match more closely the content that was covered in the classroom" (p. 203). This choice of tests therefore, offered the opportunity to develop items that have coverage in learning outcomes of taught contents.

The instruments were designed and developed by the researchers. Test items were constructed based on treated topics selected from the 2019/2020 academic year's scheme of work in the Municipality based on the teaching syllabi (MOE, 2012) of Mathematics and Integrated Science. In order to ascertain the coverage of each topic in the tests, the researchers round to collate the topics from the respective subject teachers of each sampled school including the piloted school.

The collated topics taught by the teachers in conjunction with the Bloom's Taxonomy were used to construct the test plans for each test just as noted by (Etsey, 2012). The test plan aided to avoid lop-sidedness and inadequacy of sampling from the topics during the construction of the items. This ensured content representativeness and relevance for each instrument just as Nitko (2001) noted. The procedure guaranteed that, the assessment tasks reflected the learning outcomes specified in the syllabus and ascertained the content validity of the tests' scores.

In writing the individual tasks for the instruments, the test plan was strictly adhered to. The restricted response type of task was used to construct the tasks. The purpose was to aid students to interpret each item as intended and to discourage biasness when scoring. This improved the scoring reliability and hence validity. During the writing of the test tasks, the language, operational level of the students and clarity of the tasks were considered. These aided in conveying uniform meaning and single interpretation by all. Each task of the instruments was immediately accompanied by its scoring key at the construction stage (Amedahe & Asamaoh-Gyimah, 2016; Crocker & Algina, 1986; Nitko, 2001; Quaigrain & Arhin, 2017).

The Mathematics instrument consists of six on-demand tasks while the Integrated Science composes of seven on-demand tasks. The tasks engulfed higher-order thinking skills from analysis, synthesis and to evaluation (creating) which required the students of 'what to do with the facts'; that is to first understand the facts, then connect the facts, categorize the facts, manipulate the facts and use them in new situations to obtain new solutions to problems (Nursalam, Angriani, Darmawati, Baharuddin, & Aminuddin, 2018). The time allotted for students to attempt all tasks in each test was 30 minutes. The tests' time apportioned was in alignment to suggestions made by Nikto (2001). Also, clear and concise instructions were spelt out for the entire tests. This encouraged students to attempt tasks as expected in order to attract the required marks and enhance students' interest in the entire tests.

The instruments/tests were reviewed to ascertain the content and construct validity of the written tasks for the intended purpose. This began with the reviewing of each

instrument after constructing each test and putting it aside for three weeks. Tasks were critically scrutinized to detect flaws after which necessary corrections were made. Tasks were examined again by six experienced professional Mathematics and Integrated Science teachers at the JHS level. Thus, three teachers examined each subject's instruments. They helped to confirm that the tasks carry the appropriate meaning of the constructs that were intended to measure. An example was, whether the tasks required students to use deep thought in manipulating materials to arrive at answers. Further, to confirm that clarity of each task matched the ability of the students, the teachers reviewed the test contents, language aspects of tasks and scoring keys as well. The reviewers' judgements and feedbacks were subsequently taken in good faith and all corrections were appropriately made. The inter-rater correlation reliability coefficient (Hallgren, 2012) of 0.97 and 0.95 for the Mathematics and Integrated Science respectively were ascertained to achieve a valid and reliable test instruments which would measure the achievement of the students.

Determination of High and Low Ability Groups

Before the data collection commenced, the Ability Determination Test (ADT) instrument was first administered to all students as the initial test. Its purpose was to determine each student's academic ability. The test tasks were scored and individual scores ranked and used to categorise the students into high and low ability groups. The groups were created using the twenty-seven per cent (27%) each of upper (U) and lower (L) scores as the high and low ability students respectively as suggested by Crocker and Algina (2006) and Nitko, (2001). Of the 488 students from the sampled schools, 434 students representing 88% took all the tests and hence participated in the study. However, 234 students forming a total of the high and low ability groups; was used for the data analysis and were used to determine the students' achievement in the respective subjects (Mathematics and Integrated Science).

Data Processing and Analysis

The instruments (Performance Assessment for Mathematics [PAM] and Performance Assessment for Integrated Science [PAIS]) had total score of 30 marks for each subject. For PAM tasks 1, 2, 3, 4, 5 and 6 were subjectively scored 3, 3, 12, 6, 3 and 3 marks respectively. The PAIS tasks 1, 2, 3, 4, 5, 6, and 7 were also scored subjectively as 5, 1, 3, 6, 5, 5 and 2 marks respectively. The scores of each instrument were converted into percentages so that a unified unit could be attained for easy comparison of the low and high ability groups. The data (tests scores) were analysed using SPSS version 23. The analysis involved the computation of descriptive statistics, specifically, means and standard deviations and one-way multivariate analysis of variance (MANOVA) to test the hypothesis at .05 significance level.

Results and Discussion

In testing the hypothesis, the data were analysed using descriptive statistics. The mean scores and standard deviations obtained are presented in Table 1.

Table 1: Mean Scores of High and Low Ability Students' Performance in Mathematics and Integrated Science in Performance Assessment

Subject	Ability Group	N	Mean(M)	SD
Mathematics	High	117	85.81	10.95
	Low	117	83.50	12.38
	Total	234	84.66	11.72
Integrated Science	High	117	86.58	11.34
	Low	117	89.91	10.41
	Total	234	88.24	10.99

Source: Field survey (2020)

A close observation from Table 1 indicates that, the high ability students (M = 85.81, SD = 10.95) performed better than the low ability students (M = 83.50, SD = 12.376) in Mathematics. However, the vice versa occurred in the case of Integrated Science scores with the lead being the low ability students (M = 89.91, SD = 10.41) performing better than the high ability students (M = 86.58, SD = 11.34). In order to ascertain statistically significance in the performance between the groups, a one-way multivariate analysis of variance (MANOVA) was conducted. The results are shown in Table 2. The two dependable variables were the performance assessment scores in Mathematics (PAM) and performance assessment scores in Integrated Science (PAIS) and the independent variable was the level of ability group.

Table 2: Results of One Way MANOVA of Low and High Ability Students' Performance on Performance Assessment in Mathematics and Integrated Science

Source	Dependent Variable (Subject)	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Ability Group	Mathematics	311.54	1	311.54	2.28	.132	.010
	Integrated Science	646.67	1	646.67	5.46	.020	.023
Error	Mathematics	31683.11	232	136.57			
	Integrated Science	27494.44	232	118.51			
Total	Mathematics	1709072.00	234				
	Integrated Science	1850283.00	234				

Source: Field survey (2020)

With no violation to all preliminary assumptions, there was no statistically significant difference in the Mathematics scores but a difference occurred in Integrated Science scores at $F(2, 231) = 3.674$, $p < .027$; Wilk's Lambda = .97; partial eta squared =

.03. Dealing with the individual results of the dependable variables, the only difference attained with Bonferroni adjusted level of .025 was PAIS scores, F (1, 232) = 5.46, p = .020, partial eta squared = .02. This finding depicted that the low ability students outperformed the high ability students in Integrated Science while the difference in the Mathematics scores was just by chance and had no statistical significance in practice. Therefore, the low ability students performed better than the high ability students when assessed with performance assessment procedures.

Results from analysed data showed that the low ability students performed better than the high ability students when assessed with performance assessment procedures. Obviously, this finding stood out clearly in the study. The response rate demonstrated by the low ability students revealed the extent of advance preparation and zeal they had for high achievement in the performance assessment examinations than the high ability students. Brookhart (1997), Woodward et al. (2001) and Fastre´, Van der Klink and van Merrie¨nboer.'s (2010) studies had related findings of this current work. Fastre´ et al. (2010) explained their findings that low ability students demonstrated the requirement to expectations of the assessors to attain high scores in the performance assessment. The finding further affirmed the assertion of Meisels, Atkins-Burnett, Xue, Bickel and Son (2003) that both high and low students have opportunity of maximizing their achievement when assessed with performance tasks but performance tasks favour low ability students.

Further, these findings have shown that, high and low ability students demonstrated their acquired knowledge and skills to obtain the scores in performance assessment without memorization of procedures as perceived to occur in traditional assessment for the attainment of high scores. This aspect of the findings confirmed that of Arhin's (2015) finding which noted that performance assessment procedures make students to own the processes of tackling challenging tasks that cause the award of full marks in examinations. To explain better on the high mean scores of the performance assessments, the achievements appeared to be rooted on the exposure of students to the concrete materials, symbolic and abstract information in the performance assessment procedures especially in the Integrated Science tests. In contrast, these findings disagree with Fuchs, Fuchs, Karns, Hamlett and Katzaroff (1999) and Kim's (2005) study results that indicated that high ability students performed better than the low ability students when assessed with performance assessment tasks in tests.

Finding

The study revealed that the low ability students performed better than the high ability students when assessed with performance assessment procedures in Mathematics and Integrated Science tests.

Conclusion

The results of this study bespeak the general view that, both high and low ability students' achievements in Mathematics and Integrated Science examinations were closed to each other when assessed with performance assessment procedures. As shown in the study, low ability students performed equally and better than the high ability students in Mathematics and Integrated Science respectively. This was revealed in the students' achievement mean values in the assessment. There was no difference in their mean values in Mathematics and the difference that occurred in Integrated Science was also in favour of the low ability students. Therefore, it is important to note that performance assessment should be the tool for closing achievement gaps between high and low ability students' achievements in examinations.

Recommendations

From the study finding it is recommended teachers should adapt, implement and integrate performance assessment procedures in assessing their students. Also, assessors should know the abilities of the students to be examined in order to employ the right assessment type for such students during examinations. Finally, the authorities in Curriculum development especially National Council for Curriculum and Assessment (NaCCA) should design and implement comprehensive assessment procedures in the text books to be used for assessment purposes in the classrooms. The procedures should entail mixed items of performance tasks such as hands-on activities and the traditional test tasks.

References

Amedahe, F. K., & Asamoah-Gyimah, K. (2016). *Introduction to measurement and evaluation.* Hampton Press.

American Educational Research Association, American Psychological Association & National Council on Measurement in Education. (1999). Standards for educational and psychological testing. American Educational Research Association.

Arhin, A. K. (2015). The effect of performance assessment-driven instruction on attitude and achievement of senior high school students in Mathematics in Cape Coast metropolis, Ghana. *Journal of Education and Practice, 6*(2), 109-116.

Ary, D., Jacob, L. C., & Sorensen, C. (2010). *Introduction to research in education.* Cengage Learning.

Birenbaum, M., & Feldman, R. (1998). Relationship between learning patterns and attitudes towards two assessment formats. *Educational Research, 40*(1), 90-98.

Brookhart, S. M. (1997). Effects of the classroom assessment environment on mathematics and science achievement. *The Journal of Educational Research, 90* (6), 323-330.

Brooks, L. A. (1999). *Adult ESL student attitude towards performance-based assessment*. [Unpublished master's thesis, University of Toronto].

Cohen, L., Manion, K., & Morrison, L. (2007). *Research methods in education*, (6th ed.). New York: Routledge Taylor & Francis Group.

Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory*. New York. Harcourt Brace Jovanovich.

Demirel, Ö. (2007). *Planning and evaluation in teaching*. Pegem A Publishing.

Dzulkifli, M. A., & Alias, I. A. (2012). Students of low academic achievement their personality, mental abilities and academic performance: how counsellor can help? *International Journal of Humanities and Social Science, 2*(23), 220 – 225.

EMIS Report. (2019). Educational directorate report. Ahanta West

Etsey, Y. K. A. (2012). *Assessment in education*. [Unpublished document, University of Cape Coast, Cape Coast].

Fastre´, G. M. J., Van der Klink, M. R., van Merrie¨nboer, J. J. G. (2010). The effects of performance-based assessment criteria on student performance and self-assessment skills. *Advance in Health Science Education*, 15, 517–532.

Frey, B., & Schmitt, V. (2010). Teachers' classroom assessment practices. *Middle Grades Research Journal, 5*(3), 107-117.

Fuchs, L. S., Fuchs, D., Karns, K., Hamlett, C. L., & Katzaroff, M. (1999). Mathematics performance assessment in the classroom: Effects on teacher planning and student problem solving. *American Educational Research Journal, 36*(3), 609-646.

Gagné, F. (2005). From gifts to talents: The DMGT as a developmental model. In R. J. Sternberg, & J. E. Davidsson (Eds.), *Conceptions of giftedness* (pp. 98- 119). Cambridge University Press.

Hallgren, K. (2012). Computing inter-rater reliability for observational data: An overview and tutorial. *Tutorial in Quantitative Methods for Psychology, 8*(1), 23-34.

Herman, J. L., Aschbacher, P. R., & Winters, L. (1992). *A practical guide to alternative assessment*. Association for Supervision and Curriculum Development.

Jaglois, O. M., & Kitchel, A. (2014). Characteristics of high achieving students enrolled in an online economics course. The National Association for Business Teacher Education.

Khalanyane, T., & Hala-hala, M. (2014). Traditional assessment as a subjectification tool in schools in Lesotho. *Educational Research and Reviews, 9*(17), 587-593.

Khattri, N. Kane, M. B., & Reeve, A. L. (1995). How performance assessments affect teaching and learning. *Educational Leadership, 53*(3), 80-83.

Kim, S. (2005). *Effects of implementing performance assessments on student learning: Meta-analysis using HLM*. [Unpublished doctoral thesis, The Pennsylvania State University].

Koné, K (2015). *The impact of performance-based assessment on university ESL learners' motivation.* [Unpublished masters' thesis, Minnesota State University, Mankato].

Lane, S. (2010). Performance assessment: The state of the art. (Scope student performance assessment series). Stanford University, Stanford Center for Opportunity Policy in Education.

Lambert, D., & Lines, D. (2000). *Understanding assessment*. Routledge Falmer.

Leon, S. & Elias, M. (1998) A comparison of portfolio, performance, and traditional assessment in the middle school. *Research in Middle Level Education Quarterly, 21*(2), 21-37.

Madaus, G., & Dwyer, L. (1999). A short history of performance assessment: Lessons learned. *Phi Delta Kappan, 80*(9), 688-695.

Maier, A., Adams, J., Burns, D., Kaul, M., Saunders, M., & Thompson, C. (2020). Using performance assessments to support student learning: How district initiatives can make a difference. Palo Alto, CA: Learning Policy Institute-learning. https://learningpolicyinstitute.org/product/cpac-district-initiatives-assess-student.

McCoach, D. B., & Siegle, D. J. A. E. (2001). A comparison of high achievers' and low achievers' attitudes, perceptions, and motivations. *Gifted Child Quarterly*, 2, 71-76.

Meisels, S. J., Atkins-Burnett, S., Xue, Y., Bickel, D. D., & Son, S. (2003). Creating a system of accountability: The impact of instructional assessment on elementary children's achievement test scores. *Educational Policy Analysis Archieves,* 11(9), 1-18.

Mellroth, E. (2014). *High achiever! Always a high achiever? - A comparison of student achievements on mathematical tests with different aims and goals*. [Unpublished licentiate thesis, Karlstad University Studies]

Meyer, C. A. (1992). What's the difference between authentic and performance assessment? *Educational Leadership, 49*(8), 39–40.

Ministry of Education (2012). Integrated Science for junior high school. CRDD.

Ministry of Education (2012). Mathematics for junior high school. CRDD.

Mueller, J. (2016). Authentic assessment tool box. http://jfmueller.faculty.noctrl.edu/toolbox

Mussawy, S. A. J. (2009). Assessment practices: Student's and teachers' perceptions of classroom assessment. [Unpublished master's thesis, University of Massachusetts].

Nitko, A. J. (2001). *Educational assessment of students* (3rd ed.). Merrill Prentice Hall Inc.

Nursalam, Angriani, A. D., Darmawati, Baharuddin, & Aminuddin (2018). Developing test instruments for measurement of students' high-order thinking skill on Mathematics in junior high school in Makassar. 2nd International Conference on Statistics, Mathematics, Teaching, and Research: Journal of Physics Conf. Series, 1028(012169) 1742-6596/1028/1/012169.

Oosterhof, A. (2001). *Classroom application of educational measurement* (3rd Ed.). Prentice Hall Inc.

Osterlind, S. (2006). *Modern measurement: Theory, principles, and applications of mental appraisal.* Pearson Prentice Hall.

Project Appleseed, the National Campaign for Public School Improvement (2018). Testing, testing [Brochure]. St. Louis, MO. http://www.projectappleseed.org/#!assessment/cwvf

Quaigrain, K. & Arhin, A.K., (2017). Using reliability and item analysis to evaluate a teacher-developed test in educational measurement and evaluation. *Educational*

Assessment & Evaluation. 4(1301013), 1 – 11.
Stevens-Fulbrook, P. (2019). 15 Lerning theories in education (A complete summary). http/ Home Best of teacherofsci.com
Tamakloe, E. K., Atta, E. T., & Amedahe, F. K., (2005). *Principles and methods of teaching*. Black Mask.
Taylor, C., Kokka, K., Darling-Hammond, L., Dieckman, J., Santana Pacheco, V., Sandler, S., & Bae, S. (2016). Student engagement: A framework for on-demand performance assessment tasks. Stanford Center for Opportunity Policy in Education.
Wiggins, G. (1989). Teaching to the (authentic) test. *Educational Leadership, 46*(7), 41-47.
Wilson, M. S., & Peterson, P. L. (2006). Theories of learning and teaching what do they mean for educators? National Education Association.
Woodward, J., Monroe, K., Baxter, J. (2001). Enhancing student achievement on performance assessments in Mathematics. *Learning Disabilities Quarterly,* 24(1), 33-46.
Yaduvanshi, S. & Singh, S. (2019). Fostering achievement of low-, average-, and high-achiever's students in biology through structured cooperative learning (STAD method). *Education Research International, 1*(4), 1-10.