European Journal of Education Studies
ISSN: 2501 - 1111
ISSN-L: 2501 - 1111
Available on-line at: www.oapub.org/edu

DOI: 10.46827/ejes.v8i4.3697 Volume 8 | Issue 4 | 2021

ASSESSING THE QUALITY ASSESSMENT PRACTICES ON STUDENTS' EFFORT AND ACHIEVEMENT IN SOME GOVERNMENT SECONDARY SCHOOLS IN MFOUNDI DIVISION, YAOUNDE, CAMEROON[i]

Chaffi W. Cyrille Ivan[ii]
PhD, Faculty of Education,
University of Yaounde I,
Cameroon

Abstract:
The purpose in this study is to investigate the correlation between assessment issues as a predictor of students' effort and achievement in some Government Secondary Schools in Mfoundi Division-Yaounde. A simple random sample of 201 students and 10 volunteer teachers drawn from four Secondary Schools was used. Pearson Product Moment Correlation Analysis were used to answer the research questions and to test the stated hypotheses at 0.05 level of significance. The reliability of the questionnaire gave an index of Cronbach's Alpha = .952. The data were analysed using content analysis for qualitative data and statistics with SPSS version 21 for quantitative data. In-depth of analysis of data obtained from questionnaire and interview guide focus on the perceptions that respondents have relating to assessment, revealed that there is a strong positive correlation between assessment issues and students' effort and achievement. Based on the outcomes of the study it was recommended that school authorities to train and recycle teachers on how to assess students, this to guarantee the validity of the content, the reliability of the assessment and to avoid bias against any results.

Keywords: assessment quality, assessment practices, assessment validity, assessment reliability

Résumé :
Le but de cette étude est d'étudier la corrélation entre les problèmes d'évaluation en tant que prédicteur de l'effort et des résultats des élèves dans certaines établissements d'enseignement secondaire publiques du Département du Mfoundi-Yaoundé. Un échantillon aléatoire simple de 201 élèves et 10 enseignants volontaires de quatre

[i] EXAMINER LA QUALITE DES PRATIQUES D'ÉVALUATION COMME PREDICTEUR DE L'EFFORT ET DES RÉSULTATS DES ÉLÈVES DANS CERTAINS ÉTABLISSEMENTS D'ENSEIGNEMENT SECONDAIRE PUBLIQUES DU DÉPARTEMENT DU MFOUNDI-YAOUNDÉ.

[ii] Correspondence: email cyrille.chaffi@univ-yaounde1.cm

établissements secondaires a été utilisé. L'analyse à travers le test de corrélation de Pearson a été utilisée pour répondre aux questions de recherche et pour tester les hypothèses énoncées avec α = 0,05 comme seuil de significativité. La fiabilité du questionnaire a donné un indice de l'alpha de Cronbach = 0,952. Les données ont été traitées à l'aide de l'analyse de contenu pour les données qualitatives et des statistiques avec SPSS version 21 pour les données quantitatives. L'analyse approfondie des données obtenues à partir du questionnaire et du guide d'entretien sémi-directif, centrée sur les perceptions des répondants concernant l'évaluation, a révélé qu'il existe une forte corrélation positive entre les problèmes d'évaluation, les efforts et les résultats des élèves. Sur la base des résultats de l'étude, il a été recommandé aux autorités scolaires de former et de recycler les enseignants sur le processus d'évaluation des élèves, afin de garantir la validité du contenu, la fiabilité de l'évaluation et d'éviter tout biais dans les résultats.

Mots-clé: qualité de l'évaluation, pratiques d'évaluation, validité de l'évaluation, fiabilité de l'évaluation

1. Introduction

The assessment of student learning is one of the essential parameters in school activities. Generally, the quality of an assessment has a significant influence on the learner's performance in class as well as during national examinations (ISO, 1986; Stiggins, 2001). Assessment in literacy work involves understanding how assessments are made, what type of assessments answer what questions, and how the data from assessments can help teachers, students, parents, and other stakeholders make decisions about teaching and learning. Assessment designers strive to create assessments that show a high degree of fidelity to the following five traits or concepts of this study's variables; Validity, Reliability, Fairness, Consequential Relevance and Student Engagement and Motivation. According to Evans (2013), a student might be engaged in the assessment but still not enjoy the activity of being assessed. Engagement speaks to the student's effort into the assessment since this is an instrument that influences students' effort and achievement (Gruender, 1996). For assessment to be effective and of good quality, it should be aligned with the curriculum objective, address learners' competencies through an emphasis on valid and reliable items that will facilitate the examination of students' abilities to achieve outcomes. In that view, Brown, Bull and Pendlebury (1997: 6) declare that "*assessment defines for students what is important, what counts, how they will spend their time and how they will see themselves as learners. If you want to change student learning, then change the methods of assessment*". Consequently, the assessment should be more valid, accurate with consistent results and fair to the teaching and learning process.

2. Statement of the problem

The purpose of education is to equip citizens with skills, values, knowledge to reshape their society and eliminate inequality (Kimani, Kala, & Njagi, 2013). On this basis, learners put in considerable time and finances to undergo training in schools with the expectation to come out with good results and abilities that will permit them to further their education. They expect that the system of evaluation should be reliable, valid, fair and relevant at all time for their performances to be well measured. However, assessing students' abilities seems to be losing quality within secondary education institutions in Yaounde.

Research and observation have shown that in Cameroon, most teachers give students a surprise test; this test mostly focuses only on one aspect taught and may sometimes be considered for end of sequence exams. The exams sometimes are too ticklish for students to decipher, with unconventional techniques that probably misdirect students from focus (Gronlund, 2006). Such exams may be set on what was not even done in class, without a focus to test a particular skill; some are just a transfer of past exams with any seasonal modification or update. Despite all the measures that the government had put in place to train and educate teachers during seminars just to ensure they give out a good quality assessment to students during evaluation so that their performance will be outstanding as compared to previous years when these trains were not put in place (Otaya, Kartowagiran & Retnawati, 2020). Yet, some teachers are not still able or willing to apply the technique of what they were taught to assess their students. These practices are detrimental to student's performances, and the results that have might not be linked to the learner's fundamental competencies.

The Classical Psychometric theory (Traub, 1997) is relevant to this research since its centres on the type of testing and the scores and their impacts on the learner. Classical test theory helps to observed test scores (TO), which are composed of a true score (T) and an error score (E) where the true and the error scores are independent. This function helps to create the relationship between variable as students' performances are easily measured through the scores. It helps us avoid errors in the testing process and provide a concept of correlation on indexing it. This theory creates a relationship between the two variables. The testing items determine student's performances. This theory has made it clear that if teachers don't make errors during testing, students' true scores will be obtained because the teachers have avoided this error in the cause, thought's known that errors in testing make it challenging to get students' true scores. Therefore, it is only when test conditions are taking into consideration like, how to construct a good test, how it will be scored and how and to whom it has to be administered that students' individual performance will be read as a valid score (Van der Scheer, Bijlsma & Glas, 2019).

2.1 General research hypothesis

Quality of assessment practices significantly affects students' effort and achievement in government secondary schools in Mfoundi Division.

2.2 Specific research hypotheses

RH1: Ha. There is a significant relationship between Evaluation Validity and students' effort and achievement in Some Government secondary schools in Mfoundi Division.
RH2: Ha. Evaluation Reliability is significantly linked to students' effort and achievement in some Government secondary schools in Mfoundi Division.

3. Methodology

The researcher had four schools that he worked with, and these schools are found in Yaounde in the Centre Region of Cameroon, in the Mfoundi Division. Those schools selected through a cluster sampling technique are GBHS. Etoug-Ebe, GBHS. Mendong, GBPHS, LBA and GBHS. Nkol-Eton. These schools are all bilingual schools with English and French sections (French/English). Each of these schools has about five classes, each from Forms One to Four and at least 2 or 3 classes each for Form 5, Lower and Upper Sixth. The data were collected through a questionnaire for 201students (74 male, (36,8%) and 127 females, (63,2%) constructed along with the pattern of the Likert scale: Strongly agree (SA), Agree (A), Strongly disagree (SD), Disagree (D), and interview guide for teachers (10). The items in the questionnaire design as follows: Assessment validity (10 items), Assessment reliability (09 items), demographic data (5 items). The interview guide contains five items to explore the perceptions that teachers have relating to assessment and equally to seek out teachers experiences on assessment and what is implicit about the issue of students' performance. The reliability of the questionnaire gave an index of Cronbach's Alpha = .952. The data were analysed using content analysis for qualitative data and statistics with SPSS version 21 for quantitative data.

4. Results

The first phase of the results deals with data presentation regarding students' opinions on assessment validity, reliability, and impact on their effort and achievement. The second phase comprises the test of hypotheses with Pearson moment product correlation.

4.1 Assessment validity

The results obtained from this item (Assessment Validity) present the variation in the perception of evaluation validity on quality assessment practices in schools following the given items. Generally, the participants, on these items, perceive that evaluation validity is not necessary for education for quality assessment practices because they believe that no matter how a test is given, if you have read, you will make it successfully; that is why we have (M=3.69; SD=1.066). Specifically, they believed that tests are based not only on what is taught in class since there are various ways they are testing with (M=3.97; SD=.916). Also, it is perceptive that most of our teachers do not often test students based on what is taught at the beginning of the class, with reasons that there were many lessons taught; therefore, they have many other areas that they can test them on (M=3.27;

SD=1.179). Equally, some students think it is unnecessary that at the beginning of each section of a test, instruction should be given on how to answer questions since most of the questions themselves already tell you what you need to do based on their setting (M=4.26; SD=.863). Also, they believe that marks allocated per question do not correspond to the number of elements needed as answers to the questions, which causes them to write more and at times get caught up by time shortage during exams (M=3.51; SD=1.221). Finally, each test does not have to be constructed based on all that has been taught in class because if that is done, the implication is that other evaluation methods would not be useful in education (M=3.45; SD=1.153).

4.2 Assessment reliability

Following this item on students' perceptions of test reliability on quality assessment practices, generally, students perceive these items on test reliability to be of greater importance as far as quality assessment practices in education are concerned. They believe that some teachers give tests objectively, care about, and ensure that all tests given are reliable (M=3.02; SD=1.301). Specifically, they think that they cannot have the same result if the same test with the same conditions is given because they believe that their teachers cannot test them twice on the same thing due to other areas of lessons taught in class, whereas they have the tendency of reading all that was taught (M=2.37; SD=1.313). Equally, it is usually challenging for them to succeed since their performance in a test is different when their scripts are marked by other teachers who are not used to the student's way of writing, and so could mark things that are right wrong and some wrong, right (M=2.95; SD=1.399). They believe that a conducive environment for exams is unnecessary for students to get good results if they write a test in such an environment because they believe that environment does not have a big part to play in exam success. They also think that if a student has studied well for an exam, no matter where he will write it, that student will succeed (M=3.49; SD=1.327). Testing students in different groups is not practically possible because some weak students will hide behind the intelligent ones, and this will not help advance education because such students will be blocked ahead and will be forced to drop out or use malpractice to succeed. Therefore, group testing is not the best to be implemented in some schools, especially if the teachers do not know their students' ability (M=2.51; SD=1.311). From the above, students are always tested using different testing methods that cover almost all the lessons taught, which makes tests very difficult for them since the teacher is not versed with a particular method in testing them (M=3.78; SD=1.157). In this light, Teacher X **(**He is a Physics teacher teaching the first and second cycles) declares: "*Yes, I evaluate students following the school program; however, I decided to give them some kind of evaluation to see how the lessons have gone through based on their strengths and weaknesses on the concept taught. I always evaluate them in groups theoretically and sometimes practically because, as far as Physics is concerned, we do have practical sessions during which students receive instructions to follow to collect data that I mark to see if they have gotten the concepts. Generally speaking, most of my exams or tests are always done in-group"*.

4.3 Students' effort and achievement

Concerning the above item where students perceive their effort and output to be of no importance to impact education in our country, generally, students perceive their performance not to be useful in education because it does not impact it, as we fight to get marks rather than the lessons taught actually to understand them (M=3.69; SD=1.095). We observe that it was not easy for them to have excellent results during the last sequence test, given that education today does not reflect that of yesterday; today's youths are blown away by the pleasures of the world rather than staying calm and taking their time to prepare for and succeed in exams (M=3.29; SD=1.177). Sometimes, it is dangerous for students to get good results through motivation and praise because when some receive that, they tend to perform poorly instead because they relent their effort knowingly or unknowingly (M=3.73; SD=1.148). Some see no real tendency of them performing better when exam conditions are favourable since they believe that whatever condition you write under does not matter but rather what and how you wrote (M=3.83; SD=1.123). When instructions are clearly given, students always obtain good results from the test, which is one of the best ways to test knowledge; clear instructions can make students have a good performance (M=3.81; SD=1.022). Some students perform well each time the teacher informs them of a test ahead of time. This is not important; students are different; some can succeed even without pre-notification about the test since they read continuously (M=3.82; SD=1.006). These results confirmed consistently linked to the assertion of teacher Y, who declares that: "*Before we run out a test item, we make sure that the instructions are clearly given to avoid errors in carrying out a certain concept. Of course, mark allocations are also there to guide the students from the answers required to give from the questions that I ask them in the test. After the test constructions, we sit down as a team of Physics teachers and multiply the questions and the time for the test by 3 given that they have different sections to evaluate them*".

4.4. Verification of Hypotheses

RH1: Ha. There is a significant relationship between Evaluation Validity and students' academic performance in Some Government secondary schools in Mfoundi Division.

Here, the correlation between evaluation validity and students' performance were examined according to the respondents' views. The first hypothesis is addressed here which was stated two-fold (Ha, Ho) thus: There is no relationship between evaluation validity and students' performance on the one hand and on the other hand, there is a relationship between evaluation validity and students' performance, to be verified and ascertain its relationship and its level of significance. To confirm this relationship, the hypothesis was tested with the Pearson correlation, and the following results were obtained and presented on the tables below.

Table 1: Descriptive statistics of evaluation validity and students' performance

Descriptive Statistics			
	Mean	**Standard deviation**	**N**
Evaluation Validity	3.6935	.58001	201
Students' Performance	3.6975	.63501	201

Source: Field Work (2020).

Table 2: Correlation matrix between evaluation validity and students' performance

Correlation			
		Evaluation Validity	**Students' Performance**
Evaluation Validity	Pearson's Correlation	1	.099
	Sig. (bilateral)		.162
	N	201	201
Students' Performance	Pearson's Correlation	.099	1
	Sig. (bilateral)	.162	
	N	201	201

Source: Field Work (2020).

Table 2 shows that there is no correlation between evaluation validity and students' performance (r (201) = .099, p-value = .162). So, hypothesis 1 is rejected. This is rejected because students did not see evaluation validity as a strong variable to play on students' performance in exams just because they are not versed with it or simply because their teachers do not consider when giving them a test. That is why they find it difficult to understand its use for them to get good performance thanks to the quality of assessment practices put in place.

Concerning the item link to evaluation validity, the interview results show that it is acceptable for the students to be instructed on how to answer the different items. Without proper instruction, it will not be easy to evaluate them because, during the evaluation, nothing is taken for granted, and they need to know to enable them to answer the scripts' questions. Therefore, the question itself must be obvious to help students in answering the questions. As said by an English teacher of GBHS Nkoleton, instructions guide the teacher and the student. Each subject or sections of the test scripts have an instruction to guide both the teacher and the student since they are divided into different sections. For example, in the English Language, which has five other areas; Grammar may be allocated 10 marks, Vocabulary 10 marks, Reading Comprehension 10 marks, listening 10 marks and Writing 10 marks, giving a total of 50 marks. An example of Grammar instruction could be reading the questions carefully and answering the questions that follow, for each question carries its marks. Mark allocation is crucial during testing because it aids and orientates the students to know the length of the answers that are required from him/her. For instance, a question with one mark tells the student that he/she does not need to write much, whereas that of 5 marks tells you the length of the answer required from you. On the other hand, a script without mark allocation makes answering and marking very difficult because not all students will write what is required of them.

RH2: Ha. Assessment Reliability is significantly link to students' academic performance in some Government secondary schools in Mfoundi Division.

Here, the researcher examines the correlation between evaluation reliability and students' performance according to the respondents' views. The second hypothesis is addressed here which was stated two-fold (Ha, Ho) thus: There is no relationship between evaluation reliability and students' performance on the one hand and on the other hand, there is a relationship between evaluation reliability and students' performance, to be verified and ascertain its relationship and its level of significance. To confirm this relationship, the hypothesis was tested with the Pearson correlation, and the following results were obtained and the following tables.

Table 3: Descriptive statistics of test reliability and students' performance

Descriptive Statistics			
	Mean	**Standard deviation**	**N**
Test Reliability	3.0194	.68167	201
Students' Performance	3.6975	.63501	201

Source: Field Work (2020).

Table 4: Correlation matrix between test reliability and students' performance

Correlations			
		Test reliability	**Students' Performance**
Test Reliability	Pearson's Correlation	1	.238**
	Sig. (bilateral)		.001
	N	201	201
Students' Performance	Pearson's Correlation	.238**	1
	Sig. (bilateral)	.001	
	N	201	201
**. The correlation is significant at 0.01 (bilateral).			

Source: Field Work (2020).

Table 4 shows that there is a correlation between test reliability and students' performance ($r(201) = .238$, p-value = .001). So, hypothesis 2 is accepted. This result implies that students who perceive tests as reliable get good marks, showing that test reliability is a powerful variable used to get the quality of assessment practices that have a role to play in students' performance. This is to say, for us to achieve this in schools where our students will be performing excellently well, this variable has to be applied and put in place during testing in schools. Therefore, this stood to contribute to this work's success because it helps us understand why students' results are what they are.

Table 5: Regression between test reliability and students' performance

$R^{2\ adjusted}$ = .052; F (1.199) = 11.918, p-value = .001						
Coefficients						
Model		**Non-standardised Coefficients**		**Standardised Coefficients**	**T**	**Sig.**
		A	Standard Error Margin	Beta		
1	(Constant)	3.029	.199		15.257	.000
	Test Reliability	.221	.064	.238	3.452	.001
a. Dependent Variable: Student' Performance						

Source: Field Work (2020).

Table 5 shows that there is a significant regression between test reliability and students' performance. Therefore hypothesis 2 is accepted. This means that test reliability is a good predictor of students' performance as seen on the p. value, which is .001; that is why it can be explained here where almost 25% of students' performance variances was (β = .238, t (200) = 15.257, p< .0001). This variable is accepted here still under regression to tell us how important it is to apply that in any test we want to give to the students, so their marks will reflect their performance if this is considered before testing.

The results from the interview show that most of the teachers gave a negative response. These responses were justified, probably because the students might not be expecting the same test or questions to be repeated to them in the next test. Some might score the same marks or above what they achieved in that test, while others might not or score below their first mark just because they were not prepared for it, thereby lowering their performance in the test.

Another group of teachers assured us that usually, it is better, but difficult to give the same test twice because they hardly do that except under some exceptional circumstances or exam revision. That notwithstanding, looking at the CBA in which much is taught, they do not see why we should test students on the same things when there are many other areas to test them on. We can do that only if we want to determine their level of attainment from the lessons taught. But there are cases where you can test even three times: the first time 18, the second time ten and the third time 5. This could be due to the other, maybe inattentiveness in class, mostly with the junior level. However, it is unnecessary because you might give a test now, and the scores will not be the same when you give it the second time.

5. Discussion

Rr1: There is a significant relationship between Assessment Validity and students' academic performance in Some Government secondary schools in Mfoundi Division.

From the questionnaire, Assessment validity and students' performance (r (201) = .099, p-value = .162). So, hypothesis 1 is rejected. The findings indicate that assessment validity negatively influences students' performance in secondary schools in Mfoundi Division. This, therefore, is interpreted in this work as: if the teachers and the

management put all other resources together to improve or ensure the validity of the examination given to students in secondary schools, their performance will improve. Assessment validity was examined here in relation to having the test items measure what it is intended to measure (Mbua, 2003). This implies that teachers who designed assessments focused on specific skills, topics or expecting to see certain behaviour; the assessment items should be focused on that particular aspect.

Furthermore, the results also imply that according to the respective schools where the data were generated, the rate at which test items test exactly what it is planned to test does not have any impact on students' performance in school. This result supports Traub (1997) Classical theory in which he states that there were generally always going to be errors in test measurements that these errors are random variables, and finally, that they could be correlated and indexed (Mohajan, 2017).

Validity occurs when the assessment can provide truthful information about what the students have learned (Barak & Haick, 2020; Black & Wiliam, 2018, Broohart, 2001).). The task itself should be assessing what you intend it to assess (EST 1084). An example would be that if you want to assess students' ability to scrutinise information, but the test only assessed students' recall of facts, this lacks validity. Valid assessments can measure the depth of understanding from learning outcomes. Ensuring that students are equipped with the skills needed to do an assessment task by teaching them what you will assess is the most effective method of achieving this outcome. The same situation occurs with the findings of this research wherein the respondents' results show that validity may have an important role to play in test item selection, but it does not influence students' performance in any way. The concept of validity in assessment could be more valid with teachers' ability to select the right questions that test the right skills they wish to test; on the other hand, the students' role may be to respond to the respective items using the knowledge learned in the classroom, thus allowing their teachers to ascertain understanding and knowledge application (Van der Scheer, Bijlsma & Glas, 2019, Morad, Ragonis & Barak, 2021). The majority of the interviews gave corresponding responses. There is a need to sort out the respective processes to ensure quality assessment in secondary school. This is because the assessment is used in the promotion process or as an indicator of school quality (Carnoy, 1999). Continuous assessment is also used to provide teachers with feedback about students' performance and achievement.

Rh2: Assessment Reliability is significantly link to students' academic performance in some Government secondary schools in Mfoundi Division.

The results showed a significant relationship between evaluation reliability and students' performance in the selected secondary schools in Mfoundi Division. This implies that when the evaluation process is reliable (without any error), the students tend to perform better than not. This result ties with Spearman's declaration that if students take the same test several times, they should get about better scores each time. So, assuming that the conditions are the same, they will get better scores on a test because the test itself is well designed. A variety of factors could influence observed test scores,

including how the items were generated, tests were administered, and items were scored (especially if they were free responses scored by raters). The impact of assessment reliability on learners' performance is significant; this implies that there is a teacher need more time to study studying of evaluation to organise evaluations in their classrooms effectively. In this light, Airasian (1994), Carey, (1994); O'Sullivan & Chalnick, (1991)) opined that classroom assessment had received increased attention from the measurement community in recent years. Since teachers are primarily responsible for evaluating instruction and student learning, there is widespread concern about classroom assessment quality (Barak & Haick, 2020; Black & Wiliam, 2018). The literature on classroom assessment has delineated the content domain in which teachers need to develop assessment skills. The teacher is called upon to master the science in measuring test items (Saoul, 2013). A case that best explains the process of assessment reliability could be understood here as if a person weighs himself severally in the course of a day, he would expect to see a similar reading. Scales that measured weight differently each time would be of little use. The same analogy could be applied to a tape measure that measures inches differently each time used. It would not be considered reliable. If findings from the research are replicated consistently, they are reliable. A correlation coefficient can be used to assess the degree of reliability. If a test is reliable, it should show a high positive correlation. The more experienced and knowledgeable a teacher is in pedagogy (teaching and evaluation) and motivated to implement the methods, the test items will be reliable and students' performance will be on the constant rise. Once learners perform will throughout the learning institutions, they have more chances of succeeding in future. They are open to scholarships, good jobs, advantaged in the competitive world when it comes to certificate evaluation. They also make good alumni for their institutions and showcase the almamater's name to society. These make prove of quality in an educational system.

According to results from the interviews, Teacher A from Etoug-ebe High School confirmed this result as she said that students perform better when the students meet the same or expected tests or questions in an evaluation. She added that she personally endeavours to revise past questions and get other questions from different schools in order to get her students acquainted and familiar with all the possible questions that may be asked in the taught curriculum. Others confirmed that repeating the same test to students is largely a way to enhance performance.

6. Conclusion

The quality of assessment practices plays a significant role in students' academic performance. The conditions under which assessments are conducted can ensure yes or not that questions are well-written, unambiguous, and have been vetted for accuracy. From the results of this study, it will be very important for teachers to be trained and have tailored coursed on how to assess, this to guarantee the validity of the content, the

reliability of the assessment, that will generate interest on the students will consequential relevance with pertinent and valid results.

Conflict of Interest Statement

The authors declare no conflicts of interests.

About the Author

Chaffi W. Cyrille Ivan (PhD), University of Yaounde I, Faculty of Education, Department of Curriculum and Evaluation, Senior Lecturer, email: cyrille.chaffi@univ-yaounde1.cm.

References

Airasian, P. W. (1994). *Classroom assessment*. New York: McGraw-Hill.

Barak, M., Watted, A., & Haick, H. (2020). Establishing the validity and reliability of a modified tool for assessing innovative thinking of engineering students. *Assessment & Evaluation in Higher Education*, *45*(2), 212-223.

Black, P. & William, D. (2018). Classroom assessment and pedagogy. *Assessment in Education: Principles, Policy & Practice*, *25*(6), 551-575.

Broohart, S. M. (2001). Successful Students. Formative and Summative use of Assessment Information. *Assessment in Education, 8,153-169.*

Brown, G., Bull, J. and Pendlebury, M. (1997). *Assessing Student Learning in Higher Education.* London and New York: Routledge.

Carey, L. M. (1994). *Measuring and evaluating school learning*. Boston: Allyn and Bacon.

Chan, L. L., & Idris, N. (2017). Validity and reliability of the instrument using exploratory factor analysis and Cronbach's alpha. *International Journal of Academic Research in Business and Social Sciences*, *7*(10), 400-410.

Davies, A. (2000). *Making Classroom Assessment Work*. Merville, BC: Connections Publishing.

Gronlund, N. E. (2006). *Assessment of student achievement* (8th ed.). Boston: Pearson.

Gruender, C. D. (1996). Constructivism and learning: a philosophical appraisal. *Educational Theory, 36:21-29.*

Kavanaugh, B. C., Tuner, O. F., & Wexler, B. E. (2019). Measuring and improving executive functioning in the classroom. *Journal of Cognitive Enhancement*, *3*(3), 271-280.

Mohajan, H. K. (2017). Two criteria for good measurements in research: Validity and reliability. Annals of Spiru Haret University. Economic Series, 17(4), 59-82.

Morad, S., Ragonis, N., & Barak, M. (2021). The validity and reliability of a tool for measuring educational innovative thinking competencies. *Teaching and Teacher Education*, *97*, 103193.

O'Sullivan, R. G. & Chalnick, M. K. (1991). Measurement-related course work requirements for teacher certification and recertification. Educational Measurement: Issues and Practices.

Otaya, L. G., Kartowagiran, B., & Retnawati, H. (2020). The construct validity and reliability of the lesson plan assessment instrument in primary schools. *Jurnal Prima Edukasia*, *8*(2), 126-134.

Souza, A. C. D., Alexandre, N. M. C., & Guirardello, E. D. B. (2017). Psychometric properties in instruments evaluation of reliability and validity. *Epidemiologia e Serviços de Saúde*, *26*, 649-659.

Stiggins, R. J. (2001). Student-involved classroom assessment. 3rd edn. New Jersey: Prentice Hall.

Tabuena, A. C. (2020). Effectiveness of classroom assessment techniques in improving performance of students in music and piano. *Global Researchers Journal*, *6*(1), 68-78.

Traub, R. (1997). Classical Test Theory in Historical Perspective. *Educational Measurement: Issues and Practice 16 (4), 8–14.*

Van der Scheer, E. A., Bijlsma, H. J., & Glas, C. A. (2019). Validity and reliability of student perceptions of teaching quality in primary education. *School effectiveness and school improvement*, *30*(1), 30-50.