

**AN ANALYSIS OF THE QUALITY OF TEACHER – MADE MULTIPLE
CHOICE TESTS USED AS SUMMATIVE ASSESSMENT FOR ENGLISH
SUBJECT AT SMP NEGERI 5 SINGARAJA**

Oleh :

Tikvi Yulia Wahyu Saputri

English Language Education

Ganesha University of Education Singaraja, Indonesia

e-mail: yulia.wahyu.saputri@undiksha.ac.id

ABSTRACT

The teacher – made multiple choice test that used by the teachers to assess the students' level must have a good quality by following certain norms in constructing it. This study was conducted to analyze the quality of the teacher – made multiple choice tests used as a summative assessment especially for middle test in the first semester at SMP Negeri 5 Singaraja. This study used content analysis design. The data were obtained from 4 different teacher – made multiple choice tests with the total 100 items. Document analysis was used to compare the teacher – made multiple choice tests with the norms in making a good multiple choice test to decide the quality of each item. Then, the interview was done to get additional information. The result shows that all of the items of teacher – made multiple choice tests fulfilled more than 75% norms, means that they belong to very good quality. However the students' National Examination score was low. Thus, it can be concluded that the quality of teacher – made MCT is not the only factor that affects the students' National Examination score.

Keywords : Teacher – made multiple choice test, instrument quality, norms, summative assessment

1. Introduction

According to article 70 section 3 and 5 Government Regulation number 19 about National Education Standards states that English is one of the subjects that will be tested in National Examination in Junior and Senior High School. It means that English is one of the main subjects that must be taught in Junior High school students. According to Reece and Walker (1997), there are four aspects that must be exist in teaching and learning process and one of those aspects is assessment. An academic course is not complete without an assessment plan (Rehman et al, 2018).

According to Freeman and Lewis (1998), assessment is a process aimed at assessing students' learning. According to Rodriguez (2014), assessment affects the students through the practices that are used by the teachers to assess the students. Moreover, according to Angelo and Cross (1988), in order to evaluate what students are learning in the class and how well they learn, classroom assessment can be used by the teachers as an approach to help them. There are two kinds of assessment that can be used, there are formative and summative assessment. Chappius&Chappius (2007) state that formative assessment delivers information while the learning process, it is ongoing and dynamic process before the summative assessment has done. On the other hand, Summative assessment is used to measure the students about certain material that have been taught in the class in order to know how far they learn, and it is done in a certain time that is determined, such as in the middle of the semester or at the end of the semester (Kibble, 2016). By assessing the students, the teachers can diagnose the strengths and the weaknesses of their students and the outcome also can be a reflection for the teachers to improve their teaching performance.

Assessment is regulated in 2013 Curriculum. It is regulated in *PerMendikbud* (The Ministry of Education and Culture Regulation) No.23/2016 about Educational Assessment Standard. It is used as the basis for 2013 curriculum assessment standard. Based on the regulation in (Article 3 section (1)) There are three main points that should be assessed, those are: attitude, knowledge and skill. In line with this regulation, (Article 9 section (1) point (c)) states that the student knowledge aspects can be assessed through written tests, oral tests, and assignments based on competency that they want to achieve. Under the regulation, teachers can test students' knowledge through written tests and one type of written tests is multiple choice test.

Multiple choice test is a form of test that commonly used in education in Indonesia (Kusumawati and Hadi, 2018). Besides, according to Roberts (2006), Multiple choice tests have been used extensively for many years for academic assessment purposes. The most common examples of tests that use multiple choice tests are TOEFL, IELTS, and TOEIC. Multiple choice test is good for

measuring knowledge, comprehension, and can be designed to measure application and analysis (Hameedetal, 2005). The multiple choice test must fulfill certain standards in making a good multiple choice test in order to have a good quality because it is used as an instrument to assess the students' knowledge.

According to Burton et al (1991), the multiple choice tests' quality can be determined from the norms that are implemented while constructing it. In line with this statement, there are four theories that support this statement. 1)Haladyna (2004),2)Hall and Marshal (2013), 3) Zimmaro (2016), 4) *Puspendik Kemendikbud 2019*. Haladyna (2004) states that a certain guidelines or norms must be implemented in writing the items of multiple choice test. Those are 31 norms including 4 categories: content guidelines, style and format concern, writing the stem, and writing options;Hall and Marshal (2013) state that there are 12 norms that must be concerned in making good multiple choice test;Zimmaro (2016) states that there are 50 norms including 5 categories: procedural, content related, stem construction, general option development, and distracter development rules. *Puspendik Kemendikbud 2019*states that there are 16 norms including 3 categories: material, construction, and language.

In reality, based on the observation data showed that there are 7 English teachers in SMP Negeri 5 Singaraja. 6 teachers used multiple choice test but there are only 4 different multiple choice tests that were made by the English teachers in SMP Negeri 5 Singaraja. It was used to assess the students in the middle test especially in the first semester as summative assessment for seventh, eighth, and ninth grade. There are four basic competencies for seventh grade, those are: basic competency point 3.1 is about greeting, saying good bye, thanking and apologizing. Point 3.2 is about asking and giving for information about self identity and 3.3 about asking and giving for information about the name of the day, month and years. The basic competencies for eighth grade are: basic competency point 3.1 is about asking and giving for attention, understanding, and opinion. Point 3.2 is about asking and giving for information about ability and willingness. 3.3 is about asking and giving for information about compulsion and prohibition and point 3.4 is about asking and giving for information about

commanding, inviting, and permission. For ninth grade there are also four basic competencies that should be achieved: Point 3.1 is about the spoken and written interpersonal interaction in stating wishes and congratulate someone for the achievement. Point 3.2 is about asking and giving for information about agreement in doing an activity. Point 3.3 is about asking and giving for information relate to the label in medicine or food. The last one is point 3.4 about text procedure.

The test is very useful in order to know whether the students already achieve the learning objectives based on the basic competencies or not, because later on the students have to face Ujian Nasional (National Examination). The test items in National Examination reflect the basic competencies that the students' have learned before. However, in this case the students in SMP Negeri 5 Singaraja academic year 2018/2019 could not achieve the minimum standard score of English subject. The students' score average was 46,53 while the minimum standard score is 55,00 (*PuspendikKemendikbud, 2019*).

According to Black and William (1998a), good assessment practice can improve the students' achievement level. This case shows that the assessment practice in the school is not maximal yet. One of the indicators of good assessment practice is that the teachers are able to construct a good instrument to assess their students' achievement. The instrument that is used by the teachers in the school is multiple choice test. In constructing the multiple choice test there are certain norms that should be followed. Crockett & Churches (2016), state that the norms or guidelines in making a good test must be followed by the teachers when they make it. Thus, the norms is a starting point in constructing good multiple choice test, it should be concerned in the beginning of constructing the items. Hence, when the norms are implemented by the teachers they will produce a good quality of multiple choice test.

Based on the situation explained above, this study was conducted in order to analyze the quality of teacher – made multiple choice tests as a summative assessment especially for the middle test in first semester at SMPN 5 Singaraja. Since the norms of making a good multiple choice test is a starting point in

making the multiple choice test this study focuses on the analysis of the congruity between each item with the norms that have been synthesized from Haladyna, Hall and Marshall, and *PuspendikKemendikbud*. The quality of the items can be seen from the result of the analysis.

2. Method

This study used content analysis design which aimed to analyze whether or not the multiple choice tests that were used by the teachers in SMP Negeri 5 Singaraja to assess their students in the middle test for English subject already have a good quality by following certain norms in making a good multiple choice test. There were some considerations in using content analysis design in this study. First, Ary et al (2010) state that content analysis focuses on analyzing visual or written material such as textbooks, letters, film, and another document. The research question will get the best answer by analyzing documents. Bengtsson (2016) also argues that content analysis can be used to analyze all types of written texts no matter where it comes from. Second, according to Elo (2007), the purpose of content analysis is to get a legit and broad description of a certain phenomena, and the result of the analysis is a concept describing the phenomena.

This study was done at SMP Negeri 5 Singaraja and took 4 teacher – made MCTs as the subject of the study and the object of the study is the quality of the teacher – made MCTs based on the congruity between each item of MCTs with the 18 norms of making a good MCT. Document analysis and interview were used as the methods to collect the data in this study with the assistance of checklist and interview guide as the instrument. Checklist was used to examine whether or not the items of the MCT follow the norms in making a good multiple choice test and the interview guide was used as the guidance to get more information related to the result of the document analysis.

The quality of the teacher made MCT was analyzed by comparing the total items of MCTs that were used by the teachers with the 18 norms in making a good

MCT that are suggested by Haladyna (2004), Hall and Marshall (2013), and *PuspendikKemendikbud* (2019). In this case, checklist was used to compare the items with the norms to get the data that will be analyzed further statistically. After that, the data was analyzed by using formula that is suggested by Nurkencana&Sunartana's (1992). Then, the results were calculated and classified to some classifications. The classifications were determined by using the following formulas:

Tabel 1
Data Classification Formula

Interval	Criteria
75% x	Very Good
58% $x < 75%$	Good
42% $x < 58%$	Sufficient
25% $x < 42%$	Poor
$x < 25%$	Very Poor

Based on the table above, it can be seen that there are five classifications. Those are very good, good, sufficient, poor, and very poor. If the percentage of teacher-made MCT is more than or equal to 75% it belongs to very good. If the percentage is more than 58% or less than 75% it belongs to good. If the percentage is more than 42% or less than 58 it belongs to sufficient. If the percentage is more than 25% or less than 42% it belongs to poor and the last one is very poor if the percentage is less than 25%.

3. Findings and Discussion

In this study the total 4 teacher – made multiple choice tests were analyzed. There was an instrument or 1 multiple choice test for seventh grade (include 25 items), 2 multiple choice tests for eighth grade (25 and 30 items), and 1 for ninth grade (include 25 items). Thus, the total number of the items were 100. The quality of those items were analyzed based on the congruity between the items

and the 18 norms that were synthesized from Haladyna (2004), Hall and Marshall (2013), and *Puspendik Kemendikbud* (2019). In this case *Microsoft Excel* was used to analyze the data.

The percentage of each item toward followed norms can be divided into 5 categories. Those are: 1) 100% which means that the item totally followed 18 norms, 2) 94 % there is a norm unfulfilled, 3) 89 %, two norms unfulfilled, 4) 83%, three norms unfulfilled, 5) 78%, the lowest fulfilled norms which is four norms unfulfilled. There are 4 items that completely followed the 18 norms which is 100% (4%), 27 items are 94% fulfilled the norms (27%), 44 items are 89% fulfilled the norms (44%), 23 items are 83% fulfilled the norms (23%), and 2 items are 78% fulfilled the norms (2%). From 78% to 94% fulfilled norms have the same mistake in the use of punctuation and capitalization.

There were 4 items that completely followed all of the norms in making a good MCT. Those items were correct in grammar, spelling and punctuation. The items also reflected the basic competencies, not using opinion based question, independent or not depend on the other items and have a clear instruction. Moreover, the items followed the norms about punctuation, clues, word repetition, and overlapping. The options were plausible, homogeneous and formatted vertically. None of the items used double negatives and “*all of the above*”.

There were 27 items that 94% fulfilled the norms or there was only an unfulfilled norm. The norm that was being unfulfilled is norm number 9 about the style and format concern. In this division the use of punctuation and capitalization are arranged. All of 27 items have a mistake in this norm, such as:

1. There are more than three full stops at the blank space in the beginning and the options are ended with a full stop. It should not be so.
2. The options are capitalized and finished with full stop when the blank space is in the middle of the stem. It should be not capitalized and not finished with full stop.
3. Using more than four full stops at the end of the stem and starting the options with capital letter. When the blank space is at the end of the stem, it should be only four full stops and the options are not capitalized.

4. There are three or more full stops before a question mark and there is no full stop in the options. It should be no full stops before question mark and the option should be ended with a full stop.

Besides punctuation and capitalization there are 5 norms that were unfulfilled by a few of items. First, the item not giving a clear focus on what is being asked. In the stem, the students are asked to arrange a sentence based on the number in every word, but there is no number under the word, it will make the students confuse because the item is not clear (VIII A. 1) . Besides, there is an item that do not provide a dialogue even in the stem talking about the expression in the dialogue (VIII A. 3). Second, the item gives a clue to the correct answer. The clue appears in the option, for example the correct answer is the longest or shortest option and sometimes it is one space ahead than the other options (VII. 20). Third, incorrect spelling. For example the stem is asking about the time but the word time is spelled incorrectly became “tinme” (VII. 17). Forth, the options are not homogeneous. It happens because there is a different form in the options. For instance, 2 options are in the form of interrogative sentence (asking for opinion) but there is an option that is talking about giving opinion (VIII A. 3). Fifth, options are not placed in logical and numerical order. The option should be formatted from the shortest to the longest, but in this case there are some options put the longest answer in the middle of the option. Moreover, when using number in the option it should be placed from the highest to the lowest or vice versa but there is an item not followed it (VIII A. 1).

There were 66 items fulfilled above 80% norms. Those items unfulfilled 2 – 3 norms. 44 items unfulfilled two norms and 23 items unfulfilled three norms. The norms that were unfulfilled are varied such as what are mentioned above and another norm. Those were grammar, options and the distracters. 15 items had a grammar mistakes, it was found in the stem and sometimes in the options. In the options there were two problems, such as the options were not formatted vertically and the options were not in the same length. All of the options in Teacher – made MCT for ninth grade were formatted horizontally. The last one was the distracters. There were some distracters that are not plausible. It means

that the distracters in the options are not in the same context with what is being asked in the stem.

There were 2 items that had the lowest percentage of fulfilled norms which is 78%, means that 2% of the items unfulfilled 4 norms. The item has 4 problems. The first problem is incorrect grammar. For instance, the sentence “*What is his father job?*” the grammar structure in this sentence is not correct because the word “*father job*” needs a possession to make it correct and clear that what is being asked is his father’s job. The stem should be added with “ ‘s” and it becomes “*What is his father’s job?*”. Second is problem in punctuation and capitalization especially issue number 4, the options should be ended with a full stop because there is a question mark in the stem. Third, the option D is not in the same length with other options. The last or forth one is that the distracter is not plausible. The distracter D “*He is Mahendra’s father*” father is not a job, while the other distracters state about a job.

Based on the result of the item analysis above, it can be seen that all of the items have very good quality. Meanwhile, there are some norms that were unfulfilled. In this case, the most unfulfilled norm is the norm number 9 about the use of punctuation and capitalization. Based on the analysis only 9 items (9%) that followed the norm while 91 items unfulfilled it. For the rest of the norms are fulfilled by more than 70 items with the percentage start from 71% - 100%. The percentage of fulfilled norms of each item from the multiple choice tests were also analyzed. It can be seen on the table below.

Table 2

The Percentages of Fulfilled Norms

Norms Number	Norms’ Description	Number of Item Fulfilling Norm	Percentages of Fulfilled Norms (%)
1	Reflecting basic	100	100%

Norms Number	Norms' Description	Number of Item Fulfilling Norm	Percentages of Fulfilled Norms (%)
	competencies		
2	Not depending on the previous option	100	100%
3	Giving clear focus	91	91%
4	Avoiding opinion-based items	100	100%
5	Not giving clue to the correct answer	94	94%
6	Being grammatically correct	85	85%
7	Having correct spelling	99	99%
8	Options are formatted vertically	71	71%
9	Taking concern on the use of the	9	9%

Norms Number	Norms' Description	Number of Item Fulfilling Norm	Percentages of Fulfilled Norms (%)
	punctuation and capitalization		
10	Not containing double negatives	100	100%
11	Options are homogenous	98	98%
12	Having one correct answer	100	100%
13	Options have about the same length	88	88%
14	Options are placed in logical and numerical order	93	93%
15	Options do not repeat the same words or phrases	100	100%
16	Options are not	100	100%

Norms Number	Norms' Description	Number of Item Fulfilling Norm	Percentages of Fulfilled Norms (%)
	overlapping		
17	Distractors are plausible	80	80%
18	Not using "none of the above" or "all of the above"	100	100%

The result of the percentage of fulfilled norm is described on the table 2. It can be seen that the most unfulfilled norm is norm number 9 which is about the use of punctuation and capitalization. There are only 9 out of 100 items that fulfilled the norm. In writing the item of MCT this norm should be considered in order to avoid misunderstanding the meaning. Conelly (2009) states that inaccurate punctuation leads to the ambiguity. In regard to this statement, Samhon and Abdall (2016) argue that the wrong use of punctuation makes the idea and the meaning different from what is meant to be. Based on the interview with the test makers shows that the knowledge in using punctuation and capitalization especially in the process of making MCT is still less. When they were asked about this norm three of the test makers said that *"I do not really know about this."* while a test maker said *"Maybe I've learned about it but I forgot, So I did not really focus on it"* In line with the result of the interview, Kurniawan et al (2014) state that the teachers' comprehension in the use of punctuation is still low.

The norm about punctuation and capitalization is considered as the most unfulfilled norm, on the other hand there are 8 norms that are totally followed by all of the items (100%). They are reflecting the basic competency, not depending on the other items (Independency), opinion based items, double negatives, having

only one correct answer, word repetition, overlapping and not using “*none of the above*” or “*all of the above*”. From those 8 norms there are some norms that are understood by the teacher, so that they really take concern about it while the other norms are not.

Relating to the norm of reflecting the basic competency. Based on the interview, the teachers said that the first thing that must be considered in making questions is the congruity between every item with the basic competencies. The items were made to assess the students’ understanding towards the material given. One of the teachers also stated that one of the norms in making a good MCT is every item must follow the basic competencies that want to be achieved. Then, the independency. The teachers believe that when an option have a relation with another item, it can be a clue for the students and if there is a word repetition in the option the students will be confuse to choose the best answer. Using double negatives and “*none of the above*” or “*all of the above*” also can be a clue for the students to choose the answer. Besides, opinion based item is also avoided by the teachers because when the teachers use opinion based items, the students will answer it based on their opinion not their knowledge about the material.

The quality also can be seen from the other 2 norms which are about the location of the correct answer and the instruction. 3 out of 4 teacher – made MCT put the correct answer randomly, the correct answer does not make a pattern. However, in the teacher – made MCT for ninth grade the options are not placed randomly. Most of the correct answer are put in option A and C. For the instruction, the teacher – made MCT for eighth grade (VIII A) does not provide a clear instruction to the students, there is no instruction for the students to answer the questions with the best correct answer. Even though, the MCTs for VII, VIII B, and IX already provide a clear instruction such as “*Choose the best answer: A,B,C or D.*”.

Based on the interview, the test maker for ninth grade stated that he did not realize the key answer that he made is not random or the correct answer did make a pattern. Moreover, he did not check it twice because the limitation of time.

Based on the result of teacher – made MCTs quality analysis it can be said that all of the items of MCTs in SMP Negeri 5 Singaraja have a very good quality determined from fulfilled norms. It is also supported by previous studies done by Patil et al, (2016), and Hohensinn and Baghaei (2017), they state that in making a good MCT there are some norms that should be concerned. It is also confirmed by the teachers, they believe that they already made a good MCT and they are satisfied enough with their MCT.

Based on the interview, the test makers' prior knowledge in making a good MCT comes from two factors. First, the teachers got the knowledge from assessment course in college. Two of the test makers are students practice teaching. They are the students of Ganesha University who did practice teaching program at that time and they were given a responsibility to make items of MCTs to asses the students especially for the middle test. Yet, the supervisor teachers checked the MCT before it was given to the students. Second, the teachers join the workshop about making test item.

First, two out of four the test makers stated that they have the knowledge about making a good MCT from assessment course in a college. They are students in English Language Education department and Assessment is one of courses that is taught in their department. From this course they were taught how to get the information about their students' learning achievement by assessing them and also how to assess their teaching strategies as a teacher. The students can be assessed through testing, one of the instruments that can be used to test the students is MCT. Based on the interview, the test makers said that they have learned about how to make a good MCT from this course.

Second, the teachers join a workshop about making questions or test items for the students. The workshop that they joined was not specific for making a good MCT but it is a workshop of how to develop test items for the students. The teachers said that they joined the workshop that was held by the school and they have not joined a workshop held by another stakeholder.

Those are two factors that affect the quality of teacher – made MCT in SMP Negeri 5 Singaraja become very good. However, the quality of MCTs is not

in line with the result of National Examination result in academic year 2018/2019 for the English subject. The students' score average was 46,53 while the minimum standard score is 55,00. Then, it was confirmed by the teachers that they do not believe the quality of MCT will influence the students National Examination score but there are other factors that have a big role on it. Based on the interview, there are three factors that influence the achievement score of National Standard School Examination (USBN). Those are low vocabulary mastery, low motivation in learning and extra time for learning that is not effective for the students.

The first factor is low vocabulary mastery. According to Ratnawati (2006), vocabulary mastery is the ability of knowing and understanding some words in a particular language and also know their usage based on the context or situation. Based on the interview, the teachers believe that it is one of the big problems for their students because vocabulary is one of the main component in learning language. The students' vocabulary mastery should be improved because when the students have less vocabularies, they will face difficulty when they are taking the National Examination. It is in line with Daryanti (2015), states that vocabulary is an essential component that should be improved in order to improve their language skills including reading, listening, speaking and writing.

The second factor is low motivation in learning. According to Dakhi (2018), there are two kinds of motivation: intrinsic and extrinsic. Intrinsic motivation is an internal desire to learn, it naturally appears in an individual because of enjoyment and internal satisfaction while extrinsic is an external motivation that comes from the environment, means that they learn because they want to get some rewards or they avoid punishment (Dakhi, 2018). Based on the interview, the teachers stated that motivation plays an important role towards the students' achievement in learning. However, the students at SMP Negeri 5 Singaraja still have a low motivation in learning whether it is intrinsic or extrinsic. It can be seen from their performance in class, most of the students do not participate well during the learning process. Highly motivated students tend to be active students in class and also have a good achievement, such as good score in a test. The statements from the teachers also revealed by some related research had

done before by Baranek (1996), Abdurrahman (2014), Muhamad et al. (2016), and Tokan (2019). They believe that there is a positive relation between motivation with the students' learning achievement in class.

The third factor is the extra time for learning (Pengayaan) that is not effective enough. This activity actually is meant to be a preparation for the students before taking National Examination. The subjects that are taught in this class are Mathematics, BahasaIndonesia, Science, and English. The students are given the material related to the basic competencies or indicators for National Examination. It is done because the time for teaching and learning on school hours is not enough. However, the teacher said that the time that is chosen does not suit for the students. It is because they have to join this class after regular class, so they have to spend more time at school in the afternoon. This teachers' statement is supported by Gusneli (2011) states that when the school runs in the afternoon or even evening, then the students conditions will not be in their best. Moreover, Mursidin (2017), the most effective time for learning process is in the morning, it is because the physical and psychological conditions of the students is still fresh so they can easily acquire the material given by the teacher. Besides, Lestari (2015) states that some students tend to be more focus and easy to concentrate in the morning while in the afternoon the students are tired and sleepy. Thus, the learning process will be more effective in the morning.

Thus it can be concluded that the quality of teacher made MCT is not the only factor that affect the students' score but there are another, such as what were explained above. In this case those factors need to be concerned in order to improve the students' score in National Examination besides the quality of the MCT.

CONCLUSION

it can be concluded that 1) The quality of teacher – made MCT in SMP Negeri 5 Singaraja is very good. 2) 100% of the items have very good quality. 27 items are 94% followed the norms, 44 items are 89% followed the norms, 23 items are 83% followed the norms, and 2 items are 78% followed the norms. 3)

The most unfollowed norm is about punctuation and capitalization, only 9% of the items followed this norm. 4) The quality of the teacher – made MCT is not the only factor that affect the students’ achievement. All of the items already have very good quality but the average of students’ National Examination score was under minimum score.

SUGGESTION

Based on the result of this study, some suggestions will be directed to the teachers, stakeholders, and other researchers. They are as follow:

The teachers are expected to always reread and recheck their MCT before it is given to the students. It should be done in order to avoid some mistakes such as type wrong and not clear instruction. It was found that the teacher forgot to put number in the item and dialogue. The time management in making the MCT needs to be noticed, so the teachers will have enough time to recheck their MCT. Besides, they are also expected to find out and learn the norms in making a good MCT and then try to apply it while making MCT.

The stakeholders at SMP Negeri 5 Singaraja are expected to conduct a workshop for the teachers. The workshop of making a good MCT is needed in order to give the knowledge and give an opportunity to the teachers to practice how to make a good MCT based on the norms in making a good MCT.

For the other researchers, the researcher expects that they can do research in the same field which is MCT analysis. In this study the researcher analyze the quality of the MCT based on 18 norms that have been synthesized from 3 different sources. It was found out that the quality of the teacher – made multiple choice test was not the only factor that affect the students’ score in National Examination. For further research, the other researchers can conduct a study which try to analyze the MCT based on another standard of constructing a good MCT or they can analyze further the factors that affect the students’ National Examination score.

REFERENCES

- Abdurrahman, et al. (2014). THE IMPACT OF MOTIVATION ON STUDENTS' ACADEMIC ACHIEVEMENT IN KEBBI STATE JUNIOR SECONDARY SCHOOL MATHEMATICS. *International Journal of Advance Research, IJOAR* .org Volume 2, Issue 12, December 2014, Online: ISSN 2320-9143
- Ary, D., Jacobs, L.C., Sorensen, C, & Razavieh. (2010). *Introduction to research in education*. Canada: Wadsworth, Cengage Learning.
- Baranek, Lori Kay. (1996). *The Effect of Rewards and Motivation on Student Achievement*" (1996). Masters Theses. 285. <http://scholarworks.gvsu.edu/theses/28>.
- Black, P. & William, D. (1998a). "Assessment and Classroom Learning." *Assessment in Education: Principles, Policy, and Practice*. 5(1). 7-73.
- Bengtsson, Mariette. (2016). *How to Plan and Perform a Qualitative Study Using Content Analysis*. Faculty of Health and Society, Department of Care Science, Malmö University, SE 20506 Malmö, Sweden.
- Burton, S. J., Sudweeks, R. R., Marrill, P.F., & Wood, B. (1991). *How to prepare better multiple-choice test items: Guidelines for university faculty*. Brigham Young University Testing Services and The Department of Instructional Science.
- Chappius, Stephen and Chappius, Jan. (2007). *The Best Value in Formative Assessment*.
- Connelly, M. (2009). *Get writing: Sentences and paragraph*. Cengage Learning.
- Crockett, L.W. & Churches, A. (2016). *Mindful assessment*. Indiana: Solution Tree Press.
- Dakhi, S., & Damanik, I.S. (2018). Students' Motivation in Reading English Text: A Qualitative Study in EFL Context. *Journal of English Teaching* VOLUME 4, NUMBER 2, JUNE 2018, pp. 81-93.
- Daryanti, T. (2015). *THE CONTRIBUTION OF VOCABULARY MASTERY TOWARD READING COMPREHENSION*. Universitas Negeri Yogyakarta.
- ELO, S. & Kyngas, H. (2008). The qualitative content analysis process. *Journal of Advanced Nursing* 62(1), 107–115 doi: 10.1111/j.1365-2648.2007.04569.x.
- Gusneli, F.K. (2011). *Efektifitas pemanfaatan waktu belajar peserta didik di SMA Negeri 11 Padang*. STKIP PGRI Sumatera Barat.

- Haladyna, T. M. (2004). *Developing and Validating Multiple-Choice Test Items*. Mahwah, New Jersey, London: Lawrence Erlbaum Associates.
- Hall, and Marshall. 2013. *A Guide for Developing Multiple Choice and Other Objective Style Questions*. Centre for Academic Development, Victoria University of Wellington, New Zealand.
- Hohensinn, C. & Baghaei, P. (2017). Does the position of response options in multiple-choice tests matter? *Psycologica*, 38, 93-109.
- Kurniawan, A. (2016). *Graduate Students' Lived Experience in Writing Academic Papers*. Sanata Darma University: Yogyakarta.
- Kusumawati & Hadi. (2018). An Analysis of Multiple Choice Questions (MCQs): Item and Test Statistics from Mathematics Assessments in Senior High School, 4(1), 70-78.
- Kibble, Jd. (2016). *Best Practices in Summative Assessment*. College of Medicine, University of Florida, Orlando, Florida. *Adv Physiol Educ* 41: 110–119, 2017; doi:10.1152/advan.00116.2016.
- Lestari, Indah. (2015). PENGARUH WAKTU BELAJAR DAN MINAT BELAJAR TERHADAP HASIL BELAJAR MATEMATIKA. *Jurnal Formatif* 3(2): 115-125 ISSN: 2088-351X.
- Muhammad, et al. (2015). IMPACT OF MOTIVATION ON STUDENTS' ACADEMIC PERFORMANCE: A CASE STUDY OF UNIVERSITI SULTAN ZAINAL ABIDIN STUDENTS. *The American Journal of Innovative Research and Applied Sciences*. 2015; 1(6): 221-226.
- Mursidin, Rudi. (2017). PENGARUH PENEMPATAN JAM PELAJARAN MATEMATIKA TERHADAP PRESTASI BELAJAR MATEMATIKA SISWA KELAS V DI SEKOLAH DASAR NEGERI KONGPATI TAHUN PELAJARAN 2017/2018. Universitas Negeri Mataram.
- Nurkencana & Sunartana. (1992). *Evaluasi Hasil Belajar*. Surabaya: Usaha Nasional.
- Patil, R., Palve, S. B., Vell, K., & Boratne, A. V. (2016). Evaluation of multiple-choice questions by item analysis in a medical college at Pondicherry, India. *International Journal of Community Medicine and Public Health*, 3(6), 1612–1616.
- PerMendikbud. (2016). *Peraturan Menteri Pendidikan dan Kebudayaan nomor 23 tahun 2016*.
- Puspendik Kemendikbud. (2019). Average score of 2016/2017 school year grade. [online]. Available: <https://hasilun.puspendik.kemendikbud.go.id/>

Reece, I. & Walker, S. (1997). *Teaching, training and learning: A practical guide*. Sunderland: Business Education Publisher.

Rehman, Abdul. et al. (2018). *Item Analysis of Multiple Choice Questions*. Pakistan Oral and Dental Journal. Vol 38, No 2 (April – June 2018).

Rodriguez, Michael C. (2014) *The Role of Assessment in Student Performance on TIMSS*. University of Minnesota.

Tokan, Moses K. & Imakulata. (2019). *The effect of motivation and learning behaviour on student achievement*. South African Journal of Education, Volume 39, Number 1, February 2019.

Zimmaro. (2016). *Writing Good Multiple Choice Exam*. University of Texas – Austin. Last Updated December 1, 2016.