

P-ISSN
2548-6063

KURIKULA: JURNAL PENDIDIKAN
VOLUME, 7 No 2 TAHUN 2023
<https://ejournal.iaingawi.ac.id/index.php/kurikula/index>

E-ISSN
2746-4903

How is the Quality of Reading Test Items on English Subject?

Yulia Agustina^{1*}, Nila Hayati², M. Adib Nazri³

Universitas Hamzanwadi¹, Indonesia
Universitas Hamzanwadi², Indonesia
Universitas Hamzanwadi³, Indonesia

rahestin@gmail.com¹
hayatisyahidani@hamzanwadi.ac.id²
adibnazri88@gmail.com³

Article history

Submitted
21/11/2022

Accepted
16/2/2023

Published
21/03/2023

ABSTRACT : The objective of the study is to ascertain the quality of the reading test items on English subject for the eighth students at SMPN 12 Tasikmalaya. Those items were analyzed based on difficulty level items, discrimination power, and distractor effectivity. This study employed descriptive quantitative. It is implemented to describe the quality of reading test item of eighth students of SMPN 12 Tasikmalaya. Whereas, quantitative data used as the data or information in the form of numerical data and analyzed it by using Item and Test Analysis (ITEMAN) program 3.0 version. The researchers had already conducted the study on reading in previous study, so in collecting the data, the researcher used the documentation. It used to know the quality of reading test items that had been made by the first researcher. The results showed that: 1) based on difficulty level of items indicated 14 (23%) items in easy, 32 (53%) items in medium, and 14 (23%) items in hard. Dealing to discrimination power of items, it showed that 51 (85%) items in very good, 5 (8%) in good, 0 (0%) in enough, and 4 (7%) in poor, must be discarded. Meanwhile, the distractor effectivity on reading test items on English subject of eighth students of SMPN 12 Tasikmalaya were: 44 (73%) items in function, 12 (420%) items in enough, 2 (7%) items in unfunctional, and 0 (0%) items in poor. Overall, the quality of reading test which had been made by the first researcher in previous study was in good quality.

Key Words: Language learning, learning quality, item analysis, reading test, ITEMAN program

ABSTRAK: Tujuan dari penelitian ini adalah untuk mengetahui kualitas butir-butir tes membaca pada mata pelajaran bahasa Inggris untuk siswa kedelapan di SMPN 12 Tasikmalaya. Butir-butir tersebut dianalisis berdasarkan butir soal tingkat kesukaran, daya pembeda, dan keefektifan pengecoh. Penelitian ini menggunakan deskriptif kuantitatif. Hal ini dilaksanakan untuk mendeskripsikan kualitas butir tes membaca siswa VIII SMPN 12 Tasikmalaya. Sedangkan data kuantitatif digunakan sebagai data atau informasi berupa data numerik dan dianalisis dengan menggunakan program Item and Test Analysis (ITEMAN) versi 3.0. Peneliti sudah melakukan penelitian membaca pada penelitian sebelumnya, sehingga dalam pengumpulan datanya peneliti menggunakan dokumentasi. Digunakan untuk mengetahui kualitas butir tes membaca yang telah dibuat oleh peneliti pertama. Hasil penelitian menunjukkan bahwa: 1) berdasarkan tingkat kesukaran butir soal menunjukkan 14 (23%) soal mudah, 32 (53%) soal sedang, dan 14 (23%) soal sulit. Berkaitan dengan daya pembeda item, menunjukkan bahwa 51 (85%) item dalam sangat baik, 5 (8%) dalam baik, 0 (0%) cukup, dan 4 (7%) dalam buruk, harus dibuang. Sedangkan efektivitas pengecoh pada soal tes membaca pada mata pelajaran bahasa Inggris siswa VIII SMPN 12 Tasikmalaya adalah: 44 (73%) butir soal berfungsi, 12 (420%) butir soal cukup, 2 (7%) butir soal tidak berfungsi, dan 0 butir soal. (0%) item dalam miskin. Secara keseluruhan, kualitas tes membaca yang telah dilakukan oleh penelitian pertama pada penelitian sebelumnya dalam kualitas yang baik.

Kata Kunci: Pengajaran bahasa, kualitas pembelajaran, analisis butir soal, tes membaca, program ITEMAN

A. BACKGROUND

Today, people cannot deny that language has very important role in daily life communication allaround the world. Nowadays, they are using the language not only as a tool of communication but also, more than that, language is kinds of human needs. In every aspect of life, language plays an important role in the economy aspect such as trade and business. Dealing with technology, humans would not be able to keep hand in hand with current technology without mastering the language, especially English. However, in education nowadays people can easily continue their education with language as the main big modal. Accordance with the previous fact, that it is very necessary to master English and use it as thetool to transfer knowledge in teaching learning process.

Reading is one of skills to be taught in every school. By reading, the students can update their knowledge, increase vocabularies, add the information, knowing the author's writing style, ect. Reading is a process to decode the written symbol which involves a reader in understanding and attribute the information from a text to build meaning as a piece of communication between the reader and writer (Agustina, 2013). Further, she explains that there are five aspects of reading that lead to the indicators: (1) find main idea; (2) find explicit information; (3) find implicit information; (4) find word references; and (5) find meaning of certain word based on the context. In the same way, Catherine (2002: 11); Mullis (2017) state reading comprehension defines as the process of simultaneously extracting and constructing meaning through interaction and involvement with written language. In addition, reading is to recall information from the text, to answer questions,

and to apply the knowledge obtained from the text (Van den Broek, 2011). It can be concluded that reading is an activity to get the informations by extracting main idea, explicit information, implicit information, word references, and meaning of certain word based on the context from the written symbol.

Learning achievement test is used to know the students' students reading skill. In conducting a test, the teacher can use two kinds of tests namely standardized tests and tests made by the teacher himself. Test in this study is made by the researcher herself when doing a research. A good test will expose the real condition of students, and uneffective test will not expose the real condition of students. Test is used for measuring students' skill after following education activity in certain amount of time using methods and rules that have been determined. (Arikunto, 2013). Whereas, Sudijono (2008) states a test is method or procedure in measuring and evaluating in the field of education, in the form of assigning tasks both questions or instructions. In different opinion comes from Muchtar in Arikunto (2013) a test is conducted to find out learning achievement of a student or group of students. It can be concluded that the test is a procedure to know students' skill in learning achievement.

In general, there are two functions of test: as a tool to measure students' learning achievement and as a tool to measure the success of teaching programs. Thus, this study focus on analyzing item on reading test which intends to know the students' reading skill. This study was conducted to know the quality of reading test by computing the difficulty level items, discrimination power, and the effective distractor by using ITEMAN program. The fact showed that most of the teachers are rarely doing item analysis. Analysis items is very necessary to improve items' quality that deliver to students. Those items will be analyzed in order to know whether the items are effective or not. A good item could be a model to measure and reference in doing the next test. For negative items, it must be revised so that it will not give bad effect for the students. The result showed that test was in good criteria, so that the researcher will distribute this reading test to SMP or MTs as bank of items test. Further, it can be used as the reference for conducting the next other test.

Analyzing the level of difficulty items mean that studying the difficulty level each of the items so that it can be categorized into easy, medium and hard level. The level of difficulty items can be determined from the ability of the participants or students in answering the questions not from the ability of the teacher when he/ she composes the items (Bagiono, 2017). Generally speaking, an item can be categorized as an effective item if it is not either not too easy or too hard. The items which are correctly answered by the students or participants can be categorized into the effective items, the same things happened in contrary. For those two kinds of categorizes, it must be revised if it will be

use in any other time. The assumption is in order to get the good quality of items so it must be in balance between its difficulty items (Bagiono:2017).

Discrimination power items refers to ability of an item to differentiate the group of students or participants. One of the aims is to determine the items' ability in categorizing the students or participants with low ability and high ability. As stated by Sudijono (2011: 385) that discrimination power items is the ability of an item test to differentiate between testee with high level of ability and testee with low ability. It is very important to know the discrimination power items, because it will be used as the main guidance in constructing learning achievement test. Assumption which says that the ability of the testee vary one to another (Amaliata and Widayati, 2012). If an item has the positive level, so the item can be assumed that item has discrimination power. The students with have high ability are able to answer more correct questions and the same happened in contrary. It can be concluded that a good item is the item in medium criteria, either not too easy or too hard.

In multiple choice test, there are some optional answers. Among those optional answers there is only one correct answer. The distractor of an effective item is averagely chosen by students and in contrary, unfunctional item is averagely chosen by students. This statemen is hand in hand with Sudijono (2011) the distractor enrolls its function if it is chosen by 5% of students of the test, and most of the choice comes from students who does not understand the material yet. An effective distractor can be used on the next test. The effectivity of distractor can be seen from the wrong optional items since it is able yo distract the students who does not know the correct answer. The more of the participants chose the distractor item, so the distractor is succeeded. (Amaliata and Widayati, 2012).

This study concerns on using Item and Test Analysis (ITEMAN) program 3.0 version. It is a software that made certainly for analyzing test items in which an empirical data analysis with a classical approach model that is useful for determining the quality of items or a test. The result of item analysis as what explained above: difficulty level of items, discrimination power, and distractor effectivity. In addition to produce test item statistics, the program also produces test statistics which include reliability test, measurement errors or standard errors and score distribution. It has many functions in applying this program: 1) file data analysis; 2) scoring and analysing multiple choice items and Likert's scale; and 3) analysing a test consisting 10 scales (subtest) and giving the information the items validity.

The relevant studies come from Amalia and Widayanti (2012) and Wahyuningsih (2015) who conduct the research on test items analysis on Economic and Accounting subjects. Here the reason of why the researchers selects this study for ascertaining the quality of reading test on English subject for eighth grade of senior high school that made by the first researcher. It is intended to improve the quality of test items made by teachers or researchers which are able to make students understand on the reading context well. If

the test items are on low quality, it is recommended that teachers or researchers rewrite the questions. Therefore, it is very important to know the quality of the questions made by the teachers or researcher whether they are in the good category or not. From elaboration above, the researchers construct the research questions are as follows:

1. How is the quality of the reading test items on English subjects for the eighth students of SMPN 12 Tasikmalaya in terms of difficulty level items?
2. How is the quality of the reading test items on English subjects for the eighth students of SMPN 12 Tasikmalaya in terms of discrimination power?
3. How is the quality of the reading test items on English subjects for the eighth students of SMPN 12 Tasikmalaya in terms of distractor effectivity?

B. METHOD

This study employed descriptive quantitative. It's the integration between descriptive method and quantitative approach. Descriptive study is the method to attempt describing and interpreting an object based on the reality (Ary et al., 2010). While, quantitative gives a pressure in the systematic analysis, using statistical analysis toward quantitative data. Descriptive quantitative implemented in this study because the quantitative data analysis was accounted descriptively. This study intended to describe the quality of reading test of eighth students of SMPN 12 Tasikmalaya consisting 60 items test in multiple choice form. Whereas, quantitative data used as the data or information in the form of numerical data and analyze it by using Item and Test Analysis (ITEMAN) program. Reading test items construct from five reading indicators in understanding: meaning of certain words based on the context, references, main ideas, implicit information and explicit information. Each indicator consists of 12 items; such as follows:

Table 1 Blue Print of Reading Test

No	Aspects	Number of items	Total
1	Understanding the meaning of certain word based on the context	4, 6, 15, 17, 20, 28, 39, 47, 50, 52, 56, 57.	12
2	Understanding references.	5, 8, 13, 16, 27, 29, 30, 37, 42, 46, 54, 58.	12
3	Understanding main ideas.	1, 14, 19, 24, 31, 34, 41, 44, 48, 51, 55, 60.	12
4	Understanding implicit information.	7, 9, 10, 12, 18, 22, 25, 26, 35, 40, 45, 49.	12
5	Understanding explicit information.	2, 3, 11, 21, 23, 32, 33, 36, 38, 43, 53, 59.	12
Total		60	60

C. RESULT AND DISCUSSION

RESULT

Difficulty Level of Items

As for analysis results of students' reading test items, the researchers attached on the table 2 below:

Table 2 Blue Print of Reading Test

No. Items	Difficulty Level		No. Items	Difficulty Level	
	Criteria	Meaning		Criteria	Meaning
1	0.676	Medium	31	0.541	Medium
2	0.730	Hard	32	0.838	Hard
3	0.838		33	0.459	Medium
4	0.514	Medium	34	0.595	
5	0.784	Hard	35	0.243	Easy
6	0.676	Medium	36	0.892	Hard
7	0.676		37	0.216	
8	0.892		38	0.297	Easy
9	0.784	Hard	39	0.216	
10	0.784		40	0.568	Medium
11	0.189	Easy	41	0.784	Hard
12	0.459		42	0.649	Medium
13	0.405		43	0.081	Easy
14	0.351		44	0.432	Medium
15	0.351	Medium	45	0.595	
16	0.405		46	0.270	Easy
17	0.622		47	0.459	Medium
18	0.351		48	0.297	Easy
19	0.378		49	0.622	Medium
20	0.108	Easy	50	0.737	Hard
21	0.784	Hard	51	0.189	Easy
22	0.757	Hard	52	0.027	
23	0.757		53	0.811	Hard

24	0.649		54	0.730	
25	0.378		55	0.270	Easy
26	0.351	Medium	56	0.405	
27	0.405		57	0.595	Medium
28	0.216	Easy	58	0.757	Hard
29	0.514	Medium	59	0.162	Easy
30	0.649		60	0.568	Medium

In difficulty level of items, the results showed that 14 (23%) items in easy, 32 (53%) items in medium, and 14 (23%) items in hard. As stated by Sudjana (2011) generally, criteria of difficulty level items should be in the medium category, 0.31-0.70. In this interval, the information of students' skill will be maximally obtained because the items are in the category that are not too easy and not too difficult. This study was in line with previous research conducted by (Amalia and Widayati, 2012; Wahyuningsih, 2015) said that good items are items that have levels in medium difficulty.

From those explanations above, it can be inferred that difficulty level items on reading test on English subject of eighth students of SMPN 12 Tasikmalaya were in the middle of criteria, medium. For medium difficulty level, the items can be used in another time and used as the bank of reading test items.

Discrimination Power

Discrimination power items indicate the difference of high and low students' groups. Its criteria stated in proportion. The higher distinguish items criteria, the abler to distinguish students from the high and low groups. As the result of ITEMAN analysis in distinguishing item, the researchers used following criteria:

Table 3 Criteria Guidelines of Discrimination Power

Criteria	Meaning
$D \geq 0.40$	Very Good
$0.30 \leq D \leq 0.39$	Good
$0.20 \leq D \leq 0.29$	Enough. need revision
$D < 0.19$	Poor. must be discarded

(Source: Ebel & Frisbie. 1979 quoted by Nurung. 2008: 371)

Based on ITEMAN analysis, the followings are the discrimination power of each item on students' reading test:

Table 4 Discrimination Power Items

No. Items	Discrimination Power		No. Items	Discrimination Power	
	Criteria	Meaning		Criteria	Meaning
1	0.881		31	0.657	Very Good
2	0.669		32	0.340	Good
3	0.931		33	0.708	
4	0.762		34	0.613	
5	0.487		35	0.418	
6	0.700	Very Good	36	0.664	
7	0.592		37	0.651	Very Good
8	0.690		38	0.773	
9	0.603		39	0.684	
10	0.611		40	0.708	
11	0.364	Good	41	0.685	
12	0.476	Very Good	42	0.181	Poor
13	0.360	Good	43	0.833	
14	0.699	Very Good	44	0.454	
15	0.092	Poor	45	0.482	
16	0.515		46	0.778	
17	0.430		47	0.628	Very Good
18	0.868		48	0.647	
19	0.454		49	0.500	
20	0.364		50	0.524	
21	0.850		51	0.569	
22	0.639	Very Good	52	0.048	Poor
23	0.516		53	0.530	
24	0.664		54	0.530	Very Good
25	0.836		55	0.442	Good
26	0.496		56	0.392	Good
27	0.459		57	0.482	Very Good
28	0.132	Poor	58	0.323	Good
29	0.677	Very Good	59	0.733	Very Good
30	0.573		60	0.437	

Dealing to discrimination power of items in this study were: 51 (85%) terms in very good, 5 (8%) in good, 0 (0%) in enough, and 4 (7%) in poor, must be discarded. A good discrimination power was able to distinguish students who have the ability in answering the test correctly and who do not. While, a bad discrimination power was not able to distinguish students who have the ability in answering the test correctly and who do not. It is different from the research conducted by (Wahyuningsih, 2015), her result showed that test items of final examination odd semester on Economic subject of X class of SMAN 1 Mlati in the academic year 2013/2014 was in a matter of lack of quality. It means that the items yet can distinguish between students who have high ability with students who have low ability. 24 items (48%) from 50 items, are classified as good items. In short, those items cannot perform its function.

Regarding to the result of discrimination power in this study, it can be concluded that mean score of discrimination power items on reading test on English subject of eighth students of SMPN 12 Tasikmalaya in very good criteria. It is evidenced by biserial mean, 0.563. This matter means that items were able to distinguish students who have high and low ability. The results of this study are strengthened by the theory according (Arikunto, 2021) the discrimination power is items' ability to distinguish between students who have high and low ability. Moreover, a good item was answered correctly by students who have high ability only (Wahyuningsih, 2015). The discrimination power can be used to increase the items quality based on empirical data from items analysis. Its criteria indicate whether the items: good, need revise, or rejected.

Distractor Effectivity

Dissemination of distractor could be a basic modal in study the items so that we are able to know whether the items are effective or not. An effective distractor if it is chosen by 5% of students of the test, and most of the choice comes from students who does not understand the material yet. According to the result of the analysis, it states that effective distractor if it is chosen, at least, by 5% from the total number of the participants of the test. The distribution of the distractor for each item can be seen in the following table:

Table 5 Key Answers on Each Item

No Item	Key Answer	Distractor			Explanation
		Good ($\geq 5\%$)	Not Good ($< 5\%$)	No Voter (0%)	
1	B	A.C.D	-	-	
2	A	B.C.D	-	-	Functionate
3	C	A.D	B	-	Enough
4	D	A.B.C	-	-	Functionate
5	C	A.B.D	-	-	
6	B	A.B	D	-	Enough
7	D	A.B.C	-	-	Functionate
8	A	C	B.D	-	Unfunctional
9	A	C.D	B	-	Enough
10	B	A.C.D	-	-	
11	B	A.C.D	-	-	
12	C	A.B.D	-	-	
13	D	A.B.C	-	-	
14	C	A.B.D	-	-	
15	Check C to D	A.B.D	-	-	Functionate
16	D	A.B.C	-	-	
17	A	B.C.D	-	-	
18	C	A.B.D	-	-	
19	D	A.B.C	-	-	
20	B	A.C.D	-	-	
21	A	C.D	B	-	
22	A	C.D	B	-	Enough
23	C	A.B.D	-	-	Functionate
24	C	A.B	D	-	Enough
25	B	A.C.D	-	-	
26	C	A.B.D	-	-	
27	A	B.C.D	-	-	Functionate
28	B	A.C.D	-	-	
29	D	A.B.C	-	-	
30	C	A.D	-	B	Enough
31	A	B.C.D	-	-	Functionate
32	B	A	C	D	Unfunctional
33	D	A.C	B	-	Enough
34	A	B.C.D	-	-	
35	A	B.C.D	-	-	Functionate
36	B	A	D	C	Unfunctional
37	B	A.C.D	-	-	
38	B	A.C.D	-	-	
39	A	B.C.D	-	-	Functionate
40	B	A.C.D	-	-	
41	A	C.D	B	-	Enough
42	A	B.C.D	-	-	Functionate

43	C	A.B.D	-	-	
44	C	A.B.D	-	-	
45	C	A.B.D	-	-	
46	D	A.B.C	-	-	
47	C	A.B.D	-	-	
48	B	A.C.D	-	-	
49	D	A.D	B	-	
50	C	A.D	B	-	Enough
51	B	A.C.D	-	-	Functionate
52	Check D to C	A.B.C	-	-	Unfunctional
53	D	A	C	1	
54	D	A.C	B	-	Enough
55	A	B.C.D	-	-	
56	A	B.C.D	-	-	
57	A	B.C.D	-	-	
58	C	A.B.D	-	-	Functionate
59	B	A.C.D	-	-	
60	D	A.B.C	-	-	

On the distractor table above, almost items are functioned properly in each item. For 15 and 52 items asked to be checked up the key answers for alternative goal. There are two possibilities in checking up these answers: 1) If it is turn out to incorrect item then do back the item analysis; and 2) if it is correct there may be errors on the students' ability. On the following was the percentage of distractor effectivity:

Table 6 Percentage of Distractor Effectivity on Each Items

Criteria	Total	Percentage
Functionate (all distractors $\geq 5\%$)	44	73%
Enough (2 distractors $\geq 5\%$)	12	20%
Unfunction (1 distractor $\geq 5\%$)	4	7%
Poor (no distractors $\geq 5\%$)	0	0%

The table above explains 44 (73%) in functionate, 12 (20%) in enough, 4 (7%) in unfunctional, and 0 (0%) in poor. It can be inferred that, all items on this reading test were properly work which mean the reading test items can be used for next other test.

DISCUSSION

There are still many English teachers who have not evaluated their English teaching. This is evidenced by the existing studies that several scholars examined on evaluating their instructions in subjects other than English. This study can provide a reference for English teachers to determine the quality of the reading test items made. Therefore, when they've already known about the quality of the reading test items, the English teacher will always provide good quality reading test items.

Based on the research result, in difficulty level of items, the result showed that 14 (23%) items in easy, 32 (53%) items in medium, and 14 (23%) items in hard. As stated by Sudjana, Nana (2001:137) generally, criteria of difficulty level items should be in the medium category, 0.31-0.70. In this interval, the information of students' skill will be maximally obtained because the items are in the category that are not too easy and not too difficult. This study was in line with previous research conducted by Amalia and Widayati (2012); Wahyuningsih, (2015), said that good items are items that have levels in medium difficulty.

From those explanation above, it can be inferred that difficulty level items on reading test on English subject of eighth students of SMPN 12 Tasikmalaya were in the middle of criteria, medium. For medium difficulty level, the items can be used in another time and used as the bank of reading test items. While, for hard level items can be discarded or reexamined, tracked, and traced then it can be known the causative factors of difficult items answered by students. Last, for the easy items was the same as the treatment in the hard criteria items.

Dealing to discrimination power of items in this study were: 51 (85%) items in very good, 5 (8%) in good, 0 (0%) in enough, and 4 (7%) in poor, must be discarded. A good discrimination power was able to distinguish students who have the ability in answering the test correctly and who do not. While, a bad discrimination power was not able to distinguish students who have the ability in answering the test correctly and who do not. It is different from the research conducted by Wahyuningsih, (2015), her result showed that test items of final examination odd semester on Economic subject of X class of SMAN 1 Mlati in the academic year 2013/2014 was in a matter of lack of quality. It means that the items yet can distinguish between students who have high ability with students who have low ability. 24 items (48%) from 50 items, are classified as good items. In short, those items cannot perform its function.

Regarding to the result of discrimination power in this study, it can be concluded that mean score of discrimination power items on reading test on English subject of eighth students of SMPN 12 Tasikmalaya in very good criteria. It is evidenced by biserial mean, 0.563. This matter means that items were able to

distinguish students who have high and low ability. The results of this study are strengthened by the theory according Arikunto (2013) the discrimination power is items' ability to distinguish between students who have high and low ability. Moreover, good items were answered correctly by students who have high ability only (Wahyuningsih, 2015). The discrimination power can be used to increase the items quality based on empirical data from items analysis. Its criteria indicates whether the items: good, need revise, or rejected.

The findings showed that the distractor on reading test items on English subject of eighth students of SMPN 12 Tasikmalaya, they were: 44 (73%) items in functionate, 12 (420%) items in enough, 2 (7%) items in functionate, and 0 (0%) items in poor. As stated by Daryanto (2007) the effective distractor if it is chosen, at least, by 5% from the total number of the participants of the test. Otherwise, the research conducted by Wahyuningsih (2015) explains that the distractor effectivity of test items on Economic subject of X class of SMAN 1 Mlati in the academic year 2013/2014 were in lack quality because there are no items that qualify as either very good or good. The distractor which was not chosen at all by the teste means that distractors were so bad, too conspicuously misleading. Conversely, a distractor can be said functionate well if the distractor has great appeal for the students who have low competency understanding (Arikunto, 2013). The number of items with poor quality indicate that the distractor cannot function properly yet, it too conspicuous, misleading, and tend to be heterogeneous (Wahyuningsih, 2015). In sum, the distractor has not great appeal for the students who have low understanding or lack mastering the material on English subject.

D. CONCLUSION

It can be concluded that reading test items that was done by researcher in previous study, there were in good quality. It was proved by the students' answer on the reading test items and the findings above. As a good result, this reading test can be used as a reference by English teachers in conducting reading test for eighth students of SMP/MTs level. Moreover, it can be used as reading test items bank as exercise for eighth students in improving their ability in comprehending the text. Then, when creating a test, it must be considered of adapting of difficult level of items which does not match the projected level of early difficulties, eliminating the items that have too easy items, and adapting the items that have low discrimination of power. In addition, the results of this study have an indirect impact on students' writing skills, where the good quality of the reading test items will increase the quantity of students' English vocabulary definitely will flow into their writing. As stated by Amna and Rizki (2018) there is a significant correlation between students' reading habit and their writing ability.

E. REFERENCES

- Amna & Zakaria. (2018). Kontribusi Kebiasaan Membaca dan Motivasi Belajar terhadap Kemampuan Menulis Teks Eksposisi Siswa Kelas X Sma Negeri 2 Padang. *Menara Ilmu*, 7(6).
<https://jurnal.umsb.ac.id/index.php/menarailmu/article/viewFile/833/744>
- Agustina, Yulia. (2013). *The Effectiveness of Semantic Mapping to Teach Reading Viewed from Students' Intelligence*. Unpublished Thesis. Surakarta: Sebelas Maret University.
- Amaliata, Ata Nayla., & Widayati, Ani. (2012). Analisis Butir Soal Tes Kendali Mutu Kelas Xii Sma Mata Pelajaran Ekonomi Akuntansi Di Kota Yogyakarta Tahun 2012. *Jurnal Pendidikan Akuntansi Indonesia*, 10(1).
<https://journal.uny.ac.id/index.php/jpakun/article/view/919>.
- Arikunto, Suharsimi. (2012). *Prosedur Penelitian Suatu Pendekatan Praktek*. Jakarta: RinekaCipta.
- Arikunto, Suharsimi. (2013). *Dasar-dasar Evaluasi Pendidikan*. Jakarta: PT Bumi Aksara.
- Ary, et al. (2010). *Introduction to Research in Education*. Wadsworth: Cengage Learning.
- Bagiono. (2017). The Analysis of Difficulty Level and Discrimination Power of Test Items of Radiography Level 1 Examination, 16(1).
<http://jurnal.batan.go.id/index.php/widyanuklida>
- Catherine, Snow. (2002). *Reading for Understanding*: published by RAND: Pittsburg.
- Daryanto. (2007). *Evaluasi Pendidikan*. Jakarta: PT Rineka Cipta.
- Gay (1992). *Education Research. Competencies for Analysis and Application*. New York: Macmilan Publishing Company.
- Mullis, I. V. S., Martin, M. O., Foy, P., & Hooper, M. (2017). ePIRLS 2016 International Results in Online Informational Reading. Retrieved from Boston College, TIMSS & PIRLS International Study Center website:
<http://timssandpirls.bc.edu/pirls2016/international-results/>
- Sudjana, Nana. (2011). *Penilaian Hasil dan Proses Belajar Mengajar*. Bandung: Rosda Karya.
- Sudijono. Anas. (2008). *Pengantar Evaluasi Pendidikan*. Jakarta : PT Raja Grafindo Persada.
- Sugiono. (2011). *Metode Penelitian Kuantitatif dan Kualitatif dan R&D*. Bandung: Alfabeta.
- Van den Broak, P., Kendeou, P., Lousberg, S., & Visser, G. (2011). Preparing for Reading Comprehension: Fostering Text Comprehension Skills in Preschool

and Early Elementary School Children. *International Electronic Journal of Elementary Education*, 4 (1), 259-268.2011.

Wahyuningsih, Erwin Tri. (2015). Analisis Butir Soal Tes Objektif Buatan Guru Ulangan Semester Ganjil Mata Pelajaran Ekonomi Kelas X Di Sma Negeri 1 Mlati Tahun Ajaran 2013/2014. Unpublish Thesis. Yogyakarta: UNY.