Test items Analysis Constructed by EFL Teachers of Private Senior High School in Kupang, Indonesia

Marselus Yumelking

Lecturer of Nusa Nipa University, Indonesia

Abstract— *This research deals with the analysis of test items constructed by English as a foreign language (EFL) teachers of private senior high schools in Kupang, Indonesia. It aimed to know discrimination power, difficulty level and distracter of test items constructed by*
the EFL teachers. The design used in this research study is descriptive qualitative. Data resource is semester test items of grade XII students of language program of Sint Carolus Kupang private senior high school in the academic year 2017/2018. The result of this study showed that discrimination power has 52,5 % test items that could not be used to test the students, 25% test items that must be revised and 22,4% test items that are really good to be used to test the students. Difficulty index involved level of easy test items as much as 35 %, 37 % test items as difficult level, and 27,5% test items as moderate level. Distracter involved 11,92 % very poor distracters, 49,66% poor distracters, 13,90 % fair distracter, 13,90% good distracter and 10,59% very good distracter.

Keywords— *Test items, Analysis, EFL Teachers.*

I. INTRODUCTION

Instruction is the way to reach the goals of education. Instruction has a role on the process of teaching and learning that explicitly results on the reach of the cognitive, affective and psychomotor goals of education. As result of this, the implementation of instruction is designed by some components to be implemented. Djiwandono (2014:2) said that instructional design has some components such goal, activity and evaluation. These three components encompass the main basic of the implementation of instruction that relate to each other. The previous components decide the existence of the subsequent components.

On line with the goals of instruction, a framework of activities is designed to reach its goals. To know that those goals reached, thus, process of instruction required to be evaluated. Djiwandono (2014: 10) said that generally, evaluation in the implementation of instruction is defined as process of collecting information about the implementation of instruction as the main basic of making judgment. Judgment to know someone passes or certifies. The process of collecting information can be reached through a process of measurement.

Measurement includes language tests that used to determine someone passes or certifies. As a result of this, a language test must be considered to be a good achieved by the test maker. To reach this objective, a test maker needs

to know how to construct a good test. A good test must have a good test items. A good test items must consider some criteria such qualitative analysis and quantitative analysis. As stated by Puspendik (2013:2) that there two kinds of test items analysis; qualitative analysis such as analysis of material, construction, language, while quantitative analysis such as discrimination power, level of difficulty, and distracter. As a result of this, the writer would like to study the quantitative analysis of the test items of a private senior high school by doing an analysis study under the title "Test Items Analysis Constructed by EFL Teachers of Private Senior High School in Kupang". The aims of this research are to know discrimination power, difficulty levels and distracter of the test items constructed by the EFL teachers.

According to Arikunto (2013:207) item analysis aims to identify dropped, corrected and good test items. With item analysis, can be reached information about dropped item analysis and a lesson to correct the test items. To know a good test items, it is required to know three kinds of analysis related to the item analysis such level of difficulty, discrimination power, and distracter. Daryanto (2010:179) argued that a good test items is an item that is not too easy and too difficult. A too easy item does not stimulate students to pay more effort to finish it. Otherwise, a too difficult item will cause the students become yield and are not motivated to try again because it

is out of their reaches to finish it. According to Puspendik (2013: 10) the difficulty index calculated to know the index of difficulty of both groups in the level of difficulty is 0 to 0,30, moderate is 0,30 to 0,70 while easy for more than 0,70.

Arikunto (2013:211) said that discrimination power is the ability of an test items to differentiate upper level and lower level of the students. The number that shows discrimination power is called index of difficulty or index of discrimination about 0,00 to 1,00. According to Puspendik (2013:8) the result of discrimination power of test items analysis of 0 and smaller than 0 means that, it's test items is dropped, while more than 0 to 0,25 is corrected, and finally more than 0,25 means that the test items is good achieved .

Sudijono (2011:409) said that in objective test of multiple choice items, there are some possible answers or what it called options or alternative answers. Those options involve four to five. From those options or possible answers, one of them is correct answer while the others are incorrect answers. Those incorrect answers are called distracters. In this test items analysis, there are five options, thus, distracter consists of four. There is one correct answer and four distracters of each test items. To know the result of computation of distracter , the level of distracters are standardized as very good, good, fair, poor and very poor. According to Nitko, 1996 (in Lawa 2015: 20-21) distracter range of 0,76 to 1.25 is very good, 0.51 to 0.75 and 1.26 to 1.50 are good, 0.26 to 0.50 and 1.51 to 1.75 are fair, 0 to 0.25 and 1.76 to 2.00 are poor, more than 2.00 is very poor.

II. METHOD

The design used in this research study is descriptive qualitative. This research study was carried out in Sint Carolus Kupang private senior high schools which is located at Adisucipto street Kupang tengah subdistrict, Kupang mayor. This school is located near the researcher's place so it was easy to be reached. The researcher has conducted this research for 1week on 19th May 2017 to 26th May 2017. The data resource of this research study is semester test items of XII grade students of language program of Sint Carolus Kupang private senior high school in the academic year 2017/2018.

This research study used some procedures of research as the following; first, the researcher scored the students' answers. Second, those scores then divided into two groups such high level and low level of the students. Third, the researcher counted the discrimination power to know whether each test items is dropped, corrected or good. Then, levels of difficulty were counted to know whether

the test items had an easy, moderate or difficult level. Finally, the researcher counted distracter of each test items to know very poor, poor, fair, good, and very good of distracter.

There are some steps of analyzing the data that used in this research study. It is described as follow:

1. Counting the level of difficulty using the formula:
   Difficulty level or difficulty index $\text{DI} = \dfrac{C}{N}$

DI = Difficulty Index

C = The number of students who have correct answer

N = Number of analyzed students

Puspendik 2013: 9)

The difficulty index calculated to know the index of difficulty of both groups in the level of easy, moderate or difficult. It standardized as the following:

0< DL ≤ 0,3   : Difficult

0,30< DL ≤ 0,70 : Moderate

> 0,70   : Easy

(Puspendik 2013: 10)

2. Counting the discrimination power using the formula:
   $$\text{DP} = \dfrac{2\,(HC\text{-}LC)}{N}$$

DP   = Discrimination Power

HC   = The number of students of high group who have correct answer

LC   = The number of students of low group who have correct answer

N   = Number of analyzed students

(Puspendik 2013: 7)

To know the test items is good, corrected or dropped; discrimination power has the certain standard to decide it. We can see in the following formula:

DP ≤ 0   :   Dropped

0 < DP ≤ 0,25 :   Corrected

DP> 0,25   :   Good

3. Counting the distracter using the formula:
   $$\text{DtI} = \dfrac{nDt}{(N\text{-}nC)/(Alt\text{-}1)}$$

Dtl   : Distracter Index

nDt   :Total students who choose the distracter

N   :Total test takers

nC   : Total correct answer

Alt   : Total number option

To know the result of computation of distracter above, it presented the level of distracter as the following table:

Table 2.1 Distracter range and level

Distracter Range	Distracter Level
0,76 to 1.25	Very good
0.51 to 0.75 and 1.26 to 1.50	Good
0.26 to 0.50and 1.51 to 1.75	Fair
0 to 0.25 and 1.76 to 2.00	Poor
More than 2.00	Very poor

Nitko, 1996: 310 (in Lawa 2015: 20-21)

III. RESULTS

In this section, the writer would like to analyze each test items constructed by EFL teachers of Sint. Carolus Kupang private senior high schools in the academic year 2017/2018. There are 40 test items analysis which were taken from first semester test of grade XII of language program. This analysis considered to know to what extend the test items chosen discriminates the higher group and lower group, the level of difficulty of both groups and distracter of each items are working for higher group and lower group. Three points of items analysis above can be described as the following:

3.1 Discrimination power

The discrimination power of test items constructed by EFL teachers of Sint. Carolus Kupang private senior high schools in the academic year 2017/2018 is presented in the table 3.1.

Table 3.1 represents the discrimination power of items test which involved 21 of 40 test items or 52,5% that must be dropped or cannot be used to test the students. It spread out in the test items of number. 1, 2, 6, 9, 10, 11, 12, 15, 18, 22, 23, 25, 26, 27, 29, 31, 34, 35, 36, 40. In addition, there are 10 test items that must be corrected or revised. It involves 25% that spread out in the test items number 3,7,8,14,19,21,32,37,38,39. Thus, only 9 of 40 test items that is really good to be used to test the students. It includes 22,4% that spread out in the test items number 4,5,13,16,17,20,24,28,30.

Table 3.1. Discrimination Power

	DP	DECISION
1	0,00	Dropped
2	-0, 17	Dropped
3	0, 17	Corrected
4	0,33	Good
5	0,67	Very good
6	-0,17	Dropped
7	0,17	Corrected
8	0,17	Corrected

9	0,000	Dropped
10	-0,67	Dropped
11	0,17	Dropped
12	0,17	Dropped
13	0,50	Very good
14	0,17	Corrected
15	0,000	Dropped
16	0,33	Good
17	0,50	Very good
18	-0,17	Dropped
19	0,17	Corrected
20	0,33	Good
21	0,17	Corrected
22	-0,17	Dropped
23	0,000	Dropped
24	0,33	Good
25	0,00	Dropped
26	-0,17	Dropped
27	0,00	Dropped
28	0,33	Good
29	0,00	Dropped
30	0,50	Very good
31	-0,17	Dropped
32	0,17	Corrected
33	-0,17	Dropped
34	-0,17	Dropped
35	0,00	Dropped
36	0,00	Dropped
37	0.17	Corrected
38	0,17	Corrected
39	0,17	Corrected
40	0,00	Dropped

3.2 Difficulty Index

The index of difficulty is represented in the following table:

Table 3.2 Difficulty index

NO	DI	DECISION
1	1,00	Very easy
2	0,42	Moderate
3	0,92	Easy
4	0,08	Difficult
5	0,50	Moderate
6	0, 42	Moderate
7	0,08	Difficult
8	0,08	Difficult
9	1,00	Very Easy
10	0,33	Moderate
11	0,92	Easy
12	0,25	Difficult

13	0,58	Moderate
14	0,75	Easy
15	0,00	Difficult
16	0,83	Easy
17	0,75	Easy
18	0,92	Easy
19	0,92	Easy
20	0,83	Easy
21	0,92	Easy
22	0,25	Difficult
23	0,83	Easy
24	0,33	Moderate
25	0,00	Difficult
26	0,42	Moderate
27	1,00	Easy
28	0,50	Moderate
29	0.00	Difficult
30	0,75	Easy
31	0,25	Difficult
32	0,25	Difficult
33	0,33	Moderate
34	0,42	Moderate
35	0,00	Difficult
36	0.00	Difficult
37	0.08	Difficult
38	0,58	Moderate
39	0,08	Difficult
40	0,00	Difficult

Table 3.2 describes the difficulty index of test items constructed by EFL teachers of Sint. Carolus Kupang private senior high scools with level of easy items as much as 14 of 40 test items. It involves 35 % that spread out in the number of 1,3,9, 11,14,16, 17,18,19,20,21,23,27,30. There are 15 test items involved difficult level or 37, % of 40 test items. It spread out in the test items of number 4,7,8,12, 15,22,25,29,31,32,35,36,37,39,40. Thus, there are only 11items test that involved as moderate level or 27,5% of 40 test items. It spread out in the number of 2,5,6,10,13,24,26,28,33,34,38.

3.3 Distracter

The result of distracter analysis of 40 test items constructed by EFL teachers of Sint. Carolus Kupang private senior high schools in the academic year 2017/2018 is in the following analysis:

Item number 1

$$DtI = \frac{nDt}{(N-nC)/(Alt-1)}$$

Option B = $\frac{0}{12-11/5-1}$

$$\frac{0}{1/4} = \frac{0}{0,25} = 0$$

Option C = $\frac{0}{12-11/5-1} = 0$

$$\frac{0}{1/4} = \frac{0}{0,25} = 0$$

Option D = $\frac{0}{12-11/5-1} = 0$

$$= \frac{0}{1/4} = \frac{0}{0,25} = 0$$

Option E = $\frac{0}{12-11/5-1} = 0$

$$= \frac{0}{1/4} = \frac{0}{0,25} = 0$$

Test items number 2 until 40 are analyzed as the example of the analysis above. It presented on the following table:

Table 3.3. Distracter

N	A	B	C	D	E
1	*	0	0	0	0
2	*	1.33	0	0	0
3	4	*	0	0	0
4	1	1	1	1	*
5	2	0.67	0.67	*	0.67
6	2.29	0	*	0.57	1.14
7	0.36	1.09	0.73	*	2.18
8	0	0	2.91	1.09	*
9	0	0	0	0	*
10	0	*	3.5	0.5	0
11	*	1	0	0	0
12	*	0.44	0.44	0	3.11
13	0.8	0.8	0.8	1.6	*
14	*	0	0	2.67	1.33
15	1.33`	2	0.33	*	0.33
16	2	*	0	2	0
17	1.33	1.33	1.33	*	0
18	0``	*	0	4	0
19	*	4	0	0	0
20	2	*	0	0	2
21	4	0	0	0	*
22	*	2.22	1.78	0	0
23	*	0	0	0	0
24	2.5	*	1	0.5	0
25	*	0	0	3.67	0.33
26	*	0	1.14	0.57	2.29

27	*	0	0	0	0
28	0	*	0.67	0.67	3.33
29	0.33	*	1	2	0
30	0	*	0.75	1.33	2.67
31	2.67	*	0.44	0.44	0.44
32	0.44	1.78	0	*	1.78
33	*	1.6	0.44	0.89	0
34	0	3.43	0.57	*	0
35	1.33	2	1	0	*
36	0	3.33	0.33	*	0.33
37	0.36	0.73	2.18	*	0.73
38	1.6	*	0	0	2.4
39	0	2.55	1.45	0	*
40	0.33	3.33	0.33	0	*

The result of distracter analysis above is interpreted by using standardize level of distracter that is presented in the following table

Table 3.3.1. Distracter level

Level	A	B	C	D	E	∑	%
Very Poor	4	2	3	3	6	18	11,92
Poor	13	11	17	18	16	75	49,66
Fair	6	1	6	4	4	21	13,90
Good	3	4	7	4	3	21	13,90
Very Good	2	4	6	3	1	16	10,59
	28	22	39	32	30	151	100

Table 3.3.1 shows that there are 18 of 151 or 11, 92 % very poor distracters, 75 of 151 or 49,66% poor distracters, 21 of 151 or 13,90 % fair distracter, 21 of 151 or 13,90% good distracter and 16 of 151 or 10,59% very good distracter.

IV. DISCUSSION

In this section, results of data findings are described and discussed. All data finding analysis through some standardize formulations are explained. Thus, it involves discussion of discrimination power, difficulty index and distracter of items constructed by EFL teacher of SMA Sint. Carolus Kupang in the academic year 2017/2018. The results of data analysis are explained in the following sections:

4.1 Discrimination Power

Discrimination power distinguishes student's academic ability of high group and low group in answering the items test. It is computed by the answers of both groups. It is used to decide the test items are well used, dropped or must be revised. The result of the discrimination power of

test items constructed by the EFL teachers are represented by the following chart

Chart 4.1 Discrimination power
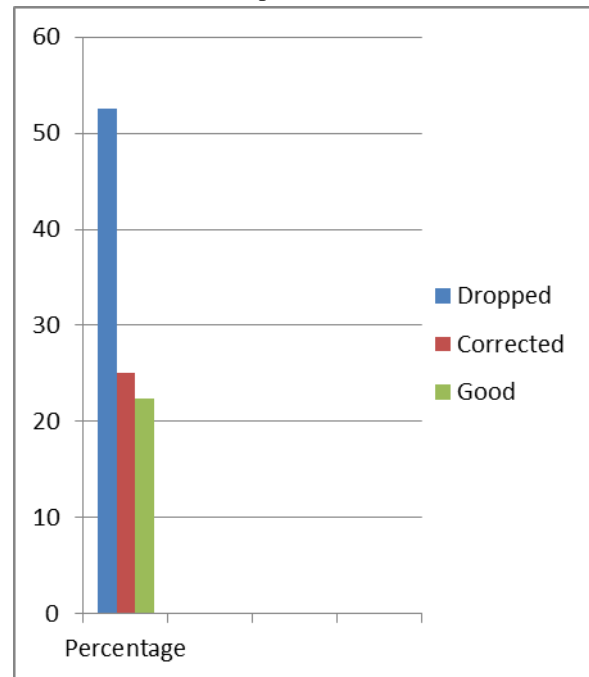
The chart shows that, there are 21 of 40 test items constructed by the EFL teachers that must be dropped. It involves 52,5 % test items that could not be used to test the students. In addition, there are 10 test items or that must be corrected or revised. It involves 25%. Thus, there are only 9 of 40 test items that are really good to be used to test the students. It includes 22, 4%.

4.2 Difficulty Index

Difficulty index refers to the proportion of correct answers of both groups; high group and low group. It is used to measure whether the each test items is in level of difficult, moderate or easy. The level of difficulty index of test items constructed by the EFL teacher is represented by the following chart.

Chart 4.2 Difficulty Index
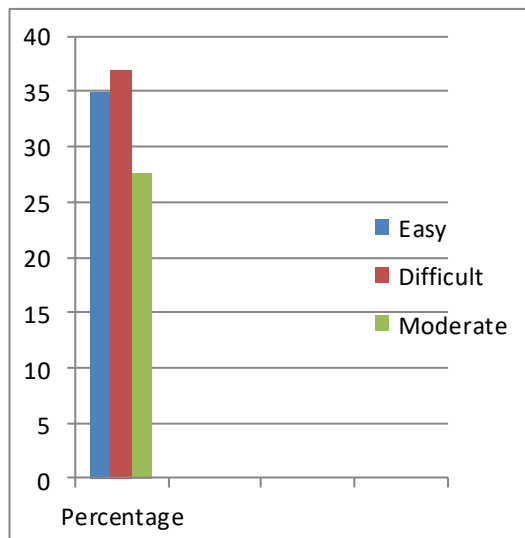
The difficulty index chart above shows that level of easy test items constructed by the EFL teachers is 14 of 40 test items. It involves 35 %. In addition, there are 15 test items involves difficult level or 37 % of 40 test items. Thus, there are only 11items test that involved as moderate level or 27, 5% of 40 test items.

4.3 Distracter

Distracter deals with the incorrect options that spread out in an test items. Distracter can be clearly seen in difficulty level and discrimination level. The weak distracters can cause test questions have poor discrimination or an undesirable level of difficulty. Thus it involves very poor, poor, fair, good and very good distracters. Distracters of test items constructed by the EFL teachers are described in the following chart
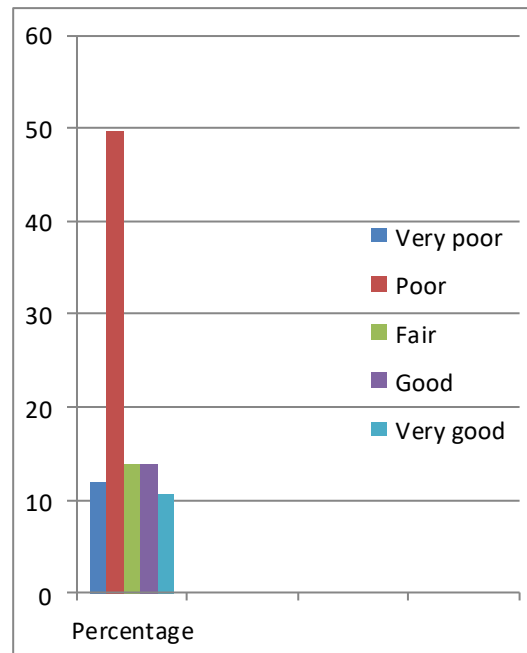
The chart above shows that distracters of test items constructed by the EFL teachers are 18 of 151 or 11, 92 % very poor distracters, 75 of 151 or 49,66% poor distracters, 21 of 151 or 13,90 fair distracter, 21 of 151 or 13,90% good distracter and 16 of 151 or 10,59% very good distracter.

V. CONCLUSION

This study was limited only on the analysis of the test items constructed by the EFL teacher of Sint.Carolus private senior high schools Kupang, Indonesia in the academic year 2017/2018. Considering the result of item analysis above, it can be concluded that there were only 9 or 22,4% of 40 test items constructed by English foreign language teacher of Sint Carolus Kupang private senior high schools in the academic year 2017/2018 that were really good to be used to test the students. Therefore as a test maker, we should pay attention to the good requirements of discrimination power, difficulty index and distracter. The advantage of this study is the researcher has already used the national standard requirement of discrimination power, difficulty index and distracter which was composed by Indonesian center of education (PUSPENDIK). As a result of this, this reference is a possible application for the test maker to analyze the test items.

REFERENCES

[1] Arikunto, S. 2013. Dasar – Dasar Evaluasi Pendidikan. Ed.2. Jakarta. PT BumiAksara.

[2] Daryanto, H. 2010. Evaluasi Pendidikan. Ed.2 Jakarta. Rineka Cipta.

[3] Lawa.S.T.M.N. 2015.The quality of English Items Constructed by EFL Teachers at Private Senior high School in Kota Kupang in the Academic Year 2014/2015. Unpublished thesis. Kupang. Nusa Cendana University

[4] Puspendik 2013. Analis soal (Manual).Pusat Penilaian Pendidikan, Badan Penelitian dan Pengembangan, Departemen Pendidikan Nasional

[5] Sudijono, A. 2011. Pengantar Evaluasi Pendidikan. Jakarta. PT Raja GrafindoPersada