CHAPTER IV

RESEARCH FINDING & DISSCUSSION

This chapter is to explain about the result of the research. The researcher analyzed English language testing that had been constructed by an English teacher of second grade in Sekolah Indonesia Kuala Lumpur, Malaysia based on characteristic of a good test, validity, and item analysis which include index of difficulty, index of discrimination and distracters.

Furthermore, all data that helped the researcher to analyze were conducted from documents such as, the English teacher interview result, the item test, and the answer keys.

A. Description of the English Final Test

The English final test had been conducted by the English teacher group of Sekolah Indonesia Kuala Lumpur. The final tested, second, and third grade of this school. Here, the researcher focused on second grade. The number of the students at second grade was thirty six students. The number of students is from both Science class and Social class. The researcher took both of class because the limited number of students in each class. The test was

50

objective test that consisted of forty multiple choice questions and had five multiple choice objections.

The content of the test is reading test. The English test was held on June 4th 2013, they started the test from 08.00 to 09.30 am Malaysia time.[1]

B. The result of the Final Test

The subject of the study was second grade of Sekolah Indonesia Kuala Lumpur Malaysia. The number of the student was thirty six students. This table below is classification score from upper and lower students' score. The score is divided into two groups. This table help the analysis of test item easier to understand.

Table IV. 1

Score List of Eleventh class

Upper Group		lower Group	
No	Students' score	No	Students' score
1	94	1	65
2	91	2	64
3	91	3	64
4	90	4	64

---

[1] Dra Rumiyati Atik, English Teacher in Sekolah Indonesia Kuala Lumpur, *private interview,* Kuala Lumpur, 20 September 2013, 12:45.

5	90	5	63
6	88	6	59
7	86	7	59
8	85	8	58
9	84	9	58
10	84	10	57
11	80	11	56
12	80	12	56
13	79	13	50
14	78	14	49
15	77	15	49
16	71	16	45
17	70	17	45
18	71	18	45

Table above is used to make it easy to classify the upper and lower students of science class. The first column is the student's number who got high score, the second column is the score of students who is classified in upper group, the third column is the student's number who got low score, and the last is the score of students who is classified in lower group.

C. Research findings

After classifying the students to the upper and the lower group, the next step is analysing the validity and item analysis. The researcher takes two kinds of validity; include face validity and content validity. Item analysis includes index of difficulty, index of discrimination and distracters.

1. Face validity

To show the result of face validity of English Test for second grade in Sekolah Indonesia Kuala Lumpur, the researcher took two steps. First step was classifying the matter of the test. Second step was analysing test based on criteria in the table below.

First, the test is printed in A4 paper. The test consisted of seven pages. The first page was used to cover. The cover of the test had; a logo of the school, the subject of test, date and time to do the test as well. Second page until six pages was used to reading section that consisted of forty items. The last pages contained essay test, but in this research, the researcher only analyzed the multiple choice test.

Second, analyzed the test based on the table below.

Table below shows the result of the analysis of the face validity of

the test.[2]

<p style="text-align:center">Table IV. 2</p>

Step	Aspect of test and questions	Explanation
1	Test appearance • How is the cover of test? • How is the letter used in the test? • How is the test layout • How is the size of test paper used?	- The test cover used black colours and suitable font, can be read easily - The size of the letter in the test used 12. Can be read easily. - The test had good layout. The picture used in the test are understandable - The paper in this test used A4 paper
2	The direction • How is the general	- The general instructions of

[2] Adopted from: Sri Utami, *An analysis teacher made English UKK test for academic years 2012 – 2013 for seventh graders of Muhammadiyah 9 surabaya,* Thesis S1, (Unpublished)

	instruction of the test?	this test are understandable.
	• How is the specific instrument of the test?	- This test had no a specific instrument.
	• How is the instruction for going on to text section in the next page?	- This test had no instruction going to the next section/ending.
3	Test items types • How many types of the test have been chosen? • How are the text presented?	- This test had 2 types of test. The test had multiple choice test and essay test. - This test is quite well presented in the layout or arrangement.

2. Content Validity

To show the result of the analysis of content validity of the English test for the second grade of high school of Sekolah Indonesia Kuala Lumpur Malaysia, the researcher uses Standard of Graduates competencies of 2012 to know the connection the test with the standard competencies. The analysing of content validity used table specification (see appendix 3)

There are seventh columns in that table. The first column contains of standard competence, second column contains of basic competencies, the third column contains of indicators, the forth column contains of learning experience, the fifth column contains of item test that is appropriate with the basic competencies, the next column contains of the number of items test ($\sum$) and the last column contains of the percentage of total numbers of particular items represent the elated basic competence.

According to J.B Heaton, the test can be said had a good content validity if it covers all the contents as stated in the curriculum. Based on the result of analysing content validity in appendix 3, this test just covers two criteria, the percentage of every aspect of learning content is concluded as follows:

1. There are 45% 0r 18 items for reading which focused on narrative, hortatory exposition, and spoof.

2. There are 40% or 16 items for linguistics which focused on simple past, past tense, and adverbs.

3. There are 15 % or 6 items unsuitable because it focused on descriptive and present future tense.

Based on the result above, we can conclude that English test in second grade in Sekolah Indonesia Kuala Lumpur high school Malaysia is good since 85% items test represents all materials. It is more than 50%,

according to Bloom if the agreement of the test is 50% or more, it can be concluded that the test had high content validity[3].

Moreover, there are 6 items or 15% of the test did not cover the materials, they are the items test number 7, 8, 9,10,30,35. Those item are unsuitable with the indicator of standard and basic competencies and were not taught in this semester.

3. **Analysing of item test**

   a. **Analysing index of difficulty**

   To get the data of index of difficulty, the researcher divided the class into 2 groups (see table IV.1). The first group was upper group, who were students who get a good score. The second is lower group, who were student who get the bad score.

   After the researcher got the data, she did the analysis using formula as follows:

   $$FV = \frac{R}{N}$$

   Note : FV : index of difficulty

   R : the number of correct answer

   N : the number of students taking test

   There are eight columns in the table. First column contained the number of English test items. The second, it was contained the

---

[3] Bejamin Bloom S, *Handbook on Formative and Summative Evaluation of Student Learning*, 1981, (NewYork: McGraw-Hill Book Co,), page 73.

score of the upper group which answer correctly of each English test items. The third contained the score of the lower group who answer correctly of each English test items. The fourth column contained total of upper group and lower group who answer correctly of each items. The fifth column contained the value of index of difficulty. The sixth column contained upper group minus lower group who answer correctly of each items. The seventh column contained the value of index of discrimination. The eight columns contained comment for each item of index difficulty and index discrimination. See appendix 4.

The researcher did analysis of the English test in second grade of high School Sekolah Indonesia Kuala Lumpur, the class that the researcher use to collect the data was second grade of Science class, and second grade of Social class. The total numbers of student in those classes were thirty six students. The numbers of student of two classes were taken as a sample of this research. The researcher used those classes because of in that school the second grade presently contains of two classes, Science class and Social class. Each major of the school presently had one class and in each class contained less than 20 students, so the researcher uses those two classes to taken a sample of this research. The students divided into two groups as the upper group consist of eighteen students and lower group consist eighteen students.

After analysing the index of difficulty, the next step is machining the result with the criteria of index of difficulty according to Arikunto. The analysis is organized in the following table.

Table IV. 3

Criteria of index difficulty

Index of difficulty	Criteria	Item number	Total of item
0,00 – 0,30	Difficult	34,35,40	3
0.31 – 0,70	Moderate	3,4,5,7,8,9,10, 11, 12,13,14,15, 16,17,18,19, 20,21,22,24,25,26,27,28, 29,31,32, 36,37,38,39.	31
0,71 – 1,00	Easy	1,2,6,23,30,33	6

The table above shows that there are 31 items are moderate level. There are 3 items are difficult level. There are 6 items are easy level. Almost test items are moderate. It means that those items are good to be given to the students. The English test items for second graders high school of Sekolah Indonesia Kuala Lumpur have acceptable index of difficulty.

b. Index discrimination

Index discrimination is tools to differentiate between students who are in the upper group (achieved well) and the lower group (who did not achieve well). To analyze index discrimination, the researcher arranged student in the upper group and the lower group, same as analysing the index difficulty. After arranging the upper and the lower group then the researcher computed data of the index of discrimination.

There are six eight in the table of analysis index of difficulty and index of discrimination. The first columns contained of the number of items. The second column contained the score of student in upper group who answer correctly of each item. Third column contains the score of the lower group who answer correctly of each item. Fourth column contained of total of the upper group and the lower group who answer correctly each items. Fifth column contained the value of index of difficulty. The sixth column contained the numbers of students in the upper group minus the number of students in lower group who answer correctly of each items. Seventh column contained the value of index of discrimination. The eight columns contain comment for each item of index difficulty and index discrimination. See appendix 4.

To calculate the index discrimination for each item number, the

formula used:

$$D = \frac{Correct\,U - Correct\,L}{n}$$

D   = index of discrimination

Correct U   = the number of students in upper group who answer

the item correctly

Correct L   = the number of students in lower group who answer

the item correctly

N   = number of candidate of one group

Table IV. 4

Criteria of index discrimination

Index of discrimination	criteria	Item number	Total of item
0,00 – 0,20	Poor	1,2,4,5,6,7,8,11,12, 14,16,18,22,23,24, 26,27,30,31,32,33, 34,35,39,40	25
0.20 – 0,40	satisfactory	3,9,10,13,15,17,19,20, 21,25,28,29,36,37	13

0,40 – 7,00	Good	38	1
0,70 – 1, 00	Excellent	-	-
-0	Wrong	-	-

Based on the table, the result of index of discrimination shows that there are 25 items had poor index of discrimination, there are 13 items had satisfactory, and there are 1 items had good index of discrimination. Almost students index of discrimination are poor. It means that those items are categorized poor. It means that the English test must be revised.

c. **Analyzing the Effectiveness of Distracters**

Item distractors are the incorrect options in the multiple choices which district the testee from the correct answer. A good distractor will attract more students from the lower group than the upper students. Thus, if there are more able students chosen the distracters, it means that the item does not function as expected in it must be revised.

This English test contains of forty items test and each items test had five answer options. This appendix contains of four columns each items. The first column is the number of items test. Second column is the total correct answer of items from the upper group.

Third column is the total correct answer from the lower group. The last column is the comment of the distracter. See appendix 5.

According to Arikunto, if the distracter was chosen at least by 5% of student who take the test, it is called a good test. (5% from testee = 5% x 60 students = 3 students)[4]. In this case 5% of the total student is 2 students (5% x 36 students).

The result of distracters shows that most of all distractors had good criteria because the distracters have been chosen by more the lower group than the upper group. So the English test had good distracters.

D. Discussion

1. Face Validity

The result of face validity above shows that the test had the criteria of good test. From the cover of the test, it had clear font and colour. The test also had fine letter size to be read. In addition, the test had the acceptable paper size, the test used A4 paper. From the instructions, the instructions are simple and clearly understandable. The first instruction contains date of the test, time to do the test, and how the test must be done. The instruction of each section used unclear instructions. The instruction

---

[4] Suharsimi Arikunto, *Dasar –dasar Evaluasi Pendidikan*, page 238

of each section had items number without explanation about the section. The last criteria are about the kind of the test. The test contain 40 items multiple choice, and 10 essay. From the explanation above the researcher concluded that the English test of second grade of high school Sekolah Indonesia Kuala Lumpur Malaysia had acceptable quality, not actually well other than acceptable to the students.

2. Content Validity

Based on the result of content validity above, the test had 85 % items that covered the indicators of Standard of competencies of 2012. It is 15% items that did not cover the indicators of Standard of Graduates Competencies.

The 45% of the test is focused on reading skill. The test had narrative text, hortatory exposition and spoof text. However in this test there is one descriptive test that not thought in this semester.

Then, 40 % from the test is focused on linguistics. There are simple past, past tense, and conjunction. But in this test there is one of items test that use present future that not thought in this semester.

According to Bloom, if the test agreement is 75% or more, then it can be said that the test had high content validity. On the other hand, if

agreement is less than 50% the rest is considered having low content validity[5]

From the explanation above the researcher concluded that the English test that used for second grade Sekolah Indonesia Kuala Lumpur high school Malaysia had good content validity which had 85% covered indicator of Standard Competencies.

3. Analysing of items test

a. Index of difficulty

Based on the table index of difficulty, the result reported that there are 3 items that classified in difficult criteria. In the index difficulty, 3 out of 40 had difficult value had 0,00 until 0,30. These items must be revised because it is too difficult to be done by the students. This test had 6 out of 40 of easy criteria. The easy value of index difficulty is from 0,71 until 1,00. These items are too easy to the students. This items tests it must be revised.

It can be concluded that most of items or 31 out of 40 items are moderate. The moderate criteria had value from 0,31 until 0,70. These items tests were acceptable for the students. This items test did not to be revised.

---

[5] Bejamin bloom S, *handbook on formative and summative Evaluation of students Learning*, 1981, (New York: McGraw- Hill Book Co)p 73.

b. Index of discrimination

From the result of index discrimination, it was explained that the test have 1 items of the 40 items in good criteria. The good discrimination value is 0, 40 until 0, 70. This item is not to be revised but it must be put in more items to make a good test. Then, this test had 13 out of 40 items that include in satisfactory discrimination value. These items must be revised because it is not a big number as a poor criteria.

Then the researcher concluded that the test is poor criteria. The tests have 25 out of 40 items. In this discrimination value, the poor criteria had 0, 00 – 0, 19 value. These criteria must be revised because the number of poor criteria is major of the total of test items.

c. Analyzing of effective distracter

Item distracters are the incorrect options in the multiple choice which is can amuse the student who do the test from the actual answer. A good distractor will attract more students from the lower group than the upper students.[6]

In this English test of second grade of Sekolah Indonesia Kuala Lumpur Malaysia had forty multiple choice items. Those are number

[6] Sri Utami, *An analysis teacher made English UKK test for academic years 2012 – 2013 for seventh graders of Muhammadiyah 9 surabaya,* Thesis S1, (Unpublished)

one until number forthy, and each items number was contained of five options. So the item distracters of this test was 160. See appendix 5.

From the result above, 7 out of 160 was bad distracters because those item distracters was chosen less that 5% of the total students who take the test. The distracters items must be revised. Besides that, there are 153 out of 160 was good item distracters because the items was chosen by 5% or more of tthe total of the students who take the test .

In Addition, according to Nurgiyantoro the data for analyzing the effectiveness of distracters in the appendix 5 showed that there are 4 out of 160 non function distracters since none from both the upper group and the lower group of students chosen those distracters. Besides, there are 8 out of 160 distracters categorized as adequate, because they had same amount of voters from the upper and and the lower group. Moreover, there are 4 out of 160 malfunction distracters since those items attracts more students in the upper group than students in the lower group, which is good distractors must been chosen by more the lower group than the upper group. These items must be revised. However, there are 144 distractors are good since worked properly to the students. That is concluded that the test had good distracters and not to be revised.