

## CHAPTER IV

### FINDINGS AND DISCUSSION

#### A. Findings

The research is conducted toward the final examination test at Language Development Center in IAIN Antasari Banjarmasin. The total of items are 50 questions. Those are consist of 30 multiple choices, 10 True or False test, and 10 reading comprehension short text (see appendix), but this research only focus to multiple choice items. The test was held on Saturday, 16 December 2015 with the given time 120 minutes.

The analysis quality of test items in final examination at Language Development Center of IAIN Antasari Banjarmasin 2015/2016 based on validity, reliability, item difficulty, item discrimination, and distractor efficiency will be described as follow:

##### 1. Test Validity

The items' items validity count by using Biserial Point Correlation formula:

$$r_{pbis} = \frac{Mp - Mt}{St} \sqrt{\frac{p}{q}}$$

Sample of students in this research are (N=288) students in significant 1% and 5%. If  $r_{pbis} > r_{table}$ , the item is valid. based on the data on appendix, validity for item number 1 is :

$$Mp = 17,74$$

$$Mt = 17,89$$

$$SD = 4,422$$

$$p = 0,163$$

$$q = 0,837$$

$$r_{pbis} = \frac{17,745 - 17,899}{4,422} \sqrt{\frac{0,1632}{0,8368}}$$

$$r_{pbis} = \frac{-0,154}{4,422} \sqrt{0,195}$$

$$= -0,035 (0,442)$$

$$= -0,015$$

Thus, the biserial point correlation for test item number 1 is  $-0,015$ .

On the table of “r” Product Moment (Madeamin, 2013) with significance 1% and 5% for 288 subjects are not available. Thus, to determine  $r_{table}$  for  $N=288$  which exist between 200 – 300, the writer used Interpolation Linier Formula as follow:

$$\frac{y-y_1}{y_2-y_1} = \frac{x-x_1}{x_2-x_1}$$

$$1\% \quad : 0,181 \quad \quad 1\% \quad : 0,148$$

$$5\% \quad : 0,148 \quad \quad 5\% \quad : 0,113$$

$$N=200$$

$$N=300$$

From the data above, we know:

$$X = 200 = 0,191 \quad \quad Y = 200$$

$$X_1 = 288 = ? \quad \quad Y_1 = 300$$

$$X_2 = 300 = 0,148 \quad \quad Y_2 = 288$$

$$\frac{200-288}{300-288} = \frac{0,181-x_1}{0,148-x_1}$$

$$\frac{-88}{12} = \frac{0,181-x1}{0,148-x1}$$

$$-88(0,148 - x1) = 12 (0,181-x1)$$

$$-13,024 + 88x1 = 2,172 - 12 x$$

$$100x = 15,196$$

$$x = \frac{15,196}{100}$$

$$= 0,15196$$

$$= 0,152$$

So, “r” table in significance 1% for N = 288 is 0,152. Furthermore, “r” table in significance 5% for N = 288 using the same formula is 0,116. Biserial point correlation for item number 1 is  $-0,015 < r_{table}$  for significance 1% and 5%. Thus, item number 1 is invalid. Next, the results of next items (Item number 2 to item number 31) use the same formula as shown in the following table below:

**Table 4.1 The Results of Test Validity Analysis**

Items Number	$r_{pbis}$	$r_{table}$	Result	Interpretation
1	- 0,015	1% = 0,152	$r_{pbis} < r_{table}$	Invalid
		5% = 0,116	$r_{pbis} < r_{table}$	Invalid
2	0,326	1% = 0,152	$r_{pbis} > r_{table}$	Valid
		5% = 0,116	$r_{pbis} > r_{table}$	Valid
3	0,465	1% = 0,152	$r_{pbis} > r_{table}$	Valid
		5% = 0,116	$r_{pbis} > r_{table}$	Valid
4	0,338	1% = 0,152	$r_{pbis} > r_{table}$	Valid
		5% = 0,116	$r_{pbis} > r_{table}$	Valid
5	0,467	1% = 0,152	$r_{pbis} > r_{table}$	Valid
		5% = 0,116	$r_{pbis} > r_{table}$	Valid
6	0,469	1% = 0,152	$r_{pbis} > r_{table}$	Valid
		5% = 0,116	$r_{pbis} > r_{table}$	Valid
7	0,344	1% = 0,152	$r_{pbis} > r_{table}$	Valid
		5% = 0,116	$r_{pbis} > r_{table}$	Valid

8	0,246	1% = 0,152	$r_{pbis} > r_{table}$	Valid
		5% = 0,116	$r_{pbis} > r_{table}$	Valid
9	0,168	1% = 0,152	$r_{pbis} > r_{table}$	Valid
		5% = 0,116	$r_{pbis} > r_{table}$	Valid
10	0,370	1% = 0,152	$r_{pbis} > r_{table}$	Valid
		5% = 0,116	$r_{pbis} > r_{table}$	Valid
11	0,284	1% = 0,152	$r_{pbis} > r_{table}$	Valid
		5% = 0,116	$r_{pbis} > r_{table}$	Valid
12	0,448	1% = 0,152	$r_{pbis} > r_{table}$	Valid
		5% = 0,116	$r_{pbis} > r_{table}$	Valid
13	0,393	1% = 0,152	$r_{pbis} > r_{table}$	Valid
		5% = 0,116	$r_{pbis} > r_{table}$	Valid
14	0,243	1% = 0,152	$r_{pbis} > r_{table}$	Valid
		5% = 0,116	$r_{pbis} > r_{table}$	Valid
15	0,232	1% = 0,152	$r_{pbis} > r_{table}$	Valid
		5% = 0,116	$r_{pbis} > r_{table}$	Valid
16	0,156	1% = 0,152	$r_{pbis} > r_{table}$	Valid
		5% = 0,116	$r_{pbis} > r_{table}$	Valid
17	0,337	1% = 0,152	$r_{pbis} > r_{table}$	Valid
		5% = 0,116	$r_{pbis} > r_{table}$	Valid
18	0,431	1% = 0,152	$r_{pbis} > r_{table}$	Valid
		5% = 0,116	$r_{pbis} > r_{table}$	Valid
19	0,476	1% = 0,152	$r_{pbis} > r_{table}$	Valid
		5% = 0,116	$r_{pbis} > r_{table}$	Valid
20	0,279	1% = 0,152	$r_{pbis} > r_{table}$	Valid
		5% = 0,116	$r_{pbis} > r_{table}$	Valid
21	0,481	1% = 0,152	$r_{pbis} > r_{table}$	Valid
		5% = 0,116	$r_{pbis} > r_{table}$	Valid
22	0,336	1% = 0,152	$r_{pbis} > r_{table}$	Valid
		5% = 0,116	$r_{pbis} > r_{table}$	Valid
23	0,455	1% = 0,152	$r_{pbis} > r_{table}$	Valid
		5% = 0,116	$r_{pbis} > r_{table}$	Valid
24	0,176	1% = 0,152	$r_{pbis} > r_{table}$	Valid
		5% = 0,116	$r_{pbis} > r_{table}$	Valid
25	-0,0523	1% = 0,152	$r_{pbis} < r_{table}$	Invalid
		5% = 0,116	$r_{pbis} < r_{table}$	Invalid
26	0,526	1% = 0,152	$r_{pbis} > r_{table}$	Valid
		5% = 0,116	$r_{pbis} > r_{table}$	Valid
27	0,039	1% = 0,152	$r_{pbis} < r_{table}$	Invalid
		5% = 0,116	$r_{pbis} < r_{table}$	Invalid
28	0,435	1% = 0,152	$r_{pbis} > r_{table}$	Valid
		5% = 0,116	$r_{pbis} > r_{table}$	Valid
29	0,261	1% = 0,152	$r_{pbis} > r_{table}$	Valid
		5% = 0,116	$r_{pbis} > r_{table}$	Valid

30	0,593	1% = 0,152	$r_{pbis} > r_{table}$	Valid
		5% = 0,116	$r_{pbis} > r_{table}$	Valid

Based on the data above, the validity of the test can be described specifically in the table as follow:

**Table 4.2 The Percentages of Test Validity Analysis**

Criteria of Validity	Number of Items	Total	Percentage
Valid	2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 26, 28, 29, 30	27	90%
Invalid	1, 25, 27	3	10%

Based on the table above, it is known that the test items in English final examination at Language Development Center IAIN Banjarmasin Academic year 2015-2016 are 90% or 27 items (2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 26, 28, 29, and 30) as valid items and there 10% or only 3 items (1, 25, and 27) as invalid items in significance 1% and 5%

## 2. Test Reliability

The result of reliability for English final examination at Language Development Center IAIN Antasari Banjarmasin is 0,743. The data is taken from this calculation:

$$r_{11}^{22} = \frac{n\sum x_1x_2 - (\sum x_1)(\sum x_2)}{\sqrt{n\sum x_1^2 - (\sum x_1)^2}(n\sum x_2^2 - (\sum x_2)^2)}$$

$$r_{11}^{22} = \frac{288(23947) - (2786)(2369)}{\sqrt{(288(28380) - (2786)^2)(288(21609) - (2369)^2)}}$$

$$r_{11}^{22} = \frac{6896736 - 660034}{\sqrt{(8173440 - 7761796)(6223392 - 5612161)}}$$

$$r_{22}^{11} = \frac{296702}{\sqrt{(411644)(611231)}}$$

$$r_{22}^{11} = \frac{296702}{\sqrt{251609573764}}$$

$$r_{22}^{11} = \frac{296702}{501606,99}$$

$$r_{22}^{11} = 0,5915029214$$

$$r_{22}^{11} = 0,591$$

After that, the next step is using the next formula to find the coefficient of reliability:

$$r_{11} = \frac{2r_{22}^{11}}{1+r_{22}^{11}}$$

$$r_{11} = \frac{2(0,591)}{1+0,591}$$

$$r_{11} = \frac{1,182}{1,591}$$

$$r_{11} = 0,7429289755$$

$$r_{11} = 0,743$$

Based on the calculation about index of reliability of Spearman-Brown with ANATES, the result of reliability showed that the coefficient of reliability of the English test for group A Language Development Center IAIN Antasari Banjarmasin is 0.743 for Multiple Choice test. It means that the English final examination at Language Development Center IAIN Antasari Banjarmasin has high level of reliability because the coefficient of reliability  $0,743 > 0,700$ . In other words, the test items in English final examination at Language Development Center IAIN Antasari are reliable.

### 3. Item Difficulty

Test items which are accepted have index of difficulty 0.30 – 0.70. As a result, test items with item difficulty  $< 0.30$  cannot be accepted because the test items are too difficult, and also the test items with item difficulty  $> 0.70$  cannot be accepted because the test items are too easy.

For instance, in test items number 1, the students who can answer correctly are 47 students from 288 students. Thus, the index of difficulty :

$$P = \frac{Np}{N}$$

$$P = \frac{47}{288} = 0,16$$

From the calculation, the test items difficulty for number 1 is 0,16. It means that the test item number 1 is very difficult test items where the index of difficulty is  $0,16 < 0,30$ . The next results of next items (2 – 30) use the same formula as shown in the following table below:

#### 4.3 The Result of Item Difficulty Level Analysis

Item Number	Np	N	P	Interpretation
1	47	288	0,16	Too difficult
2	118	288	0,41	Moderate
3	190	288	0,66	Moderate
4	99	288	0,34	Moderate
5	168	288	0,58	Moderate
6	166	288	0,58	Moderate
7	202	288	0,70	Moderate
8	142	288	0,49	Moderate
9	266	288	0,92	Too easy
10	79	288	0,27	Too difficult
11	248	288	0,86	Too easy
12	183	288	0,64	Moderate
13	203	288	0,71	Too easy
14	77	288	0,27	Too difficult
15	274	288	0,95	Too easy
16	256	288	0,89	Too easy
17	202	288	0,70	Moderate

18	164	288	0,57	Moderate
19	139	288	0,48	Moderate
20	201	288	0,70	Moderate
21	194	288	0,67	Moderate
22	221	288	0,77	Too easy
23	168	288	0,58	Moderate
24	76	288	0,26	Too difficult
25	46	288	0,16	Too difficult
26	157	288	0,55	Moderate
27	280	288	0,97	Too easy
28	225	288	0,78	Too easy
29	159	288	0,55	Moderate
30	205	288	0,71	Too easy

Based on the data above, the item difficulty level can be described more specific in the table as follow:

**Table 4.4 The Percentages of Item Difficulty Analysis**

<b>Criteria of Item Facility</b>	<b>Number of Items</b>	<b>Total</b>	<b>Percentage</b>	<b>Result</b>
Too difficult (< 0.30)	1, 10, 14, 24, 25	5	16,67%	Not Accepted
Moderate (0.30 – 0.70)	2, 3, 4, 5, 6, 7, 8, 12, 18, 19, 20, 21, 23, 26, 29	15	50%	Accepted
Too Easy (> 0.70)	9, 11, 13, 15, 16, 17, 22, 27, 28, 30	10	33,33%	Not Accepted

Based on the table above, it showed that 15 items (50%) of Multiple Choice test belong to moderate items which have good level in difficulty such as item number 2, 3, 4, 5, 6, 7, 8, 12, 18, 19, 20, 21, 23, 26, and 29. There are 5 items (16,67%) which belong to difficult items such as item number 1, 10, 14, 24, and 25 where they have the range level of difficulty less than 0,30. At last, the items number 9, 11, 13, 15, 16, 17, 22, 27, 28,



and 30, belong to very easy item because it has the level of difficulty more than 0,70.

#### 4. Item Discrimination

Index discrimination power has a range between -1.00 until +1.00. Test items which are accepted have Index of item discrimination more than 0.20. Test items which have discrimination power 0.30 – 0.40 are accepted but still need revision, and the test items which have discrimination power 0.20 – 0.30 must be revised. As a result, the test items with item difficulty  $< 0.20$  cannot be accepted because the test items cannot divide ability of high-ability students and low-ability students (Brown & Priyanvada, 2010, p.70).

To know the Discrimination Power Index use the Formula, as follow :

$$D = P_A - P_B$$

$$\text{Where } P_A = \frac{Ba}{Ja} \text{ and } P_B = \frac{Bb}{Jb}$$

For item number 1, Ba (numbers of the higher group who answer correctly) is 0 and Ja (number of high-ability students) are 78. So, proportion of the higher group ( $P_A$ ) is :

$$P_A = \frac{11}{78} = 0,141$$

Then, number of lower group who can answer correctly (Bb) are 3 and numbers of student who follow the test (Jb) are 20, So, proportion of the lower group ( $P_B$ ) is :

$$P_B = \frac{10}{78} = 0,15$$

$$\text{Then, } D = P_A - P_B$$

$$D = 0,141 - 0,128$$

$$D = 0,013$$

From the result, it is known that item test number 1 has bad discrimination power index. It is  $0,013 < 0,20$ . Next, the results of next items (2 – 30) use the same formula as shown in the following table below :

**Table 4.5 The Results of Item Discrimination Power Analysis**

Item Number	P <sub>A</sub>	P <sub>B</sub>	B <sub>a</sub>	J <sub>a</sub>	B <sub>b</sub>	J <sub>b</sub>	D	Result
1	0,141	0,128	11	78	10	78	0,013	Poor
2	0,628	0,218	49	78	17	78	0,410	Very Good
3	0,885	0,372	69	78	29	78	0,513	Very Good
4	0,513	0,179	40	78	14	78	0,333	Very Good
5	0,846	0,269	66	78	21	78	0,577	Very Good
6	0,846	0,256	66	78	20	78	0,590	Very Good
7	0,885	0,474	69	78	37	78	0,410	Very Good
8	0,615	0,333	48	78	26	78	0,282	Very Good
9	0,987	0,872	77	78	68	78	0,115	Good
10	0,513	0,128	40	78	10	78	0,385	Very Good
11	1	0,743	78	78	58	78	0,256	Good
12	0,923	0,359	72	78	28	78	0,564	Very Good
13	0,897	0,462	70	78	36	78	0,436	Very Good
14	0,397	0,141	31	78	11	78	0,256	Good
15	1	0,885	78	78	69	78	0,115	Enough
16	0,962	0,859	75	78	67	78	0,103	Enough
17	0,885	0,5	69	78	39	78	0,385	Very Good
18	0,846	0,295	66	78	23	78	0,551	Very Good
19	0,808	0,179	63	78	14	78	0,628	Very Good
20	0,872	0,526	68	78	41	78	0,346	Good
21	0,936	0,346	73	78	27	78	0,590	Very Good
22	0,923	0,577	72	78	45	78	0,346	Very Good
23	0,885	0,295	69	78	23	78	0,590	Very Good
24	0,372	0,154	29	78	12	78	0,218	Enough
25	0,103	0,192	8	78	15	78	-0,090	Poor
26	0,885	0,231	69	78	18	78	0,654	Very Good
27	0,987	0,987	77	78	77	78	0	Poor
28	0,961	0,5	75	78	39	78	0,462	Very Good
29	0,705	0,410	55	78	32	78	0,295	Very Good
30	0,987	0,308	77	78	24	78	0,679	Very Good

The result of the index of discrimination can be seen in the table as follows:

**Table 4.6 The Percentages of Item Discrimination Power Analysis**

<b>Criteria of Discrimination Value</b>	<b>Number of Items</b>	<b>Total</b>	<b>Percent age</b>	<b>Result</b>
Very Good ( $D > 0.40$ )	2, 3, 4, 5, 6, 7, 8, 10, 12, 13, 17, 18, 19, 21, 22, 23, 26, 28, 29, 30	20	66,67%	Accepted
Good( $0.30 < D \leq 0.40$ )	9, 11, 14, 20	4	13%	Revise
Enough ( $0.20 < D \leq 0.30$ )	15, 16, 24	3	10%	Revise
Poor ( $D \leq 0.20$ )	1, 25, 27	3	10%	Not Accepted

From the table above, it showed that most of the items in English final test at Language Development Center IAIN Antasari are good. It showed that 20 items (66,67%) of Multiple Choice test have been accepted as very good items such as item number 2, 3, 4, 5, 6, 7, 8, 10, 12, 13, 17, 18, 19, 21, 22, 23, 26, 28, 29, and 30. There are 4 items which accepted as a good test but it still needs a revision such as 9, 11, 14, and 20. There are 3 items worse than the 4 items above but those still use and should be revised such as number 15, 16, and 24. At last, there are 3 items which are not accepted such as number 1, 25, and 27.

#### 5. Distractor Efficiency

A distractor is effective if all of distractors are effective, the item will be accepted. If one of distractors in test item are not effective, the item must be revised. And if all distractor are not effective, the item are not accepted.

To know the whether the distractors have good function or not, use the formula below :

$$D = \frac{\text{number of students who choose the distractor}}{\text{number of students who follow the test}} \times 100\%$$

**Table 4.7 The Results of Distractor Efficiency Analysis**

Item Number	Key Answer	Distractor in %					Result
		A	B	C	D	*	
1	A	16,67	7,29	32,29	43,40	0,35	Accepted
2	D	25,35	13,89	19,44	40,97	0,35	Accepted
3	D	13,19	8,68	11,46	65,97	0,69	Accepted
4	C	13,54	48,26	34,38	<b>3,13</b>	0,69	Revised
5	B	17,01	59,38	<b>4,51</b>	18,40	0,69	Revised
6	B	34,03	57,64	<b>3,82</b>	<b>4,51</b>	0	Revised
7	B	21,88	70,14	<b>3,82</b>	<b>3,82</b>	0,35	Revised
8	C	<b>4,86</b>	10,76	48,6	35,42	0,35	Revised
9	B	<b>0,35</b>	93,40	<b>4,86</b>	<b>1,39</b>	0	Discard
10	B	11,46	27,08	22,92	38,19	0,35	Accepted
11	A	86,46	<b>4,51</b>	7,99	<b>1,04</b>	0	Revised
12	C	35,07	<b>0,35</b>	63,89	<b>0,35</b>	0,35	Revised
13	C	27,78	<b>0,69</b>	70,49	<b>1,04</b>	0	Revised
14	D	44,10	23,96	5,21	26,39	0,35	Accepted
15	A	95,49	<b>1,39</b>	<b>1,39</b>	<b>1,74</b>	0	Discard
16	C	7,29	<b>3,47</b>	88,54	<b>0,69</b>	0	Revised
17	C	14,93	9,38	70,49	5,21	0	Accepted
18	C	11,46	11,11	57,64	19,79	0	Accepted
19	A	48,26	8,33	31,60	11,81	0	Accepted
20	B	18,06	69,79	<b>3,82</b>	7,99	0,35	Revised
21	A	67,36	8,68	8,68	14,93	0,35	Accepted
22	A	77,08	10,76	5,21	6,94	0	Accepted
23	C	13,54	23,96	58,33	<b>3,13</b>	1,04	Revised
24	B	45,83	26,39	14,24	12,50	1,04	Accepted
25	D	45,49	5,91	31,60	15,97	1,04	Accepted
26	A	54,51	<b>0,35</b>	10,07	34,72	0,35	Revised
27	B	<b>1,74</b>	96,88	<b>0,69</b>	<b>0,69</b>	0	Discard
28	C	<b>2,78</b>	11,46	77,78	6,25	1,74	Revised
29	A	55,21	28,82	14,93	<b>0,69</b>	0,35	Revised
30	A	71,88	13,54	8,68	5,56	0,35	Accepted

Note: the items with bold is percentage of distractor less than 5%

To determine the efficiency of distractor, the writer consider if the distractor less than 5%, it means that the distractor is bad. There are 3 categories in classifying the distractor. First, if all distractor are good , it

means that the distractors of the item are accepted. Second, if some distractors are bad, it means that the distractors should be revised. At last, if all distractors are bad, it means that the item test is bad item that should be discarded.

The result of the distractor efficiency can be seen in the table as follows:

**Table 4.8 The Percentages of Distractor Efficiency**

Number of Item	Total	Percentage	Result
1, 2, 3, 10, 14, 17, 18, 19, 21, 22, 24, 25, 30	13	43,33%	Accepted
4, 5, 6, 7, 8, 11, 12, 13, 16, 20, 23, 26, 28, 29	14	46,67%	Revised
9, 15, 27	3	10%	Discarded
Total	30	100%	

Based on the table, the distractors efficiency of multiple choice test item are 13 (43,33%) items as effective and accepted item, 14 (46,67%) as ineffective and need to be revised item, and 3 (10%) items need to be discarded. The effective test items are 1, 2, 3, 10, 14, 17, 18, 19, 21, 22, 24, 25, and 30. Furthermore, test items which are not effective are 4, 5, 6, 7, 8, 11, 12, 13, 16, 20, 23, 26, 28, and 29. But, the test item which has all ineffective distractors are 9, 15, and 27.

## **B. Discussion**

After all of the findings are described above, the writer will also discuss them. The result of analysis will determine whether the quality of the test items are good or not.

### 1. Analysis of Validity

Based on the result of validity using biserial point correlation, as shown on the table 4.1 and 4.2, there are 27 items (90%) are valid and 3 items (10%) valid for significant 1% and 5%. As a result, it can be assumed that the test items in English Final examination at Language Development Center IAIN Antasari Banjarmasin are mostly valid. There are only 3 item need to be revised.

The writter found the data that the reasons why the items are valid because when it is compared with the objectives (see appendix), for entire test items were representative almost all objectives. It means that the test is appropriate with used syllabus in English Course at Language Development Center.

### 2. Analysis of Reliability

Based on the calculation by using ANATES software. The reliability of test items in English final examination at Language Development Center IAIN Antasari Banjarmasin are 0,743. In this research, the reliabilty between 0.60 until 0.80 is in high level of reliability. Because of that, from the reliability result, the test items entirely can be accepted as a good test, reliable and has a good consistency.

### 3. Analysis of Item Difficulty Level

There are some differences level of test items based on index of difficulty, those are very easy item, moderate item, and very difficult item. In this research the index of difficulty the test items are sufficient. It has

more moderate level and some very easy and very difficult test without far space among them. The test is known from the data where totally 50% of test items are moderate level. Very easy items consist of 10 items, moderate items consist of 15 items, and very difficult items consist of 5 items. Thus, the test items entirely are not too difficult or too easy for the students. But, based on the theory, the very easy and very difficult test should be revised or discarded.

a. Very Easy Items Difficulty

For very easy items, there are some reasons why they cannot be accepted as a good test. As an example the test item number 9 as follow:

9. DON'T FEED THE ANIMAL!

Where do you find this notice?

- a. at home
- b. at the zoo
- c. at the hospital
- d. at the market

The percentage based on the answer of 10 of high-ability students and 10 low-ability students:

**Table 4.9 The Difficulty Level Analysis for Number 9**

No.	Level of students	A	B	C	D	Result
9.	High-ability students	-	10	-	-	Too easy
	Low-ability student	-	9	-	1	
	Total	0	19	0	1	
	Percentage	0%	95%	0%	5%	

The key answer is B. Based on the table above, the correct answer test item is chosen by 100% of 10 of high-ability students and 90% of low-ability students., totally 20 students from both level students 95% choose the correct answer. It showed that the test is very easy to do. In addition, the index of difficulty at number 9 is 0,92. It can be concluded that the test item cannot be accepted because it is too easy. As a result, the test items which included as too easy items should be revised or discarded.

b. Very Difficult Items Difficulty

For the very difficult test items, they also have some reasons why they cannot be accepted as good test items. As an example the test item number 25 as follow:

25. My brother \_\_\_\_\_ very hard at the moment, because some of his officemates are off sick.

- a. works
- b. work
- c. is working
- d. worked

The percentage base on the answer of 10 of high-ability students and 10 low-ability students:

**Table 4.10 The Difficulty Level Analysis for Number 25**

No.	Level of students	A	B	C	D	Result
25.	High-ability students	5	1	3	1	Too difficult
	Low-ability student	5	1	3	1	



	Total	10	2	6	2	
	Percentage	50%	10%	30%	10%	

The key answer is C. Based on the table above, the test item is chosen only by 30% of 10 of high-ability students and 30% of low-ability students, totally 20 students from both level only students 30% choose the correct answer. It showed that the test is very difficult to do. In addition, the index of difficulty at number 25 is 0,16. It can be concluded that the test item cannot be accepted because it is too difficult. As a result, the test items which included as too difficult items should be revised or discarded.

#### c. Moderate Test Item Difficulty

For moderate level test, there are some reasons why the test can be accepted. For example the test item number 2 as follow:

2. Which one is the wrong statement?

- a. He is never glad to see me.
- b. He never forgets to do his homework.
- c. Usually, I drink milk before going to bed.
- d. I hardly ever don't sing for public.

The percentage base on the answer of 10 of high-ability students and 10 low-ability students:

**Table 4.11 The Difficulty Level Analysis for Number 2**

No.	Level of students	A	B	C	D	Result
2.	High-ability students	0	0	0	10	Good item
	Low-ability student	1	3	4	2	
	Total	1	3	4	12	
	percentage	5%	15%	20%	60%	

The key answer is D. Based on the table above, the test item is chosen only by 100% of 10 of high-ability students and only 10% of low-ability students, totally 12 students from both level only students 60% choose the correct answer. It showed that the test is not too easy but also not too difficult to do. In addition, the index of difficulty at number 2 is 0,41. It can be concluded that the test item is accepted because it is not too difficult and also not too easy. As a result, the test items which included as moderate items can be accepted as a good test.

#### 4. Analysis of Discrimination Power

##### a. Very Good Level of Item Discrimination

The descriptions of test item which has very good level in discrimination for example number 2 as follows:

**Table 4.12 The Discrimination Power Analysis for Number 2**

No.	Level of students	A	B	C	D	Result
2.	High-ability students	0	0	0	10	Accepted as good item
	Low-ability student	1	3	4	2	
	Total	1	3	4	12	
	Percentage	5%	15%	20%	60%	

Based on the table above, it showed that option A has been chosen by 1 low-ability student (5%), option B has been chosen by 3 low ability students (15%), option C has been chosen by 4 low-ability students (20%), and option D as the correct answer has been chosen by 12 students (60%) who consist of 10 high ability students (100%) and 2 student low ability students (20%). This test item is accepted as very good level of discrimination power because it test item can divided

between high-ability students and low ability students. In addition, if a test item has index value more than 0.40, it means that the test is very good. This test item's index discrimination for number 2 is 0.410. So, based on the data, this test has very good quality in index of discrimination.

b. Good Level of Item Discrimination

The descriptions of test item which has good level in discrimination power:

**Table 4.13 The Discrimination Power Analysis for Number 17**

No.	Level of students	A	B	C	D	Result
17.	High-ability students	2	0	8	0	Good
	Low-ability student	1	1	8	0	
	Total	3	1	16	0	
	Percentage	15%	5%	80%	0%	

Based on the table above, it showed that option A has been chosen by 3 students (15%) who consist of 2 high-ability students (20%) and 1 low-ability student (10%), option B has been chosen by 1 low ability students (5%), option D has been not chosen by any student (0%), and option C as the correct answer has been chosen by 16 students (80%) who consist of 8 high ability students (80%) and 8 student low ability students (80%). From the data, it is known that the test item is accepted as good level of discrimination power because it test item can divided between high-ability students and low ability students. But, this test item cannot be accepted as very good level in discrimination because this test

item's index discrimination is 0.385 only where it means that this test item has index score less than 0.40. Thus, based on the data, this test has good quality in index of discrimination.

c. Enough Level of Item Discrimination

The descriptions of test item which has moderate level in discrimination power:

**Table 4.14 The Discrimination Power Analysis for Number 14**

No.	Level of students	A	B	C	D	Result
14.	High-ability students	4	0	0	6	Enough
	Low-ability student	3	3	0	4	
	Total	7	3	0	10	
	Percentage	35%	15%	0%	50%	

Based on the table above, it showed that option A has been chosen by 7 students (35%) who consist of 4 high-ability students(40%) and 3 low-ability student (30%), option B has been chosen by 3 low ability students (15%), option C has been not chosen by any student (0%), and option D as the correct answer has been chosen by 10 students (50%) who consist of 6 high ability students (60%) and 4 student low ability students (40%). From the data, it is known that the test item is accepted as enough level of discrimination power because it test item can divided between high-ability students and low ability students. But, this test item only has index discrimination for number 14 is 0.256. If a test has index of discrimination  $> 0.20 - \geq 0.30$ , that test item is categorized as enough level in index of discrimination. Thus, based on the data, this test still accepted as good quality test in index of discrimination.

d. Poor Level of Item Discrimination

The descriptions of test item which has poor level in discrimination power:

**Table 4.15 The Discrimination Power Analysis for Number 27**

No.	Level of students	A	B	C	D	Result
27.	High-ability students	0	10	0	0	Poor
	Low-ability student	0	10	0	0	
	Total	0	20	0	0	
	Percentage	0%	100%	0%	0%	

Based on the table above, it showed that option A, C, and D have been not chosen by any students (0%), but option D as the correct answer has been chosen by 20 students (100%) who consist of 10 high ability students (100%) and 10 student low ability students (100%). From the data, this test item cannot accepted as a good test, because this test cannot divided between high-ability students and low ability students. In addition, the index of discrimination power for this test item is 0.00 where if a test has index discrimination less than 0.20, it is called poor index discrimination power.

Test items in English final test at Language Development Center IAIN Antasari are very good. It is known that 14 items of Multiple Choice test have been accepted as very good items, 6 items which accepted as a good test but it is still need a revision, 4 items that must be revised and only 6 items (20%) which are not accepted. The total of all items that can be accepted or need revision are 24 items (80%) from 30 items of Multiple choice test.

From the analysis of item discrimination power above, the test items are very good. It is mean that the test items can divide between high-ability students and low-ability students. If the item discriminatoin power is very good, the group of smart students easily can be known based on the result of test.

## 5. Analysis of Distractor

### a. Accepted Distractor

The descriptions of good distractor are as follow:

**Table 4.16 The Discrimination Power Analysis for Number 2**

No.	Level of students	A	B	C	D	Result
2.	High-ability students	0	0	0	10	Accepted as good item
	Low-ability student	1	3	4	2	
	Total	1	3	4	12	
	Percentage	5%	15%	20%	60%	

A distractor called effiience when it has been chosen by minimally 5% test-taker. In this item, all options have been chosen by minimally 5% of students. It is means that all distractors have done their function to attract the student to wrong answer.

2. which one is the wrong statement?
  - a. He is never glad to see me.
  - b. He never forgets to do his homework
  - c. Usually, I drink milk before going to bed
  - d. I hardly ever don't sing for public

The key answer is D. The options have similar length. The options of the test item are homogen where the content use adverbial time

complication to test students' ability in comprehending the correct grammar.

b. Revised Distractor

The descriptions of bad distractors are as follows:

**Table 4.17 The Discrimination Power Analysis for Number 14**

No.	Level of students	A	B	C	D	Result
14.	High-ability students	4	0	0	6	Accepted as good item
	Low-ability student	3	3	0	4	
	Total	7	3	0	10	
	Percentage	35%	15%	0%	50%	

A distractor called efficient when it has been chosen by minimally 5% test-taker. In this item, option C has been not chosen by any student. That is one point which shows that the distractor should be revised. Besides, the option A has attracted 4 high-ability students, it should be revised to make the test to be a good item.

14. Zaky : Why \_\_\_\_\_ on the floor?

Elma : Because there aren't any chairs to sit on.

a. do you sit

b. are you sit

c. you sit

d. are you sitting

There are some discussion about this test item. The key answer is D. The options have similar length. The options of the test item is not homogen where the option C use statement form while another option in

question form, it makes 0% of 20 students who consist of 10 high-ability students and 10 low ability students chose the option C.

The indicator of this test is how to ask in simple present continuous tense to test students' ability in comprehending the correct grammar.

c. Discarded Distractor

The descriptions of very bad distractors are as follows:

**Table 4.18 The Discrimination Power Analysis for Number 27**

No.	Level of students	A	B	C	D	Result
27.	High-ability students	0	10	0	0	Poor
	Low-ability student	0	10	0	0	
	Total	0	20	0	0	
	Percentage	0%	100%	0%	0%	

In this item, option A, C, and D have been not chosen by any student. It means that for high-ability students and low-ability students this question is too easy and it is known all the distractor cannot attract them. Those are the points which shows that the distractor should be discarded.

27. Patrick : Hi, Bob. What \_\_\_\_ you \_\_\_\_?

- a. Is – does
- b. are – doing
- c. are – do
- d. is – doing



There are some discussions about this test item. The key answer is B. The options have similar length. The indicator of this test is how to use in simple present continuous tense and to be correctly, the test is to test students' ability in comprehending the correct grammar. The options of the test item is homogen where all of the options first as to be and second as verb. However, it still makes 100% of 20 students who consist of 10 high-ability students and 10 low ability students chose the option B. Therefore, the test item has very bad quality of distractor and it should be discarded.

A distractor can be accepted as good distractor if 5% of test-taker choose it. A distractor which is not chosen by all of test-taker will be a bad distractor. It also happen if a distractor is chosen by all test-taker. In this research, the total of 43,33% or almost a half of test items have good distractors. It also showed that 46,67% or a half of test items have bad distractor. But, there are 10% of test items which has ineffective distractor. based on the data, the efficiency of distractor are in sufficient level or moderate level where most of the distractors in the test still need revised for make test as a good test.

## 6. Grouping of test items quality based on Item Difficulty and Item Discrimination

Novitasary (2010, p. 93) mention that, a good test item is based on the item difficulty and item discrimination. If in a test item the difficulty and the discrimination power are good, it means that the test item is good. But, if one of item difficulty and item discrimination is bad, it means that the test item should be revised. However, if item difficulty and item discrimination are bad, it means the test items is bad and should be discarded. After analysing the quality of test items, the writter will grouping the test items quality based on item difficulty and item discrimination. The description will be explain in the table as follow:

**Table 4.18 Group of Test Items After Analyzing Process**

Categorize	Number of item	Total	Percentage
Good	2, 3, 4, 5, 6, 7, 8, 12, 18, 19, 21, 23, 26, and 29	14	46,67%
Revised	9, 10, 11, 13, 14, 15, 16, 17, 20, 22, 24, 28, and 30	13	43,33%
Discarded	1, 25, and 27	3	10%
Total		30	100%

Based on the table above, there are 3 items which is categorized as bad or error item from 30 test items in the multiple-choice test. It means the test items of English final examination at Language Development Center IAIN Antasari Banjarmasin are very good because the test items only have 10% of bad items or error. It is appropriate with the result of reliability which showed that the test items have good validity and high level of reliability.

There are 14 items as good item such as item number 2, 3, 4, 5, 6, 7, 8, 12, 18, 19, 21, 23, 26, and 29. Those can be used in the next test or as a references of good test. Next, there are 13 items that should have revision because there are some problems in item difficulty and item discrimination. They are items number 9, 10, 11, 13, 14, 15, 16, 17, 20, 22, 24, 28, and 30. At last, there are totally 3 items as bad items that should be discarded or have review in depth, those are items number 1, 25, and 27.