

AN ITEM ANALYSIS OF THE ENGLISH SUMMATIVE TEST FOR THE SEVENTH GRADE STUDENTS OF SMPN 12 TANGERANG SELATAN

Sabilar Rosyad

Universitas Muhammadiyah Jakarta

Email: sabilarrosyad@umj.ac.id

Abstract

The objective of this study is to find out whether or not the English summative test for the Seventh Grade Students of SMPN 12 Tangerang Selatan can be categorized as a good test based on the difficulty level, the discriminating power and the effectiveness of distracters. The respondents of this study were the seventh grades students of SMP Negeri 12 Tangerang Selatan. The respondents of this study were taken from two classes. The number of this population was 68 students. The students of each class consist of 38 and 30 students. The number of sample of this study was 34 respondents, 17 respondents of the upper group and 17 respondents of the lower group. The method of the study was descriptive quantitative and qualitative method. In analyzing the difficulty level and the discriminating power, the writer used the formula suggested by Heaton to analyze the difficulty level and the formula suggested by Hopkin and Antes to analyze the discriminating power. After analyzing the difficulty level, there have been 24 items which categorized in the easy level, 24 items which categorized in the medium level, and 2 items which categorized in the difficult level. In terms of the discriminating power, there have been 37 items can be accepted, 10 items should be revised, and 3 items should be discarded. Based on the results of the difficulty level and the discriminating power, it can be concluded that the test fulfills the criteria of a good test.

Keywords: *Evaluation, Test, The Criteria of Test, Item Analysis, The Criteria of Item Analysis.*

A. INTRODUCTION

English is an international language. Many people use it for communication and interaction with other people. English is also used as a foreign language and taught at school. By learning English, students could achieve four basic skills, e.g. listening, speaking, reading and writing. In order to know if the students achieve the skills, the teachers need to give evaluation.

In school, evaluation is part of teaching and learning process. Through evaluation the teacher are able to find out the effectiveness of a method and the students' achievement in mastering the lesson. According to Gronlund (1981:5-6) stated that evaluation may be defined as a systematic process of determining the next extent to which instructional objectives are achieved by pupils.

Evaluation cannot come to the end if it has no information or data. The data must be collected from the students in order to be analyzed to decide whether or not the instructional objectives are achieved by students. In this case, one of the instruments that can be applied is test. According to Brown (2004:3) stated that a test is a method of measuring of person's ability, knowledge, or performance in a given domain.

In order to get accurate measurement of the students' ability, test must be well constructed by good items. According to Harris (1969:13) said that there are three criteria of a good test. They are reliability, validity and practicality. By having the three qualities, the test can be said a good test. Moreover, through test the teacher can get information of the students

accurately, so that the teacher can make right judgment about the level of the students' achievement and the effectiveness of the teaching method used in the classroom.

Besides concerning about these criteria of validity, reliability and practicality, the quality of a good test can be known by identifying the quality of individual test items. For knowing the quality of the test items, the teachers must do the procedure which is usually called item analysis.

Arikunto (2009:206) said that item analysis is used to know the quality of a good test item. Her native language said: *Analisis soal bertujuan untuk mengadakan identifikasi soal-soal yang baik, kurang baik, dan soal yang jelek. Dengan analisis soal dapat diperoleh informasi tentang kejelekan sebuah soal dan "petunjuk" untuk mengadakan perbaikan.*

This explanation means that through an item analysis, the teachers can identify whether the items is good or poor, and the items must be kept, revised, or discarded.

Ideally, every school must do the item analysis for knowing the quality of the test items, because teachers must make right judgment in their teaching. For making a right judgment, they need data accurately. By using the item analysis, the teachers can measure the items through the difficulty level, discriminating power, and the effectiveness of the distracters, so the teachers can know whether the items is good or poor, and the items must be kept, revised, or discarded.

Schools always hold the summative test in the end semester. One of the schools which hold the summative test is SMPN 12 Tangerang Selatan. SMPN 12 Tangerang Selatan is the favorite school in Tangerang Selatan. The writer sees that many people are interested to study there. The writer also wants to know the quality of the individual items. According to the information above, the English summative test in that school had never been analyzed, so the students and the teachers do not know whether the test has a good quality or not.

Based on the information that SMPN 12 Tangerang Selatan had never been analyzed, the writer is interested in making analysis in that school. Therefore, the writer concentrates on the discussion of "An Item Analysis of English Summative Test for the Seventh Grade Students at SMPN 12 Tangerang Selatan."

The question of the problem as follow: "Do the English Summative Test Items for the Seventh Grade Students for SMPN 12 Tangerang Selatan in the 2011/2012 Academic Year have a good criteria of test based on difficulty level, discriminating power and the effectiveness of the distracters?"

The Objective of the study is to find out the empirical answer to the question which has been formulated in the statement of the study whether the test can be categorized as a good test based on difficulty level, discriminating power and the effectiveness of distracters.

B. LITERATURE REVIEW

1. The Understanding of Evaluation

In teaching and learning process, evaluation is one of the important things for the teachers and the students. By evaluation, the teachers can determine the ability of the students. According to Gay (1991:6) stated that evaluation is the systematic process of collecting and analyzing data in order to make decisions. The key terms of this definition are systematic process, collecting, analyzing data, and make decisions. *Systematic process* means evaluation has a fixed set of procedures or steps. *Collecting* is getting the data or information. *Analyzing* is examining the data or information. *Make decision* means derive judgment. Based on the meanings of the key terms above, the definition of evaluation of Gay is meant as a systematic process of making judgment through getting and examining the data or information.

Parallel to the above meaning, Gronlund and Linn (1990:5) said that evaluation is the systematic process of collecting, analyzing, and interpreting information to determine the extent to which pupils are achieving instructional objectives. The key terms of this the

definition are systematic process, collecting, analyzing, interpreting, determine the extent, and instructional objective. The term *interpreting* means to explain the data based on the students' achievement. *Instructional objective* means the aims that must be achieved by the students in learning and teaching activity. Based on the key terms of Gronlund above, evaluation means to make judgment based on three procedures or process of getting, examining, and explaining the data of the students' achievement the aims in learning and teaching activity. This definition is related with Gay' statement, because the term is about getting and examining the data for judging or making decision.

Bachman and Palmer (2010 : 21) added that evaluation involves making value judgments and decisions on the basis of information, and gathering information to inform such decisions is the *primary* purpose for which language assessments are used. The key terms of this definition are making value judgments, basis of information, gathering information, the primary purpose, and language assessments. In this case, *making value judgment* is the making decision, *gathering information* means collecting information, *basis of information* means the most important of the data or the information, *the primary purpose* means the main objective to do something, and *language assessment* means the situation to decide or judge in the language. Based on the explanation of the key terms above, evaluation has the main objective to get the data or the information for making decision in the language.

Based on the definition above, the writer summarizes that evaluation is a systematic process of making judgment or decision through three steps; getting, examining, and explaining the data or the information to determine how far the students' achievement in their learning.

2. The Understanding of Test

To measure the students' achievement, the teachers need instruments that is used for collecting the data or the information. One of the evaluation instruments is a test.

Hopkins and Antes (1990:130) say that test is an instrument, device, or procedure that proposes a sequence of task to which a student is to respond, the results of which are used as measure of a specified trait. The key terms of this definition are: instrument or device, sequence, task, to respond, measure, and a specific trait. *Instrument* or *device* means a tool, *sequence* means a series, *task* is assignment. *to respond* means to do or to finish, *to measure* means to know how far the students' achievement of instructional objective, and *a specific trait* means a particular characteristic. Based on the meanings of the key terms above, a test is a tool consisting of assignments that must be done by the students for measuring their specific characteristic.

Furthermore, the explanation of test is added by Hughes (2003:1). He said that: A test was devised which was based directly on an analysis of the English language needs of first year undergraduate students, and which included tasks as similar as possible to those which they would have to perform as undergraduates (reading, textbook materials, taking notes during lectures, and so on).

The important terms in definition are devised and included tasks. *Devised* means instrument. *Tasks* mean assignments which usually come from reading, textbook materials, and so on, it can be used as an instrument. Based on the meanings of the key terms above, test is an instrument consisting of some assignments for measuring the students' ability. This definition is related with Hopkins and Antes' statement. Both use the term instrument for test. Hopkins and Antes' statement that test is used to measure the students' characteristic while Hughes' statement that it is used to measure the students' ability.

According to Bachman (2004:9) stated that a test is a particular type of measurement that focuses on eliciting a specific sample of performance. The key terms of this definition are measurement, eliciting and a specific sample of performance. *Eliciting* means to get something especially the information. *A specific sample of performance* means a particular sample about

the students' ability. Based on the meanings of the key terms above, test is used to measure the students' ability in attaining the objective in learning process.

Based on the definitions and the elaboration of the key terms above, the writer summarizes that test is an instrument or tool consisting of some assignments that must be done by the students, and the objective is to get the measurement or scores of their ability in teaching and learning process.

3. The Criteria of a Good Test

A good test must have criteria. All good tests have three qualities: Validity, Reliability, and Practicality. It means that test must have the criteria above that is used to know the quality of the test which is made by the teachers.

a. Validity

The first criteria of good test is validity. According to Brown as quoted by Hudson and James (2002 : 212) stated that the validity has been defined as the degree to which a test measures what it claims, or purports, to be measuring.

b. Reliability

The second criteria of good test is reliability. According to Gronlund (1981:93) stated that reliability refers to the *consistency* of measurement – that is, to how consistent test scores or other evaluation results are from one measurement to another.

c. Practicality

The third criteria of good test is practicality. According to Brown (2004 : 19) said that an effective test is practical.

4. The Understanding of Item Analysis

Item analysis is an important thing which is useful to identify good items, less item and poor items. So, by item analysis can be gotten information about “Weakness” an item and “Instruction” to do the revision.

Purwanto (1994:118) stated that the objective of item analysis is to find whether the item which is good or not, and why the item can be said good or not. His native language stated “*Tujuan khusus dari items analysis ialah mencari soal tes mana yang baik dan mana yang tidak baik, dan mengapa item atau soal itu dikatakan baik atau tidak baik*”. The key terms of this definition are to identify and whether the item is good or not. *Identify* is used to recognize some item which is good or not. Based on the meanings of the key terms above, through item analysis, the teachers can recognize the test item which is good or not.

Furthermore, the explanation of test is added by Madsen (1983:180). He said that: Each question needs to function properly; otherwise it can weaken the exam. Fortunately, there are some rather simple statistical ways of checking individual items. This procedure is called “item analysis”. An item analysis has basically three things: how difficult each item is, whether or not the question “discriminates” or tells the difference between high and low students, and which distractors are working as they should.

The key terms of this definition are simple statistical way, checking individual items, the difficulty, discriminates, and working distractors. *Simple statistic way* means the item analysis used easy statistic method. *Checking individual items* mean the examination of the individual items. *The difficulty* is used to know whether the item is easy or not. *Discriminates* is used to divide the difference between upper and lower students. *Working distractor* is used to know the individual items have distractor to distract the correct answer effectively. Based on the meanings of the key terms above, item analysis is statistical method to examine the individual item by knowing the item is easy or not, dividing the difference between upper and lower students, and the effectiveness the distractor of the individual items. This definition of item analysis is different with the Purwanto' statement which telling about the aim of the item

analysis without mentions about the difficulty level, the discriminating power, and the effectiveness of the distractor.

According to Lado (1962 : 342) stated that item analysis is the study of validity, reliability, and difficulty of test items taken individually as if they were separate tests. The key terms of this definition are validity, reliability, and difficulty. *Validity* means the data or the information is able to be accepted based on the truth, and *reliability* means the data or the information can be believed. Based on the meanings of the key terms above, item analysis is used to examine the data or the information is able to be accepted, believed, and difficult of test items which taken individually. This definition of test here is different from Madsen' statement in the key terms; the difficulty level, the discriminating power, and the effectiveness of the distractor.

Based on the definition above, the writer summarizes that item analysis is statistical method to examine the test item which is good or not by knowing the item is easy or not, dividing the difference between upper and lower students, the effectiveness of the distractor of the individual items.

5. The Criteria of Item Analysis

a. The Difficulty Level

According to Heaton (1990 : 178) said that the index of difficulty (FV) is generally expressed at the fraction (or percentage) of the students who answered the item correctly. It is calculated by using the formula:

$$FV = \frac{\text{Correct } U + \text{Correct } L}{2n}$$

Notes :

FV: The difficulty index

U : The number in the upper group who answered correctly

L : The number in the lower group who answered correctly

2n : The total of the test takers in upper and lower groups

The result of the difficulty index is interpreted by using the following categories from Sudjana (2005 : 137).

0,00 to 0,30 the item is categorized "difficult"

0,31 to 0,70 the item is categorized "medium"

0,71 to 1,00 the item is categorized "easy"

b. The Discrimination Power

Madsen (1983:182) stated that the discriminating level is how well it differentiates between those with more advanced language skill and those with less skill. The discrimination power can be calculated with the formula below:

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

Notes :

D = The discrimination index

U = The number in the upper group who answered correctly

L = The number in the lower group who answered correctly

n = The number in one of the criterion groups

According to Arikunto (2009:218) said that the classification of the discrimination power is

D ; 0.70-1.00	: Excellent
D ; 0.40-0.70	: Good
D ; 0.20-0.40	: Satisfactory
D ; 0.00-0.20	: Poor
D ; < 0,00 or negative	: Very Poor

C. METHOD

The study was conducted at SMP Negeri 12 Tangerang Selatan, which is located at Jalan Jurang Mangu Barat No.62 Kecamatan Pondok Aren Kelurahan Jurang Mangu Barat Tangerang Selatan. The study was held on March 29th, 2012. The respondents of this study were the seventh grades students of SMP Negeri 12 Tangerang Selatan. The respondents of this study were taken from two classes. The number of this population was 68 students. The students of each class consist of 38 and 30 students. The number of sample of this study was 34 respondents, 17 respondents of the upper group and 17 respondents of the lower group. To obtain the data, the writer used the summative test paper and students' answer sheets as the instrument of the study. For analyzing and recording the data, the writer developed the tables and analyzed it.

The method of the study was descriptive quantitative and qualitative method. The quantitative method was used to analyze the data to get the difficulty level and the discriminating power by using mathematic statistic formula. Furthermore, the qualitative method was used to analyze the weakness of the items.

The steps to follow in item analysis are:

1. Going to school and asking permission for the headmaster of SMPN 12 Tangerang Selatan. Besides that, the writer consults with English teacher in that school about the summative test.
2. Asking permission to copy the learner answer sheets, the test paper and the answer key.
3. Re-checking the answer key and the English summative test.
4. Scoring the answer key and arranging based on the rank scores, from the highest to the lowest group.
5. Dividing the answer sheet into three groups by taking:
25% of the highest rank as the Upper Group (UG)
50% of the medium rank as the Medium Group (MG)
and 25% of the lowest rank as the Lower Group (LG).
In this matter, the writer does not use the medium group, which is about 50%, because in the item analysis only the upper group and the lower group are needed.
6. Tabulating the answer of each students, who answer correctly from each groups in the different tables.
7. Calculating the index of the difficulty level and the discriminating power from each item.
8. Identifying the item based on the discriminating power, which are needed to revise or to reject.
9. Making the alternative of the revision.

D. RESULTS AND DISCUSSION

1. The Description of the Data

a. The Difficulty Level

- Easy items number : 1, 2, 3, 4, 5, 6, 8, 10, 15, 17, 18, 19, 20, 24, 28, 33, 37, 39, 40, 41, 43, 45, 46, 47 = 24

$$\frac{24}{50} \times 100 \% = 48 \%$$

- Medium items number : 7, 9, 11, 12, 13, 14, 16, 21, 22, 23, 25, 26, 27, 29, 30, 31, 32, 34, 36, 38, 42, 44, 48, 50 = 24

$$\frac{24}{50} \times 100 \% = 48 \%$$

- Difficult items number : 35, 49 = 2

$$\frac{2}{50} \times 100 \% = 4 \%$$

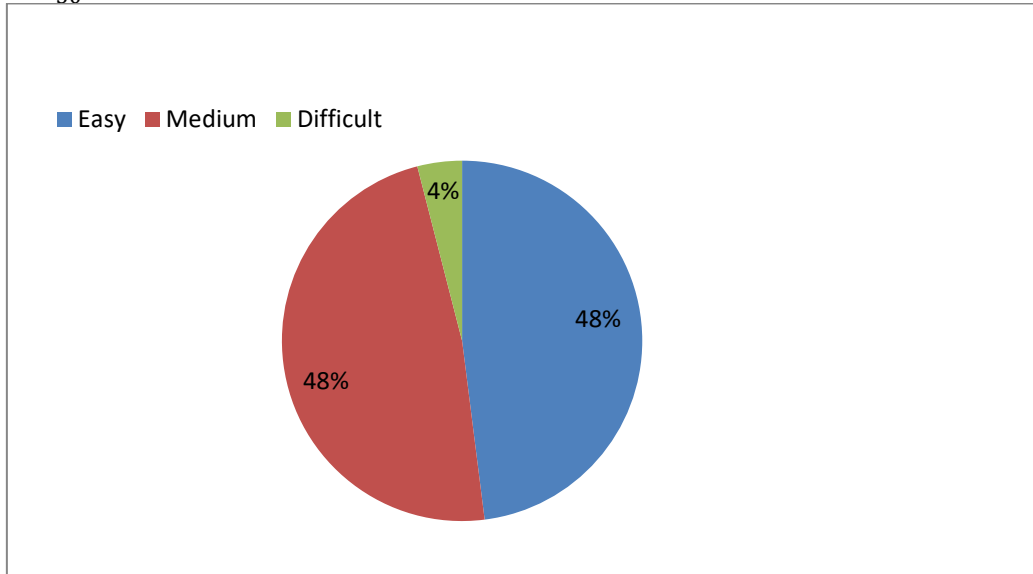


Figure 1. The Pie Chart of Difficulty Level

b. The Discriminating Power

Based on the discriminating power, the writer finds out that the test contains 3 items should be discarded; 10 items should be revised; and 37 items should be accepted.

- Discarded items number : 15, 33, 34 = 3

$$\frac{3}{50} \times 100 \% = 6 \%$$

- Revised items number : 1, 3, 4, 6, 19, 24, 28, 40, 47, 50 = 10

$$\frac{10}{50} \times 100 \% = 20 \%$$

- Accepted items number : 2, 5, 7, 8, 9, 10, 11, 12, 13, 14, 16, 17, 18, 20, 21, 22, 23, 25, 26, 27, 29, 30, 31, 32, 35, 36, 37, 38, 39, 41, 42, 43, 44, 45, 46, 48, 49 = 37

$$\frac{37}{50} \times 100 \% = 74 \%$$

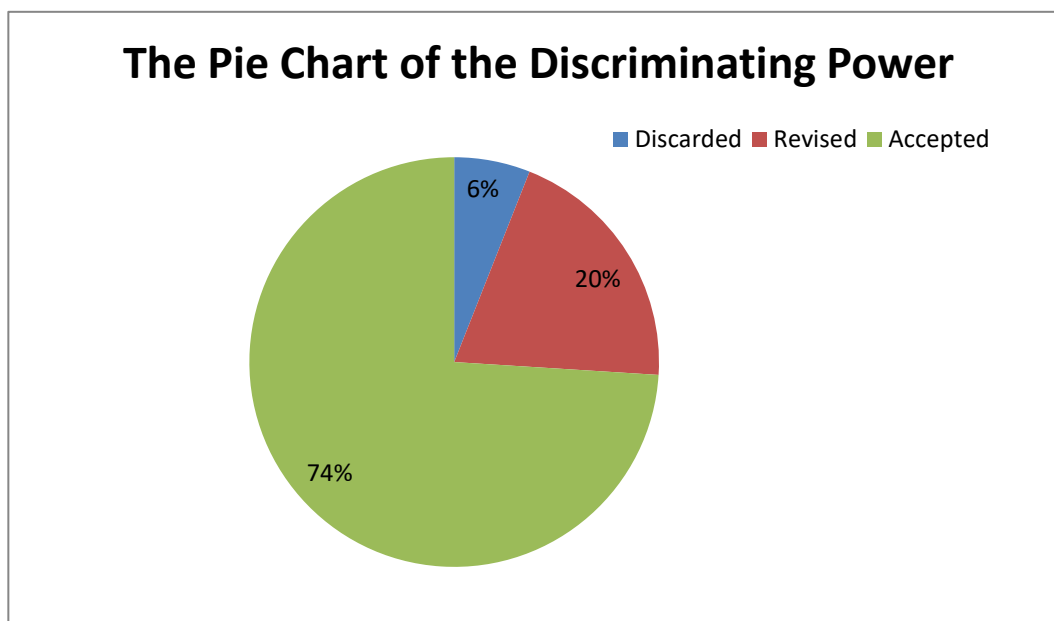


Figure 2. The Pie Chart of the Discriminating Power

Based on the percentage of the result of difficulty level, it can be seen that easy level has the same percentage with medium level about 48%. The last rank is difficult level with 4%. Based on the percentage of the result of discriminating power, it can be seen that the high rank is accepted items with 74%. The second rank is revised item with 20%. The last rank is discarded items with 6%. The revisions are presented below:

In this case, the writer has made the revisions of the weak items. The items that should be revised are numbers 1, 3, 4, 6, 19, 24, 28, 40, 47, and 50. These items have some weaknesses and they are discussed as follows:

1. Item Number 1

a. Copy of the item

1. Choose the right answer to complete the dialog.

Bob: Good morning. Rob

Roby : Good morning. ... ?

Bob: I'm fine.Thanks

a. how about you c. And you

b. how are you d. how old are you

b. The Competence Measured

The item measures the students' abilities in identifying kinds of expression of "greeting" to others.

c. The Data of the Item

Item Number	Groups	Options				Total	FV	DP
		A	B*	C	D			
1	Upper Group	3	14	0	0	17	0.74	0.18
	Lower Group	3	11	1	2	17		

* The correct answer

d. The description of the weakness

The item is easy so that cannot discriminate the upper group and the lower group. Most of the upper group and the lower group can answer it correctly. The writer finds out that the weaknesses come from:

First, there have been the most mistake happened in the text. The performance of the test, especially the typing in the stem of the question is to much spacebar in the first and the third sentences. And don't forget the capital letter in the beginning sentences in all options.

Second, some students do not know the point of the dialog in this stem. This statement is supported by the evidence that there had been several students who chose option A, "How about you?". The expression in the option A and B have the similar purpose to ask someone, but the situation in the sentence is different. The option A is used when suggesting or offering to someone, while the option B is used for asking someone about the health or the feeling.

Third, several options do not work effectively, because they cannot distract the students. Most of upper and lower group can answer correctly and only some students from upper group and lower group who cannot answer it. The option B as the correct option is easy to find, so other options must be changed. The option C "And you" because this sentence is used for asking back someone. It should be changed into "How do you do?". "How do you do" is used in a formal greeting for someone that you have not met before. Based on the dialog, it can be seen that Bob has know Roby before, so by changing the option, it might distract the students. The option D "How old are you?" because this dialog do not ask about the age. It should be changed into "What happen to you?".

e. Revision

Text

1. Choose the right answer to complete the dialog.

Bob : Good morning, Rob.
 Roby : Good morning,

Bob : I'm fine. Thanks

a. How about you ?
 b. How are you ?
 c. How do you do ?
 d. What happen to you ?

2. Item Number 3**a. Copy of the item**

3. Y : What's your name?
 X : My name is Rama
 Y :
 X : R - A - M - A

a. How are you? c. How about you?
 b. How do you spell it? d. Nice to meet you?

b. The Competence Measured

The item measures the students' abilities in identifying kinds of expression "spelling the letter" to other.

c. The Data of the Item

Item Number	Groups	Options				Total	FV	DP
		A	B*	C	D			
3	Upper Group	0	17	0	0	17	0.97	0.06
	Lower Group	0	16	0	1	17		

* The correct answer

d. The description of the weakness

The item is easy so that cannot discriminate the upper group and the lower group. Most of the upper group and the lower group can answer it correctly. The writer finds out:

First, the weaknesses come from the options do not work effectively, because they cannot distract the students. Most of upper and lower group can answer correctly and only one students from lower group who cannot answer it. The option B as the correct option is easy to find, so other options must be changed.

Second, the option A “*How are you?*” is impossible to distract the students, so the students do not answer this option because this sentence is used to ask someone about the health, so this expression does not relate with the dialog. This option should be changed into “*How do you say it?*”. By changing the option might distract the students.

The option C “*How about you?*”, is similar with the option A which is used to ask someone, so this option should be changed into “*How do you tell it?*”.

e. Revision

3. Y : What's your name ?	
X : My name is Rama	
Y :	
X : R – A – M – A	
a. How do you say it ?	b. How do you tell it ?
b. How do you spell it ?	c. Nice to meet you ?

3. Item Number 4**a. Copy of the item**

4. Rudy : he works in the post office and usually he sends letters from home one home. what does he do?	
Iwan : He is a	
a. nurse	c. postman
b. secretary	d. waiter

b. The Competence Measured

The item measures the students' abilities in identifying the information about the job on the dialog.

c. The Data of the Item

Item Number	Groups	Options				Total	FV	DP
		A	B	C*	D			
4	Upper Group	0	0	17	0	17	0.94	0.12
	Lower Group	0	2	15	0	17		

* The correct answer

d. The description of the weakness

The item is in the easy level so that cannot discriminate the upper group and the lower group. It may be caused by several possibilities which is made the item is poor. The writer finds out:

First, there have been the most mistake happened in the text. The performance of the test, especially the typing in the stem of the question is too much spacebar in the first sentence. And don't forget the capital letter in the beginning sentences in the dialog and all options.

Second, there is not need for overall revision in the text but just a little touch for the second sentence. The sentence “ *usually he sends letters from home one home*” should be changed into “ *He works in the post office and sends the letters from one home to other home*”.

Third, the option A “*Nurse*” should be changed into “*Salesman*” because it does not relate to the post office. It might distract the students. The option D “*Waiter*” should be change into “*Writer*”.

e. Revision

4. Rudy : He works in the post office and
sends letters from one home to other home.
What does he do ?

Iwan : He is a ...

a. Salesman	c. Postman
b. Secretary	d. Writer

4. Item Number 6**a. Copy of the item**

Read the text to answer for no 5 – 7

My name is Sila. I want to tell about my teacher. My teacher's name is Miss Rika. She is kind and patient teacher. She cares with her students well. Most of her lesson are interesting. She loves to help slow students to improve in their work. She often tells us jokes when we do our work. I hope she is to be my teacher again next year.

6. The teacher has

a. bad character	c. selfish character
b. good character	d. emotional character

b. The Competence Measured

The item measures the students' abilities in identifying the information in the text.

c. The Data of the Item

Item Number	Groups	Options				Total	FV	DP
		A	B*	C	D			
6	Upper Group	0	17	0	0	17	0.91	0.18
	Lower Group	1	14	1	1	17		

* The correct answer

d. The description of the weakness

The item is in the easy level so that cannot discriminate the upper group and the lower group. Most of the upper group and the lower group can answer it correctly. It can be caused by several factors that make this item should be revised. The writer finds out:

First, There are many mistakes in typing the text. In the second sentence need more spacebar from “*Iwant*” into “*I want*”. For using the punctuation mark do not attended in the third sentences. The full stop is put near the last word in the sentence, after that give one space. In the fifth sentence, the word “*is*” must be ommitted because the word “*cares*” is “verb”. The sixth sentence, the word “*lesson*” must be added “*s*” in the last of the word, it becomes “*lessons*”. The seventh sentence, the writer thinks that the use of the word “*slow*” is not suitable, so it should be changed into “*weak*”.

The stem of the question “*The teacher has*” should be changed into “*The character of Miss Rika are true, EXCEPT ...*”. By changing the question, the writer thinks it might distract the students because the first question is easy to answer by the students.

The writer changes all of the option. The option A “*Bad character*” should be changed into “*Patient*” because it is stated in the text. The option B “*Good character*” should be changed into “*Selfish*” because it is opposite with the character of Miss Rika, so it is the correct answer. The option C “*Selfish character*” should be changed into “*Caring*” because it is stated in the text. The option D “*Emotional character*” should be changed into “*Kind*” because it is stated in the text.

e. Revision**Read the text to answer for no. 5 - 7**

My name is Sila. I want to tell about my teacher. My teacher’s name is Miss Rika. She is kind and patient teacher. She cares with her students well. Most of her lessons are interesting. She loves to help weak students to improve in their work. She often tells us jokes when we do our work. I hope she is to be my teacher again next year.

6. The character of Miss Rika are true, EXCEPT

- | | |
|------------|-----------|
| a. Patient | c. Caring |
| b. Selfish | d. Kind |

5. Item Number 19

a. Copy of the item

Read the text to answer no 18 – 22

Our class studies in the laboratory. Mrs . Wadarti our Biology teacher explains the amphiby family. She uses a frog as the example. She takes a dead frog and cut off the stomach. She shows us the parts of its explains their function one by one. We stand in front of her. We also have to draw the parts of the body because we have a test about it next week. The bell rings students leave the laboratory to have break.

19. When do the student have test?

a. Next week

c. Today

b. Now

d. everyday

b. The Competence Measured

The item measures the students' abilities in identifying the information in the text.

c. The Data of the Item

Item Number	Groups	Options				Total	FV	DP
		A*	B	C	D			
19	Upper Group	17	0	0	0	17	0.97	0.06
	Lower Group	16	1	0	0	17		

* The correct answer

d. The description of the weakness

The item is easy so that cannot discriminate the upper group and the lower group. Most of the upper group and the lower group can answer it correctly. The writer finds out that the weaknesses come from:

First, the text in the third sentence “.... *and cut off the stomach*”. The word “*cut*” should be added “*s*” in the end of the word into “*cuts*” so the sentence becomes “....*and cuts off the stomach*”. The text in the seventh sentence “..... *because we have a test about it next week*”. It should be changed into the future form “..... *because we will have a test about it next week*”. It is caused by there is “*next week*” in this sentence.

Second, the stem of the question “*When do the student have a test?*” should be change into “*Will the students have a test next week?*” because it related to the text. By changing the question, the writer thinks it might distract the students because the first question is easy to answer by the students.

Third, all of the options should be change into the future tense because the question needs yes/no answer. The option A “*Next week*” should be change into “*Yes, they will*”. The option B “*Now*” should be change into “*No, they will not*”. The option C “*Today*” should be change into “*Yes, we will*”. The option D “*Everyday*” should be change into “*No, we will not*”.

e. Revision

Read the text to answer no. 18 - 22

Our class studies in the laboratory. Mrs.Wadarti our Biology teacher explains the amphiby family. She uses a frog as the example. She takes a dead frog and cuts off the stomach. She shows us the parts of its explains their function one by one. We stand in front of her. We also have to draw the parts of the body because we will have a test about it next week. The bell rings students leave the laboratory to have break.

19. Will the students have a test next week?

- a. Yes, they will c. Yes, we will
b. No, they will not d. No, we will not

6. Item Number 24**a. Copy of the item**

Complete the blank of the text for no 23 -26

Dona . . . (23) as a . . . (24) in a new department store. She never . . . (25) late every morning. She is very . . . (26) worker. She usually types 12 letters on her computer.

24. a . doctor c . nurse
b . secretary d . student

b. The Competence Measured

The item measures the students' abilities in choosing the suitable word in the form of fill in the blank.

c. The Data of the Item

Item Number	Groups	Options				Total	FV	DP
		A	B*	C	D			
24	Upper Group	0	17	0	0	17	0.91	0.18
	Lower Group	2	14	0	1	17		

* The correct answer

d. The description of the weakness

The item is easy so that cannot discriminate the upper group and the lower group. Most of the upper group and the lower group can answer it correctly. The writer finds out:

First, related to the answer key B "*Secretary*", commonly the job of "*secretary*" is not in the department store but in the office. It makes the the first sentence in the text should be changed from "*Dona works as a secretary in a new department store*" into "*Dona works as a secretary in a new office*".

Second, all of the option should be changed into the occupations which are in the office. The option A "*Doctor*" should be changed into "*Director*". The option C

b. The Competence Measured

The item measures the students' abilities in identifying kinds of expression "asking the time" to others.

c. The Data of the Item

Item Number	Groups	Options				Total	FV	DP
		A	B	C*	D			
40	Upper Group	0	0	17	0	17	0.94	0.12
	Lower Group	0	2	15	0	17		

* The correct answer

d. The description of the weakness

The item is in the easy level so that cannot discriminate the upper group and the lower group. It can be caused by several factors that make this item should be revised. The writer finds out:

The weaknesses come from the options do not work effectively, because they cannot distract the students. Most of upper and lower group can answer correctly and only one students from lower group who cannot answer it. The option C as the correct option is easy to find, so other options must be changed. The option A "Where is it" should be changed into "What time do you have". The option B "What is it" should be changed into "What's the time". The option D "How is it" should be changed into "Do you have the correct time". By changing the option, it might distract the students.

e. Revision

40. X : ? Y : It is five past ten a. What time do you have b. what's the time c. What time is it d. Do you have the correct time

9. Item Number 47**a. Copy of the item**

47. Y : ... ? X : I am ten years old a. How are you b. What are you c. How old are you d. Where are you
--

b. The Competence Measured

The item measures the students' abilities in identifying kinds of expression "asking the age" to others

c. The Data of the Item

Item Number	Groups	Options				Total	FV	DP
		A	B	C*	D			
47	Upper Group	0	0	17	0	17	0.94	0.12
	Lower Group	2	0	15	0	17		

* The correct answer

d. The description of the weakness

The item is in the easy level so that cannot discriminate the upper group and the lower group. It may be caused by several possibilities which is made the item is poor. The writer finds out:

The option B “*What are you*” should be changed into “*How do you do*”. The option D “*Where are you*” should be changed into “*How about you*”. By changing the options, It might distract the students.

e. **Revision**

47. X : ? Y : I am ten years old a. How are you b. How do you do	c. How old are you d. How about you
--	--

10. **Item Number 50**

a. **Copy of the item**

50. A person who repairs the leaking tap is ... a. a plumber b. bricklayer	c. tailor d. carpenter
--	---------------------------

b. **The Competence Measured**

The item measures the students’ abilities in identifying the job

c. **The Data of the Item**

Item Number	Groups	Options				Total	FV	DP
		A*	B	C	D			
50	Upper Group	7	8	0	2	17	0.32	0.18
	Lower Group	4	1	12	0			

* The correct answer

d. **The description of the weakness**

The item is in the medium level but it failed to discriminate the upper and the lower group. It can be caused by several factors that make this item should be revised. The writer finds out:

The question in the stem “*A person who repairs the leaking tap is...*” is difficult to answer for the students because the writer thinks that the students do not know the meaning of “*the leaking tap*” which is meant in Indonesia language (*kebocoran kran*), so the question should be changed with the familiar words into “*A person who repairs plumbing is*”.

There is no need the revision to the option. They work effectively, because they might distract the students.

e. **Revision**

50. A person who repairs the plumbing is ... a. Plumber b. Bricklayer	c. Tailor d. Carpenter
---	---------------------------

E. CONCLUSION

Item analysis is statistical method to examine the test item which is good or not by knowing the item is easy or not, dividing the difference between upper and lower students, the effectiveness of the distractor of the individual items to find out whether the items need to be accepted, revised or discarded.

After analyzing the items based on the difficulty level the writer finds out that there are 24 items can be categorized in the easy level or 48%, then there are 24 items can be categorized in the medium level or 48% and the last, there are 2 items can be categorized in the difficult level or 4%. Ideally, the proportion of a test should be 25% of easy items, 50% of medium items, and 25% of difficult items. According to the result of the difficulty level, the test is not proportional.

Related to the result of the discriminating power, there are several numbers of item should be revised, discarded and accepted. Based on the data the writer finds out that there are 10 items should be revised with percentage 20%, there are 3 items or should be discarded with percentage 6% and there are 37 items should be accepted with percentage 74%. According to the result of discriminating power, there are only few of discarded items. Based on the results above it can be concluded that the English Summative test items for seventh grade of SMPN 12 Tangerang Selatan fulfill the criteria of a good test.

REFERENCES

- Arikunto, S. (2009). *Dasar-dasar Evaluasi Pendidikan Edisi Revisi*. Jakarta: Bumi Aksara.
- Bachman, L. F., & Palmer, A. (2010). *Language Assessment in Practice: Developing Language Assessment and Justifying their Use in the Real World*. New York : Oxford University Press.
- Bachman, L. F. (2004). *Statistical Analyses for Language Assessment*. New York: Cambridge University Press
- Brown, D. H. (2004). *Language Assessment Principles and Classroom Practices*. New York: Longman Inc.
- Brown, J. D., & Hudson, T. (2002). *Criterion-referenced Language Testing*. New York: Cambridge University Press.
- Gay, L. R. (1991). *Educational Evaluation and Measurement: Competencies for Analysis and Application*. New York: Machmillan, Inc.
- Gronlund, N. E. (1981). *Measurement and Evaluatin in Teaching*. New York: Macmillan.
- Gronlund, N. E., & Linn, R. L. (1990). *Measurement and Evaluation in Teaching*. New York: Macmillan
- Harris, D. P. (1969). *Testing English as a Second Language*. New York: Tata McGraw-Hill.
- Heaton, J. B. (1990). *Writing English Language Tests*. London: Longman.
- Hopkins, C. D., & Antes, R. L. (1990). *Classrom Measurement and Evaluation*. Itasca: F. E. Peacock Publisher, Inc.
- Hughes, A. (2003). *Testing for Language Teachers*. New York: Cambridge University Press.
- Lado, R. (1964). *Language Testing*. New York: McGraw-Hill.
- Madsen, H. S. (1983). *Techniques in Testing*. New York: Oxford University Press.
- Nunnally, J. C. (1972). *Educational Measurement and Evaluation Second Edition*. New York: McGraw-Hill, Inc.
- Purwanto, N. M. (1994). *Prinsip-prinsip dan Teknik Evaluasi Pengajaran*. Bandung: Remaja Rosdakarya.
- Sudjana, N. (2005). *Penilaian Hasil Proses Belajar Mengajar*. Bandung: Remaja Rosdakarya.