

## CHAPTER II

### THEORETICAL FRAMEWORK

#### A. The Scope of English Test at Junior High School Level

##### 1. Definition of Test

In the learning process, the teacher is required to make a lesson plan before teaching. Lesson plans are needed to make it more accessible for the teacher to deliver material by educational standards. The implementation includes preliminary, core and closing activities. By gives material, the teacher evaluates the students. One of them is through a test to know student understanding among learning process.

A test is an instrument or systematic procedure for observing and describing a student's characteristics using either a numerical scale or a classification scheme.<sup>1</sup> Brown stated that a test is "a method of measuring a person's ability, knsowledge or performance in a given domain."<sup>2</sup> Through testing, the teacher can measure student achievement and teacher performance in the classroom. A test gives the data and the information to the teacher about how well the student mastered the materials taught in the learning process.

##### 2. English Test at Junior High School

The level of school that is taken after completing elementary school is called junior high school. Many subjects have been selected to be taught. One of them is a

---

<sup>1</sup> Anthony J Nitko and Susan M Brookhart, *Educational Assessment of Students. Sixth Edition* (Edinburgh Gate: Pearson, 2014), 27.

<sup>2</sup> Brown, *Language Assesment Principles and Classroom Practices* (USA: Longman, 2004), 3.

foreign language. Nelly and Karen stated that foreign language is currently optional at junior high school. Still, in practice, almost all schools in junior high school environment choose a foreign language to be taught.<sup>3</sup> English as a foreign language to be one subject in junior high school that uses the textbooks are authorized. Therefore, every semester, the teacher makes the test a final term to know how far the student understanding, whether it has been taught in the textbook or not, which called the summative test. The English summative test that examined in junior high school are:

#### **a. Speaking**

Speaking is a study about how students communicate with each other. Speaking also shares student expression, idea, information and others with their partner or listener. Baker and Westrup stated that speaking the students repeat sentences or dialogues to create an encourage environment where students can practise expressing themselves.<sup>4</sup> The student usually practicing the dialogue with their friends in their class. In the exam, the student must complete the dialogue with an appropriate word or phrase and determine the meaning of the expressions as the written test.

#### **b. Reading**

Reading is a process of learning about writing symbols and analyzing the meaning of the text. Nation declared, "Reading is a source of learning and a source

---

<sup>3</sup> Nelly P. Stromquist and Karen Monkam, *Globalization and Education: Integration and Contestation across Cultures* (USA: Rowman and Littlefield, 2000), 246.

<sup>4</sup> Joanna Baker and Heather Westrup, *Essential Speaking Skills: A Handbook for English Language Teachers* (London: Continuum, 2003), 7.

of enjoyment. It can be a goal in its own right and a way of reaching other goals”.<sup>5</sup> It means reading is a process of understanding the contents of the text either loudly or silently. The process of knowing the success of reading is by evaluation. The evaluation can give through the test. In reading, the test usually consists of determining stated information, determining implied information, determining specific information, determining moral value and determining the meaning of a word, phrase or sentence.

### **c. Writing**

Writing is learning to put down the graphic symbol that represents the utterance meaning.<sup>6</sup> Sometimes, writing has a different meaning from structured sentences and utterance meaning. Writing is essential to build student skills in understanding the text, whether written by other people in the coursebook or by students self in their book. The teacher gives an exam designed to rearrange jumbled sentences into a meaningful paragraph and arranging jumbled words into complete sentences to know the student's writing ability.

### **d. Vocabulary**

Ur declares, "vocabulary can be defined, roughly as the words we teach in the foreign language."<sup>7</sup> Hatch and Brown define vocabulary as the setlist of the words

---

<sup>5</sup> I. S. P. Nation, *Teaching ESL/EFL Reading and Writing* (New York: Routledge, 2009), 49.

<sup>6</sup> Robert Lado, *Language Teaching: Scientific Approach*, (New York: Mc. Graw Hill, Inc. 1964), 35.

<sup>7</sup> Penny Ur, *A Course in Language Teaching Practice and Theory* (Cambridge: Cambridge University Press, 1996), 60.

for specific language speakers uses.<sup>8</sup>Order researchers Kamil and Hiebert also explained that vocabulary is a kind of word that students need to achieve the knowledge and meaning to increase the comprehension of a text.<sup>9</sup>

From the definition above, vocabulary is a series of words that can learn to communicate and write or speak. The test is usually designed to identify the profession correctly, things, place artist correspond to the description, determining correctly characteristic of someone (adjective) correspond to the description, and determining the correct particular place for short functional text.

#### **e. Grammar**

The rules for writing and speaking with an actual or reasonable language are defined as grammar. Sean stated that grammatical concepts are like a tool in writing, making the student writing more descriptive, clear or exciting, such as an artist that uses a marker or paintbrush as their tools.<sup>10</sup> It means grammar has an essential function in arranging the text. In general, the test item of grammar is identified about determining social function, the structure of text and language features, and complete the gaps with the right preposition.

---

<sup>8</sup> Evelyn Hatch and Brown, *Vocabulary, Semantic, and Language Education* (Cambridge: Cambridge University Press, 1995), 1.

<sup>9</sup> Michael L.Kamil and Elfrieda H. Hiebert, *Teaching and Learning Vocabulary: Bringing Research to Practice* (New Jersey: Lawrence Erlbaum Associates, 2005), 3.

<sup>10</sup> Sean Ruday, *The Common Core Grammar Toolkit: Using Mentor Texts to Teach the Language Standards in Grades 6-8* (New York: Routledge, 2014), 153.

## **B. Test Item Analysis**

A powerful tool for test improvement and accumulating the quality of the test is the analysis of student response to an objective test item.<sup>11</sup> The Activity that aims to analyze the items on the test questions whether it complies with the criteria and function is called item analysis. Nitko and Susan stated that item analysis is the process of collecting, summarizing and using information from student responses to the item for making decisions.<sup>12</sup> The test item can analyze after the teacher scored the student answer sheet. It is done to know the student's ability to answer the test and the effectiveness of the test that is examined test. The manner to analyze the item also must be appropriate with their procedures. There are three norm-referenced for item analysis procedures:<sup>13</sup>

1. The difficulty of the item
2. The discrimination power of the item
3. The effectiveness of each alternative

Thus, the three norm-referenced can show the differences between each item on the test. It can show whether each item is too easy or too hard for the students, how well the item discriminates the student scores between lower and higher scores, and whether all of the alternatives work like functions.

---

<sup>11</sup> Robert L. Ebel and David A. Frisbie, *Essentials of Educational Measurement*. Fifth Edition (New Delhi: Prentice-Hall of India, 1991), 225.

<sup>12</sup> Nitko and Brookhart, *Educational Assessment*, 254.

<sup>13</sup> Norman E. Gronlund, *Constructing Achievement Test*. Second Edition (USA: Prentice-Hall, Inc, 1977), 110.

### **C. Level Difficulty**

The percentage that shows the difficulties and easier of the question is mentioned as the level of difficulty. Boopathiraj and Chellamani stated item difficulty is also referred to as the  $p$ -value. The higher the value, the easier the item.<sup>14</sup> If the student scores higher dominant in the upper group than the lower group, the test item can mention as easier. But if it does not, then the item is difficult.

To determine the difficulty level, item difficulty has an index ranging from 0 to 1. The analysis of item difficulty functioned to know the quality of each item. It is also functioned to know whether the student prepares the material well or not. Is the item categorized as easy, moderate or difficult based on the index ranging.

### **D. Discrimination Power**

The percentage that shown the differences in student ability is mentioned as discrimination power. It is discriminate between upper and lower groups. According to Osterlind "Discrimination power is an index to determining differences among individual examinees on the subject matter or psychological construct being assessed".<sup>15</sup> it means the test item must be analyzed to know the index criteria of each item.

The discrimination power functioned to prove the effectiveness of the test item. If most of the students from the upper group answer correctly, it means the test item is excellent and good. But, if most students from the lower group answer correctly,

---

<sup>14</sup> C. Boopathiraj and DR. K. Chellamani, "Analysis of Test Items on Difficulty Level and Discrimination Index in the Test for Research in Education", *International Journal of Social Science & Interdisciplinary Research*, Vol. 2, No. 2, (February 2013), 189-193.

<sup>15</sup> Steven J. Osterlind, *Constructing Test Item* (London: Kluwer Academic Publisher, 1989), 282.

it means the test item is poor. It means the item is not enough effectiveness and the item must be revised or discarded.

#### **E. Criteria of Interpreting of Level Difficulty and Discrimination Power**

In analyzing the level of difficulty and discrimination power of test items, interpreting is needed. It is adopted for the analysis process by following the procedure. Item analysis procedure of multiple-choice are:<sup>16</sup>

1. Arrange the scored test answer sheets in score order from highest to lowest
2. Identify an upper group and a lower group separately. The upper group is the highest, scoring 27% of the group and the lower group is the lowest, scoring 27% of the group.
3. For each item, count the number of examinees in the upper group that chose each response alternative. Do a separate, similar tally for the lower group.
4. Record these counts on a copy of the test at the end of the corresponding response alternative.
5. Add the two counts for the keyed response and divide this sum by the total number of students in the upper and lower groups. The result is an estimate of the index of item difficulty.
6. Subtract the lower group count from the upper group count for the keyed response. Divide this difference by the number of examinees in a group (either group since both are the same size). The result expressed as a decimal is the index of discrimination.

---

<sup>16</sup> Ebel and Frisbie, *Essentials of Educational Measurement*, 225.

After knowing the procedure above, the criteria of discrimination power and level of difficulty on analysis test items will be mentioned. The Criteria of interpreting of difficulty level classified as below:

**Table 2.1 The Classification of Item Difficulty Level<sup>17</sup>**

<b><i>P</i></b>	<b>Interpretation</b>
0.00 – 0.30	Difficult
0.31 – 0.70	Moderate
0.71 – 1.00	Easy

In interpreting the criteria, the formula is needed for computing item of level difficulty for analyzing multiple-choice is as follow:<sup>18</sup>

$$P = \frac{B}{T}$$

Where:

*P* = the index level of difficulty

*B* = the number of test takers in the total group who answered correctly

*T* = total number of test takers in the group

Criteria of interpreting discrimination power classified as below:

**Table 2.2 The Classification of Item Discrimination Power<sup>19</sup>**

<b>Discrimination Index</b>	<b>Qualification</b>	<b>Interpretation</b>
Negative Value	Very poor/ Worst	Definitely discard

---

<sup>17</sup> Kunandar, *Penilaian Autentik (Penilaian Hasil Belajar Peserta Didik Berdasarkan Kurikulum2013): Suatu Pendekatan Praktis* (Jakarta: Raja Grafindo Persada, 2013), 241.

<sup>18</sup> Kunandar, *Penilaian Autentik*, 240.

<sup>19</sup> Anas Sudijono, *Pengantar Evaluasi Pendidikan*. Ed. 1. Cet.15 (Jakarta: Rajawali Press, 2016), 389.



<0.20	Poor	Deemed not to have good discriminating power and discard
0.20 – 0.40	Satisfactory	Has sufficient discriminating power and need to review
0.40 – 0.70	Good	Has good discriminating power and possibilities for improvement
0.70 – 1.00	Excellent	Has excellent discriminating power and retain

In interpreting the criteria, the formula is needed for computing item discriminating power for analyze multiple-choice is as follow:<sup>20</sup>

$$D = \frac{Ru - Rl}{\frac{1}{2}T}$$

Where:

$D$  = the index of discriminating power

$Ru$  = the number in the upper group who answered the item correctly

$Rl$  = the number in the lower group who answered the item correctly

$T$  = total number of students included in the item analysis

## F. Previous Of Study

A previous study is a research of previous works related to the title and research of the researcher to avoid plagiarism or duplication and find the validity of previous works regarding the effectiveness of the methods used. The previous study gives inspiration to the researcher on research this title. Based on the results of the previous

---

<sup>20</sup> Gronlund, *Constructing Achievement Test*, 112-113.

study, the researcher found the research result that has similarities with research that the researcher did but had differences in the substance of their contents.

The first research by Alif Fadillah A, "An Analysis on Difficulty Level and Discriminating Power of English USBN Test 2018 (A descriptive Analysis of SMPN 10 Kota Tangerang Selatan in Academic Year 2017/2018)". This research uses quantitative by using descriptive analysis as a method, which aims to analyze and describe the quality of English USBN test items of difficulty and discriminating power at SMPN 10 Kota Tangerang Selatan. The population of this research is 80 students from nine grade and conducted a purposive sampling technique. A finding of this research showed that English USBN items had 65% (26 test items) included to be moderate level in term of difficulty level and 45% (18 test items) included to be good quality in terms of discriminating power.<sup>21</sup>

The second research by Eni Larianti, "The Analysis of Final School Exam Items on English Subject at SMPN 14 Seluma in Academic Year 2018/2019". This research focuses on the final English test in multiple-choice. The method of this research is descriptive quantitative. This research analyzed the difficulty level, discrimination power and item distractor in nine grades. This research finding showed that there are 0 excellent items, 6 good items, 15 satisfactory items and 15 poor items and 4 very poor items in discrimination power. There are 3 very difficult items, 16 difficult items, 18 desirable items, 3 easy items and 0 very easy item in

---

<sup>21</sup> Alif Fadillah A, "An Analysis on Difficulty Level and Discriminating Power of English USBN Test 2018: A descriptive Analysis of SMPN 10 Kota Tangerang Selatan in Academic Year 2017/2018" (Skripsi, UIN Syarif Hidayatullah Jakarta, 2018), 1-47.

difficulty level. In distractor, there are 6 very good items, 15 good items, 10 average items, 2 poor items and 6 very poor items.<sup>22</sup>

The last research by Hisbullah, "The Analysis of Discrimination Power of English Summative Test at MA Muhammadiyah Tengnga Lembang Sinjay." This research aims to find out the discrimination of English summative tests for third-grade students at MA Muhammadiyah Tengnga Lembang in Sinjay Regency. The method of this research is quantitative descriptive. The instruments are the teacher made-test, student's answer sheet and answer key. The test consisted 40 numbers of the multiple-choice test. The finding of this research showed that discrimination power of the multiple-choice test had 7.5% (3 test items) in a good category, 37.5% (15 test items) in a satisfactory category, 27.5% (11 test items) in a poor category, and 27.5% (11 test items) in a very poor category.<sup>23</sup>

The researcher took a different place, time and population. The researcher took an English final term examination in the academic year 2020/2021 to analyze the items at MTs, while the other took a USBN test at SMP, mid-term test at SMP and summative test at SMA to analyze. The similarity of those studies above with the researcher's study is the analysis of test item on level difficulty and discrimination power which took only multiple choices question and quantitative method. As a result, all of those researches give a contribution to education research. Therefore, the researcher got inspiration from those previous researches to analyze the test items

---

<sup>22</sup> Eni Larianti, "The Analysis of Final School Exam Items on English Subject at SMPN 14 Seluma in Academic Year 2018/2019" (Skripsi, IAIN Bengkulu, 20099), 1-61.

<sup>23</sup> Hisbullah, "The Analysis of Discrimination Power," 1-65.

in level difficulty and discrimination power of English final term examination at seventh grades of MTs Al-Khairiyah Kamasan.

### **G. Thinking Framework**

To know the student's progress overall in a semester, the teacher should give the test to the student, which is called a summative test. The test is needed to measure the student's ability to capture the lesson which has been taught. It is also functioned to evaluate the learning process in every semester.

Exactly, before the students continue their lesson in the second semester, they have to do the final term examination in the first semester in all subjects, including English subjects. It is used to measure the student knowledge competencies in learning English material in one semester using an analysis test item.

The analysis of test items must be appropriate with their procedures: the difficulty of the item, the discrimination power of the item and the effectiveness of each alternative. By analyzing the test item, the teacher can find out the difficulty level in terms of difficult, moderate and easy on test items and find out how well the test can discriminate between the upper and lower groups. Through that thing, the teacher also can find out whether the item test is reused for next exam, discarded or changed.