

## CHAPTER I

### INTRODUCTION

#### ***A. Background***

One way for teachers to find out the extent to which students understand what they have learned is to provide an appropriate evaluation of the learning. Evaluation is an important component of the teaching-learning curriculum. A significant application of evaluation is to continuously monitor learning activities to provide feedback to students and teachers.<sup>1</sup>

Suharsimi argues on Dewi Manalu that evaluation could be a handle of collecting information to decide the degree in terms of what and how he instructive objectives have been accomplished.<sup>2</sup> Evaluation shows the level of achievement in a certain thing after studying it. In education, evaluation is very important after completing a lesson to find out the level of effectiveness of the strategy or method used by the teacher through the results of student evaluation.

Based on Sridadi's mind who is also on Dewi Manalu that there needs to be an assessments as a form of decision making on something and the type of education as a form of educational responsibility. One of the components that must be evaluated is student learning outcomes. Assessment of student learning

---

<sup>1</sup>Gyata Mehta and Varsha Mokhasi, 'Item Analysis of Multiple-choice Questions- An Assessment of the Assessment Tool', *Historical Aspects of Leech Therapy*, 4.7 (2014), p. 1-3, accessed on 23<sup>rd</sup> December 2020.

<sup>2</sup>Dewi Manalu, Kammer Tuahman Sipayung, and Febrika Dwi Lestari, 'An Analysis of Students Reading Final Examination By Using Item Analysis Program on Eleventh Grade of SMA Negeri 8 Medan', *JETAL: Journal of English Teaching & Applied Linguistic*, 1.1 (2019), p. 13, accessed on 23<sup>rd</sup> December 2020.

outcomes is intended as an effort to collect a variety of continuous and complete information about the process and learning outcomes that students have achieved through teaching and learning activities as a basis for determining the next steps.<sup>3</sup>

There are methods used by a teacher to find out the extent of the success of the learning process and the extent to which the level of understanding of students in understanding the subjects given by a teacher. Among these methods is the test form evaluation method. The word test is familiar to students at both lower and higher levels.

In English learning it is called a language test. The English test is a person's assessment where the focus of the assessment is English, either to see the perspective on English or to see the achievement or ability of someone in English.<sup>4</sup> The purpose of a language test is to determine a person's knowledge or ability in language and to differentiate that person's ability from the abilities of others. These abilities may be of a different type, achievement, skill or talent. The test, unlike a scale, consists of certain tasks by which language skills are acquired. In Tomlinson's view, the main purpose of language testing is to provide learning opportunities, both for students who are being tested, and for professionals who are carrying out exams.<sup>5</sup>

---

<sup>3</sup>Dewi Manalu, Kammer Tuahman Sipayung, and Febrika Dwi Lestari, 'An Analysis of Students Reading Final Examination By Using Item Analysis Program on Eleventh Grade of SMA Negeri 8 Medan', p. 13-14.

<sup>4</sup>Alderson J C, Clapham C, Wall D, *Language Test Construction and Evaluation* (Cambridge language teaching library, 1995), p. 14-15.

<sup>5</sup>Brian Tomlinson, 'Testing to Learn: A Personal View of Language Testing', *ELT Journal*, 59.1 (2005), p. 39-42.

Language test is a type of measurement that aims to test students' competence and achievement in listening, speaking, reading and writing. The Language test measures students' language competence and performance, not a mechanical structure test. Language tests are used to check what students can do with the language being studied. The students are tested not only in their language skills, but also in their language area.<sup>6</sup> Currently multiple-choice questions are a form of evaluation that is often or commonly used by teachers to their students. However, it is said that multiple-choice questions emphasize remembering factual information rather than conceptual understanding and conceptual interpretation.<sup>7</sup>

A teacher in giving tests to students must be in accordance with what students have learned and the teacher must analyze the tests that will be given to students to get the right questions to give to students. Analyzing the test referred to here is analyzing multiple-choice tests which are called item analysis. Item analysis is the act of analyzing student responses to individual exam questions with the intention of evaluating exam quality. It is an important tool to uphold test effectiveness and fairness.<sup>8</sup>

Item analysis is likely something teachers do both consciously and unconsciously on a regular basis. In fact, grading literally involves studying student responses and the pattern of student errors, whether to a particular

---

<sup>6</sup>Brian Tomlinson, 'Testing to Learn: A Personal View of Language Testing', p. 41.

<sup>7</sup>Gyata Mehta and Varsha Mokhasi, 'Item Analysis of Multiple-choice Questions- An Assessment of the Assessment Tool', p. 7.

<sup>8</sup>Cyril J Weir, *Language Testing and Validation An Evidence-based Approach* (England: Roehampton University, 2005), p. 8.

question or particular types of questions. But when the process is formalized, item analysis becomes a scientific method through which tests can be improved, and academic integrity upheld.

Constructing good objective test items requires a great deal of time and effort. Before the test items are used, it is necessary to identify which items are answered correctly by the more able students taking the test and badly by the less able students. The identifications of certain difficult items in the test, together with the knowledge of the performance of the individual distractors in multiple-choice items, can prove just as valuable in its implication for teaching as for testing.

All items should be examined from the point of view of their difficulty level and their level of discrimination as well as the distractors for multiple-choice test items.<sup>9</sup>

The essence of the Language Test is designed to measure the competence and actual performance of students in the language that has been taught. To do this step, the teacher must know the type of exam that will be given to them based on the knowledge that has been taught. There are four skills that must be included in the English Language Test such as listening, speaking, reading and

---

<sup>9</sup>Baso Jabu, *English Language Testing* (Makassar: UNM Publisher, 2008) p. 7.

writing. It also integrates with a grammar test and vocabulary test. Teachers have to do tests based on the syllabus.<sup>10</sup>

There are several potential errors that may result in inaccurate measurements. Based on these tests, some of the potential errors that lead to inaccurate measurements are invalid, unreliable, non-discriminatory. But, in this research, the researcher only focus on analyzing each item on multiple-choice questions. From the application of item analysis of multiple-choice questions, the teacher can see the good questions, the questions that must be fixed, the questions that must be eliminated.

Tests are very important in education, with tests we can find out the results of the achievement of student competencies in certain learning. Based on what often happens in schools that the tests given by the teacher are sometimes inversely proportional to the learning given. Sometimes students do well in class, but fail the test. Perhaps most people, especially teachers, think that students fail the test because the strategies, approaches, and methods used are not right with the learning they are given. Based on the researcher's view that the failure of these students could be because the tests given are not in accordance with what they were learning.

---

<sup>10</sup>Abdul Rehman, Aslam Ayesha, and Hassan Syed Hammad, 'Item Analysis of Multiple-choice Questions', *Pakistan Oral & Dental Journal*, 38.2 (2018), p. 93, accessed on 25<sup>th</sup> December 2020.

Therefore, based on the background above, the researcher wanted to see the responses and apply the item analysis to the English test to see the quality of the test made by the English teachers of SMP Negeri 4 Watampone.

### ***B. Problem Statements***

Based on the explanation in the background above, this research is conducted to answer the following questions:

1. How do the English teachers of SMP Negeri 4 Watampone apply item analysis to the English multiple-choice tests they made?
2. How are the results of analysis item of English multiple-choice tests made by English teachers of SMP Negeri 4 Watampone?

### ***C. Operational Definition***

The researcher provides some explanations of some key terms which appear in this research. The explanations of those key terms purpose are to make the readers easily understand about the meanings of those key terms.

To get comprehension and avoid misunderstanding, it is prudent to examine each part of this research. The title of this research is “Using Item Analysis as an Examining English Test Made by English Teacher of SMP Negeri 4 Watampone”.

#### **a. English Test**

English test is designed to measure the competence and actual performance of the learners in the English they have learned.<sup>11</sup> English Test is a test that the

---

<sup>11</sup>Baso Jabu, *English Language Testing*, p. 8.

focus of assessment is on the English language itself by paying attention to applicable rules in four skills (listening, speaking, reading, and writing), grammar, and structure.

b. Item Analysis

Based on Komaruddin's mind that analysis is the activity of thinking to describe a whole into components so that you can recognize the signs of the components, their relationship to each other and the functions of each in one integrated whole.<sup>12</sup>

Item Analysis is the examination of individual items on a test, instead of the test as an entire. Item analysis is valuable in making a difference test creators decide which items to keep, adjust, or dispose of on a given test, and how to finalize the score for a understudy.<sup>13</sup> It means item analysis is the activity to analyze the items in order to determine the quality of the items. By doing item analysis we can see the good items, the items that need to be fixed, and the items that need to be replaced.

Having understood the terms above, it can be known that the research discuss how to assess the quality of tests, especially English tests to get the right test for students by using item analysis program.

---

<sup>12</sup>Komaruddin, 'Definition of Analysis' (2001), *accessed on 25<sup>th</sup> December 2020*.

<sup>13</sup>Unknown, 'Item Analysis for Teachers', *State University of NewYork*, 2016, p. 2, *accessed on 28<sup>th</sup> December 2020*.

#### ***D. Objective and Significance of the Research***

##### 1. Objective of the Research

The following is objective of this research, explicitly:

- a. To find out how the application of the item analysis carried out by the English teacher of SMP Negeri 4 Watampone to the English multiple-choice tests they made.
- b. To find out the results of analysis item of English multiple-choice tests made by the English teacher of SMP Negeri 4 Watampone.

##### 2. Significance of the Research

###### a. Theoretical

The results of this research are expected to make a meaningful contribution in increasing understanding and become a guidance for future researchers conducting research on item analysis.

###### b. Practice

The results of this study are expected to provide understanding and motivation for the authors themselves, teachers in providing multiple-choice test to their students, and other researchers in conducting item analysis research.

#### ***E. Previous Related Research Findings***

Several researchers have previously attempted to find the results of applying the item of analysis to multiple-choice questions to obtain a good test.



From the writings and the results of the study, the researcher was inspired to reexamine the effect of doing item analysis on multiple-choice questions.

The first research was written by Dian Nurliyanto entitled "*Test Item Analysis of the Final Examination on Economics Subject in Grade XII IPA SMA Negeri Banyumas Academic Year 2014/2015*". This research is a descriptive quantitative research. The research subjects were 112 economics students of class XII IPS at SMA Negeri Banyumas. The data was collected using the documentation method. However, in this research, it was not only the analysis of multiple-choice or objective tests but also subjective tests that could be called difficult to measure. The results of this analysis indicated that there are more items that need to be repaired or removed than items that need to be retained. Among the bad items were due to being invalid, unreliable, and too difficult for students.<sup>14</sup>

The second research was written by Nana Erna, Baso Jabu, and Kisman Salija. This thesis focuses on the quality of multiple-choice tests made by English teachers. This research was conducted on 2 teachers and 40 questions in junior high schools. Mixed methods were used as the method of this research. Data collection was done through interviews and validation sheets that have been validated by experts. Quantitative data were analyzed using ANATES version 4.0 program and qualitative data were analyzed using validation sheets. The results

---

<sup>14</sup>Dian Nurliyanto, 'Test Item Analysis of the Final Examination on Economics Subject in Grade XII IPS SMA Negeri Banyumas Academic Year 2014/2015' (Yogyakarta State University, 2015), p. 60-64, accessed on 28<sup>th</sup> December 2020.

of the interviews indicated that the teacher's procedures for compiling the tests were divided into three main stages. They prepare test specifications, write, and analyze. Then the results of the analysis are based on the principle of writing tests using a validation sheet in the form of multiple-choice questions in accordance with the material, construction, and language. There are several questions that still need revision. Analysis findings based on analysis items such as the reliability obtained are not reliable. The test has a difficulty level in the categories too easy 15%, easy 35%, medium 42.5%, and difficult 7.5%. The discrimination index is only 40%, it can be categorized as good. Then, the switching power works 50%.<sup>15</sup>

The third research was written by Manalu, Sipayung, and Lestari entitled "*An Analysis of Students Reading Final Examination by Using Item Analysis Program on Eleventh Grade of SMA Negeri 8 Medan*". This research is qualitative-quantitative research. The subject of research are the grade XI of SMA Negeri 8 Medan. This research aims specifically at the process of analyzing multiple-choice questions in reading skills to improve the quality of the items. Twenty five multiple-choice questions were tried out to prove the quality of the questions. The item is one type of correct answer, has a bar and four options, one of which is correct and the other three are "bullies". Based on student responses, the items were analyzed using ANATES Software. Anates is a statistical

---

<sup>15</sup>Nana Erna, Baso Jabu, and Kisman Salija, 'An Analysis of English Teacher-Made Tests' (2000), p. 7, accessed on 29<sup>th</sup> December 2020.

calculation software that is very effective for calculating reliability, item difficulty, item discrimination strength, and level of interference. The result of this research is that item analysis has provided useful information about the characteristics of the items in one of the tests studied. Some items, once analyzed, can be revised, changed or even removed. Based on the above analysis, it is known that the test has high reliability with a coefficient of 0.90. Based on these findings, many items were categorized as marginal and poor in terms of difficulty level, discrimination power and level of interference.<sup>16</sup>

The similarity of the previous research above with the research of the researcher is that the research both applies or practices the item analysis on a test. The test is analyzed to see the quality of the test, so that we can find out which items are good, items that need to be repaired, and items that are replaced or removed. From some of the research above have several differences with the research that was carried out by researcher. Among others, that the research has a different version of applying the item of analysis to a test to determine the quality of the test, but it is basically the same. Then, from some of the studies above, there is that than analyze subjective test, while the research that was carried out by researcher is only to analyze the objective test, namely the English multiple-choice test made by the English teacher of SMP Negeri 4 Watampone.

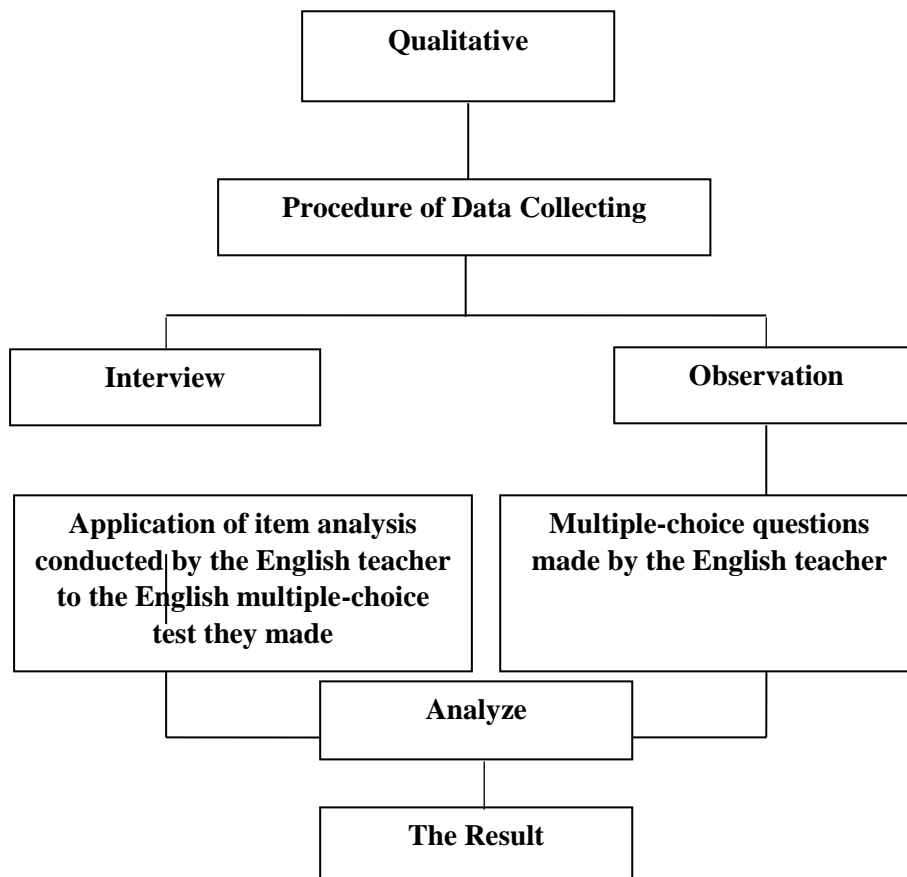
---

<sup>16</sup>Dewi Manalu, Kammer Tuahman Sipayung, and Febrika Dwi Lestari, 'An Analysis of Students Reading Final Examination By Using Item Analysis Program on Eleventh Grade of SMA Negeri 8 Medan', p. 17-18.

## F. *Conceptual Framework*

Based on the research objectives previously described, the conceptual framework proposed by the researcher as follows:

**Figure 1.1**



## G. *Methodology*

### 1. Research Design

Based on the conceptual framework in figure 1.1 above, it can be understood that this research is a qualitative research. In procedure of data collecting, researcher used interview and observation. The researcher gave a

short interview to find out how the application of the item analysis carried out by the English teacher to the English multiple-choice tests they made. Then, the researcher conducted a observation to observe the English multiple-choice test that has been made by the English teacher by item analysis to see the quality of the test.

## 2. Location and Population

### a. Location

This research is conducted at SMP Negeri 4 Watampone. It is located on Letijen D.I Panjaitan street, Bone Regency, South Sulawesi Province.

### b. Population

The population in this research is the English teachers of SMP Negeri 4 Watampone who teach in class VIII and the multiple-choice questions. In total, about 2 teachers and 20 items.

## 3. Data and Data Sources

Based on the title, the researcher uses a data source. There are two types of data sources. They are primary data and secondary data. Primary data in this study are teachers of SMP Negeri 4 Watampone and the English test questions they made. Secondary data of this research are all relevant things from this research is from some journals and books.

## 4. Instruments

The researcher uses two different instruments, they are:

a. Observation

Observation is making direct observations on the object being the target. The observation guide contains a list of the types of activities that may arise and will be observed. Observation guidance is needed especially if the researcher applies focused observation in the data collection process. In focused observation, the researcher focuses his attention only on some aspects of the behavior or phenomenon that is the object of the target.<sup>17</sup> This instrument is used to provide valuable background information about the environment where a research project is being undertaken. Observation guides and field notes describe, as accurately and as comprehensively as possible, all relevant aspects of the situation. They contain two basic types of information: (1) descriptive information about what the observer has directly seen or heard on-site through the course of the study and (2) reflective information that captures the researcher's personal reactions to observations, the researcher's experiences, and the researcher thoughts during the observation session.<sup>18</sup>

b. Interview

The interview is in the form of a structured interview, where the researcher has several short questions that get the same information from

---

<sup>17</sup>Muhammad Arifin and Khoirudin Asfani, *Instrumen Penelitian Kualitatif, Kuantitatif dan Pengembangan* (University of Malang, 2014), p. 13-15.

<sup>18</sup>L.R. Gay, et al., *Educational Research: Competencies for Analysis and Applications* (Tenth Edition; New Jersey: Pearson Education Inc., 2012), p. 382.

the respondent. The contents are questions about the teacher's experience in applying item analysis to the multiple-choice questions made.

## 5. Procedure of Data Collecting

This procedure of collecting data of the research followed the procedure as follows:

### a. Observation

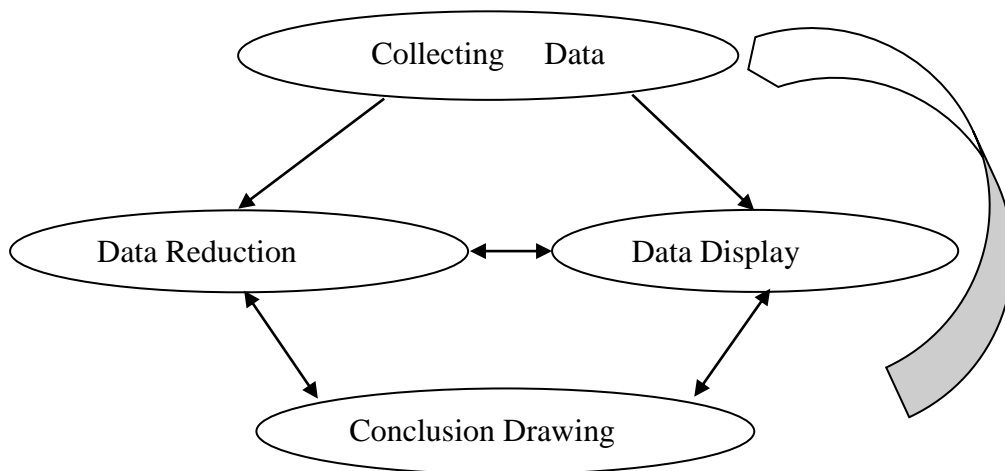
Observation is a technique that can be used when data cannot be collected in any other way, or that collected through other means has limited value or is difficult to validate. For example, in an interview, participants may be asked how they behaved in a certain situation, but there is no guarantee that they actually did what they said. Therefore, observing the situation is more valid. There are two types of observing human behavior. Namely participant observation and non-participant observation. Participant observation is an observer who takes part in the situation he observes while non-participant observation observes the situation, openly or secretly but does not participate. In this research, the researcher only used participant observation.

### b. Interview

This data collection procedure took one session of 10 minutes. The researcher met directly with the English teachers who make multiple-choice questions for analysis. The researcher had good conversations with teachers. Previously, the researcher would get acquainted with the

teacher. Furthermore, the researcher asked about the teacher's experience in making multiple-choice questions.

## 6. Technique of Data Analysis



**Figure 2.1 Data Analysis from Miles and Huberman in Keith**

The qualitative data analysis suits this research is from Miles and Huberman in Keith.<sup>19</sup> The approach is familiarly called ‘transcendental realism’ and the main components of their analysis are data reduction, data display, drawing and verifying conclusions interacting throughout the analysis as shown in the frame above. The first is the collecting data, all the data was collected to analyze. The second is the data was analyzed through data reduction. It occurred continually throughout the analysis as a main part and can not be separated from the analysis. This analysis happened throughout conceptualizing and explaining. The third is data display, data

---

<sup>19</sup>Keith F. Punch, *Introduction to Research Methods in Education* (Singapore: SAGE, 2009), p. 174.



display and data reduction are really related, because the data that has reduced was displayed through many different ways of displaying data graphs, charts, networks diagrams of different types and anyway that moves the analysis forward is appropriate. The last, the data concluded, this part it still relate with the data reduction and display data. This part may be not noted early in the analysis, but they may be vague a dill-formed at this stage. They are not finalized until all the data are in, and have been analysis.

The explanation of the main components is explained as follows:

a. Data Reduction

Data reduction occurs continually throughout the analysis. It is not something separate from the analysis; it is part of the analysis. In the early stages, it happens through editing, segmenting and summarizing the data. In the middle stages, it happens through coding and memorizing, and associated activities such as finding themes, clusters, and patterns. In the later stages, it happens through conceptualizing and explaining, since developing abstract concept is also a way of reducing the data. In both quantitative and qualitative analysis, the objective of data reduction is to reduce the data without significant loss of information. In qualitative analysis, an additional important component of not losing information is not to strip the data from their context.

b. Data Display

Data display organizes compress and assemble information. Because qualitative data are typically voluminous, bulky and dispersed, displays help a tall stage in the analysis. Miles and Huberman regard displays an essential, often using the phrase 'You know what you display'. They have no doubt that better displays are a major avenue to valid qualitative analysis. There are many different ways of displaying data-graphs, charts, networks diagrams of different types and anyway that moves the analysis forward is appropriate. Displays are uses data stages, since they enable data to be organized and summarize, they show what stage the analysis has reached and they are the basis for further analysis.

c. Drawing Conclusions

The reasons for reducing and displaying data are to assist in drawing conclusions. While drawing conclusions logically follows reduction and display of data, infect it takes place more or less concurrently with them. Thus possible conclusions may be noted early in the analysis, but they may be vague a dill-formed at this stage. They are held tentative pending further work and sharpened during it. They are not finalized until all the data are in, and have been analysed. Conclusions were in the form of propositions and once they have been drawn.