CHAPTER II

REVIEW OF RELATED LITERATURE

This chapter provides two sub-sections, first is the review of related literature that describes the basis theory of assessment which is used in this research. The second one is review of previous study that explains the differences of this research with other research which is done by another researcher.

A. Review of Related Literatures

1. Cognitive and Knowledge Assessment

Assessment is defined as the process of gathering information which is used to propose decision regarding the students, curriculum program, educational policy, and method or education instrument.¹ It is an integral part of learning, whereas it is used to check the students' understanding related to the material that has been given by the teacher. The indicator of the achievement of learning outcomes can be reflected from it.

In Indonesia, there are two curriculums which are applied by the government; there are KTSP (Kurikulum Tingkat SatuanPendidikan) and K-13 (Kurikulum 2013). In line with it, in this semester, in practice teaching 1 (PPL 1) the lecturer applies both of curriculums, since some of school still use KTSP as the curriculum and some of them have implemented K-13.² It is

¹Anthony J - Nitko, *Educational Assessment of Students*(Ohio: Merrill, an Imprit of Prentice Hall, 1996), 4.

²Kementerian Pendidikan dan Kebudayaan, *Surat Keputusan Penetapan Sekolah Pelaksana Uji Coba Kurikulum 2013 No 233/C/KR/2015* (Jakarta: Kementerian Pendidikan dan Kebudayaan, 2015)

aimed to let the student teachers experience both of it, so it can be their background knowledge in facing their internship program (PPL 2).

Curriculum regulates the process of teaching and learning process, including designing the lesson plans, developing the material, designing assessment and designing follow up program.³ Assessment as part of the curriculum plays role as the way to collect an important variable of teaching and learning process. It becomes the evidence of the result in teaching and learning process. The evidence is the achievement of students' competence which shows whether the objective of the lesson can be fulfilled or not. Related with the assessment design, in KTSP and K-13, the concept of the assessment is similar. In KTSP the domain or the competency which is assessed is cognitive domain, affective domain, and psychomotor.⁴ While in K-13 there are attitude competency, theologies competency, knowledge competency and skill competency.⁵ Both of the curriculums have the same concept of the domain,⁶ whereas cognitive domain is related with knowledge competency and psychomotor domain with skill competency. For attitude competency, actually, it is similar with affective domain. The difference is, in K-13, attitude competency is divided into two categories; theologies aspects and social aspects.

³Kunandar, *Penilaian Autentik* (Jakarta: Raja Grafindo Persada, 2014), 31.

⁴Sudaryono, *Dasar-dasar Evaluasi Pembelajaran* (Yogyakarta: Graha Ilmu, 2012), 43.

⁵Kunandar, *Penilaian Autentik...*, 103.

⁶Kunandar, *Penilaian Autentik...*,165.

In the case of assessing attainment of the students with the lesson, this research is focused on the assessment instrument of cognitive domain for KTSP and knowledge competency for K-13. In KTSP, cognitive is the step whereas the students understand something, and apply it in psychomotor domain.⁷ While in K-13, knowledge becomes the main competency (KI 3), whereas the process of transferring knowledge or material happens. Both of these domains have the same characteristics whereas it deals with how well the students reflect their knowledge that they have gotten during the lesson theoretically. There are 6 domains of knowledge competency⁸ and so does cognitive.⁹

- a. Knowledge: it is the stage whereas the students are able to remember the learning material in the previous meeting.
- a. Comprehension: the ability of grasping the meaning of the material.
- b. Application: the student's ability in applying general idea, theory, formula, principal and every material in a new and concrete situation.
- c. Analysis: the student's ability in explaining and elaborating detail part of the lesson and understanding the connection from part to part.
- d. Synthesis: it is the stage whereas students are able to fuse every part of the lesson logically and structure it all.

⁷Sudaryono, *Dasar-dasar Evaluasi Pembelajaran...*, 255.

⁸Kunandar, Penilaian Autentik...,168.

⁹Sudaryono, Dasar-dasar Evaluasi Pembelajaran..., 43.

e. Evaluation: the student's ability in considering idea, situation, value and method of guideline and the certain criteria.

So it can be seen, both of cognitive and knowledge domain has the same concept and indicator to assess the students. It is used to reflect the achievement of the students learning result that they got through teaching and learning process.

2. Cognitive and Knowledge Assessment Techniques

There are three techniques for assessing cognitive and knowledge assessment; written test, oral test and assignment.¹⁰ Written test can be in the form of multiple choice, descriptive test, short answer, true-false, matching and so on. While oral test is in the form of question lists, and assignment is in the form of homework which is done individually or group.

Furthermore, the subject that student teachers' teach is English. They teach based on four English skills; listening, speaking, reading, and writing. As stated in the background the researcher focused on reading. The only assessment techniques which suit with reading is written test. Written test as the technique for assessing reading has its own instrument which is regulated in Indonesian curriculum and in English assessment instrument. Thus, considering and combining both of instrument are possible to be done in case of getting reliable reading assessment.

¹⁰Kementerian Pendidikan dan Kebudayaan, *Model Penilaian Hasil Belajar Peserta Didik* (Jakarta: Kementerian Pendidikan dan Kebudayaan, 2013), 9.

3. Assessment Instrument

a. Assessment Instrument in Indonesian Curriculum

Teacher assesses the knowledge competency through the instrument that has been designed. The assessment instrument here is called test. It is a task or set of tasks that elicits observable behavior from the test taker.¹¹ Because test is a form of assessment, test also answers the question of how well the individual perform either in comparison with others or in comparison with a domain of performance task. Test is commonly used to upgrade the teaching and learning process. It is because the students' answer becomes the best sources for knowing the achievement of the teaching aim, the students' comprehension of the lesson and the effectiveness of the teaching process.

Both of cognitive and knowledge competency has its own test instrument. In cognitive the test is called objective test and subjective test,¹² while in knowledge competency it is called written test.¹³ The detail of the test as the assessment instrument can be seen below.

¹¹Ered Genesee- John A Upshur, *Classroom-Based Evaluation Second Language Education*, (New York: Cambridge University Press, 2007), 141.

¹²Hamzah B Uno - Satria Koni, Assessmen Pembelajaran (Jakarta: Bumi Aksara, 2013), 112.

¹³Kunandar, Penilaian Autentik ..., 125.

Table 2.1 Types of the Assessment instrument				
Cognitive Domain (KTSP)		Knowledge Competency (K-13)		
Objective Test	Subjective Test	Written Test		
Multiple choice	Short answer	Multiple choice		
True-false	Descriptive test	True-false		
Matching		Matching		
Fill in the blank		Fill in the blank		
		Short answer		
		Descriptive test		

Table 2.1 Types of the Assessment Instrument

As it is stated in table 2.1, the kinds of the assessment instrument technique from both of the cognitive domain and knowledge competence are similar. The difference is only on the name of its classification. To make it clear, the descriptions of each technique is described below.

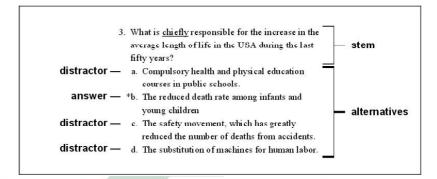
1) Multiple choice

Multiple choices are a form of assessment in which a question needs to be answered by selecting the best possible answers out of the choices from a list.¹⁴ Every multiple choice item consist of **stem** (the beginning of the item) and either three, four, or five **answer options** (with four options being the most common format), one and only one of the options are correct, and that is called the **key**, the incorrect options are called **distractors.**¹⁵ Here is the example.¹⁶

¹⁴Kunandar, Penilaian Autentik..., 120.

¹⁵Kathleen Bailey, Learning about Language Assessment (America: Heinle Publishers, 1998), 131.

¹⁶Steven J. Burton, et all, *How to Prepare Better Multiple-Choice Test Items: Guidelines for University Faculty* (England: Brigham Young University, 1991), 5.



The advantages of creating multiple choice tests is the amount of the question are expected to represent the entire competency that are going to be assessed. Besides, it will ease both of the teacher and the test-taker, for the teacher, they will be easily grading and analyzing the question.¹⁷ While for the test taker, it will ease them since they only need to respond by making a mark on the paper.¹⁸ Instead of the advantages, the difficulties of creating multiple choices are also existed. For example, guessing might be done by the test taker whereas it will impact on test scores.¹⁹ Since the answer choice is existed, the students tend to choose it randomly when they do not get the right answer. Moreover, it might facilitate the students to cheat. As a result it can be difficult to analyze the question and the result of the student's test.

Designing multiple choices as the test or task to measure the student's competence has to follow some certain criteria. It will assess

¹⁷Kunandar, *Penilaian Autentik* (Jakarta: Raja Grafindo Persada, 2014), 187.

¹⁸Arthur Hughes, *Testing for Language Teacher*(New york: Cambridge University Press, 2007), 76.

¹⁹Arthur Hughes, *Testing for Language Teacher...*, 76.

the assessment instrument whether it will be able to measure what is going to measure or not. There are three categories that need to be considered when designing multiple choice tests; there are material, construction and language.²⁰ Each category have its own indicator, the detail of the indicator is below.

- a) Material
 - (1) The questions are appropriate with the indicator.
 - (2) The limitation of the question and the answer are clear.
 - (3) The material is measured based on the competence.
 - (4) The question is suitable with the students' level or class level.
 - (5) The answer choices are logic, the distractor does not looktotally wrong.
- b) Construction
 - (1) The main question (stem) is proposed clearly.
 - (2) The question and the answer choices are formulated in brief and clear way.
 - (3) The questions do not show the key answer.
 - (4) The questions do not use double negative statement.
 - (5) If the negative statement is used, it must be in italic form or underline.
 - (6) The answer choices are homogenous.

²⁰ Sri Wahyuni, Assessmen Pembelajaran Bahasa (Bandung: Refika Adi Tama, 2012), 54-56.

- (7) The answer choices do not use the statement of "all of answer are correct" and "all of answer are wrong".
- (8) The length of the answer choice is relatively similar.
- (9) The answer choice in time and numeric form is formed based on the amount and chronological events.

(10) Picture, graph, table and diagram are clear and useful.

- (11) The questions do not depend on the previous questions.
- c) Language
 - (1) It use communicative language.
 - (2) The statement use good and correct English.
 - (3) The questions do not cause double interpretation or misunderstanding.
 - (4) Use common language, not local language or taboo.
 - (5) The question do not consist word that can insult others' feeling.

Each of the indicators above has to be fulfilled, so the test or task will be able to meet the objective of the test and measure the student's competence.

2) True-false

Items in which the test taker has merely to choose between yes and no, or between true and false, are effectively multiple choice items with only two options.²¹The true word will be symbolized by (T) and (F) for false. Here is the example.

Surabaya is the capital city of east javaT/FAlbert Einstein is the discoverer of lampT/F

The students' duty is to decide whether the statement is true or false. The characteristic of this test is, it is easily to arrange and grade. In contrast, for the students, it might facilitate them in guessing the answer, so it will be difficult to know their understanding of the material through that test. To overcome it, sometimes this test is modified by adding the reason of the students in stating the true and false of the statement.

Using true and false as the technique for testing the students is not only about creating the questions and asking the students to answer it. Nevertheless, the test design has to meet the requirement of standardization test. There are 12 criteria that need to be aware when designing this technique.²² Here are the criteria.

- a) The questions are appropriate with the indicator.
- b) The questions do not reflect the answer.
- c) The question is in positive sentences, if it use negative sentence, it is not more than two sentences.

²¹Kunandar, *Penilaian Autentik...*, 79.

²²Kunandar, Penilaian Autentik..., 206-207.

- d) The question is not similar with the book sentences.
- e) Do not use the question that use to distract the students' answer
- f) Do not use the words that are able to disturb the answer, (usually, never, no one, only will disturb true answer), (actually, maybe and sometimes will disturb the false answer).
- g) Use brief sentences for the question.
- h) Use recognizable word for the students.
- i) The amount of true answer and the false answer are balanced.
- j) Do not use ambiguous statement "some of them are correct and some of them are false".
- k) Do not use qualitative or quantitative language which is easily to guess.
- 1) The arrangement of true and false statement is arranged randomly.

By paying attention on the criteria or indicator above, it will help the teacher to meet the need of the assessment objectives, so the assessment will be able to assess the students' ability well.

3) Matching test

Matching test is written test which consist of question and answer category. The category group is written as the questions and the left side will be the answer category.²³ The question can be in the

²³Sudaryono, Dasar-dasar Evaluasi Pembelajaran..., 115.

form of interrogative sentence, or statement. Then, the answer will be in the form of word, number, picture or symbol. Here is the example.

A. Narrative	1. To tell the past experience
B. Procedure	2. To tell how to make something
C. Recount	3. To amuse the reader

Here the students are intended to match the word or the statement in right group to the statement in left group. Dealing with the use of matching test, it can be useful for assessing student's competence in simple ways, like identifying information based on simple relationship, or identifying the correlation between two things.

In addition, when a teacher set a test or task for their students, it is not about deigning the question and answer. They must pay attention on their instrument whether it is appropriate or not for their students. In this case they have to follow the indicator of the assessment technique that they use. In matching technique, there are nine categories or indicator which needs to be aware.²⁴ Here they are.

- a) The questions are appropriate with the indicator.
- b) The problem in the question is in the same genre, for example; time with time, place with place (connect each other).
- c) The words in the question are not more than 15 words.
- d) The words of the answer are more than the words in the question.

²⁴Kunandar, Penilaian Autentik..., 208.

- e) The question is proposed in the brief sentences.
- f) The question is placed on the left with number, while the answer is placed on the right with alphabet.
- g) The question and answer are arranged in systematic, if it is used number, it is arranged from the biggest number, and then if it is a word, it is arranged based on alphabet.
- h) The statement is written in a page.
- i) The score is calculated from the right answer.

The indicator above will help the teacher in designing the right test, so the test objectives are able to be fulfilled.

4) Completion test (fill in the blank)

Completion test requires the student to answer a question or to finish an incomplete statement by filling in a blank with the correct word or phrase.²⁵ Here is the example.

Declaration of independence was the first to be _____ by the United States²⁶

From that example, it can be seen that the answer might be based on the reading that the students had been read, or it based on the student's knowledge. The limited context provided by the sentence

²⁵Nuhad Yazbik Dumit, *Writing Effective Questions: True & False, Matching, & Completion Test Items* (USA: American University of Beirut, 2011), 29.

²⁶Principles of Test Creation A Self-Instructional Handbook for Byu Educators (Hawai: Brigham Young University Press, 2003), 19.

allows the teacher to focus the item on the one thing she wishes to assess.²⁷ Then related to the implementation of fill in the blank test, it is essential that the statement is clearly stated and students have already told very clearly and firmly that only one word can be the answer and put it in the blank.

Using fill in the blank as the test technique might look simple, since the teacher only need to remove some of words from the text that has been given to the students. In contrast, it is not about remove the words from the text, there are some indicator that need to be aware related with the design of this technique. There are nine categories or indicators that will lead this technique to meet the objectives of the assessment.²⁸

- a) The questions are appropriate with the indicator.
- b) The question is not taken directly from book.
- c) The question is clear, it does not cause ambiguity.
- d) Use the words that do not have important meaning regarding to the answer.
- e) The answer consists of one or two words.
- f) Put the blank part of the answer in the middle of the question.

²⁷Kathleen M.Bailey, *Learning about Language Assessment* (America: Heinle & Heinle Publishers, 1998), 59.

²⁸Kunandar, PenilaianAutentik..., 209.

- g) The length of the full stop for the blank answer is similar from one question to another.
- h) There is possible key answer under the question list.
- i) The score of the answer is counted based on the amount of the blank.
- 5) Short answer

In this test, the teacher gives the question to the students which need a short answer. It is commonly used in listening and reading test. Moreover, it is also suitable in assessing knowledge aspects especially in memory domain. Here is the example.²⁹

What does "it" in the last sentence refer to?

How old was Hannibal when he started eating human being? Here, the focus of the question is the word **it** in the sentence, so there will be one possible answer. Instead of it, it might be difficult for the students who have difficulties in using the language, since they need to produce language in order to respond. Then for the teacher by having short answer test, it will minimize the possibility of guessing and cheating. So the grading will show the real result of the students' understanding.

As stated above, using short answer as the technique for assessing the students might help the teacher to minimize the

²⁹ Arthur Hughes, *Testing for Language Teacher* (New York: Cambridge University Press, 2007), 79.

possibility of cheating, since the students need to answer it by producing the language. However, designing the question for short answer technique might become a challenge for the teacher. They have to pay attention whether the questions have already met the assessment objectives or not. Therefore, they need to fulfill the indicator of short answer technique. There are six indicators that need to be aware when using short answer techniques.³⁰

- a) The questions are appropriate with the indicator.
- b) Use interrogative sentence.
- c) The answer consists of one or two words.
- d) The answer sheet is separated from the question list.
- e) The question is not similar with the book sentences.
- f) There is only one correct answer.
- 6) Descriptive test

Descriptive test is an assessment instrument which demands the students in remembering, understanding, and organizing idea that have been learned by stating it with their own words. Descriptive test is divided into two categories:

a) Open-ended test

The question is given based on main problem which is discussed without giving any answer instruction. For example

³⁰Kunandar, Penilaian Autentik ..., 205.

"Why does Indonesia support Palestine to achieve their independence?"

b) Closed test

The question is directed into certain problem, so the answer of the students will be based on the question claim. For example. "*mention three reasons of Michael living his family*!"

In designing descriptive question, the teachers have to follow the indicator of each categories in this technique; material, construction and language.³¹ The indicator will assess the test before it is used to test the students. It will analyze whether the test is able to measure that the teacher is going to measure, so that the test objectives can be fulfilled. The categories and indicator is divided as follows.

(1) Material

- (a) The questions are appropriate with the indicator.
- (b) The limitation of question and answer is clear.
- (c) The contain of the material is suitable with the assessment purpose.
- (d) It is suitable with the student's level or class level.
- (2) Construction
 - (a) The question is in the form of interrogative sentence or command which demands descriptive answer.

³¹ Sri Wahyuni, Assessmen Pembelajaran Bahasa ..., 53-54.

- (b) There is clear instruction about how to answer the question.
- (c) The question does not depend on the previous question.

(3) Language

- (a) Use communicative language.
- (b) Use the language which is appropriate in English form.
- (c) The question forms do not cause double interpretation or misunderstanding.
- (d) Use common language (English form) not local language or taboo.
- (e) The question do not consist word that can insult others feeling.

b. English Assessment Instrument

1) Assessment and testing

Assessment is ongoing process which cover wider domain.³² Whenever a teacher asks their students, offer their comment, or try a new word together, at that moment, the teacher assesses their students' performance. When the teacher tries to assess the students' performance, it is not always incidental or intended. They might assess it while teach them.

³²Douglas Brown, *Language Assessment* (San Fransisco: Longman, 2003), 4.

Testing is one of form of assessment.³³It helps the teacher to discover information related with the students' achievement toward the lesson. Moreover, it can be the way for a teacher to analyze students' strength and weakness point.Test is designed based on its purpose, especially test which is used to test various abilities; reading, listening, speaking and writing. Each of those abilities has its own test techniques. Since the researcher focuse on reading, so the further explanation of reading test techniques will be described below.

2) Testing Reading Skill

Teaching is the primary activity in teaching and learning process, while testing will inform the result of the teaching process. It will be needed to provide information about the student's achievement.³⁴ There will be some techniques to test the students learning result. The difference of each technique is caused by the purpose of the test. In case of assessing reading, there will be some of certain techniques which are able to ease the teacher to test their students. The technique below is proposed from Arthur Hughes' book entitled testing for language teachers and J Charles Aldersons' book entitled assessing reading.

³³Arthur Hughes, *Testing for Language Teacher* ..., 5.

³⁴Arthur Hughes, *Testing for Language Teacher* ..., 4.

a) Multiple choice

Multiple choice is a common techniques to test students' comprehension toward the text. Students will answer multiple choice questions by eliminating the distractor by their logical analysis that they have gotten by reading the text³⁵. For example:

Memorizing is easier when the material to be learned is

a) In a foreign language

b) Already partly known

- c) Unfamiliar but easy
- d) Of no special interest.

Here, they students will chose an answer based on the choices above.

b) Matching technique

Here, two sets of question and answer have to be matched against each other,³⁶ for example matching the vocabulary and its meaning, matching the heading of the paragraph to their corresponding paragraph. For example:

- A. Narrative
- e 1. To tell the past experience
- B. Procedure 2. To tell how to make something
- C. Recount
- 3. To amuse the reader

 ³⁵Charles Alderson, *Assessing Reading* (New York: Cambridge University Press, 2000), 211.
 ³⁶Charles Alderson, *Assessing Reading*..., 215.

They will be intended to match the word on the left to the meaning or explanation in the right.

c) Ordering tasks

In ordering tasks students are given scrambled set of words, sentences, paragraph or texts and have to put them into their correct order.³⁷ For example:

order i of enamples	
A it was called the last waltz	1 <u>D</u>
B the street was in total darkness	2
C because it was one he and Richard had learnt at school	3
D Peter lo <mark>ok</mark> ed outside	4
E he recognized the tune	5
F and it seemed deserted	6
G he thought he heard someone whistling	7

d) Dichotomous items

This technique is called true and false techniques whereas the students are presented with a statement which is related with the text and decide whether the statement is true or false.³⁸ This test is intended to test the ability of infer meaning. For example:

- 1. John is the youngest in his family T/F
- 2. John sister always comes home late T/F

³⁷Charles Alderson, Assessing Reading..., 219.

³⁸Charles Alderson, Assessing Reading..., 222.

Here the students will decide whether the statement above is true or false based on the text that they have read.

e) Short answer test

In this technique the students are asked a question which requires a brief response. The short answer technique will work well to test the ability to identify referents, predict the meaning of unknown words from context, and write items related to the structure of text.³⁹ For example:

What does the word 'it' (line 26) refers to? _____

Which town listed in table 4 has the largest population?

f) Gap filling

This technique is particularly useful in testing reading. It can be used to get the students response. Students read the text and in the same time they read the summary of the text whereas some of the keywords have been removed. Their task is to restore or fill the missing keyword; consequently they have to understand the main ideas of the original text to find the right keywords.⁴⁰ For example:41

To support his claim that the mafia is taking over Russia, the author points out that the sale of _____ in

 ³⁹Arthur Hughes, *Testing for Language Teacher...*, 145.
 ⁴⁰Charles Alderson, *Assessing Reading...*, 240.

⁴¹Arthur Hughes, *Testing for Language Teacher*..., 149.

Moscow has increased by _____ per cent over the last two years.

c. English Assessment Instrument in Indonesian Curriculum

Dealing with English assessment instrument, there are two kinds of way to assess it. From Indonesian curriculum technique or from the original techniques of assessing reading based on Arthur and Charles books, both of the technique in Indonesian curriculum or English assessment is similar, it can be seen from the table below.

Tuble 2.2 Teeningu	
Assessment Technique in	English Assessment Instrument
Indonesian Curriculum	Based on Arthur Hughes and J
	Charles Alderson
Multiple choice	Multiple choice
True-false	Dichotomous Items(T/F)
Matching	Matching Technique
Fill in the blank	Gap filling
Short answer	Short answer
Descriptive test	Ordering tasks

 Table 2.2 Technique in Assessing Reading

Based on the chart 2.2, technique in assessing reading is similar both in Indonesian curriculum and in English assessment instrument. It has the same types and characteristic. The important point is both of the techniques are used to assess the students' knowledge, since there is no best method for testing something. It depends on the purpose or objective of the teaching and learning.

4. Testing the Assessment Instrument

Test is not only about testing the students but also it is about how the test can measure what is going to measure. It is essential since testing is the only way to monitor and systematically rank the students. In measuring and evaluating the test, it has to consider on its validity and reliability. In this case, reliability will be omitted, since it can be measured if the test is already used to examine the student's competence. While in this research, the researcher takes the assessment instrument which is designed by student-teachers during their practical teaching, whereas the test is used to examine other student-teachers. It means that the students are not the real students; therefore the reliability cannot be measured. As a result, validity will be the focus to measure and evaluate the test.

According to Scarvia B Anderson in his book "Encyclopedia of Educational Evaluation" stated that "a test is valid it measures what it purpose to measure". That is why the validity will be the concern before the test is conducted. It is used as the quality control of the test. In addition, there are two areas which need to be considered when discussing the validity of the test.⁴²

a. Consider how closely the test performance resembles the performance we expect outside the test.

⁴²Owen, *Testing* (Birmingham: The Center For Language Studies, 1997), 13.

 b. Consider to what extent evidence of knowledge about the language can be taken as evidence of proficiency.

In this research, validity is classified into two criteria. First is content validity, this validity is defined as any attempt to show that the content of the test is the representative sample from the domain which needs to be tested.⁴³ It is related with the compatibility of the test use. For example the teacher design reading test for their students, here the content validity can be seen from the content of the test; if it is appropriate with the level of the student then, this test has its content validity. In addition, content validity shows how far the assessment instruments are able to represent the sample behavior which is used as the teaching purposes.⁴⁴ It is related with the material of the test, if the material in the test has been taught by the teacher, it means that the material in charge with the content validity. Moreover, content validity is also known as curriculum validity whereas it will be related with the syllabus and lesson plan. In this case the other professional teacher is involved in assessing the content validity.

The next one is construct validity; this validity is the most possible one to be done by the teacher itself before conducting the test. It is concerned with the level of accuracy a construct within a test is believed to

⁴³Glen Flucher - Fred Davidson, *Language Testing and Assessment* (USA: Routledhe, 2007). 4.

⁴⁴Sudaryono, *Dasar-dasar Evaluasi Pembelajaran...*, 138.

measure.⁴⁵Construct validity can be fulfilled if the test are able to measure what's going to measure, so it will be related with the correlation between the indicator and the purpose of the study. The requirement of proposing a good test can be the requirement to test the test, the instrument of the test is required to reflect the requirement of the standardization. It can be seen from the material, construction and language criteria. Then, between the content validity and construct validity is related to each other.

B. Review of Previous Study

Here, the researcher reviews some researchers which were related to this research, as follows:

It was a research which was done in 2009 conducted by IffahMursyidahMayangsari, entitled "an analysis of UAS English Test of Second Semester 2008/2009 by Teacher Made English Test at SMA 2 MuhammadiyahSidoarjo."⁴⁶ In this research, Iffah analyzed test by analyzing its validity, reliability, and its index of difficulty and discrimination. Descriptive qualitative is implemented as the methodology to answer the problem and she found that the teacher made test which is designed by SMA Muhammadiyah 2 Sidoarjo is good since it fulfills the characteristics of a good test.

⁴⁵Douglas Brown, *Language Testing: A Critical Survey and Practical Guide* (San Fransisco:Prentice Hall, 1994), 256.

⁴⁶Iffah Mursyidah Mayangsari, Thesis: An Analysis of UAS English test of second semester 2008/2009 by teacher made English test in SMA 2 Muhammadiyah Sidoarjo. (Surabaya: Perpustakaan IAIN Sunan Ampel Surabaya. 2010)

The other similar study was conducted by AbidatulKhoiro in 2012, entitled "An Analysis of Teacher-Made English Try out Test form of National Examination 2010-2011 for the Third Graders of MAN Sidoarjo."⁴⁷ The object of the study is similar with the research which is done by Iffah. Here Abidatul did the research dealing with the teacher made test and the problem that she was going to find is also similar, it was about validity, index of difficulty and index of discrimination. By using descriptive method as her methodology, she found that there are several indicators which were not able to be covered by the test. In conclusion, although the result of the student test was good but the test has low quality as a test.

Still in 2012, another similar research was done by ItaFaradillah, entitled "An Analysis of Essay Test on English Final Test for Grade Eleven Students of SMAN 1 Lamongan."⁴⁸ Such as another previous research which was done by Abidatul and Iffah, this research also uses qualitative research to find the content validity, index difficulty, and index discrimination of the final test. As the result of the research, the researcher found that the content validity is in good level whereas 80% items represent all materials; Both the Index of difficulty and the Index discrimination is acceptable.

⁴⁷Abidatul Khoiro, Thesis: An Analysis of Teacher-Made English Try out Test form of National Examination 2010-2011 for the Third Graders of MAN Sidoarjo.(Surabaya: Perpustakaan IAIN Sunan Ampel Surabaya. 2012)

⁴⁸ Ita Faradillah, Thesis: An Analysis of Essay Test on English Final Test for Grade Eleven Students of SMAN 1 Lamongan. (Surabaya: Perpustakaan IAIN Sunan Ampel Surabaya. 2012)

In short, three of the previous study above were commonly focus on analyzing the assessment which had been used to assess the. It analyzed the content validity, index difficulty, and index discrimination of the test. While in this research the researcher focusing on the assessment instruments which were not used to test the students and assessed it based on test standardization. Therefore by assessing or analyzingthe test before it is used to test the students, the test is expected to measure what it is going to measure.

Another similar study related with reading test was done by SukmaliaYashintaNovansa in 2012, entitled "The Implementation of Reading Aloud Test in Reading Comprehension on the Grade X1 Class Students of SMA Negeri 3 Lumajang."⁴⁹ Quasi experimental is used as the methodology. In this research Sukmalia found that the student response of reading aloud test in reading text materials was better when using reading aloud; it is proven with the comparison of the mean score. Aloud" pretest 63, 0625 and posttest 79,375, usual 63, 75 post 67, 5. Here sukmalia focused on the implementation of testing reading using reading aloud test, while in this research the researcher focuses on types of reading test technique in written form; multiple choice, descriptive text and others.

⁴⁹ Sukmalia Yashinta Novansa. Thesis: *The Implementation of Reading Aloud Test in Reading Comprehension on the Grade X1 Class Students of SMA Negeri 3 Lumajang*. (Surabaya: Perpustakaan IAIN Sunan Ampel Surabaya. 2012)

The newest research was done by ArisBahariRizki in 2014 with the research entitled "A Study of Formative Assessment under the 2013 Curriculum at SMP WachidHasyim 7 Surabaya".⁵⁰ By using descriptive qualitative research, here, Aris tries to find the challenges of formative assessments under the 2013 curriculum and the implementation of it. He found that the implementation of formative assessment through self and peer assessment are well implemented. Here the subject of the research is the real teacher.

Seeing from the study that has been conducted before, the researcher concludes that all the previous studies have the similarity and different areas of study. Those previous studies could be the foundation of conducting this research. The previous study mostly focusing on the implementation of the assessment technique and the assessment which was designed by the real teacher, the content validity, index discrimination and index of difficulty, while in this research, the researcher focused on the quality of the test based on the standardized test and the causing factor.

In addition, some of the setting of the research above is school and the real teacher as the subject, the real teacher might be understood well of how to design the good test. Consequently, the significance of the research was not really impacted. While in this research, the subject of the research was the teacher candidate, whereas they need to know of how to design the good test as their

⁵⁰Aris Bahari Rizki. Thesis: A Study of Formative Assessment Under the 2013 Curriculum at SMP Wachid Hasyim 7 Surabaya. (Surabaya: Perpustakaan UIN Sunan Ampel Surabaya. 2014)

experience for being the real teacher someday, so the significance of the research could have more impact.

Moreover, related to the object of the research, some of the object of the research above is the assessment that the format is summative assessment or assessment which is done in the end of semester, while in this research, the researcher focused on the daily assessment instrument or known as formative assessment, whereas it is going to deal with the student every day to know their real progress, not just the latest result of their learning process.

