

REVIEW OF RELATED LITERATURE

A. Review of Related Literatures

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 Satuan Pendidikan) 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

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

- 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, *Assesmen Pembelajaran* (Jakarta: Bumi Aksara, 2013), 112.

¹³Kunandar, *Penilaian Autentik ...*, 125.

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 look totally 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.

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 java T/F

Albert Einstein is the discoverer of lamp T/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.

- The questions are appropriate with the indicator.
- The questions do not reflect the answer.
- 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.
- l) 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.

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, *Penilaian Autentik...*, 209.

- (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.

2) Testing Reading Skill

³³ Arthur Hughes, *Testing for Language Teacher* ..., 5.
³⁴ Arthur Hughes, *Testing for Language Teacher* ..., 4.

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

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 _____

³⁹ Arthur Hughes, *Testing for Language Teacher...*, 145.

⁴⁰Charles Alderson, *Assessing Reading...*, 240.

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

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.

Table 2.2 Technique in Assessing Reading

Assessment Technique in Indonesian Curriculum	English Assessment Instrument 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

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.

- 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: Routledge, 2007). 4.

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

measure.⁴⁵ Construct 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.

⁴⁵ Douglas Brown, *Language Testing: A Critical Survey and Practical Guide* (San Francisco: 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 analyzing the 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 XI 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

⁵⁰ Aris 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.