

Cholilulloh, HerlinaSofi. 2016.Item Test Analysis on Short-Answerin Narrative Reading Assessment atthe Tenth Grade Students of SMAN 1 Jenangan in Academic Year 2015/2016.Advisor: WiwinWidyawati, M.Hum.

Key Words:Short answer, narrative, reading assessment.

One of the complex problems in English classroom is students' difficulty to understandthe instruction and answer the questions about the text.It makes the students' achievement bad. Especially in English reading narrative text. Only 20% of 21 students can fill the KKM. The teacher must create the most effective test item to assess students' reading skill especially in narrative text. One of them is short answer test because it help students to answer question easier.

There were two statements of the problem;How is item test validity on short answerin narrative reading assessment at the tenth grade students of SMAN 1 Jenangan in academic year 2015/2016 and How is item test reliability on short answer in narrative reading assessment at the tenth grade students of SMAN 1 Jenangan in academic year 2015/2016. The significance wereto give contribution of knowledge to develop and explore about short-answer test, to get more knowledge about education especially on students narrative reading assessment andto make students more autonomous in learning reading. Particularly students at the tenth grade of SMAN 1 Jenangan in academic year 2015/2016.

This research used descriptive quantitative research design. There were 89 populations and 21 samples from tenth grade students of SMAN 1 Jenangan in academic year 2015/ 2016. The data of the researchwere collected by documentation and test. The researcher used SPSS 19.00 version for windows 8 to analyze the data.

The result showed that item test validity of short answer in narrative reading assessment at the tenth grade students of SMAN 1 Jenangan in academic year 2015/2016 was good. It was drawn from the validity result analysis with $r_{table}=0,433$ (df=19 and the significant standard of 5%). 90% from 20 questions $> r_{table}$. Then, 10% from 20 questions $< r_{table}$. Item test reliability of this test wasclassified as high reliability because the alpha scale showed reliability of the test with 0,883.This test was effective test to be tested for the students. It could be said like that, because based on the result of calculation the test has good criteria of validity and reliability.

CHAPTER 1

INTRODUCTION

A. Background of the Study

Language is part of our essential human nature and was therefore neither invented nor handed down as a gift.¹ Human gets the language not only handed down as a gift from ancestors. Language is most important in communication and used as a tool of communication among the nations in all over the world.

Communication is the responsibility of both the speaker and the listener.² Communication involves both the speaker or sender and the listener or receiver. English is used to communicate, to strengthen and to fasten relationship among all countries in the world in the whole field such as: in tourism, business, science and technology, etc. Almost all countries in the world use English in international Communication. Therefore, explanation of English as a lingua franca according to Jeremy Harmer is English can be defined as a language which widely adopted for communication between two speakers whose native languages are different from each other's and where one of both speakers are using it as a "second" language although English is not the language with the largest number of native or "first" language

¹Victoria Fromkin, *An Introduction to Language* (Canada: Dryden Press, 1974), 26.

²Penny Ur, *A Course in Language Teaching* (New York: Cambridge University Press, 1996), 120.

speakers³. There are four skills on English language. They are speaking, listening, reading and writing.

Reading is fundamental skill upon which all formal education depends and about understanding written texts.⁴ Reading is a process of getting information through written text. The process involves both the writer and the reader. So for reading, the reader needs to comprehend the ideas that are conveyed by the writer, and the way to reach it is by mastering reading process.

It is important to assess students' reading ability. One of the purposes of assessing is to know students' reading ability and to measure the level of students' success in learning. Assessing the student task on reading must be suitable with objective of the material to get the accurate result.

Reading assessment is to determine student baseline performance. Students enter the classroom with different backgrounds and skills on learning. Some students enter the classroom with special needs that require review of basic skills in reading, while other students have mastered the content and teacher intends to cover. These various student levels, it is necessary to design learning instruction to meet the individual needs of each student. Individual needs can be determined by initial and ongoing reading

³ Jeremy Harmer, *The Practice of English Language Teaching* (United States: Longman, 1998), 1.

⁴ Elyzabeth S. Pang, *Teaching Reading* (Chicago: Education Practice series University of Illuonis ,2003), 6.

assessments. This assessment provides teachers with the information needed to develop appropriate lessons and improve instruction for all students.

On assessing reading, there are varieties of measures that can be used to gather data for each area of early reading. Assessment is a central element for any teacher and should be implemented regularly. Through its implementation, teachers will be able to help students access the skills and students' need from the general education curriculum. This will allow all students to achieve the highest student potential.

There are many types of reading assessments. The first involves a student reading a passage that is at an appropriate level for the student, and then having the student answer factual questions about the text. The second type involves a student answering inferential questions about implied information in the text. The third type involves a student filling in missing words from a passage. The fourth type is to have a student retell using students own words.

Reading assessments is a useful part of a reading program, as long as the aim is to help students become fluent readers who understand what they read. Reading assessments have the benefit of placing students at the appropriate reading level. So, the students do not grow frustrated with lessons that are too difficult, or bored by lessons that are too easy. Reading assessments done later along the process of learning can be used to aid

comprehension. Assessment helps the students identify the strengths, as well as weaknesses.

The most effective reading assessments help to assess what students do understand about the text, as well as what students read. This is all done with the aim of a student becoming a fluent reader. Fluent reader is not just about being able to pronounce printed words on a page. It is about understanding the content and context of those words. Teacher want students to be able read fluently, which is not just about how fast a student reads, but how students are able to show how much understand the text.

Teacher needs a test to assess the students' reading ability. There are some kinds of item test that can be used to assess students reading skill like multiple choice test, matching test, true false test, close test, and short answer test.

Short answer test is a kind of the test familiarly used because the question is easy to make and possible to interpret students' responses to see if they have really understood. According to Douglass Brown Short answer is a kind of the test can be used to assess students reading skill.⁵ Short-answer task is a reading passage is presented, and the test-taker reads questions that must be answered in a sentence or two.⁶ Short answer test doesn't need long sentences to answer the question. Short answer questions are typically

⁵ibid, 207.

⁶ Douglass Brown, *Language Assessment Principles in Classroom Practices* (San Francisco: San Francisco State University, 2000), 207.

composed of a brief prompt that demands a written answer that varies in length from one or two words to a few sentences. However, the teacher can see the reading understanding from the answer given by the students.

Short-answer questions generally require the students to remember and reproduce knowledge. Students may also be required to demonstrate the understanding of knowledge in questions. Understanding what type of response the teacher wants requires that students read and analyze the question carefully. To analyze and interpret the question accurately requires a good knowledge of the meaning of direction or question words.

In a short answer question, the student types in a word, phrase and short simple sentence to respond a question. But it must match one of acceptable answer exactly. It is a good idea to keep the required answer as short as possible to avoid missing a correct answer.

Short answer questions are quite similar to essay questions but the answer is supposed to be short. Short answer questions are easy to create, but the questions demand that students think through the possible outcomes of the question. Students have to phrase the question in such a way that the answer can be short and concise.

The question can be addressed towards the testing of a specific objective. Short answer questions have the advantage of avoiding cueing and requiring students to supply an answer, rather than selecting or guessing from options supplied. Answers are easier to mark. The form of the questions is

familiar to students. So, the students do not feel difficult in answering questions.

Short answer questions make students more flexible to explain students' understanding and demonstrate creativity. In addition, the questions used to test students understanding about the concepts of the text. Short answer test is easier to grade than essay. It helps student to the focus answer.

The researcher found that there are some problems at the school. The first is students can't answer the question well because students difficult to understand the instruction at the question. The second is students' achievement is bad. The third is students only have limited English vocabulary word. The fourth is students difficult to read and understand English text. The fifth is teacher only use the same item test technique to assess students reading.

One of the complex problems above is students were difficult to understand the instruction and answer the questions about the text. It makes the students' achievement bad. Especially in English reading narrative text. Only 20% of 21 students can filled the KKM. The teacher must create the most effective test item to assess students' reading skill especially in narrative text. One of them is short answer test because it helps students to answer question easier.

Based on explanation above, the writer intends to know the effectiveness of short-answer test in students' reading assessment. Therefore,

the researcher conduct research on SMAN 1 Jenangan with the title “**ITEM TEST ANALYSIS ON SHORT-ANSWER IN NARRATIVE READING ASSESSMENT AT THE TENTH GRADE STUDENTS OF SMAN 1 JENANAGAN IN ACADEMIC YEAR 2015/2016**”

B. Limitation of the Problem

Problem faced by the tenth grade students of SMAN 1 Jenangan in reading are very broad and complex to deal with in this study. This study was limited to the assessment of reading narrative text by using short answer test of the tenth grade students of SMAN 1 Jenangan in academic year 2015/2016. For the present study the researcher only investigates on examining reading ability of the tenth grade students of SMAN 1 Jenangan. The researcher limits only in reading narrative text which assessed by using short answer test.

C. Statements of the Problem

Based on the background stated above, the researcher formulates the question as follow:

1. How is item test validity on short answer in narrative reading assessment at the tenth grade students of SMAN 1 Jenangan in academic year 2015/2016?

2. How is item test reliability on short answer in narrative reading assessment at the tenth grade students of SMAN 1 Jenanagan in academic year 2015/2016?

D. Objectives of the Study

In line with the research question stated above the undertaking of the present investigation intends to know:

1. To analyze item test validity on short answer test in narrative reading assessment at the tenth grade students of SMAN 1 Jenanagan in academic year 2015/2016.
2. To analyze item test reliability on short answer test in narrative reading assessment at the tenth grade students of SMAN 1 Jenanagan in academic year 2015/2016.

E. Significance of the Study

The result of this study is expected to be useful for students, teachers, and readers described as follows:

1. Theoretically

The researcher hopes that this research can give contribution of knowledge to develop and explore short-answer test, to know the advantages and disadvantages of using short-answer test in assess reading at the senior high school.

2. Practically

The result of this research is expected to be beneficial for:

a. School

This study is expected to give a contribution and reference in technique on assessing reading.

b. Researcher

In order to know the effectiveness of short-answer technique on assessing reading and finding of the research will be useful since it can be used as recent data and valuable source to conduct further research of the same aspect of study.

c. Teacher

Significance of the research is expected to develop short-answer technique as the way on assessing reading.

d. Students

The significance of the research is expected to make students more autonomous in learning reading. Particularly students at the tenth grade of SMA N 1 Jenangan in academic year 2015/2016.

F. Organization of the Thesis

To make easier in writing this thesis, the discussion is grouped into five chapters and each chapter is closely related each other, which is unified completely with systematic as follows:

CHAPTER I : The introduction, talk about background of the study, limitation of the study, statement of the problem, objectives of the study, significance of the study and organization of the thesis.

CHAPTER II : Discussed review of the theoretical background that talks about: Reading, Assessing Reading, Short-Answer Test, Validity, Reliability, Narrative Text, previous study, and theoretical framework.

CHAPTER III : The general explanation of research design, population and sample, instrument of data collection, and technique of data collection and analysis.

CHAPTER IV : The explanation about the result of the study contains of research location, description, and analysis of the data and the interpretation of the result of the study.

CHAPTER V : The conclusion and recommendations.

CHAPTER II

REVIEW OF RELATED LITERATURE

A. Theoretical Background

1. Reading

Reading is grasping language patterns from their written representation.⁷ Reading is a fluent process of readers combining information from a text and their own background knowledge to build meaning.⁸ In other words, it is a process of transferring information and constructing meaning from written text. Reading is an essential skill for English as second language learners. For most of these learners it is the most important skill to master in order to ensure success not only in learning English, but also in learning in any content class where reading in English is required.⁹

From this definition, it can be said that reading is a process of decoding. It is changing the codes or verbal symbols in the form of a series of letter to the sounds of a language that can be understood. While reading for meaning is a sequence of activities to make description/organize text content, so it can be evaluated as well as to respond to what

⁷ Robert Lado, *Language Teaching a Scientific Approach* (India: Tata McGraw-Hill Publishing CO.LTD, 1976), 132.

⁸ David Nunan, *Practical English Language Teaching* First Edition (New York: McGraw-Hill, 2003), 68.

⁹ *Ibid*, 69.

is expressed or implied in the text. So reading for meaning is the ability to read for the understanding: main ideas, important details, and the whole notion.

According to fielding and Pearson 1994, in particular, reading programs having the following four components can lead to student success:¹⁰

- a. Extensive amounts of time in class for reading
- b. Direct strategy instruction in reading comprehension
- c. Opportunities for collaboration
- d. Opportunities for discussions on responses to reading

2. Assessment

Below are two definitions cited in Airasian and Nunan:

- Airasian defines assessment as a general term which includes all the ways of information is gathered.
- Nunan the set of processes through which we make judgments about learner's level of skill and knowledge".¹¹

From two definitions above assessment is a general terms to make judgments about students level and knowledge. On assessment there are four categories of test:

¹⁰ J Michael O'Malley, Lorraine Valdez Pierce, *Authentic Assessment for English Language Learners* (United States of America: Addison- Wesley publishing, 1996), 95.

¹¹ Mansyur Bennu, Agus Purnomo, *Language Assessment* (Jakarta: MGMP, 2009), 3.

- a. Placement test: placing new students in the right class in a school is facilitated with the use of placement test. Usually based on syllabuses and materials the students will follow and use once their level has been decided on, these test grammar and vocabulary knowledge and assess students' productive and perceptive skills.

Some school asks students to assess themselves as part of the placement process, adding this self-analysis into the final placing decision.

- b. Diagnostic tests: while placement tests are designed to show how good a student's English is in relation to previously agreed system of levels, diagnostic tests can be used to expose learner difficulties, gaps in their knowledge, and skill deficiencies during a course.
- c. Progress or achievement tests: these tests are designed to measure learners' language and skill progress in relation to the syllabus they have been following.
- d. Proficiency tests: proficiency tests give a general picture of a student's knowledge and ability (rather than measure progress).

They are frequently used as stages people have to reach if they

want to be admitted to a foreign university, get a job, or obtain some kinds of certificate.¹²

The term authentic assessment is to describe the multiple forms of assessment that reflect student learning, achievement, motivation and attitudes on instructionally-relevant classroom activities.¹³ In addition, it represents how far the goal of learning objectives can be reached by the students. Authentic assessment constructed by the teacher to know students ability and reflect students learning.

3. Assessing reading

Assessment is a tool used to describe the students' knowledge and ability especially in reading. It shows the areas in the reading program where learners have weakness and strengths. It shows how well students are able to extract information and analyze what students are reading. The purpose of using assessment is to test that learning objectives are being met. Forms of assessments should be based on, and reflect, what is taught in the classroom.

In spite of the wide range of reading material specially written or adopted for English language learning purposes, there are a few comprehensive systemic programs which have been constructed from a

¹²Jeremy Harmer, *The Practice of English Language Teaching* (United States: Longman, 1998), 321.

¹³J Michael O'Malley, Lorraine Valdez Pierce, *Authentic Assessment for English Language Learning* (United States of America: Addison-Wesley Publishing Company, 1996), 4.

detailed analysis of the skills required for efficient reading. Much test material is still limited to short reading extracts on which general comprehension questions are based.¹⁴ To make systemic comprehension teacher can adopt material depend on the learning purpose and objective. Material can be adopted from newspaper, magazine, internet and books.

Given an interactive transactional theory of reading, which is based on the premise that students use both top-down and bottom-up process to construct meaning, it follows that assessment should involve the kinds of reading and writing tasks that students are called on to perform in and out of school reading whole books, studying text, and writing stories or letters. Although the term assessment might conjure up visions of tests, testing is just one way to getting information about students' reading and writing performance.¹⁵

In order for reading assessment to become useful in student evaluation, teachers should consider the following:

- a. Be thoroughly familiar with developmental learning processes and curriculum.
- b. Articulate a philosophy of assessment and evaluation.
- c. Know about and have experience collecting, recording, interpreting, and analyzing multiple sources of data.

¹⁴ J. B. Heaton, *Writing English Language Test* (London: Longman: 2009), 105.

¹⁵ Thomas G. Gunning, *Assessing and Correcting Reading and Writing Difficulties* (Boston: Pearson education, inc: 2010), 65-66.

- d. Be flexible and willing to try out multiple assessment procedures.¹⁶

In reading assessment there are some basic point must be attend by the teacher. Some basic points to remember in reading assessment include:

1. Activities for assessing reading should be based on activities for teaching reading.
2. Assessment of reading, like instruction, takes planning, time, and experience.
3. Assessment of reading should include both decoding skills and reading comprehension strategies.
4. Assessment of reading should include student attitudes and feeling toward reading.
5. Assessment of reading should hold students accountable for how they use time in class for reading.
6. Assessment of reading should be conducted regularly and be ongoing.
7. Students should be actively involved in their own assessment, whether it is in setting criteria, engaging in self-assessment, or evaluating peers.
8. Teacher observations of reading should be recorded systematically.
9. Assessment of reading should consist of multiple assessments for each student in order to monitor students' progress.

¹⁶ J Michael O'Malley, Lorraine Valdez Pierce, Authentic Assessment for English Language Learners (United States of America: Addison- Wesley publishing, 1996), 97.

10. Result of reading assessment should be used to inform students, parents, and teachers of needed changes in students' performance and instruction.¹⁷

4. Short-Answer test

One of the test can be used in assessing reading is short-answer tasks. Short-answer task is a reading passage is presented, and the test-taker reads questions that must be answered in a sentence or two.¹⁸ These are questions which require the students to write down specific answers in spaces provided on the question paper. The short answer technique is useful for testing both reading and listening comprehension.¹⁹ Short answer can be used to define students reading comprehension related with the text given.

This type of question, which will require only a minimal amount of writing from the student, is particularly useful in that it encourages students to listen or read for specific pieces of information within a text.²⁰ The question is only need a simple answer consists of a word or a sentence to define student understanding about information on the text.

¹⁷ Ibid, 132-133.

¹⁸ Douglass Brown, *Language Assessment Principles in Classroom Practices* (San Francisco: San Francisco State University, 2000), 207

¹⁹ Mansyur Bennu, Agus Purnomo, *Language Assessment* (Jakarta: MGMP, 2009), 29.

²⁰ Peter Lucatoni, *Teaching and Assessing skills in English as a Second Language* (UK: Cambridge University Press, 2006), 60.

A semi-objective alternative to multiple-choice is the short answer question (which Bachman and Palmer, 1996, classify as a limited production response type). Test-takers are simply asked a question which requires a brief response, in a few words. The justification for this technique is that it is possible to interpret students' responses to see if they have really understood, whereas on multiple-choice items students give no justification for the answer they have selected and may have chosen one by eliminating others.²¹

In other words, short answer question gives more specific justification to the students answer because the question give the chance to the students to express their own word. So, the answer can represent students understanding about the text related with the question. The students can answer the questions as good as student understanding about the text. So the teacher can give more objective assessment through short answer test in order to get the most authentic assessment.

Teachers give Short-answer test in order to assess their students' comprehension. Teachers prefer to 'give short-answer' test to ascertain students' reading performance.²² Short answer test has clear justification

²¹ J Charles Alderson, *Assessing Reading* (United Kingdom: Cambridge University Press, 2000), 227.

²² Dr. Imran khan, reading assessment techniques among selected secondary school teachers in Pakistan: current trends and practices. *International Journal on New Trends in Education and Their Implications* October, November, December 2011 Volume: 2 Issue: 4 Article: 7 ISSN 1309-6249, 68.

about student comprehension on reading test. So, the teacher can give more objective assessment to the students.

Most approaches to short answer assessment are situated in an educational context. Others aim at university assessment tests in the medical domain. Another strand of approaches focuses on language teaching and learning. These entire approaches share one theme: they assess short texts written by students. These may be answers to questions that ask for knowledge acquired in a course, e.g., in computer science, or to reading comprehension questions in second language learning. While thematically related, short answer assessment is different from essay grading. Short answers are formulated by students in a much more controlled setting. Not only are they short, they usually are supposed to contain only a few facts that answer only one question.²³ In this test format teacher easy to control students' answer because student only construct the answer shortly and clearly. Therefore, Teacher can give more objective assessment and a clear judgment score to student.

Short response items are usually items that the students can answer in a few phrases or sentences. This type of item should conform to at least the following two guidelines:²⁴

²³ Ramon Ziai, Niels Ott & Detmar Meurers, Short Answer Assessment: Establishing Links between Research Strands, 190-191.

²⁴ James Dean Brown, Testing in Language Programs (New York: McGraw-Hill Companies Inc, 2005), 52-53.

1. Is the item formatted so that only one relatively concise answer is possible?

Teachers should make sure that the item is formatted so that there is one, and only one, concise answer or set of answers that they are looking for in the responses to each item.

2. Is the item framed as a clear and direct question?

Short response item should generally be phrased as clear and direct questions.

An important difference between selected and constructed response formats is that in the latter case, the student is far less certain of the kind of response that is required. Therefore, the test constructor should phrase the question in such a way that is explicit about the scope of the response expected. This includes the length of the response, the number of arguments, causes or other elements required, which aspects need to be detailed.²⁵ In short-answer question, teacher should construct the question as simple as possible.

In conclusion, teacher can use short answer test to assess students' reading. When constructing the question teacher must attend the aspect above to construct the most objective questions.

²⁵ Jaap Scheerens, Educational Evaluation, Assessment, and Monitoring (Netherlands: Swets & Zeitlinger Publishers, 2005), 102.

5. Validity

Validity is unitary concept. Valid means correct. Charles, C.M states that “Scores obtained from a test of critical thinking are valid if they represent ability to think critically, as distinct from knowledge of vocabulary or reading skill”. The correctness of the assessment is called validity and the evidence to support the correctness of the assessment is called validity evidence. The kinds of validity are:

a. Construct validity

Construct validity evidence comes from the assessment instrument used. An assessment instrument is always designed to measure specific knowledge or skills of a group of people. No single assessment instrument can be used to measure everything from everybody. So the development of an assessment instrument must always be stated from the defined construct of the knowledge or skills to be assessed. The construct defined will lead to what tasks the instrument requires students to do. The correct definition of the construct will lead to the correct selection of the task, which will result in correct data, which has strong validity.

If an assessment instrument used to measure students’ writing skills does not required the students to perform a writing activity , the scores obtained will not become valid representation of the students writing skill. These score suffer from construct validity

problem. So, to provide construct validity evident that supports the validity of the result of writing assessment, the writing assessment instrument must require the students to perform writing activity.

b. Content validity

In the content validity, the coverage of the tasks becomes the evidence. If the tasks are covering all indicators of the knowledge or the skill to be assessed, the representativeness of the indicators becomes the supporting content validity evidence.

Gronlund states that “the essence of content validation is determining the adequacy of sampling. More formally, content validation is the process of determining the extent to which a set of tests provides a relevant and representative sample of the domain of tasks under consideration”.

c. Concurrent validity

Unlike evidence of construct and content validity, evidence of empirical validity comes from another set of data resulted from another assessment for the same purpose (Helton, 1988). Charles, C.M. states that “concurrent validity is representing when an instrument is shown to a yield results that are closely related to those produced by a second instrument that enjoys high acclaim”.

In the same way, low correlation between two sets of scores shows that the scores have concurrent validity problems. High and

positive correlation between the two sets of scores means that those who get high scores from the classroom English proficiency test also get high scores from TOEFL, and similarly those who get lower scores from the classroom English proficiency test get lower scores from TOEFL

d. Predictive validity

An assessment is often conducted to predict how well someone will perform in a different time. Charles C.M states that “predictive validity is present when one measurement or set of data can predict another that occurs at a later time”.

If it turns out in the future that among those who are admitted the high scores perform better in the study than the lower score, the score of test have good prediction and supported with predictive validity evidence. But if it turns out that the higher score from the test do not perform better than the lower score in study, the score from the test does not have good prediction. This means that the scores from the test have predictive validity problem.²⁶

²⁶ Mohammad Adnan Latief, Research methods on Language learning (Malang: UM Press, 2014), 225-228.

6. Reliability

The investigation of reliability is concerned with answer and question, “how much of an individual’s test performance is due measurement error or to factor other than the language ability we want to measure?” and with minimizing the effect of these factors on test score.²⁷

Factors affecting the reliability of test are:

- a. Not the examinees’ best performance.

The condition of examinees is influence on the result score that affecting on the degree of reliability. To avoid the errors, the assessor must select the best conductive atmosphere to make sure that the examinees are in the best performance while the assessment is being administrated.

- b. Not the raters’ most objective judgment.

Like the error coming from the examinees’ physical as well as emotional constraints, error in assessment that caused the scores become underestimate the true level of the skill being assessed happen because the raters who give the judgment to the quality to the skill being assessed are not in the most natural and objective physical as well as emotional mode. To avoid the errors, the raters must in the

²⁷ Lyle F Bachman, *Fundamental Consideration in Language testing* (New York: Oxford University Press, 1990), 161.

best conducive atmosphere to make sure that the raters are not having those constraints while giving the judgment.

- c. The assessment instrument being too short.

An assessment which asks more questions and allows more time to answer the questions, the result score will have the higher reliability than the same assessment which asks fewer questions and allows shorter time.

- d. The assessment instrument content being heterogeneous.

An assessment for written integrated English skill which covers the skill of reading, writing, the knowledge of grammar, and the knowledge of vocabulary, the result score will have the lower reliability than an assessment for only one specific language skill or a knowledge of language component, like an assessment for only reading skill. In other words, an instrument with numbers of questions designed to assess many different language skills or knowledge of language components (heterogeneous) the result score will have the lower reliability than an instrument with the same number of questions designed to assess a specific language component (homogeneous).

- e. The assessment questions being too easy or too difficult.

An assessment which asks so difficult questions that only 10 percent of the examinees can answer all the questions correctly, or so

easy questions that almost all of the examinees can answer all the question correctly, the result scores will be in the lower reliability than the same assessment which asks questions with moderate difficulty that about 35 to 85 percent of the examinees can answer all the questions correctly. (Ebel & Frisbie, 1985: 85) So, the level of difficulty of the questions in the assessment instrument influences the degree of reliability.

f. The type and quality of assessment instrument.

An instrument designed to assess the knowledge of vocabulary which contains 100 multiple choice type questions could produce scores with the same degree of reliability as the same assessment instrument which contains 150 true false type questions. Or, a multiple choice type instrument with more plausible distracters can produce scores with higher reliability than the same instrument with less plausible distracters' influence the degree of reliability.

g. Cheating in the assessment.

If the examinees are not strictly watched during the assessment process, examinees might copy each other's answers or copy the note that examinees have prepared. If this cheating happens then the assessment result becomes in low reliability scores. So honesty of examinees in answering the assessment questions affects the degree of reliability.

h. Uncomfortable place and time.

An assessment conducted in an uncomfortable room, too hot, too cold, too windy, too crowded, too small, or too noisy, and in uncomfortable time at 2.00 p.m. after the examinees have worked all morning, for example, will result in the assessment scores with low reliability.²⁸

7. Narrative text

Narrative text is a story that is created in a constructive format that describes a sequence of fictional or non fictional events. Narration is the most easily recognized mode of discourse, it means simply to tell a story. A writer uses narration to relate events, either real or imagined, in chronological order.²⁹ The purpose of narrative text is to entertain, to tell a story, or to amuse the readers or listeners.³⁰ Narrative text is simple texts which tell about a story with sequence events, real or imagined in chronological order. It has the function to entertain and tell a story.

Fortunately, many experts agree that narrative text has dualistic nature of meanings: the first, it is related to the aspect of a what and the second, it is of a way. The what of narrative is in terms of narrative content containing the main elements which concern of events, actors, time and

²⁸ Mohammad Adnan Latief, *Research methods on Language learning* (Malang: UM Press, 2014), 215-218.

²⁹ Deanne Spears, *Developing Critical Reading Skills* seventh edition (New York: McGraw-Hill Companies Inc, 2006), 45.

³⁰ Nuning widyaningsih, *Kreatif Bahasa Inggris* (Klaten: Viva Pakarindo, 2013), 24.

location. The way has to do with how the narrative is told. The how means that the events, time, and location should be well-organized from the beginning up to the end in successive sequence.³¹

Generic structure of narrative text:

- a. Orientation: who were involved in the story, when and where introduces specific participant and setting.
- b. Complication: a problem arises followed by other problems. Complication explores the conflict in the story. It shows the crisis, rising crisis, and climax of the story.
- c. Resolution: it shows the situation which the problems have been resolved. In other words, resolution is a solution for the problem.³²

The grammatical Features of narrative text are:

- Use of adjectives to build noun groups to describe the people, animals or things in the story such as fresh grass, a quiet corner.
- Use of time connectives and conjunctions to sequence events through time, such as after, while, later, than.
- Use of adverbs and adverbial phrases to locate the particular events, such as into the field, one day.

³¹Dolar Yuwono, Writing 1 (Yogyakarta: Pustaka Felicha, 2015), 35.

³²Nuning widyaningsih, Kreatif Bahasa Inggris (Klaten: Viva Pakarindo, 2013), 25.

- Use of past tenses, such as he began to actually look forward to the wolf's presence.
- Use of action verbs to indicate the actions, such as took, herded, walked, and stood.
- Use of saying and thinking verbs to indicate what characters are feeling, thinking, or saying, such as moaned.³³

The following are kinds of story belong to narratives:

1. Fables are stories about animals. Fables usually point out lessons with animals as the characters. Fables entertain but they tell important truths about state the lesson or the moral at the end of the story.
2. Folktales are stories of legends forming part of an oral tradition. Folktales passed on by word of mouth rather than by writing, and thus folktales are modified by successive retellings before being written down. They also teach a moral of explain why something is the way it is.
3. Fairy tales are stories involving imaginary creatures and always have a happy ending.

³³ Bachtiar Bima, dkk, Bahasa Inggris kelas X semester 1 (Intan Pariwara, 2015), 28.

4. Personal experience is a factual past events by retelling them in order in which they occurred. It can be sad, happy, embarrassing, or funny experience.
5. Myths, legends, short stories and science fiction³⁴

B. Previous Research Finding

There are many researches that had been conducted before related to the discussion of the thesis. These are two previous research found.

The first research is started from previous research finding that is conducted by Binti Lailatul Munawaroh at year 2014 in thesis with the title An Analysis of Multiple Choice Item Test on English of Mts Ma'arif Munggun Pulung. The research concluded that multiple choice item test has low validity. It was drawn from the results of analysis item difficulty showed 48% item classified as easy item. Item discrimination showed 58% item classified as bad item. And 52% item showed item distracter not work and reliability of multiple choices test classified as high reliability. The scale Alpha showed reliability of the test with 1,99. From the result showed that, the test has low validity it is not means that the test also has low reliability.

The Second research is started from previous research finding that is conducted by Fitra Khoirul Anwar at year 2013 in thesis with the title The Effectiveness of English Item Test in Final Test on Odd Semester (item

³⁴ Ibid, 25.

difficulty and item discrimination level) at XIth IPS student of SMA Bakti Ponorogo in Academic Year 2012/2013. The research concluded that the test is effective to be tested for the students. It can be said like that, because based on the result of calculating the test has good criteria of item difficulty level and satisfactory criteria of item discrimination level.

The third research is started from previous research finding that is conducted by Anak Agung Gede Hendra 2014. In thesis with title Using Multiple Choice Test Item to Assess the Reading Comprehension of Seventh Grade Students of SMPN 2 Tampak Siring in Academic Year 2013/2014. From the results of earlier research, research concludes that the assessing of the seventh grade students of SMPN 2 Tampaksiring in using reading comprehension especially reading for information was sufficient, because 30 students from 40 students pass the passing grade.

The Fourth previous research finding that is conducted by Abdulloh H. A. Al-Qudairy 2010. In thesis with title Using Arabic (L1) in Testing Reading Comprehension in English (L2) as a Foreign Language. From the results of earlier research, research concludes that the present study aimed at investigating the effect of using Arabic (L1) as a language of questions and answers in testing reading comprehension in English (L2) by using both multiple-choice and short answer questions for upper-intermediate and post-beginner students. Similarly, there was no significant difference in the performance of the final-year secondary school students (post-beginner) in either the multiple-choice or the

short answer tests regardless of the test language. Providing the reading comprehension questions and answers in Arabic (L1) did not improve students' levels of performance.

The different of this study is using short-answer test and four previous study above using multiple choice item test and Arabic (L1) to assess students' reading. The four of researchers Binti Lailatul Munawaroh, Fitra Khoirul Anwar, Anak Agung Gede Hendra and Abdulloh H. A. Al-Qudairy give the inspiration to find out or research about the effectiveness of short-answer test on students' narrative reading assessment.

C. Theoretical Framework

Theoretical framework is a concept about theories can be related with the factors which are identified as the important problems. The researcher observes short answer test used on narrative reading assessment at the tenth grade students of SMAN 1 Jenangan in academic year 2015/2016. After that the researcher analyzes this process to know, how is short-answer test item validity and reliability on narrative reading assessment at the tenth grade students of SMAN 1 Jenangan in Academic year 2015/2016.

The researcher assumes that the short answer test is effective to assess students' reading skill on narrative text.



CHAPTER III

RESEARCH METHOD

This part discussed the methodology of the research. It included research design, population and sample, instrument of data collection, technique of data collection, and technique of data analysis.

A. Research Design

The data in this research were numeric form, the researcher categories the research as a quantitative research. Quantitative research in this study included in descriptive research. There was a definition about descriptive research, according to Deni Darmawan “penelitian ini bertujuan mendeskripsikan suatu objek atau kegiatan yang menjadi perhatian peneliti.”³⁵ It can be said that the descriptive research was a research method that addressed to describe the exits phenomena or activity and occurs in present or in past time.

B. Population and Sample

1. Population

According to Suharsimi Arikunto population is all of the research subjects.³⁶ A population is a group of individuals, objects, or items from

³⁵ Deni Darmawan, *Metode Penelitian Kuantitatif* (Bandung: PT Remaja Rosdakarya, 2014), 49.

³⁶ Suharsimi Arikunto, *Prosedur Penelitian Suatu Pendekatan Praktik*, (Jakarta: PT. Rineka Cipta, 2006), 130.

among which samples are taken for measurement.³⁷ This statement can be explained as population refers to all of subjects or objects in an area and filled the specific aspects that related with research problem, of the all of units or individual in the research scope.

Donal Ary said the larger group about which the generalization is made is called population.³⁸ In other word, population was all of group being research subject.

In this research the population was tenth grade students of SMAN 1 Jenangan is 89 students in academic year 2015/2016. It consisted of four classes, they are X 1, X2, X3 and X4. Where, the first class is 23 students, the second class is 23 students, the third class is 22 students, the fourth class is 21 students.

2. Sample

Webster said a sample can be defined as a finite part of a statistical population whose properties are used to make estimates about the population as a whole.³⁹ It can be explained that sample was a part of population.

³⁷ Kultar Singh. Quantitative Social Research Methods (New Delhi: Sage Publication India Pvt Ltd, 2007), 88.

³⁸ Donal Ary, Introduction to research in education,(United Sate of America: Northbern Illisionist University,1979), 113.

³⁹ Kultar Singh. Quantitative Social Research Methods (New Delhi: Sage Publication India Pvt Ltd, 2007), 88.

Sample is the representative of the population studied.⁴⁰ In determining the sum of sample members, the researcher can use some formulas or look at research population table scale.

Suharsimi Arikunto said that “*jika peneliti mempunyai beberapa ratus subjek dalam populasi, mereka dapat menentukan kurang lebih 25-30% dari jumlah subjek tersebut, jika jumlah anggota meliputi antara 100-150. Jika populasi kurang dari 100 maka harus dijadikan sebagai sampel*”.⁴¹ It can be said that if the researcher has population more than 100 people, the researcher can determine 25-30% from the population as subject. But, if the population was less than 100 people, the researcher used all population as sample.

Gay and Diehl said that “for descriptive research the sample is 10% from population, for correlation research at least 30 elements of population, for causal research 30 elements on every group and for experimental research 15 elements on every group”.⁴²

In this research, the researcher used Gay and Diehl theory to choose the number of sample. The sampling technique applied in this research was cluster random sampling. Cluster random sampling is determined based on the sampling areas or groups that exist in the

⁴⁰ Ibid, 131.

⁴¹ Suharsimi Arikunto, *Prosedur Penelitian Suatu pendekatan Praktik* (Jakarta: PT Rineka Cipta, 2006), 130.

⁴² Deni Darmawan, *Metode Penelitian Kuantitatif* (Bandung: PT Remaja Rosdakarya, 2014), 143.

population such as school, class, region, and not individual. Charles, C.M said that “Cluster sampling technique involves the random selection of groups that are already exists.⁴³ Because among of classes have not significance difference and choosing based on ability, the researcher chosen X⁴ class as the sample.

C. Instrument of Data Collection

Instrument was a tool for assist which was chosen and used by researcher in gathering data in order to the activities will systemic and become easy.

Instrument of data collection can be shown as the table below:

Table 3.1

Instrument of data collection

Title of Research	indicator	No. item of instrument
Item Test Analysis on Short-Answer in Narrative Reading Assessment at the Tenth Grade Students of SMAN 1	validity	1-20

⁴³ Muhammad Adnan Latief, Research Methods on Language Learning (Malang: UM Press, 2014), 185.

Jenanagan in Academic Year 2015/2016.	Reliability	1-20
--	-------------	------

D. Technique of Data Collection

1. Documentation

Documentary is kind of important technique to get data about everything or variable which is a form of notes, transcript, book, newspaper, magazine, meeting result, logger, etc.⁴⁴

In this study, the documentary technique used to get the data about students' names, condition of teachers and staffs, vision, missions, goals, facilities, infrastructure, and the structure organization of SMAN 1 Jenanagan.

2. Test

Test was some questions that given for people to show the condition or the level the people development. The fundamental use of testing in an

⁴⁴ Suharsimi Arikunto, *Prosedur Penelitian Suatu Pendekatan Praktik* (Jakarta: PT Rineka Cipta, 2006), 231.

educational program is to provide information for making decision that is for evaluation.⁴⁵

The data in this research was the result of test. The data was taken from writing test. In this study, the data was taken from the result of sample.

E. Technique of Data Analysis

Analysis data was process the data obtained by using formulas or rules that was applicable to the research or design approach taken. To analyze the item test on short answer test in narrative reading assessment at the tenth grade students of SMAN 1 Jenangan in Academic year 2015/2016, the researcher used validity and reliability.

1. Validity

Validity test was a requirement imposed on short answer test to determine the effectiveness of the test question items. The test item said to be valid if the questions on a short answer test was able to reveal something that will be measured by these short answer test.

There were some formulas to compute the validity of the data. The formula is:

⁴⁵ Lyle F. Bachman, *Fundamental Considerations in Language Testing* (New York: Oxford University Press, 1990), 54.

$$r_{xy} = \frac{n \sum XY - (\sum X)(\sum Y)}{\sqrt{(n \sum X^2 - (\sum X)^2)(n \sum Y^2 - (\sum Y)^2)}}$$

Where:

r_{xy} : index correlation

$\sum X$: Total of X

$\sum Y$: Total of Y

$\sum XY$: Total of X time Y⁴⁶

In this research the researcher used a formula provided in SPSS 19.00 for windows 8 to compute the data. HO was accepted if the correlation > than the alpha decided (5%).

2. Reliability

Reliability was a necessary characteristic of any good test. For the test must be reliable as a measuring research instrument. Louma stated that reliability is important because it means that the scores are dependable, so that we can rely on them in decision-making. Mousavi in Brown stated a reliable test is consistent and dependable. If the teacher gave the same test to the same students or matched students on two different occasions, the test should field.

⁴⁶ Retno Widyaningrum, Statistika Edisi Revisi (Yogyakarta: Pustaka Felicha, 2013), 107.

There were some formulas that can be used to compute the result of reliability such as KR-21. The formula as the following:

$$r_{11} = \left[\frac{k}{k-1} \right] \left[1 - \frac{\sum \sigma_i^2}{\sigma_t^2} \right]$$

Where:

r_{11} = Cronbach Alpha Formula

k = Number of test items

σ_i^2 = Variance of item

σ_t^2 = Total variance.⁴⁷

The variance of test was respectively computed by using the following formulas:

$$\sigma_i^2 = \frac{\sum x^2}{n} - \left(\frac{\sum x}{n} \right)^2$$

Where:

σ_i^2 = Variance of item

X = score of item

n = Total items⁴⁸

The researcher used cronbach alpha provided in SPSS version 19.00 for windows 8 to analyze the reliability of the data.

⁴⁷ Andhita Dessy Wulansari, *Penelitian Pendidikan* (Ponorogo: STAIN Ponorogo Press, 2012), 90.

⁴⁸ *ibid*, 89.

CHAPTER IV

RESEARCH FINDING

This chapter described the research location, data description, data analysis and discussion. The general description of research location included the history of establishment of SMAN 1 Jenangan, geographical location/position, school profile, vision, mission and goal of the school, curriculum, teachers' condition, students' condition, and structure organization.

The data description included the data of research subjects. The data analysis was about the analysis of the data students' narrative reading assessment test. The discussion included the discussion and interpretation of the result.

A. Research Location

1. The history of establishment of SMAN 1 Jenangan

SMAN 1 Jenangan is a State educational school institution. The history of establishment of SMAN 1 Jenangan is related to the instruction of the district government regulations of Ponorogo that every district should have at least a senior high school. The district government gave funds for jenangan sub district to establish a senior high school. Finally in 2003 a senior high school institution named SMAN 1 jenangan was established. The statistic number of this school (NSS) is 301051103004

while the founding deed number is 425/828/405.51/2003. Four years after the founding of the institute, precisely at 2007, SMAN 1 Jenangan got B (good) on accreditation status.

The first head master of SMAN 1 Jenangan was Mr. Suroto from 2003 until July 2010. The second head master was Drs. Subandi, M.Pd. He takes hold as school headmaster until March 2015. After that, the headmaster was replaced by Dra. Hj. Tutut Erliena , M.Pd as PLT until the end of 2015. And now, the headmaster is Drs. Sugeng Subagyo, M.Pd.

2. Geographical position of SMAN 1 Jenangan

The geographical position of SMAN 1 Jenangan shows at the table below:

Table 4.1

Geographical position of SMAN 1 Jenangan

School name	SMAN 1 Jenangan
Street	Jalan raya Ngebel
Village	Semanding
Sub district	Jenangan
City	Ponorogo
Province	East java

3. School profile

SMA Negeri 1 Jenangan is located on Jalan Raya Ngebel, Semanding village, Jenangan, Ponorogo, East Java. The zip code is 63 492. The phone number is (0352) 531 952. The building is located at an altitude of approximately 350 meters above sea level and is located on the slopes of Dangean, so the condition of the ground is hilly. SMA Negeri 1 Jenangan placed on Semanding village because the price of the land in this area is still relatively cheap. Before the establishment of the school, the land was used as a farm by one of the villagers.

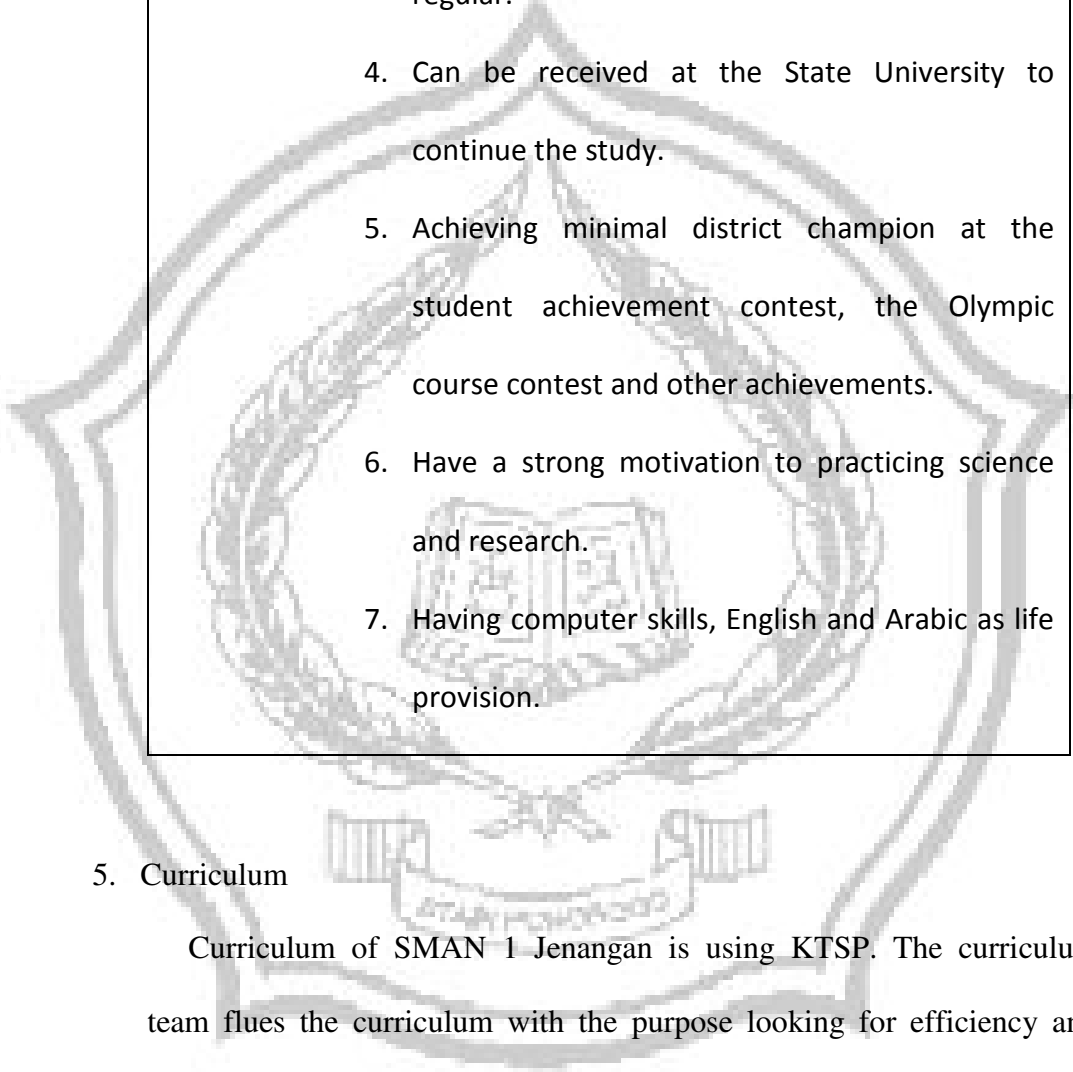
To learn more about SMA Negeri 1 Jenangan profile can use senior high school accreditation instrument that has been set by National Education Minister, Number 52 Year 2008. It consist of criteria accreditation of senior high school which include standard of contents, standard of process, standard of competencies, standard of teachers and education personnel, standard of facilities and infrastructure, standard of managements, standard of financing and standard assessment of education.

4. Vision, mission, and goal

Table 4.2

Vission, mission, and goal of SMAN 1 Jenangan

<p>Vision:</p> <p>Actualizing cultural of achievement with life skill and noble morals.</p>
<p>Mission:</p> <ol style="list-style-type: none"> 1. To built the spirit and culture of achievement both academic and non academic to the entire school community. 2. Improve understanding of the religion professed to be foundation of thinking and behaving. 3. Improving academic skills as a preparation to continue study at university. 4. Improving the communication ability using English and Arabic. 5. Develop the skills to operate a computer and internet.
<p>Goal:</p> <ol style="list-style-type: none"> 1. Have a high level of achievement in each subject.

- 
2. Can achieve NUN set above the minimum graduation.
 3. Having a strong reading culture, purposeful and regular.
 4. Can be received at the State University to continue the study.
 5. Achieving minimal district champion at the student achievement contest, the Olympic course contest and other achievements.
 6. Have a strong motivation to practicing science and research.
 7. Having computer skills, English and Arabic as life provision.

5. Curriculum

Curriculum of SMAN 1 Jenangan is using KTSP. The curriculum team flues the curriculum with the purpose looking for efficiency and relevancy. The aim of SMAN 1 Jenangan in forming students which has good moral, physical, intellectual, social, and religion that hold on to the nature of the character.

6. Teacher Condition

Teacher was a figure to be good example for the students. The teacher has to act as advisor for the students in developing creativity and self potential as motivator that help the students raising their goal and aspiration. The teacher of SMAN 1 Jenangan have educational background of under graduate (S1) and post graduate (S2) teachers from many institutes, such as graduated from IKIP PGRI Madiun, Jember university and etc.

7. Student Condition

Students were the one of important components in the education system. In the academic year 2015/2016, the total number of students at SMAN 1 Jenangan was 199 Students. It consisted of 89 students from the first grade, 68 Students from the second grade, and 42 Students from the third grade.

8. The Structure of Organization

To create good corporation in running vision and mission and also the target of education in SMAN 1 Jenangan, it was required an organization structure which has each role and function. Organization structure in institute was really important because it described each personal's duty. Organization structure of school could be seen as follow:

- a. Head master : Drs. Sugeng Subagyo, M.Pd
- b. Head of school administration : Subari
- c. Deputy of students' affairs : Bagus Nanang, S.E
- d. Deputy of facility : Drs. Sariyono, M.Pd
- e. Deputy of library : Purwindarti, S.Pd
- f. Deputy of curriculum : Drs. Puguh S
- g. Concealing : Siti Fatonah, M.Psi

B. Data Description

To introduce the research, the researcher took 21 students as the sample, those samples taken from the tenth grade of SMAN 1 Jenangan. The data descriptions in this research was short answer test on students' narrative reading assessment the questions of the test, the key answer and the answer sheets of students. These data can be seen as follow:

1. Short answer test on students' narrative reading assessment

The test could be categorized into evaluation test. It was conducted on Monday, 11th of April 2016. This test included for the tenth grade students of SMAN 1 Jenangan.

At this test, teacher gave two different texts to the students. Then, asked the students to answer the following questions. 3 points for the correct answer (spelling and grammar), 2 points for the incomplete answer, 1 point for the incomplete answer but incorrect on spelling or

grammar and 0 point if there was no answer. From the result showed that the highest score was 98 and the lowest was 45. Further information can be seen at the table below:

Table 4.3

Score of students Short answer test on narrative reading assessment



No	Name	Score
1	Wahyu P	98
2	Agus Setyaningrum	98
3	Siti A	97
4	joni P	92
5	Ida nur	90
6	Krisnatri L	88
7	Nadya Bella K	87
8	Ayu Kusumastuti	87
9	M. Taufiq	85
10	Anisa M	85
11	Septa Anggy H	85
12	Tri Ayu Larasati	83
13	Tita eka M	82

14	Susi W	82
15	Meldha A	77
16	Miftakhul Aini	67
17	Juanti	63
18	Aan purnomo	55
19	Kris Diansari	52
20	Andik Dwi S	48
21	Ridwan Jois	45

2. The question of the test.

This part described about the questions of narrative reading assessment test. It was about short answer questions. There were twenty questions. Further information can be seen at the table below:

Table 4.4

The questions of short answer test

Name:

Class:

Narrative text

Read the texts and answer the questions bellow:

The Doves

Once upon a time, there was a flock of doves that flew in search of food led by their king. One day, they had flown a long distance and were very tired. The dove king encouraged them to fly a little further. The smallest dove picked up speed and found some rice scattered beneath a banyan tree. So, all the doves landed and began to eat.

Suddenly a net fell over them and they were all trapped. They saw a hunter approaching carrying a huge club. The doves desperately fluttered their wings trying to get out, but to no avail. The king had an idea. He advised all the doves to fly up together carrying the net with them. He said that there was strength in unity.

Each dove picked up a portion of the net and together they flew off carrying the net with them. The hunter looked up in astonishment. He tried to follow them, but they were flying high over hills and valleys. They flew to a hill near a city of temples where a mouse lived that could help them free. He was a faithful friend of the dove king.

When the mouse heard the loud noise of their approach, he went into hiding. The dove king gently called out to him and then the mouse was happy to see him. The dove king explained that they have been caught in a trap and needed the mouse's help to gnaw at the net with his teeth and set them free.

The mouse agreed saying that he would set the king free first. The king instead that he first free his subjects and the king last. The mouse understood the king's felling and complained with his wishes. He began to cut the net one by one all the doves were freed including the dove king. They all thanked the mouse and flew away together united in their strength.

1. Who was flew in search of food?
2. Why did all the doves land on the ground?
3. What happened with doves after they had landed on the ground?
4. Who was approaching to the doves and carrying a huge club?
5. What was the dove's king idea?
6. Who could help doves?
7. Where did the mouse live?
8. What did the mouse do to help the doves?
9. Who was free at the first?
10. What is the moral value of the story?

THE LEGEND OF BANYUWANGI

Once upon a time, there was a local ruler named King

Sulahkromo. The king had a Prime Minister named Raden Sidopekso. The Prime Minister had a wife named Sri Tanjung. She was so beautiful that the king wanted her to be his wife.

One day, the King sent his Prime Minister to a long mission. While the Prime Minister was away, the King tried to get Sri Tanjung. However he failed. He was very angry. Thus, when Sidopekso went back, the King told him that his wife was unfaithful to him. The Prime Minister was very angry with his wife. Sri Tanjung said that it was not true. However, Sidopekso said that he would kill her.

He brought her to the river bank. Before he kill her and threw her into the river, he said that her innocence would be proven after Sidopekso killed her. He threw her dead body into the dirty river.

The river immediately became clean and began to spread a wonderful fragrance. Sidopekso said, “Banyu...Wangi... Banyuwangi”. This means “fragrant water”. Banyuwangi was born from the proof of noble and sacred love.

11. Who are the main characters of the story?
12. Who was the prime minister name?
13. Where did the king send his prime minister?

14. What did the king do when his minister was away?
15. Why was Sidopekso angry to his wife?
16. Where did Sidopekso throw his wife?
17. What was Sidopekso say before killing his wife?
18. Was Sri Tanjung lying to her husband?
19. What happened with the river after Sidopekso had thrown his wife there?
20. What is the meaning of Banyuwangi?

3. The answer key

This part described about the answer key of the narrative reading assessment test. It was about short answer questions. There were twenty questions.

Table 4.5

The answer key of short answer test

Answer key:

1. A flock of doves
2. Because they found some rice
3. They were all trapped

- 
4. A hunter
 5. To flew up together
 6. Mouse
 7. In a hill near a city of temple
 8. He gnaw the net with his teeth
 9. The king
 10. There was a strength in unity
 11. The king (Sulahkromo), Prime minister (Raden Sido Pekso), Sri Tanjung.
 12. Raden Sido Pekso
 13. To a long mission
 14. The king tried to get Sri Tanjung
 15. Because his wife was unfaithful to him.
 16. To the river
 17. He said that her innocence would be proven after killed her
 18. No, she was not
 19. The river became clean and spread a wonderful fragrance.
 20. It means "fragrant water"

4. Students' answer

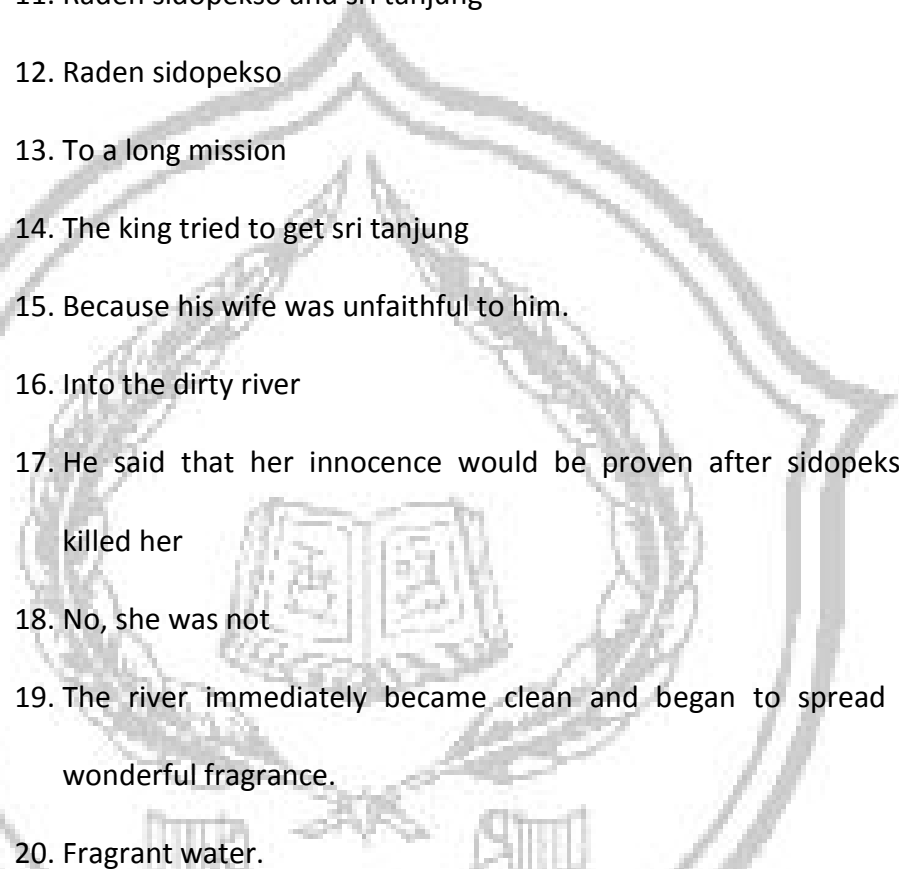
This part was about the students' answer of the narrative reading assessment test. There were 21 Students who answer the twenty questions in this research. In this section, just shows a student answer as the sample. For further information can be seen in the appendix.1

Table 4.6
Student's answer

Name : Wahyu P

Class : X4

1. A flock of doves
2. Because, the smallest dove picked up speed and found some rice scattered beneath a banyan tree
3. Suddenly a net fell over them and they were trapped.
4. Hunter
5. He advised all the doves to fly up together carrying the net with them.
6. A mouse
7. A hill near a city temple

- 
8. He gnaw
 9. His subject
 10. Together united in their strength
 11. Raden sidopekso and sri tanjung
 12. Raden sidopekso
 13. To a long mission
 14. The king tried to get sri tanjung
 15. Because his wife was unfaithful to him.
 16. Into the dirty river
 17. He said that her innocence would be proven after sidopekso killed her
 18. No, she was not
 19. The river immediately became clean and began to spread a wonderful fragrance.
 20. Fragrant water.

C. Analysis of the Data

Data that have been analyzed by the researcher was the answer sheet of the students. It was the result of evaluation test. There was short answer test type that contains 20 questions.

1. Validity analysis

The validity of each item was then analyzed by using SPSS 19.00 for Windows. After finding r_{xy} , it was consulted to r_{table} with 5% level of significance. If the value of r_{xy} was equal to or greater than the value of r_{table} , it indicated that item was valid. According to the r_{table} value for $N = 21$ on the 5% of significance level, it listed 0,433. The result of the SPSS calculation was in the following:

Table 4.7
Validity

		Total
Number 1	Pearson Correlation	.639**
	Sig. (2-tailed)	.002
	N	21
Number 2	Pearson Correlation	.835**
	Sig. (2-tailed)	.000
	N	21
Number 3	Pearson Correlation	.505*
	Sig. (2-tailed)	.020
	N	21

4	Number	Pearson Correlation	.472 [*]
		Sig. (2-tailed)	.031
		N	21
5	Number	Pearson Correlation	.861 ^{**}
		Sig. (2-tailed)	.000
		N	21
6	Number	Pearson Correlation	.469 [*]
		Sig. (2-tailed)	.032
		N	21
7	Number	Pearson Correlation	.426
		Sig. (2-tailed)	.054
		N	21
8	Number	Pearson Correlation	.471 [*]
		Sig. (2-tailed)	.031
		N	21
9	Number	Pearson Correlation	.179
		Sig. (2-tailed)	.438
		N	21
	Number	Pearson Correlation	.686 ^{**}

10	Sig. (2-tailed)	.001
	N	21
Number	Pearson Correlation	.481 [*]
11	Sig. (2-tailed)	.027
	N	25
Number	Pearson Correlation	.635 ^{**}
12	Sig. (2-tailed)	.002
	N	21
Number	Pearson Correlation	.499 [*]
13	Sig. (2-tailed)	.021
	N	21
Number	Pearson Correlation	.521 [*]
14	Sig. (2-tailed)	.015
	N	21
Number	Pearson Correlation	.679 ^{**}
15	Sig. (2-tailed)	.001
	N	21
Number	Pearson Correlation	.461 [*]
16	Sig. (2-tailed)	.036

	N	21
Number	Pearson Correlation	.839**
17	Sig. (2-tailed)	.000
	N	21
Number	Pearson Correlation	.477*
18	Sig. (2-tailed)	.029
	N	21
Number	Pearson Correlation	.705**
19	Sig. (2-tailed)	.000
	N	21
Number	Pearson Correlation	.603**
20	Sig. (2-tailed)	.004
	N	21
Total	Pearson Correlation	1
	Sig. (2-tailed)	
	N	21

3,

The table above showed that number 1, 2, 4, 5, 6, 8, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 were valid because the value in X_{total} column was greater than

the value of r_{table} , (0,433). And number 7 and 9 were invalid because the value in X_{total} Column was less than the value of r_{table} .

2. Reliability analysis

Test of reliability indicated the consistency of measurement.

Classification of reliability:

Table 4.8

Classification of reliability

value	Classification
0,800-1,000	Very high
0,600-0,800	High
0,400-0,600	Moderate
0,200-0,400	Low
0,000-0,200	Very low

From the table above, it can be said:

- a. If the calculation result of formula is less than 0,200, the result categorized as very low reliability.
- b. If the calculation result of formula is between 0,200- 0,400, the result categorized as low reliability.
- c. If the calculation result of formula is between 0,400- 0,600, the result categorized as moderate reliability.
- d. If the calculation result of formula is between 0,600- 0,800, the result categorized as high reliability.
- e. If the calculation result of formula is between 0,800- 1,000, the result categorized as very high reliability.

There were many kinds of formula that used to analyze the reliability. The calculation of reliability test by using Cronbach Alpha technique provided in SPSS 19.00 for Windows. The result of reliability calculation was presented in the following tables:

Table 4.9

Reliability

Case Processing Summary

		N	%
Cases	Valid	21	100.0
	Excluded ^a	0	.0
	Total	21	100.0

a. Listwise deletion based on all variables in the procedure.

Reliability Statistics

Cronbach's	
Alpha	N of Items
.883	20

From the result above, it can be concluded that the reliability of the test was very high, because this result more than 0,800.

D. Discussion and Interpretation

The purpose of this study was giving the picture of the item test validity and reliability of short answer test in narrative reading assessment at the tenth grade students of SMAN 1 Jenagan in academic year 2015/2016.

Referring the explanation above, the validity of the test analyze with formula provided in SPSS 19.00 for windows 8. The result of number 1, 2, 3, 4, 5, 6, 8, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 were more than r_{table} . Number 7 & 9 were less then r_{table} . It means that 90% of the result was more than r_{table} and 10% of the result was less than r_{table} .

Based on the explanation at data analysis above, the reliability test analyze with Cronbach Alpha technique provided in SPSS 19.00 for windows 8. The result of this test has very high reliability. It showed from the alpha scale 0,833. From the result the test has good validity and very high reliability.

Additionally, the test needed to be analyzed, to know the validity and reliability of the items. Some items can be used to the next test, but for some items it was useless. Moreover there were many factors that affect the quality of the result. Factors that affect the quality of the test were the teacher and the students. Other factors were administration and implementation of the test.

Finally, the teacher must analyze the test result of the exam as well as test makers, because many feedbacks must be obtained by the teacher from the result. It was also vitality important to learn about language testing, such as the issues of reliability and validity which were mentioned in this research.

However, if the teachers was opened to the possibility of constant revision of language tests, then the teacher can surely be more successful in creating test that were more effective.



CHAPTER V

CLOSING

This chapter serves the conclusion and recommendation. The conclusion came from research result and the recommendation presents the suggestions for some aspects.

A. Conclusion

1. Item test validity of short answer in narrative reading assessment at the tenth grade students of SMAN 1 Jenanagan in academic year 2015/2016 has good validity. It is drawn from the validity result analysis 90% from 20 questions $> r_{table}$. Then, 10% from 20 questions $< r_{table}$.
2. Item test reliability of short answer in narrative reading assessment at the tenth grade students of SMAN 1 Jenanagan in academic year 2015/2016 is classified as high reliability because the alpha scale shows reliability of the test with 0,883.

B. Recommendations

Based on the result of the research, the researcher would like to offer some recommendations:

1. Students

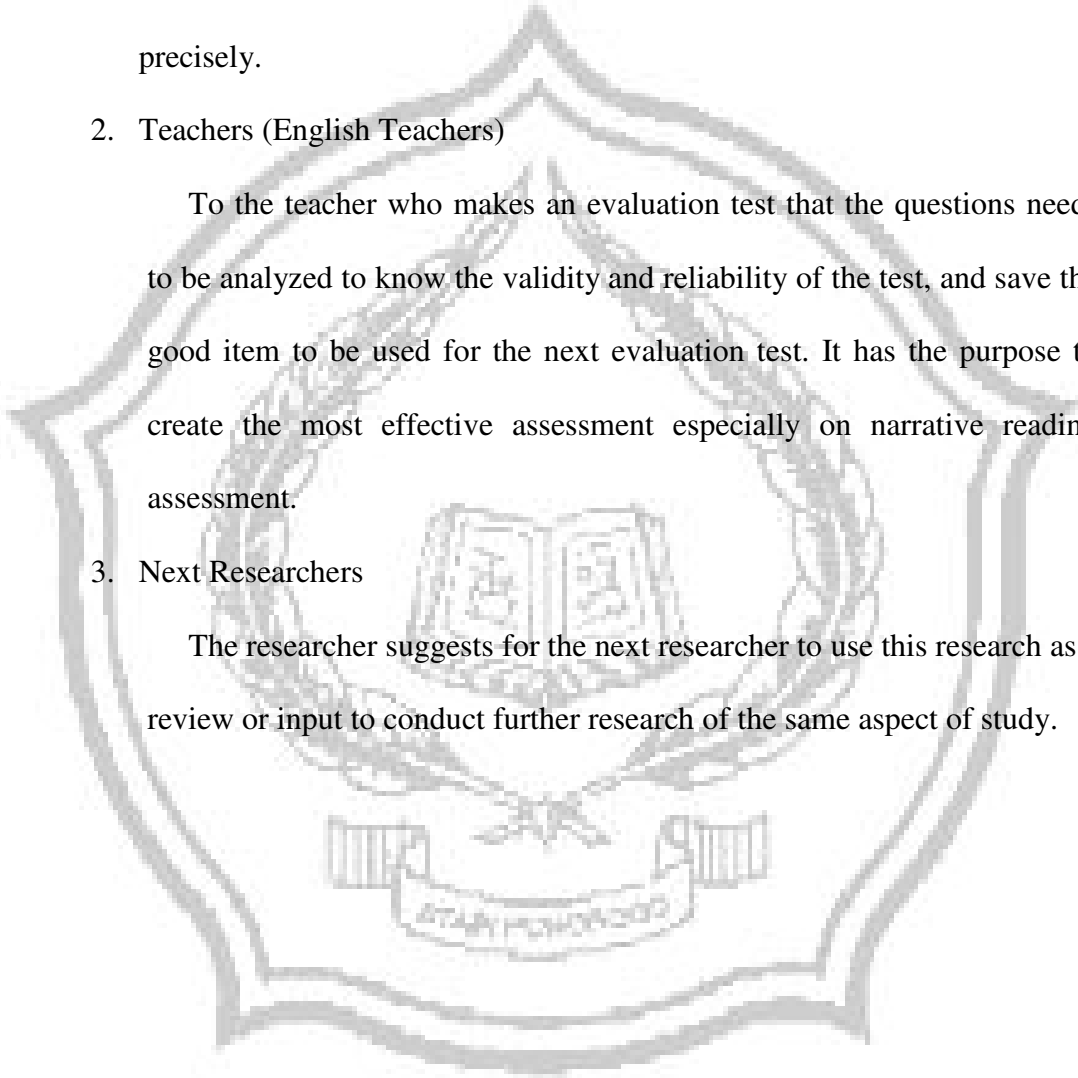
Students should train themselves to answer short answer test. Because the long question is not necessarily having long answer too, so the students should familiarize themselves to answer the question shortly and precisely.

2. Teachers (English Teachers)

To the teacher who makes an evaluation test that the questions needs to be analyzed to know the validity and reliability of the test, and save the good item to be used for the next evaluation test. It has the purpose to create the most effective assessment especially on narrative reading assessment.

3. Next Researchers

The researcher suggests for the next researcher to use this research as a review or input to conduct further research of the same aspect of study.



BIBLIOGRAPHY

- Alderson, J Charles. Assessing Reading. United Kingdom: Cambridge University Press, 2000.
- Arikunto, Suharsimi. Prosedur Penelitian Suatu Pendekatan Praktik. Jakarta: PT. Rineka Cipta, 2006.
- Ary, Donal. Introduction to research in education. United Sate of America: Northbern Illisionist University, 1979.
- Bachman, Lyle F. Fundamental Consideration in Language testing. New York: Oxford University Press, 1990.
- Bima, Bachtiar et.al. Bahasa inggris kelas X semester 1. Intan Pariwara, 2015.
- Brown, Douglass. Language Assessment Principles in Classroom Practices. San Francisco: San Francisco State University, 2000.
- Brown, Jeames Dean. Testing in Language Programs. New York: McGraw-Hill Companies Inc, 2005.
- Darmawan, Deni. Metode Penelitian Kuantitatif . Bandung: PT Remaja Rosdakarya, 2014.
- Dessy Wulansari, Andhita. Penelitian Pendidikan. Ponorogo: STAIN Ponorogo Press, 2012.
- Fromkin, Victoria. An Introduction to Language. Canada: Dryden Press, 1974.
- Gunning, Thomas G. Assessing and Correcting Reading and Writing Difficulties. Boston: Pearson education, inc: 2010.
- Harmer, Jeremy. The Practice of English Language Teaching. United States: Longman, 1998.
- Harmer, Jeremy. The Practice of English Language Teaching. United States: Longman, 1998.
- Heaton, J. B. Writing English Language Test. London: Longman: 2009.

- J Michael O'Malley, Lorraine Valdez Pierce. *Authentic Assessment for English Language Learners*. United States of America: Addison- Wesley publishing, 1996.
- Khan, Imran. reading assessment techniques among selected secondary school teachers in Pakistan: current trends and practices. *International Journal on New Trends in Education and Their Implications* October, November, December 2011 Volume: 2 Issue: 4 Article: 7 ISSN 1309-6249.
- Lado, Robert. *Language Teaching a Scientific Approach*. India: Tata McGraw-Hill Publishing CO.LTD, 1976.
- Latief, Mohammad Adnan. *Research methods on Language learning*. Malang: UM Press, 2014.
- Lucatoni, Peter. *Teaching and Assessing skills in English as a Second Language*. UK: Cambridge University Press, 2006.
- Mansyur Bennu, Agus Purnomo, *Language Assessment*. Jakarta: MGMP, 2009.
- Nunan, David. *Practical English Language Teaching First Edition*. New York: McGraw-Hill, 2003.
- Pang, Elyzabeth S. *Teaching Reading*. Chicago: Education Practice series University of Illuonis, 2003.
- Ramon Ziai, Niels Ott & Det mar Meurers, *Short Answer Assessment: Establishing Links between Research Strands*.
- Scheerens, Jaap. *Educational Evaluation, Assessment, and Monitoring*. Netherlands: Swets & Zeitlinger Publishers, 2005.
- Singh, Kultar. *Quantitative Social Research Methods*. New Delhi: Sage Publication India Pvt Ltd, 2007.
- Ur, Penny. *A Course in Language Teaching* . New York: Cambridge University Press, 1996.
- Widyaningrum Retno. *Statistika Edisi Revisi*. Yogyakarta: Pustaka Felicha, 2013.
- Widyaningsih, Nuning. *Kreatif Bahasa Inggris*. Klaten: Viva Pakarindo, 2013.
- Yuwono, Dolar. *Writing 1*. Yogyakarta: Pustaka Felicha, 2015.