

## **Analysis On Validity And Reliability Of The Reading Test At SMA Level In MBD**

<sup>1</sup>Anna Prilia Silueta, <sup>2</sup>R. F. J. Lekawael, <sup>3</sup>Fredy Meyer.

<sup>1,2,3</sup>English Education Study Program, Southwest Maluku, Pattimura University

Corresponding e-Mail: [siluetapriliana@gmail.com](mailto:siluetapriliana@gmail.com)

### **ABSTRACT**

This research aims to determine whether the reading test made by the teacher was content validity and reliability. To determine content validity and reliability, the researcher compares English tests with the syllabus, lesson plan, and students' test scores. This study used a descriptive qualitative method using document analysis in the form of a semester (summative test) and an interview with the teacher. This research collected data from two English teachers who teach at the second level at SMA NEGERI 13 MBD and SMA NEGERI 9 MBD. The results show that there are 20 multiple-choice questions about reading. From the questions, the researcher found that 17 questions were categorized by content validity, according to the syllabus and lesson plan. In contrast, three questions were classified as not having content validity. Based on the student's scores from the reading tests, the researcher found the scores reliable for the same tests given to the students in the second level and at various times. The reading tests made by the teachers were valid and reliable. From the interview results, it can be concluded that the teachers made a grid of questions before making the test. Tests made by the teachers must be related to the syllabus and lesson plan and according to the student's ability level. This shows that teachers make most of the English tests according to content validity and reliability.

*Keywords: reading, test, validity, reliability.*

### **INTRODUCTION**

Reading, one of the language arts, is a practice for thoroughly and efficiently comprehending the meaning of sentences. Nakamoto et al. (2008). Reading is a series of actions performed by a person. Reading can improve understanding of what is read by broadening knowledge and perspective. Reading is a technique that readers employ to obtain messages that they wish the author to impart to them through the use of words or written language, according to Tarigan (2015:7).

Reading comprehension tests assess your ability to read and understand written information quickly. The test is strictly timed and requires scanning passages and answering questions accurately. If you are preparing for a reading comprehension test, try the verbal reasoning test. Practicing these questions is also related to preparation (Chang, 2010; Vostal, 2011). This means that the reading test is essential to assess student's ability to know their comprehension archive in the reading test itself.

Reliability refers to the consistency of how a method measures something. A measurement is considered reliable if it consistently yields the same result using the same process under the

same conditions. Azwar (2011) said Reliability is derived from the word dependability and refers to the degree to which a measurement result is trustworthy, reliable, consistent, stable, and the comparison of several measurements of the same group of subjects. Relatively, the same results are obtained.

*Validity* is a word that refers to two things in assessment: the ability of the assessment to test what it intends to measure and the ability of the assessment to provide valuable and appropriate information for the intended purpose. According to Azwar (2011), validity means the extent to which a measuring instrument accurately performs its measuring function.

The researcher has two primary reasons for researching this title. First, this is the first researcher in this area who has researched this title. Second, based on the writer's preliminary study at SMA NEGERI 13 MBD, she found that the teacher gave reading text questions that needed to be shorter, which confused the students when reading and answering. Also, the questions given were not as taught, affecting the students' reading comprehension achievements.

Based on the previous explanation, the researcher was interested in analyzing validity and Reliability at the SMA level. In this case, the researcher intended to solve the problem through this paper entitled "Analysis on Validity and Reliability of the Reading Test at SMA level in MBD."

## **LITERATURE REVIEW**

### **Reading Test**

Reading Comprehension Testing evaluates pupils' ability to draw meaning from print English (Reza & Mahmood, 2013). In other words, reading evaluations are designed to provide feedback on the skills, processes, and knowledge resources that characterize reading ability. Reading assessments can inform researchers, teachers, administrators, and policymakers about kids' reading achievement by asking children to read grade-appropriate materials and answer questions about what they have read, Reeves (2012). However, we recognize that the difficulty of a reading test is at least a function of both passage and item difficulty.

Runge (2022) says an "item is a test question or test task; the difficulty of an item is measured simply by the proportion of candidates getting the answer appropriate compared with those getting wrong." Therefore, an item can be understood by the students if they have to understand what the questions mean so that the students can understand and answer the questions based on their knowledge. According to Arikunto (2013), the test aims to assess something in a prescribed manner and rules. In addition, according to Wahyuni and Ibrahim (2012:11), a test can be informed that it may be both a technique and tool in evaluation; hence, to conduct a test needs a set of tasks or questions; thus, it is known by the test instrument. It can be concluded that the purpose of a test is to assess the student's ability and comprehension of something they have learned.

There are several types of test items, among which are Multiple-choice and Essay Tests. *Multiple choice* is an objective assessment in which respondents must select only the correct responses from a list of possibilities. Mullen and Schultz (2012) state that the multiple-choice format is the most commonly utilized in educational testing. Essay tests are the best for evaluating higher-order thinking abilities, including analysis, processing, and assessing, since they provide pupils the most latitude to respond and create their responses. However, the most significant drawback of essays is how long they take to react to, evaluate, and identify writing potential.

Through essay tests, the teacher primarily assesses students' capacity for concept organization, expression, and defense Driana (2019).

### **Validity in the reading test**

The degree to which the test assesses what it was designed to measure is roughly described as validity, John (2015). He emphasizes that validity is typically predicated on a strong association between the study's objective and design, including a rational selection of criteria and variables, test measures, and other aspects. Furthermore, Hughes (2008) states that a test is deemed legitimate if it accurately assesses its intended measure. Besides that, Gafni (2016) mentions that there are substantial parts of validity, such as validity, which is a more critical quality any test can have; validity refers to the effectiveness of a test in measuring what it is intended to measure. Validity has many facets, takes many forms, and has been defined in many different ways. It can be demonstrated best by tryout after the test has been constructed.

### **Reliability in the reading test**

The reliability coefficient for a set of scores on an analogous reading exam acquired independently from students of the same group, E. Cho (2016). "Reliability refers to the consistency of test scores. A test cannot accurately measure anything unless it measures consistently", Bridgeman (2018). Furthermore, consistency of measurement is frequently described as reliability. There are several types of reliability in the tests. 1) Test-retest reliability: A reliability measurement is achieved by utilizing the same test twice over time to a group of people. The results from the test's stability over time can then be assessed by correlating two tests: 2) Parallel or alternate form reliability. This gauge of trustworthiness is discovered when a language teacher makes two versions of the same test by slightly altering the items. The correlation between Tests 1 and 2 results is used to describe reliability, 3) Item reliability. A reliability measure evaluates the correlation between test items that examine the same construct. The test items may not be dependable. The items may be too basic or complex, and they may need help distinguishing between intelligent and unintelligent students.

## **METHOD**

This study employed the descriptive qualitative method. According to Mohajan (2018), descriptive qualitative research signifies that the information gathered was in the form of words rather than numbers. This research used a survey to collect data, and the research instruments used were interviews and document analysis.

## **RESULTS**

### *1. How do validity and reliability aspects influence students' reading achievement?*

This discussion answers one of the research problems related to the test made by the English teacher at the second level. According to the data analysis, the test made by the teacher fulfilled the content of validity and reliability. This is valid and reliable because the question made

by the teacher in the form of multiple choice consisted of twenty numbers. Three questions were said to be invalid, and seventeen questions were said to be valid because they fulfilled the demand of the syllabus and lesson plan and such of indicators by Thoha (2003), and from the students' scores given by both English teachers, the researcher found that the scores were constants, because from the same reading tests done by the different students at different times but the results can be pretty constant. The reading tests were reliable.

Three questions were invalid, two were too easy, and the validity of this test was low. One other was ambiguous and not specific to understand, so the students could not get the appropriate answer from reading the text and the options. They differed from the syllabus, lesson plan, and other indicators Thoha (2003) described. From twenty multiple-choice questions and the student's scores given by the English teachers, the researcher found that the scores from the reading tests were reliable because the same questions were given to the different students at different times, and the scores were constant. Therefore, the reading tests are reliable.

Based on the data analysis, the researcher found that to assess the student's achievement in the reading test, the validity and reliability should fulfill the characteristics of a good test. The first characteristic of validity is the effect of measuring what is supposed to be measured. Each reading test the teacher gives must be according to what has been taught and what is in the lesson plan and syllabus. One may argue that validity concerns the degree of evaluation that is supposed to be measured in data or how well it fulfills the function for which it is being utilized (syllabus and lesson plan). Reliability is a necessary feature of any successful test and is essential to students' reading achievement. The researcher can infer that the reading test is practical if the students answer the question as accurately as possible unless the question is invalid and unreliable.

## *2. What are the factors that influence validity and reliability in reading tests?*

To answer the second research question in this study, the researcher found that based on the data collected from the teachers, some factors influenced the validity and reliability of reading tests, such as the syllabus and lesson plan. The tests should be in accordance with the syllabus and lesson, the reading text and the tests should be clear and easy to understand, and the reading text should not be longer so it can be easily read and answered.

Based on the analysis of the result of the interview with the English teachers, the researcher can conclude that the factors that influence validity and reliability in reading tests are as follows. Factors that can affect validity include the length of the test, reading vocabulary, and sentence structures that are too difficult and ambiguous. The factors that can affect reliability are the length of the test, difficulty in valuing items, test instructions, and students' scores.

## **REFERENCES**

- Athiyah Salwa, (2012). The validity, reliability, level of difficulty and appropriateness of curriculum of the english test.
- Brown, (2013). Chapter II. Test types. Perpustakaan Universitas Islam Riau.
- Brown. (2000). Principles of Language Learning and Teaching. San Francisco Longman. Brown. (2001). Teaching by Principle An Interactive Approach to Language Pedagogy. San Francisco : Longman.

Brown. (2004). *Principles and Classroom Practices*. New York : Pearson Education.

Brown & Priyanvada. (2010). *Language Assesment :Principles and Classroom Practices*. New York : Pearson Education.

Dewi Manalu, (2021). An analysis of students reading final examination by using item analysis program on eleventh grade of SMA N 8 Medan.

Ellen A. Drost, (2011). *Validity and Reliability in Social Science Research*. California State University, Los Angeles.

Indar, (2016). Analyzing the reliability and validity of reading test at the first grade SMAN 1 PATTALLASSANG.

James Dean Brown, (2000). *What is Construct Validity*. Questions and answers about language testing statistics.