Abstract

This study aims to (1) analyze the teacher-made tests so as to obtain an idea of the validity of teacher-made tests bahasa at SMA Negeri 1 Mawasangka, and (2) analyzing teacher-made tests so as to obtain an idea of the reliability of bahasa Indonesia teacher-made tests at SMA Negeri 1 Mawasangka. The method used in this research was analysis method of quantitative paradigm through the approach of evaluation with the object of study was bahasa Indonesia teacher-made tests at SMA Negeri 1 Mawasangka. The data collection was done by using the documentation by taking teacher-made tests that have been used in the midterms. Data were analyzed using product moment to calculate the validity of an essay test and the correlation formula point biserial to calculate the validity of an objective test, Kuder Richardson 20 (KR-20) to calculate the reliability of objective testing, and to calculate the reliability of the test essays using Cronbach Alpha, while the level of difficulty using the formula of proportion, and the different power using different power index analysis. This study found that: (1) the bahasa Indonesia teacher-made tests at SMA Negeri 1 Mawasangka did not qualify the validity of tests as shown in the results of the analysis of the validity of which 15 items were valid, level of difficulty grains only category of easy and medium, the different power point where 5 grains have the different power, and the persistence of test items that pengecohnya not functioning properly, and (2) the bahasa Indonesia teacher-made tests at SMA Negeri 1 Mawasangka yet did not qualify the reliability of the test as indicated by the results of the analysis of the reliability of the test, both forms of objective tests and test essay form. From the study result, it can be concluded that the bahasa Indonesia teacher-made tests at SMA Negeri 1 Mawasangka did not qualify the validity and reliability of the test was good.

Keywords: tests, bahasa Indonesia, teacher-made tests
A. Introduction

Bahasa Indonesia is one of the subjects that must be taught in every level of education in Indonesia, both at the level of basic education, medium, or high. One reason, the ability to speak (Indonesia) is the basic capabilities that must be possessed by every learner to transfer science and technology given that most of science and technology was "documented" in the form of references in a language medium. As a consequence of that, Junior High School (SMP) as one part of the basic education level, also incorporate these subjects into the curriculum, the Kurikulum Tingkat Satuan Pendidikan (KTSP).

To measure how far the objectives of teaching have been reached, can be done with the evaluation, in this case the evaluation of learning outcomes. In measuring instrument to evaluate the learning outcomes is to use the test. In its function as the actor of the students' learning, teachers should constantly keep abreast of the learning outcomes achieved by students from time to time. The information obtained through this evaluation is feedback to the teaching and learning process that can be used as reference in improving and enhancing the teaching and learning process in order to obtain maximum results. As executors of the evaluation of teaching and learning process, the teachers are required to have the ability in selecting and designing appropriate evaluation tools to be used in evaluating the teaching-learning process, and one of the evaluation tools are common and often used the teacher is a test.

Hopkins & Antes (1990: 130) stated that test is an instrument, tool or procedure that contains a number of tasks that must be answered by the students whose results can be used to measure a particular characteristic. Nitko (1996: 6) suggested that the test is an instrument or systematic procedure to observe or describe something or more characteristics of the students using the scale in the form of numbers or classification scheme. Sax (1980: 13) argues that the test is a task or set of tasks planned systematically to obtain representative information about the trait or attribute education or psychology.

The ability of teachers to devise a test that high quality is very important, because only the test of qualified teachers can measure precisely the success of the learning process that has been done and decisions taken on the results of these tests are also appropriate. The quality of teacher-made tests in schools is not known with certainty, whether the tests used in the evaluation of teachers teaching-learning process of the test is qualified or not, how the validity, reliability, level of distress, the different power and function distraktornya. This is because by teacher-made tests are not tested and analyzed the results to determine the quality of these tests before use, therefore if the test tested then is feared will be caught.

A person's ability to do something is very influenced by a knowledge test in the preparation of achievement test. Likewise, the ability of teachers in SMA Negeri 1 Mawasangka in devising tests of learning outcomes are also greatly influenced the evaluation of teacher’s knowledge and experience of teachers in preparing achievement test. Based on the fact, there are many teachers who do not perform the analysis of the tests that have been made both teachers who served in junior high or SMA/MA.

In order to determine the ability of teachers in preparing the test results to learn it is considered necessary for the research in order to obtain clear information, whether teachers in schools in various levels of education now have the ability in devising tests of learning outcomes, particularly the tests used in school exam, Measures of quality of teacher-made achievement test can be seen on the validity of the test, reliability test, have the different power tests, and the level of difficulty of the test.

B. Literature Review

1. Definition of Evaluation

Worthen dan Sanders (1973: 19) suggests that evaluation as a process of determining the value of an object of evaluation that consists of the acquisition of information to be used in determining the value of a program, product, procedure, or potential use of alternative approaches which are arranged for achieving a specific goal.

Menurut Waluyo (1987: 17) term of evaluation of education is only popular in the world of higher education, while in the community, especially the environment of elementary school to middle school is more popular with term of evaluation of learning outcomes. In general, the evaluation of learning outcomes is a process of determination of the teaching-learning process that has taken place. In particular, the evaluation of learning outcomes is an activity in an effort
to pay tribute to the students' development and progress in achieving teaching objectives have been designed.

Based on the understanding that the evaluation proposed by the experts above, it can be concluded that the evaluation is a systematic process of collecting, analyzing, and interpreting of data or information that the results are used in decision making. In the world of education, evaluation is defined as a process of assessing the success of teaching and learning process, in this case, the success of student learning and teachers teaching success. The success of this teaching process is an indicator of the success of the education. If student achievement is high, it indicates the quality of higher education, and vice versa that if student achievement is low, indicating that the low quality of education.

2. Validity

Validity means that the extent to which the precision and accuracy of a measuring instrument in performing its measurement functions. A test or measurement instrument is said to have high validity if the instruments measuring function properly execute or provide measuring results consistent with the intent to do such measurements. That is, the measurement result of the measurement is the amount of which precisely reflects the real state of facts or what measured (Sudijono, 1996: 23) is.

Validity is a product of validation. Validation is a process conducted by the author or user of the instrument to collect empirical data to support the conclusions by the score of the instrument, while the validity is the ability of a measuring instrument for measuring the measuring (Ahiri & Anwar, 2011: 243).

Arikunto (2012) suggested that validity is a measure that indicates the level of validity of an instrument. Valid instrument should have high validity. Contrary, invalid means having low validity. Test validity consists of three kinds, namely (1) the validity of the content, (2) the validity constructs, (3) the validity of the criteria.

3. Reliability

Reliability is the precision or accuracy of measurement in the assessment. (Sudjana, 1996: 16). Measuring results can be trusted when the number of times the measurement of the same subject will be obtained relatively similar results. Relative terms show that there is little tolerance for differences between the measurement results. If it was a big difference over time, the results of these measurements can not be trusted or not reliable. Measurements have high reliability are measurements that can produce reliable data.

C. Methodology

1. Place and Time of Research

This research was conducted in SMA Negeri 1 Mawasangka Central Buton regency. This research was conducted in March and April 2015. Selection of SMA Negeri 1 Mawasangka, as the location for this study based on the results of the pre-study conducted by researchers where the majority of teachers in these schools have not done the analysis of the quality of the test, so that from the findings of researchers interested in conducting research in SMA Negeri 1 Mawasangka.

2. Research variables

This research uses quantitative analysis paradigm through the evaluation approach. Through this research is intended to describe the quality of teacher-made tests Indonesian by analyzing the validity of the items, the reliability of the test, and level of difficulty, item difficulty index, the index of the effectiveness of different items and functions detractors. The variables of this research are the quality of the achievement test. Quality achievement test is measured through indicators: (1) validity, (2) the reliability of the test, (3) the level of difficulty grains, (4) different power, and (5) function detractors.

3. Population dan Sample

The population in this study is all the answer sheets teacher-made tests Indonesian who has been tested to class XII student IPS1 to XII IPS4 and class XII IPA1 to XII IPA3 by the number of test-takers is a class IPS1 until IPS4 each numbered 23 students and class XII IPA1 23 people, 23 people IPA2 XII and XII students IPA3 28 students so that the overall population of 166 students.

From the the existing population, the researchers drew a random sample or random sampling. This is in line with the Arikunto’s opinion (1986: 112) that sampling of a population
of less than 100 people, better taken all. If the subject is large, it can be 35-40% or more. Based on these opinions, then the sample is taken by 40% so that the number of samples after rounding is 66 students.

4. Technique of Data Collection

Data collection techniques in this study conducted with technical documentation, namely taking bahasa Indonesia teacher-made tests and the results of the students’ answers on the test subjects Indonesian teacher-made used at midterm, the next of the students’ answers are analyzed to determine the empirical validity test items, reliability test, the level of difficulty, different power and function tests detractors.

5. Technique of Data Analysis

The data was in the form of a score of the students’ answers on each item tests were analyzed to determine the level of validity as empirical validity, level of difficulty, different power, function humbug, and reliability to the test objective form whereas the test form of an essay analyzed with empirical validity, level of difficulty, different power and reliability. To calculate the level of validity, the product moment correlation formula was used for an essay test and the correlation formula biserial point for objective tests. Calculation of reliability of the objective form test used the formula of Kuder Richardson 20 (KR-20) while the form of an essay tests reliability used Cronbach Alpha formula.

D. Findings and Discussion

Discussion of results of research covers several aspects as stated below.

1. Test validity

The results of empirical validity analysis showed that of 20 test items objective form, the 15-point or 75% declared invalid, while 25% 5 grains or more declared invalid. Thus, when viewed from the aspect of the empirical validity of the test, the Indonesian teacher-made tests at SMA Negeri 1 Mawasangka not meet one of the requirements of quality tests. While the essay test of five test items tested, then to 5 butirnya declared invalid empirically. Quality test is a test that items are valid. Validity is very important for the items of test for validity is the ability of a test to measure the actual functions. The test that the items are invalid means such tests can not perform measuring function properly so that the measuring results obtained from the tests are not valid cannot be used as a decision-making about the success of student learning. Conversely, a valid test means that the test has been doing measuring function properly so that the measurement results can be used as decision making on the success of student learning. This is in accordance with the opinion of Sudijono (1996: 23) who argued the validity is the extent to which the precision and accuracy of a measuring instrument in doing measuring function. A test or measurement instruments said to discount high validity if the instruments measuring function properly execute or provide measuring results consistent with the intent to do such measurements. This means that the measuring result of the measurement is the amount of which precisely reflects the real state of facts or what is measured.

The results of the analysis of the level of difficulty showed that the distribution of difficulties grains teacher-made tests Indonesian at SMA Negeri 1 Mawasangka to test objective form only consists of two categories, namely easy items existing 7 items or 35%, and the category average items existing 13 items, or 65% and not one of the items that referred as difficult category. Similarly, the essay test, then the test of items is 3 test items that is classified as easy items category and 2 items were average items.

The results showed that when viewed from the aspect of the distinguishing features of 20 test items made by Indonesian teachers at SMA Negeri 1 Mawasangka. Then only 13 or 65% of test items were categorized as good, while 7 test items or 35% is classified as an items of tests that are not good because it does not have the ability to distinguish between students who study and students who have not learned that all seven test items should be discarded or revised, either partial or total revision.

An examination of the functioning distractors Indonesian teacher-made tests at SMA Negeri 1 Mawasangka not meet the criteria of a good test or quality test because the test is qualified test items that the distractors are functioning well. Criteria humbug function properly according to the experts’ evaluation is that if every man confidence chosen by 5% of test participants. According to Ahiri (2007: 327-328) that a distractor works properly chosen at least five participants from the top group and preferred by a lower group. According to Azwar (1998:
151) argue that the alternative answer is a good distractor when the distractors at least been chosen by 5% of all test participants.

Based on the findings of research and discussion, it can be said that the Indonesian teacher-made tests at SMA Negeri 1 Mawasangka not meet the standards valid, whether viewed from the aspect of content validity, empirical validity, difficulty level aspects, aspects of different power and functionality aspects of detractors. As previously stated by experts on criteria test items that either have not been achieved in test items were examined. So that the measurement results used by the Indonesian teacher-made tests at SMA Negeri 1 Mawasangka cannot be guaranteed to be good and true measure student learning outcomes.

2. Test reliability

Based on the analysis of reliability Indonesian teacher-made tests, either the test or the test objective form essay form, it is known that reliabilitas coefficient of 0.722 for the test is an objective test and 0.427 for an essay test. Thus, the Indonesian teacher-made tests are not reliable. Determination of criteria reliabilitas 0.75 refers to the opinion of Naga (1992: 129) that the reliability coefficient of 0.75 is considered reliable, but if calculated using Kuder and Richardson (K-R 20). The test is declared reliable if the $r_{count} \geq 0.75$. If the $r_{count} \leq 0.75$, then the tests are considered unreliable.

Indonesian teacher-made tests at SMA Negeri 1 Mawasangka, both objective form and the form of essays are not qualified when viewed from the aspect of reliability. This result means that the Indonesian teacher-made tests can not show the consistency of measurement. Moreover, it also does not show the accuracy of the measurement. A good test or quality test is a test that is reliable, because the test is reliable means reliable measure of student learning outcomes consistently. Haryati (2007: 17) argues that a test must meet the criteria of reliability, it relates to the consistency of the measurement results. Such assessments enable reliable comparisons and ensure consistency. Sutikno (2009: 119) states that an evaluation tool has demonstrated the reliability of reliability when measuring results. Reliability is the precision of the measuring instrument (Kerlinger: 1986: 405).

E. Conclusion

Based on the results of the discussion that has been described, it can be concluded as follows:

1. There are a few items of Indonesian teacher-made tests shaped invalid so that the objective of the content validity of the test items can not be used as a measurement of student learning outcomes. This is because the test items made by Indonesian teachers are said to have weaknesses, that the construction is not correct and incorrect answer key.

2. Indonesian teacher-made tests are not yet qualified good test or tests that have qualified from the aspects of the validity of the test. It can be seen from the results of the empirical analysis, level of difficulty, different power and function of distractor in which not all of the items having distinguishing the good, and not all function well its distractors.

3. Indonesian teacher-made tests are not yet qualified as good test or tests that have qualified from the aspects of reliability of the test. It can be seen from the analysis of reliability of the test, which only reached 0.712 for the test objective form, and 0.497 for the test essay.

4. Empirically, those items of Indonesian teacher-made test that were applied to students largely cannot be used as a tool to measure student learning outcomes as the items of test is not yet qualify as a quality test.

F. References


