

Validity and Reliability of Teacher-Made English Summative Test at Second Grade of Vocational High School 2 Palopo

Dewi Furwana

State Islamic Institute of Palopo, Indonesia

Email: dewifurwana@yahoo.com

Abstract

This research aimed to find out the validity and reliability of the English summative test for the second grade of Vocational High School 2 Palopo. The research applied mix-method design (Qualitative-Quantitative). The researcher analyzed the data using validity and reliability formula by Suharsimi Arikunto and Kuder-Richardson's. There were 40 multiple choice questions of English summative test. This research used expert judgement as validator to analyzed and commented the validity and reliability of English summative test made by the English teacher at second grade of Vocational high school 2 Palopo. The result indicated that the test items of English summative test at second grade of Vocational High School 2 Palopo had good content validity. Based on the percentage, the researcher found that the English summative test was 80% valid in terms of its conformity with the indicators. The items tests that represent the suggested indicators in the syllabus were 32 items and the inconformity items were 8 items. The reliability of English summative test was 0.6132. It means that the reliability of English summative test was in classification into the high level (0.60-0.799).

Keywords : English summative test, validity and reliability.

INTRODUCTION

There are four main skills in learning English. They are listening, speaking, reading and writing. To conduct an effective Teaching Learning Process (TLP), there are some important things that should be fully attention, for examples the teacher, curriculum, syllabus, method, facility, test, etc. According to Brown that the function of test is measure a person's ability, knowledge, and performance (Brown, 2003). Test is one of the things that must be focused in this research.

Testing and teaching are closely interrelated to each other because the success of teaching cannot be measured and known without conducting a test. So if it is related to the Teaching Learning Process (TLP), it means that test is an instrument or procedure used to measure the students' ability, to diagnose the

students' weaknesses, to get educational decision and it depends on the kinds of test conducted. Good test items should be made by considering some criteria such as reliability and validity for students.

The results of the test are used to improve the teaching and learning process and taken into account in determining grades.

The researcher did the research about English summative test which conducted at second grade of Vocational High School 2 Palopo. That there are a lot of criteria of a good test, one and other are interrelated so that with fulfill that criteria the teacher would get a good result and effective test. In this research, the researcher focused on the validity and reliability of a test, especially a content validity.

A good test should be valid and reliable. Heaton stated that The validity of a test is the

extent to which it measures what it is supposed to measure and nothing else (Heaton, 1095). The validity of a test must be considered in measurement in this case there must be seen whether the test used really measures what are supposed to measure.

Validity also stated by Hughes "the validity of a test. For it to be valid at all, a test must first be reliable as a measuring instrument" (Hughes, 2003). The test should measure what the teacher wants to measure. For example, if the teacher wants to measure the grammar ability, the teacher should give the text in form of written test, not giving oral form or recording to listen. Commonly, there are three kinds of validity. There are content, criterion-related (concurrent and predictive) and construct. From the three kinds of validity, content validity has the important roles in interpreting the test as a tool of evaluation, so that the teacher can measure student's ability effectively. Content validity depends on careful analysis of the language being tested and of the particular course objectives. The test should be so constructed as to contain a representative sample of the course (Heaton (1998). It can be understood that the content validity needs a sharp and systematic analysis because it can represent the content of the test that will be examined. The researcher will explain the content validity in the next chapter. While reliable refers to the consistency of score. For example if the same group of students took the same test twice within two days without reflecting on the first test before they sat it again-they should get the same results on each occasion. If they took another similar test, the results should be consistent.

Based on the researchers' observation at second grade of Vocational high school, the researchers' found that there are many test made by the English teacher it was copied from the Internet, the way of the teacher make the test, does not consider with the syllabus and curriculum. Based on the problem above, the

researcher interested to analyze reliability and validity of English summative tests made by the English teacher for the second semester of the second grade of Vocational High School 2 Palopo. By giving a good test, the students have an opportunity to get a good quality in learning and the result of the test are used to improve the teaching and learning process and taken into account in the determining grade.

Based on the background, the researcher formulated the research question namely: How is the validity and reliability of English summative test made by the English teacher at second grade of Vocational High School 2 Palopo?

METHODOLOGY

This researcher used mixed-method (Qualitative-Quantitative). The research applied mix-method design (Qualitative-Quantitative). The researcher analyzed the data using validity and reliability formula by Suharsimi Arikunto and Kuder-Richardson's. There were 40 multiple choice questions of English summative test. This research used expert judgement as validator to analyzed and commented the validity and reliability of English summative test made by the English teacher at second grade of Vocational high school 2 Palopo

FINDINGS AND DISCUSSIONS

Description of Data

The findings of the research were showed to describe the result of the data that analyzed statistically. It comprised of the students' score in validity and reability in making test with criteria of a good test.

The researcher analyzed the test materials whether they conform to the English syllabus of Vocational High School 2 Palopo. They were 450 students in the second grade of Vocational

High School 2 Palopo at 2017/2018. During the English summative test in the second grade of Vocational High School 2 Palopo, the students were divided into eight programs, the researcher took not all of students, there were two class took as an example to analyzed the Validity and Reliability of English Summative test. To get the data, the researcher took the questions sheet, answer sheet and the English syllabus document to be analyzed. The data that the researcher used in this research was the English summative test, the data was made by the English teacher at second grade which was called *Ulangan Akhir Semester Sekolah Menengah Kejuruan (SMK) Tahun Pelajaran 2017/2018* for

the second grade students of Vocational High School 2 Palopo. The total numbers of test items were 40 tests which consist of 40 multiple choices. The test was held on Wednesday, 5th June 2018 with the given time 120 minutes. Furthermore, this research was support by expert judgement.

Interpretation of Data

In this research, the researcher discussed the total frequency of the conformity and the inconformity of English summative test items to the syllabus based on the data of item analysis result above. Based on the data of item analysis result, can see the table as follow:

Table 4.1 The Analysis Result of the Conformity of English Final Test Items (Summative test) and the English Syllabus

Kurikulum 2013	Number of items	Total of items	Topic
Based on Syllabus			
Analyzing social function, text structure, linguistic to inquire and state about assumption if happen an event in the future, accordance to context of its use.	1,2,3,4, 5, 8,9,10	8 items	Conditional sentence
Analyzing social function, text function and linguistic some transactional text and writing about asking and giving information, related condition/ action/ activity/ accurence without mentioning the subject of scientific text, according to context of its use. Linguistic observe (passive voice).	6, 7	2 items	passive voice
Analyzing text structure and linguistic to implement text social function "factual report" to inquire and state about scientific text factual about human, animal, thing, indication and nature and social, simple, according to the context of learning in other subject.	11,12, 15,16,1 7, 20	6 items	Factual Report
Analyzing social function, text structure and linguistic on text biography of figure according to the context of its use.	21,22, 25,26, 28, 29, 30	7 items	Biographic of personage
Analyzing social function, text structure and linguistic some new items text oral and written with asking and giving information related simple news from	31,32	2 items	Advertisement

newspaper, television, radio according to the context of its use.			
Analyzing social function, text structure and linguistic interaction text, social transaction which involve action of giving and asking information related to written report.	33,34,3 5,36	4 items	Form of Report
Analyzing social function, text structure and linguistic some special text in memo, menu, schedule, and signs by giving information related school activities or place of work, according to the context of its use in working world	37,38,3 9,40	4 items	Memo
Total	32 items		

The question number 1, 2, 3, 4, and 5 were classified into Conditional sentences. Conditional sentence is one of the genres of English text that should be learnt by students. Therefore it based on the syllabus and this question was valid item to be tested.

The question number 6 and 7 were classified into Passive voice and it based on the syllabus. In this question, must be revised because the options focused on the vocabulary.

The question number 8, 9 and 10 were classified into Conditional sentences but in this question, the teachers made by Error recognition form. Therefore it based on the syllabus and valid item to be tested.

The question number 11, 12,15,16,17, and 20 were classified into Factual Report. Factual report is one of the genres of English text that should be learnt by students. The question number 13 and 19 covers the grammar aspect of pronoun reference. Therefore the question based on the syllabus.

While the question number 14 and 18 must be revised because focused on the closest meaning vocabulary.

The question number 21,22,23,24,25,26, 27,28,29,and 30 were classified into biographic of the personage and it based on the syllabus. Therefore this item was valid item to be tested. While the question number 23, 24, and 27 must be revised because focused on the meaning of vocabulary. And question number 29 has incorrect writing and must be revised to "Aceh" not "Bali".

The question number 31 and 32 were classified into Advertisement and it based on the syllabus. Therefore this item is valid item to be tested.

The question number 33, 34, 35, and 36 were classified into Form Report and it based on the syllabus. Therefore this item was valid item to be tested.

The question number 37, 38, 39, and 40 classified into Memorandum and it based on the syllabus. Therefore this item was valid item to be tested.

Based on the table above, the researcher found that there were 32 questions which conformity with the English summative test and the English Syllabus.

Table 1 The Analysis Result of the Inconformity of English Final Test Items (Summative test) and the English Syllabus

No	Kurikulum 2013	Number of items	Total of items	Topic
1.	Analyzing text structure and linguistic to implement text social function "factual report" to inquire and state about scientific text factual about human, animal, thing, indication and nature and social, simple, according to the context of learning in other subject.	13, 19	2 items	Covers the grammar aspect of pronoun reference.
2.	Analyzing text structure and linguistic to implement text social function "factual report" to inquire and state about scientific text factual about human, animal, thing, indication and nature and social, simple, according to the context of learning in other subject..	14,18	2 item	The closet meaning vocabulary
3	Analyzing social function, text structure and linguistic some new items text oral and written with asking and giving information related simple news from newspaper, television, radio according to the context of its use.	23,24,27	3 items	Meaning of Vocabulary
4.	Analyzing social function, text structure and linguistic some new items text oral and written with asking and giving information related simple news from newspaper, television, radio according to the context of its use.	30	1 item	Covers the grammar aspect of pronoun reference.
Total			8 items	

Based on the table above, the researcher found that there were 8 questions which inconformity with the English summative test and the English Syllabus.

Based on the presentation above, the researcher found that there were 80% question which conformity with the English summative test and the English Syllabus then there were 20% question which inconformity with the English summative test and the English Syllabus. The teachers of vocational high school 2 Palopo know the way how to contract a good test. They recognized that the summative test must be based on the syllabus.

Analysis of Data

In this research, the researcher would like to analyze the data from the table presentation. The researcher would like to displayed how to analyze the data by used the steps as follow: 1) The researcher calculated the right answers of each item, 2) Calculated the score of the right answers of each item. 3) The researcher would tabulate the students' result in a well-arranged form. Well-arranged form means the data presented in table. In order to keep of the privacy of the students, the names of the students would be changed into codes. The table made as follow: (Appendices 4).

Test Validity

The validity test question made by the English teacher at second grade of Vocational High School 2 Palopo year 2017/2018 as following presentation data the question used in writing test. The test question in the form of multiple choice question can be done used Suharsimi and Arikunto formula. In analyze the items qualitatively, used the review form would be helpful and facilitated the implementation procedure. The problem review format is used as a basis for analyzed each item.

In this research, multiple choice test items made by the English teacher at second grade of Vocational High School year 2017/2018 based on the validity of summative test made by the English teacher is 32 questions which conformity with the English syllabus. Based on the data presentation above, that the result of content validity of English summative test made by the English teacher is 82 and there was 8 questions which inconformity with the syllabus. Based on compared in percentage with Arikunto's criteria the validity of English summative test had a good content validity

Test Reliability

The researcher analyzed the reliability of question. Reliability referenced to the stability of multiple choice test values. In this research, the researcher used the Kuder-Richardson's formula (KR-20) in the Arikunto's book to determine the reliability coefficient of the 40 multiple choice test items.

Based on the presentation data above, the reliability of English summative test made by the English teacher was 0,6132 it means that the reliability of English summative test was in classification into the high level (0.60-0.799). (Appendices 5 and 6).

Discussion

Validity

Based on J. B. Heaton "the validity of a test is the extent to which it measure what it is supposed to measure nothing else" A test can be said as a high validity or can be declared valid, if the scores on the item concerned have a match or alignment with the total score, or with statistical language: There was a significant positive correlation between the item score with the total score. Total score here was based as dependent variable, while item score is domiciled as variable independent variable.

Thus, to arrive at the conclusion that the points that want to know validity was valid or not, can be used correlation techniques as a technique of analysis. A question matter can be declared valid if the corresponding grain score proves to have a significant positive correlation with its total score (Anas).

Indeed, the validity and reability need to be known for achievement. The good test criteria in this test were just the requirements of making the test, commonly used in making tests, multiple choice tests in this test was the type of test in making test, important parts that need to be before practice immediately.

There were differences and similarities result in research of the researcher from the previous related research findings that Masrurroh with the title "An Item Analysis on English Summative Test for Second Grade Students of Man Tulungagung 1 in Academic Year 2013/2014" research described that her test has some flaws in content validity and construct validity. In term of content validity, not all the language skills were tested in the test, listening was not tested at all. Besides, the materials in syllabus were not fully included in test items. Related to the construct validity, some of techniques used to test language skills were not relevant with the language testing theory especially writing and speaking, because these two skills were tested by used multiple choice test. The test was categorized to have fair reliability, where the test coefficient was 0.677.

While Noviyanti with the title "an analysis for content validity on English summative test for the first year students of Junior High School N 87 Jakarta" her findings of the research started that the content validity of English Summative test at SMP N 87 Jakarta for even semester of second grade student reach good validity. Areta Wulandari with the title "An Analysis on the Content Validity of English Summative Test Items at the Even Semester of the Second Grade of Junior High School" The finding of the research prove that the test items of English summative test for the even semester of the second grade students in *Al-Amanah* Junior High School have bad content validity.

Based on the percentage, the researcher found that finding of Haril Sumrotul and Wulandari it is different with finding of the researcher. The researcher found the English summative test is 80% valid in terms of its conformity with the indicators and it has good content validity same with Noviyanti's finding.

The items tests that represented the suggest indicators in the syllabus are 32 items and the inconformity items are 8 items. While the researcher found that the Summative test which administered in the second grade of English summative test of Vocational High School 2 Palopo has a good content validity.

Reliability

Reliability is the consistency of test scores across facets of the test (Anas). Reliability is necessary characteristic of any good test: for it to be valid at all, a test must first be reliable as a measuring instrument (www.socialresearchmethod.snetsnet/kb/relytypesphp/). If we give the same test to the same student or matched students on two different occasions, the test should produce similar results.

Based on the percentage, this research focused on test reliability. The reliability test analyzed display that the computational result of the KR.20 score was had a fair reliable. The

result of reability multiple choice test was 0.6132% high with 20 students.

There were some different this research with previous related findings. Wulandari and Noviyanti were not focused on reliability test but focused on the content validity. While Masruroh got the result it was had fair reliable different with this research, the researcher got the result it was had high reliable. The researcher, provided expert judgment to analyze the result of the researcher, after analyzed the expert got the same result with the researcher. The expert concluded that English summative test at second grade of vocational high school 2 Palopo had the high reliability.

CONCLUSION

The validity of English summative test made by the English teacher, the researcher found 80% valid items and it has reached a good content validity. Based on the data analysis and data interpretation, in the previous chapters, the researcher concluded that the Summative test for the second grade students of Vocational high school 2 Palopo. Academic 2017/2018 held on May 2018 had a good content validity.

The reliability of English summative test made by the English teacher, The test result reability multiple choice test was 0.6132% it meant that the reliability of English summative test was in classification into the high level (0.60-0.799).

This research provided expert judgment to analyzed the result validity and reliability of this research, and based on the expert judgment in the content validity it was not suitable with the researcher result. While in test reliability, expert judgment found that this research was reliable with researcher result.

The researcher used Mix method to design this research. The researcher used qualitative descriptive to designed validity of English

summative test, and used quantitative to designed reliability of English summative test.

REFERENCES

- Adrian, S.P., & Peter, J.M.G. (1981). *The Construct Validation of Test of Communicative Competence*, TESOL 202. Wangiston D.C.: Transit Building Georgetown University.
- Airasian, P.W. (2012). *Classroom Assessment; Concepts and Applications* (7th ed.). New York: McGraw-Hills companies.
- Arthur, H. (1989). *Testing for Language Teachers* (2nd ed.). Cambridge: Cambridge University Press.
- Bill, R.G., & Ernest P. W. (1974). *Application of Pupil Assessment Information for the Special Education Teacher*. Colorado: Love Publishing Company.
- Brown, H.D. (2003). *Language Assessment Principle and Classroom Practices*. California: Longman.
- Fred, G., & John A.U. (1996). *Classroom-based Evaluation in Second Language Education*. New York: Cambridge University Press.
- Fulcher, G., & Fred. D, (2007). *Language Testing and Assessment; an Advanced Resource Book*. Canada: Routledge.
- Gay, L.R., & Airasian. P. (2000). *Educational research: Competencies for analysis and application* (6th ed). Upper Saddle River, NJ: Prentice Hall.
- Georgia, S.A. (1964). *Measurement and Evaluation in Education, Psychology, and Guidance*. Los Angeles: Hold, Rinehart and Windston.
- Heaton, J.B. (1975). *Writing English Language Test*. New York: Longman.
- Hughes, A. (2003). *Testing for Language Teachers*, (2nd ed.). UK: Cambridge University Press <http://www.statisticshowto.com/kuder-richardson/> access on Wednesday, 18 July 2018
- Jack, R.F., & Norman E.W. (2003). *How to Design and Evaluate Research in Education* (Fifth ed.). New York: McGraw-Hill.
- Heaton, J.B. (1998). *Writing English Language test*. London and New York: Longman.
- J. Charles A., et. al. (1995). *Language Test Construction & Evaluation*. Cambridge: Cambridge University Press.
- Kenneth, D. H. (1998). *Educational and Psychological Measurement and Evaluation*. Boston: Allyn & Bacon.
- Mc, Namara. (2000). *Language Testing*. Oxford: Oxford University Press.
- Masruroh, H.Z. (2014). An Item Analysis on English Summative Test for Second Grade Students of Man 1 Tulungagung, (Tulungagung, English Education Program. State Islamic Institute (IAIN)
- Noviyanti. (2011). An Analysis on the Content Validity of the Summative Test for the First Year Student; A case study of SMP N 87 Jakarta, (Jakarta Syarif Hidayatullah State Islamic University)
- Penny, Ur. (1996). *a Course in Language Teaching: Practice and Theory*. Cambridge: Cambridge University Press.
- Peter, W. A, & Michael, K.R (2008). *Classroom Assessment*. New York: Beth Mejia.
- Riduan. (2004). *Metode dan teknik Menyusun Thesis*, Alfabeta: Bandung.
- Rosenthal, R., & Rosnow, R. L. (1991). *Essentials of Behavior Research: Methods and Data Analysis*. (2nd ed.). McGraw-Hill Publishing Company
- Sugiyono. (2014). *Metode Penelitian Kuantitatif Kualitatif dan R&D*. Bandung: Alfabeta.
- Sugiyono. (2013). *Metode Penelitian Kombinasi (Mixed Methods)*. Bandung: Alfabeta.

- Arikunto, S. (1999). *Dasar-dasar Evaluasi Pendidikan* (Rev. ed.). Jakarta: Bumi Aksara.
- William, W., & Stephent G. J. (1990). *Educational Measurement and Testing Boston*: Allyn & Bacon.
- Wilmar, T. (1988). *Evaluation of Student Achievement*. Jakarta: Departemen Pendidikan dan Kebudayaan.
- Wulandari A. (2014). An Analysis on the Content Validity of the Summative Test Items at the even Semester of the Second Grade; a Case Study of Mts Al-Amanah, Syarif Hidayatullah State Islamic University.

