

JEPAL

Journal of English Pedagogy and Applied Linguistics
<https://jurnal.masoemiversity.ac.id/index.php/englishpedagogy>

Submitted: 6 January 2022 Accepted: 28 February 2022 Available Online: 28 February 2022

AN ANALYSIS OF CONTENT VALIDITY ON SUMMATIVE TEST FOR THE 12th GRADE STUDENTS

Esti Puspitasari

SMAN 1 Cikarang Utara
estipublicemail@gmail.com

Abstract: Content validity is one important thing in evaluating test items. However, based on preliminary observation, the fact, some teachers did not consider the aspect of validity in the summative test, especially content validity. Hence, this study aims to analyze content validity on summative tests in the twelfth grade at public senior high school. The data were collected through document analysis and interviews. The result revealed that the summative test was administrated in the twelfth grade at senior high school has reached poor content validity. This finding hopefully helps the teacher to know how to design the summative better. It is recommended that further studies are conducted with lots of participants and larger classes.

Keywords: content validity, summative test, testing

INTRODUCTION

Testing is one of the crucial parts of evaluating the teaching-learning process. To provide the information about the achievement of a group of learners, giving some kinds of tests is necessary, without which it is difficult to see how rational educational decisions can be made (Hughes, 2003). Besides, the test is constructed primarily as a device to reinforce learning and to motivate the student or primarily as a means of assessing the student's performance in the language (Heaton, 1995). There are two types of testing, formative and summative tests. To know final the result of the teaching-learning process whether it is successful or not, a summative test can be the alternative.

A Summative test is the process of summing up or checking what has been learned and it is also administering at the end of teaching-learning process (Weeden, Winter & Broadfoot, 2002; Hughes, 2003). It is also used to evaluate the effectiveness of instructional programs and services at the end of an academic year or at a predetermined time (Fisher & Frey, 2007). In making good tests, teachers should consider the validity of the test. Validity is the extent to which tests measure what it is intended to measure (Alderson, Clapham, & Wall, 1995). It is a crucial factor in making a good test because the test is not only valid unless answering such a question like how do you know and for what is it valid? (Hughes, 2003).

According to Alderson, et.al. (1995), validity of a test has been divided into two: external validity and internal validity. External validity consists of concurrent and

predictive validity. Concurrent involves the comparison of the test scores with some other measure for the same candidates taken at the same time as the test (Alderson, Clapham, & Wall, 1995). Predictive validity is different from concurrent validity. Predictive validity will be gathered after the test has been given (Alderson, Clapham, & Wall, 1995).

Then, internal validity consists of face validity, response validity, and content validity. Face validity is the test's surface or public acceptability (Ingram, 1977). Response validity is the range of qualitative techniques where test takers are asked to report how they respond to test items (Henning, 1987). Last, content validity is the representativeness of sampling adequacy of the content- the substance, the matter, the topics-of a measuring instrument (Kerlinger 1973: 458). It includes gathering the judgment of 'experts': people whose judgment one is prepared to trust, even if it differs with one's own.

In fact, some teachers did not consider the aspect of validity in the summative test, especially content validity. In line with the previous research, (Fauzi, 2011) see also (Sultana, Shamim, Nahar, & Hasan, 2009; Siddiek, 2010; Adiredja, 2012; Haynes, Richard, & Kubany, 1995) said that the material of the English examination test is not appropriate on the recommended English syllabus. Hence, this test is needed to be improved. As stated by (Hughes, 2003) that there are two sources of inaccuracy tests, content, and techniques. In addition, there is another reason why the test did not consider the procedure in the guided book (Sumardi, 2008 cited in Nofiyanti, 2011). First, the time allotted in designing and adequate. Second, the teachers' competencies were still limited in designing a good test. Last, the teacher might have insufficient experiences in designing and constructing the English test.

Regarding those problems, the purpose of this study is to analyze the content validity of summative tests at third-graders at public senior high schools in Bekasi. This study is expected to contribute to the English teachers and further researchers. First, this study was expected to give meaning for teachers of English in developing a test that has content validity. Then, this study was expected to enrich the teacher's knowledge of the English proficiency test. Next, this study was expected to be a reference for the next researchers who are interested in developing a similar study.

LITERATURE REVIEW

A Summative test is the process of summing up or checking what has been learned and it is also administering at the end of teaching-learning process (Weeden, Winter & Broadfoot, 2002; Hughes, 2003). It also can be called an assessment of learning (Weeden, Winter, & Broadfoot, 2002). For teachers, it is for gauging their progress toward the course or grade level goals and benchmarks, whereas, for students, it is for grade and promotion. It is also used to evaluate the effectiveness of instructional programs and services at the end of an academic year or at a predetermined time (Fisher & Frey, 2007).

A summative test is more structured and more standardized than a formative test. A summative test is used for placing entering the next semester or the next grade. This test must be considered on syllabus and learning objectives. Therefore, it can be used for helping learning (Gardner, Harlen, Hayward, Stobart, & Montgomery, 2010). The result of this test can be used for feeding back into teaching (Gardner, Harlen, Hayward, Stobart, & Montgomery, 2010). In addition, in the formative test, validity does not become an issue, but in the summative test, validity becomes an issue.

In School-Based Curriculum (KTSP) implementation, teachers are guided to design their syllabus and material objectives based on the curriculum provided by the government (Ramadhana, 2010). Ramadhana (2010) also stated that in measuring the

students' achievements of the learning objectives, the teachers also design their test instruments. In making good tests, teachers should consider reliability, practicality, and validity. Reliability occurs when a test is given, the test user would like some assurance that the results could be replicated if the same individuals were tested again under similar circumstances (Crocker & Algina, 1986). Then, practicality is concerned with test implementation rather than the meaning of test scores. (Fulcher & Davidson, 2007). Validity is the extent to which tests measure what it is intended to measure (Alderson, Clapham, & Wall, 1995). Among the three methods, validity is a crucial factor in making a good test because the test is not only valid unless answering such a question like how do you know and for what is it valid? (Hughes, 2003). The types of validity (criterion-related, construct, and content) are simply considered to be convenient categories for accumulating evidence to support the validity of an interpretation. (Gronlund N. E., 1993). However, (Alderson, Clapham, & Wall, 1995) also stated that the validity has divided into two: external validity and internal validity.

External consists of concurrent and predictive validity. Concurrent involves the comparison of the test scores with some other measure for the same candidates taken at the same time as the test (Alderson, Clapham, & Wall, 1995). Predictive validity is different from concurrent validity. Predictive validity will be gathered after the test has been given (Alderson, Clapham, & Wall, 1995). Then, internal validity consists of face validity, content validity, and response validity. Face validity is the test's surface or public acceptability (Ingram, 1977). Response validity is the range of qualitative techniques where test takers are asked to report how they respond to test items (Henning, 1987).

Last, content validity is the representativeness of sampling adequacy of the content-the substance, the matter, the topics-of a measuring instrument (Kerlinger 1973: 458). It includes gathering the judgment of 'experts': people whose judgment one is prepared to trust, even if it differs with one's own. In other words, content validity could also be established in two stages, the development and judgment stage (Yaghmale ., 2003). The test is said to have content validity if its content constitutes a representative sample of the language skills, structure. A comparison of test specification and test content is the basis for judgments as to content validity (Hughes, 2003). This kind of validity depends on a 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, the relationship between the test items and the course objectives always being apparent (Heaton 1995 p. 160). There is a strong tendency, especially in multiple-choice testing.

METHOD

This study is a descriptive study. "Descriptive study describes and interprets what it is concerned with conditions or relationships that exist, an opinion that is held, a process that is going on, an effect that is evident or trends that are developing" (Best & Kahn, 2006). It means that words or pictures are frequently used in this study rather than numbers (Cresswell, 2008).

This study was conducted in one of the public senior high schools in Bekasi. The reason of the selecting school as the study setting as follows. First, the accessibility, this school provided easy access to the present study. Second, the feasibility of the site, this school was near to the present study's place. It makes the present study easier to conduct the study (Bogdan & Biklen, 1998) , and the last is the availability of the participants, the participants were available for this study. The participants of the

N=number of sample

The percentage will be represented with the criteria adopted from (Arikunto, 1992):

76-100% = Good

56-75% = Sufficient

40-55% = less good

≤40% = Bad

The data analysis was conducted in two steps: first, the data gained from the document, and second, at the end of the research, the data gained from the interview. There are several steps in analyzing the data from documents. The first is reading all documents. The second is selecting the documents related to the research questions. The third is merging the document selected. The fourth is that, a summative test was analyzed with formula and syllabus. The result was correlated with the data from the interview.

The data from the interview were analyzed through several steps as follows: First, transcribing the data from the interview. Second, categorizing the data from the interview. Third, reducing inappropriate data and interpreting the data from an interview in accordance with the research questions.

FINDINGS AND DISCUSSIONS

This chapter is going to describe the findings based on the research question. The analyses of those findings cover one research question. The research question was presented as follows: “does summative test relate to the syllabus?” However, not all data was presented in this chapter. Thus, the complete data are presented in appendices.

The Content Validity on the Summative Test

Summative is the process of summing up or checking what has been learned at the end of a particular stage of learning (Weeden, Winter, & Broadfoot, 2002). In other words, the summative test is administered at the end of a course of study (Hughes, 2003). The summative test was conducted on Thursday, December 1st, 2014. The summative test consists of 30 items, it contains multiple-choice which are consist of reading comprehension, cloze test, and intensive writing. Based on the data gained from the interview, those items were gained from some sources, neither teacher’s book nor student’s book. However, unfortunately, in making this summative test, the teacher did not make test specifications and question cards. The reason was the test specification and question card would not be checked, hence the teacher did not make it. Based on the data from table 4.1, in this summative test, 14 test items conform with the syllabus, whereas 16 test items are unconformity with the syllabus.

Table 1. The conformity of summative test items with syllabus

No	Summative Test Items	Number Items	Total Frequency	Percentage
1	The conformity between summative test items and English syllabus	3, 4, 6, 7, 11, 12, 14, 13, 15, 24, 25, 26, 27, 29.	14	46, 67%
2	The unconformity between summative test items and English syllabus	1,2,5,8,9,10,16,17,18, 19, 20, 21, 22, 23, 28, 30	16	53,33%

First, the present study will talk about the conformity of tests items by the syllabus. The conformities in this summative test are questions 6, 3, 7, 11, 12, 13, 14, 15, 24, 25, 26, 27, and 29. Then, in the syllabus, there are two indicators for making summative tests, which are based on indicators of reading and based on indicators of writing. Based on indicators of reading, there are nine items which are 6, 3, 7, 11, 12, 13, 14, 15, whereas based on the indicators of writing, there are five items which are 24, 25, 26, 27, 29.

Question #3 and question #6 talk about narrative text. Question #3 asks about the generic structure of a narrative text. It relates to the indicator in the syllabus which is deciding the generic structure of the narrative text. In other words, this question is included understanding the conceptual meaning, which is purpose. It is similar to the theory from (Heaton, 1995) that stated that there are some specific skills involved in reading skills; one of them is understanding the conceptual meaning. In understanding the conceptual meaning, there are six items, which are quantity and amount, definiteness and indefiniteness, comparison and degree, means and instrument, cause, result, purpose, reason, condition, addition, contrast, and concession (Heaton, 1995). In addition, based on the data from the interview, those items also relate to the material that has been taught in the classroom. Hence, test item number 3 is considered to be tested.

Question #6 asks about the moral value of the narrative text. It relates to an indicator in the syllabus that is finding the moral value from the story. It is similar to the data gained from an interview that in narrative text, the students should know the moral value from the text. The teacher said that the students should learn a lesson not only from their own life but also from the text.

Question #7 asks about the explanation text. It relates to the indicators of reading, that is identifying the text of explanation. Based on the data gained from the interview, the explanation also has been taught during the teaching-learning process. Hence, this test item is considered to be a test. In addition, this test item is given since it is one of the materials that will be existed in the national exam.

Question 11 to 15 is the cloze test. The cloze test is a good indicator of general linguistic ability, including the ability to use language appropriately based on particular linguistic and situational contexts (Heaton, 1995). This test is used for allowing the students to attend to one set of related sentences for 5 items that assess vocabulary and grammar (Brown, 2004). These items are exposition text. Further explanation related to those questions will be discussed as follows:

Question #11 asks about the usage of vocabulary in the text. This number is used for measuring the students about the usage of adjectives in the text, especially in the exposition text. This item relates to the indicator of the syllabus, which is finding the information from text.

Question #12 is the same as number 11, which is asking about adjectives that are appropriate in the sentence. This item also relates to the indicator of the syllabus, which was finding the information from the text.

Question #13, this item has a different indicator from other items, which is identifying coordinative conjunctions. It is related to the indicator in the syllabus. Based on the data gained from the interview, the teacher explained this text. When the teacher explained this text, the teacher also explained the linguistic features. It is

similar to the theory that is supposed by (Hughes, 2003) the content of the test must be related to the course that students are concerned about; hence, these items are recommended to be tested.

Question #14 is the same as numbers 11 and 12, which is asking about the vocabulary used in the text. The difference is this question asks about the usage of a noun in the text. It aims to measure the students whether the students can difference between usage of a noun, adjective, and verb or not. However, the number in the text is wrong. The number in the text is not question14, but question 13. It will make students confused to answer this item. It seems the teacher forgot to recheck this item well. Hence, this item is not valid to be tested.

Question #15 is as same as number 14. There is something wrong. It should be question 15, but it was question 50 in the text. It certain makes students confused to answer this text. It seems the teacher forgot to recheck the item. Hence, this test item is not valid to be tested.

As we can see that, there are many mistakes in this cloze test numbers 11 to 15 like wrong numbering and the deleting *n*th word. The *n*th word is deleted in the text was not consistent and clear. Whereas, according to Gestalt theory of closure (Heaton, 1995), every *n*th word is deleted in a text every fifth, sixth, or seventh word. These test items were not included good cloze although, these test items related with a syllabus. Thus, these test items should be revised first before being tested.

Besides, five items conform to this test item by syllabus based on indicators of writing, which are questions 24, 25, 26, 27, and 29. These questions included intensive writing. In intensive writing, there are seven designs: dictation and dicto, grammatical transformation tasks, picture-cued tasks, short answer, and sentence completion (Brown, 2004). In this case, these items include grammatical transformation. Since grammatical transformation discussed change the tenses in a paragraph, change full forms of verbs to reduce forms, change statements to yes/no or wh-questions, change questions into statements, combine two sentences into one using a relative pronoun, change direct speech to indirect speech, change from active to passive (Brown, 2004).

Question #24 asks about passive sentences, especially interrogative sentences. It is included the grammatical transformation task because it is about a passive sentence. This test item is clear and relates to the indicator of the syllabus, which is identifying the sentence using a passive form. Hence, this test item is valid to be tested.

Question #25also asks about the passive sentence, but this item is preferred to the positive sentence. It is also included the grammatical transformation task. It is also related to the indicator of the syllabus, which is identifying the sentences using the passive form. Hence, this test item is valid to be tested.

Question #26 is also about the passive form. This test item is the same as number 25. This item asks about a positive sentence in passive form. It also relates to the indicator of the syllabus, which is identifying the sentences using the passive form. Hence, this test item is valid to be tested.

The passive form exists in question #27 then this question is as same as question #25 and question#26. This number has the same indicator as question #25 and question #26. In other words, this number is related to the syllabus. Hence, this test item is valid to be tested.

Question 24-29 is included intensive writing. Since there are seven designs: dictation and dicto, grammatical transformation tasks, picture-cued tasks, short answer and sentence completion in intensive writing (Brown, 2004) The reason why the intensive writing was given was the teacher wanted to know how far the students knew about the pattern. She expected that students not only understand the text but also the pattern.

In this summative test, there is also unconformity between the test item and the syllabus. There are 16 items that unconformity with syllabus, which are 1, 2, 4, 5, 8, 9, 10, 16, 17, 20, 18, 19, 22, 21, 23, 28, and 30. However, most of them related to the materials that have been explained such as numbers 1, 2, 5, 8, 9, 10, 20, 18, 19, 22, 21, 23, 30.

Question 1, 2, 4, and 5 provided by the teacher are the narrative text. The teacher considered that narrative text is very good for students regarding giving them a lesson from the text. In addition, some of her students said that they were interested in this story. Further explanation related to those questions will be discussed as the following

Question #1 asks about the kinds of text. It is related to the topic in the syllabus, but it does not relate to the indicator of the syllabus. Hence, if this test item is seen by the content validity, this test item is not valid to be tested.

Question #2 asks about the narrative text in terms of the purpose of the text. The topic related to the syllabus, but it did not relate to the indicator of the syllabus. Hence, this test item was not suggested to be used.

Question #4 asks the topic of the story. Based on the data from the interview, the teacher wanted students to know the content of the text. Hence, the teacher made this item test. Unfortunately, this test item does not relate to the indicator of the syllabus.

Question #5 asks the main idea of paragraph 3. It is a good question because this question by the features of reading comprehension stated by (Brown, 2004). He explains the main idea, expressions/idioms/phrases in context, inference, grammatical features, detail, excluding facts not written, supporting idea (s), and vocabulary in context. However, this question does not relate to the indicator of the syllabus. Hence, this question is not valid to be used.

Question 8,9,10, talk about the explanation text. These questions have the same case as questions 1,2,4,5. It also relates to the topic in the syllabus, but it does not relate to the indicator of reading. Further explanations relating to those questions will be discussed as the following.

Question #8 asks the reason for the snowmelt. However, this question does not relate to the indicator of the syllabus neither reading nor writing. Therefore, this question is not valid to be applied.

Question #9 asks about the causes of the water level rise. Nevertheless, this question does not relate to the indicator of the syllabus neither reading nor writing. Therefore, this question is not valid to be applied.

Question#10 is excluding facts not written questions (Brown, 2004). This question asks about explanation text. Nevertheless, this question does not relate to the indicator of the syllabus neither reading nor writing. Therefore, this question is not valid to be applied.

Question #16 talks about the expression of encouragement in terms of the purpose of the expression. This question does not relate to the syllabus. This question relates to neither syllabus nor the materials that have been taught. In this case, the teacher seems to take this question without considering the material that has been taught and the indicator in the syllabus. Therefore, this question is not valid to be tested.

Question #17 also talks about the expression of encouragement in terms of vocabulary in this expression. This question relates to neither syllabus nor the materials that have been taught. Therefore, this question is not valid to be tested.

Question #18 includes a grammar item, namely indirect speech in terms of *wh*-clause. Nonetheless, this question does not relate to the syllabus. This question does not relate to either the syllabus or the materials that have been taught. Therefore, this question is not valid to be tested.

Question #19 also includes a grammar item, namely indirect speech in terms of *that* clause. Nonetheless, this question did not relate to the syllabus. These questions did not relate neither syllabus nor materials have been taught. Therefore, this question is not valid to be tested. Nonetheless. This question relates to neither syllabus nor the materials that have been taught. Therefore, this question is not valid to be used.

Question #20 asks about the reported speech in terms of *if/whether* items. This question included one part of intensive writing, which is a grammatical transformation task. It is used for reporting yes/no questions. Nonetheless. This question relates to neither syllabus nor the materials that have been taught. Therefore, this question is not valid to be used.

Question #21 belongs to one part of intensive writing, which is grammatical transformation tasks. This question asks about reported speech. This question is made for identifying the reported speech using tenses. Nonetheless, this question relates to neither syllabus nor the materials that have been taught. Therefore, this question is not valid to be used. In addition, this question does not present the 'problem' (i.e the option) clearly (Heaton, 1995). Therefore, this question is not valid to be used and it is not recommended.

Question #22 is one of the parts of intensive writing, which is a grammatical transformation task. This question is also about reported speech in terms of reported request then; this question is understandable since the options do not interrupt the flow of meaning in the sentence (Heaton, 1995). Nonetheless. This question relates to neither syllabus nor the materials that have been taught. Therefore, this question is not valid to be used.

Question #23 is one part of intensive writing, which is a grammatical transformation task. This question is as same as question #22, which is in terms of reported requests. This question is also preferable since this question can be read at glance (Heaton, 1995). Nonetheless. This question relates to neither syllabus nor the materials that have been taught. Therefore, this question is not valid to be used.

Question #28 is one part of intensive writing, which is a grammatical transformation task. It is about a simple present perfect. In addition, this question is also preferable since the options do not interrupt the flow of meaning in the sentence (Heaton, 1995). Nonetheless, this question relates neither to the syllabus nor the materials that have been taught. Therefore, this question is not valid to be used.

Based on the data from the interview, questions 20 to 23 and 28 were given to know students' comprehension since the teacher thought that grammar is a crucial part of learning English. In addition, from the data gained in the interview, the teacher also said that it often exists in the SNMPTN test. Hence, the grammatical transformation task was needed to be tasked. However, these items did not include in the syllabus. Hence, these test items were not suitable to be tested.

It is evidence from the questions of summative test above, it is found that in this summative test, there are fourteen items (46,67%) that were appropriate with syllabus and the material that has been taught, then there are sixteen items (53,33%) that were not appropriate with the syllabus or the material that has been taught. the percentage of conformity to the syllabus falls into the level 40-55% which means less good. The teacher knew about how to make good tests, but she should make many tests at the same time, so, the teacher did not have time to make the summative well. Hence, based on the result of the items analysis, it can be seen that the summative test administrated in the third graders at senior high school has reached less good content validity.

CONCLUSION

Based on the findings and discussion, the summative test was administrated for the third graders at senior high school held in December, 1st 2014 has poor content validity. It is proved with the conformity of this summative test is only 53.33%. Due to some items were not suitable with the indicator in syllabus and the material that has been taught.

Therefore, for making good content validity, there some criteria should be considered. First, the teacher should read the indicator in the syllabus, and then write it into the test specification, next put the test items into test cards.

REFERENCES

- Adiredja, R. K. (2012). *Content validity analysis on achievement test at a private islamic junior high school in Garut*. Bandung.
- Alderson, J., Clapham, C., & Wall, D. (1995). *Language test construction and evaluation*. Cambridge: Cambridge University Press.
- Alwasilah, A. C. (2000). *Pokonya kualitatif*. Jakarta: Pustaka Jaya.
- Arikunto, S. (1992). *Prosedure penelitian*. Jakarta: Rineka Cipta.
- Best, J. W., & Kahn, J. V. (2006). *Research in education*. United State of America: Pearson Education Inc.
- Bogdan, R., & Biklen, S. K. (1998). *Qualitative research for education: An introduction to theory and method (3rd ed)*. Boston: Allyn & Bacon.
- Brown, H. D. (2004). *Language assessment: Principles and classroom practices*. New York: Pearson Education Inc.
- Cresswell, J. W. (2008). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research*. New Jersey: Pearson Prentice Hall.
- Crocker, L. J., & Algina, J. (1986). *Introduction to classical and modern test theory*. Orlando: PL: Holt, Rinehart and Winston.

Puspitasari, E. *An Analysis of Content Validity on Summative Test for the Twelfth Grade Students*

Emilia, E. (2008). *Menulis tesis dan disertasi*. Bandung: ALFABETA.

Fauzi, s. (2011). *An analysis of the content validity of the english summative test*. Jakarta: Syarif Hidayatullah state islamic university.

Fisher, D., & Frey, N. (2007). *Checking for understanding*. Alexandria, Virginia: Assosication for supervision and Curriculum development.

Fraenkel, J., & Wallen, N. E. (1993). *How to design and evaluate in education*. Boston: McGraw Hill.

Fulcher, G., & Davidson, F. (2007). *Language testing and assessment an advance resource book*. New York: Routledge.

Gall, M. D., & Borg, W. R. (2002). *Educational research an introduction*. United States: A&B.

Gardner, J., Harlen, W., Hayward, L., Stobart, G., & Montgomery, M. (2010). *Developing teacher assessment: An introduction*. New York : Open University Press.

Gronlund, N. E. (1977). *Constructing achievement tests*. United states of America: Prentice Hall, INC.

Gronlund, N. E. (1993). *How to make achievement test and assessments*. Michigan: Allyn and Bacon.

Haynes, S. N., Richard, D. C., & Kubany, E. S. (1995). Content validity in psychological assesment: A functional approach to concepts and methods. *American Psychological Association* , 7, 238-247.

Heaton, J. B. (1995). *Writing english language tests*. New York: Longman.

Henning, G. (1987). *A guide to language testing*. Cambridge: Mass: Newbury House.

Hughes, A. (2003). *Testing for language teachers (2nd edition)*. Cambridge: Cambridge University Press.

Ingram, E. (1977). *Basic concepts in testing* . Oxford: Oxford University Press.

Kerlinger, F. (1973). *Foundations of behavioral research*. New York: Holt, Rinehart and Winston.

Nofiyanti. (2011). *An analysis of the content validity of the summative test for the first year students of junior high school*. Jakarta: state islamic university syarif hidayatullah.

Nur, M. C. (2010). *An analysis on content validity of english summative test*. Jakarta: Syarif Hidayatullah.

Ramadhana, M. A. (2010). *An analysis of the english test used in final examination at SMA Pesantren IMMIM Makasar*.

Siddiek, D. A. (2010). The impact of test content validity on language Teaching and learning. *Asian Social Science* , 6.

Silverman, D. (2005). *Doing qualitative research 2nd edition*. London: SAGE Publication.

Sultana, R., Shamim, K. M., Nahar, L., & Hasan, F. (2009). Content validiy of written examinations in undergraduate anatomy. *Bangladesh Journal of Anatomy* , 7, 14-18.

Puspitasari, E. *An Analysis of Content Validity on Summative Test for the Twelfth Grade Students*

Weeden, P., Winter, J., & Broadfoot, P. (2002). *Assessment what's in it for schools?* London and New York: Routledge Falmer.

Yaghmale F., P. (2003). Content validity and its estimation. *Journal of Medical Education*, 3.