

THE CONTENT VALIDITY OF SUMMATIVE TEST FOR THE TENTH GRADE STUDENTS

***Syahfitri Isabella Simatupang**

****Meisuri**

****Nora Ronita Dewi**

ABSTRACT

This study was a descriptive qualitative research. It aimed to describe the content validity of the English summative test items for the tenth grade students at SMK Imelda Medan. This study used documentary analysis in collecting the data. The data of this research were the English summative test items for the odd semester of the tenth grade students made by the English teacher at SMK Imelda Medan in academic year of 2018/2019. The English summative items were classified into five categorizations: interpersonal conversation, transactional conversation, short functional texts, genre, and grammar. Based on the analysis of data, 22 out of 30 English summative test items for the tenth grade students were suitable to the syllabus (73%) with the most dominant topic about transactional conversation (10 items while there were 10 out of 30 items were irrelevant (27%) to the syllabus because the items discuss about the topics that are not irrelevant to the syllabus.

Keywords: Summative Test, Content Validity, and Syllabus.

***Graduate Status**

****Lecturer Status**

INTRODUCTION

Background of the study

In teaching-learning process, there are some responsibilities that teachers have to do such as planning the materials, preparing the opening until closing of teaching-learning process, and doing assessment. Assessment is a very important activity during foreign language teaching (Qu and Zhang, 2013:335). It feeds back the information of teaching and learning, makes teachers and students see the achievements and shortcomings clearly and improve teaching and learning efficiently. Moreover, assessment provides teachers with the information about how a student understands and applies knowledge (Ing, et.al, 2015:194).

Information of students' understanding can be gained through test. A test is an evaluation education tool that has an important role in measuring students' achievement. Testing becomes an integral part of teaching because it provides significant information or inputs about the growth and achievement of learners (Desheng and Varghese, 2013:31). Test also helps the teacher measure the students' achievement after doing teaching-learning proses (Nurfiqah, et.al, 2015:2). Without testing, teachers will have difficulty in providing proof of the quality of their students.

Teacher or the other test-structor should be able to construct good test for the students. By giving a good test constructed by the teacher, the students would have an opportunity to get a good quality in learning. The teacher as a test maker therefore has a great deal of work to do if he/she wants to produce a good test (Dada and Ohia, 2014:72). It is not an easy work for teachers to construct a good test, because there are some characteristics that must be fulfilled. There are five major features of a good test, validity, reliability, objectivity, practicality, and economic (Arikunto, 2016:72).

Content validity is a particular importance for achievement test. A test score cannot accurately reflect a student's achievement if it does not measure

what the students was taught and is supposed to have taught (Gay, et.al. 2012:161). Content validity will be compromised if the test covers topics not taught or if it does not cover topics that have been taught. The fact is that not all learning objectives are equally important. For example, a high school biology teacher wants to assess her students. There are 45 learning objectives, but only five functions, and the assessment includes 50 items, then the sampling validity might be weakened. This type of threat is frequent occurrence with teacher-made test, as it is called over-sample or under-sample.

Furthermore, there are some researchers who analyzed content validity. Setiyana (2016) analyzed the content validity of the English summative test for the first grade, second grade, and third grade of senior high school. She analyzed the content validity by comparing the items to the demands of the syllabus content including standard competences and basic competencies. She found that there were 16 items which did not match English school-based syllabus; 8 items from tenth grade, 3 items from eleventh grade, and 5 items from twelfth grade. Moreover, there were 13 basic competencies which were ignored by the English teachers. They were basic competency points 4.1, 4.2, and 6.1 in the tenth grade, 3.1, 3.2, 4.1, 4.2, 5.1, and 6.1 in the eleventh grade, and 3.1, 3.2, 4.1, 4.2, 5.1, and 6.1 in the third grade.

Nurfiqah, et.al (2015) investigated the quality of English summative test for the eleventh grade students based on content validity. The summative test consisted of 50 items. The content validity was assessed by analyzing the test items with the indicator of the test. The indicator of the test was matched with each item of the test to know whether the test items and indicators were appropriate or not. She found that the content validity was 92% and only 4 items that should be revised. Based on the result, the test items fulfilled the requirement of measurement because of those items already representative with the concept.

The difference of the previous study and this research was the summative test items were categorized into 5 types of materials. If the previous research

just compared each items to the syllabus indicators that was conducted by Setiyana (2016) and similarity finding was conducted by Nurfiqah, et.al (2015) who just focused on the indicators. In this research, the researcher analyzed the content validity of summative test based on the syllabus by dividing each items into 5 types of materials, such as interpersonal text, transactional text, short functional text, genre, and grammar.

In this case, the English summative test for the odd semester of the tenth grade students at SMK Imelda Medan was chosen by the researcher, for it was necessarily important to determine the content validity level of the summative test for the tenth grade students.

THEORETICAL FRAMEWORK

Assessment

Arends (2012:217) states that assessment usually refers to the full range of information gathered and synthesized by teachers about their students and their classrooms. Information can be gathered on students in informal ways such as through observation and verbal exchange. It can also be gathered through formal means such as homework, tests, and written reports. Information about classrooms and the teacher's instruction can also be part of assessment. The range of information here can also vary from informal feedback provided by the students about a particular lesson to more formal reports resulting from course evaluation and standardized tests.

Furthermore, Yuksel and Gunduz (2017:337) state that assessment is a very important feature in determining the behavior and attitude of the student regarding learning, in determining the decisions of the instructor before education, during education, and after education in ensuring the control of the education activities. Assessment is the determinant in all decisions to be taken by the instructor. Assessment also has an important role in ensuring control in every stage of education activities.

Test

Ismail and Yoestara (2017:78) state that test have long been used in any scholastic disciplines to measure learner's ability in their cognitive achievement. Tests inform language teachers and instructors how far their students have mastered the materials taught. Then the test result is used as an assessment parameter in setting their students level of ability, whether they are high or low performance students.

Furthermore, Brown (2003:4) states that tests are a subset of assessment. He defines tests as a method of measuring a person's ability, knowledge, or performance in a given domain. A well-constructed test is an instrument that provides an accurate measure of the test-takers ability within a particular domain. Tests can be useful devices, but they are only one among many procedures and tasks that teachers can ultimately use to assess students.

Content Validity

Gareis and Grant (2015:31) believed that content validity is concerned with how an assessment samples the intended learning outcomes of an instructional unit. It means that content validity is concerned with how far an assessment is in line with the intended learning outcomes, standards, or objectives of an instructional unit. Thus, an assessment is almost always a sample from among large set of learning objectives, creating an incomplete but hopefully representative picture of students' learning.

Content validity is important when we wish to describe how an individual performs on a domain of task that the test is supposed to represent. The fact is that not all learning objectives are equally important. For example, a high school biology teacher wants to assess her students. There are 45 learning objectives, but only five functions, and the assessment includes 50 items, then the sampling validity might be weakened. This type of threat is frequent

occurrence with teacher-made test, as it is called over-sample or under-sample.

Brown (2003:22) gives an example of content validity. If you are trying to assess a person's ability to speak a second language in a conversational setting, asking the learner to answer paper-and-pencil multiple-choice questions requiring grammatical judgments does not achieve content validity. A test that requires the learner actually to speak within some sort of authentic context does. And if a course has perhaps ten objectives but only two are covered in a test, the content validity suffers.

Syllabus

Das (2013:1) states that a syllabus is a summary of topics to be covered in an education or training course. It primarily contains a schedule of topics, list of recommended books and other information necessary for completion of the course. He also adds that syllabus is a promise to the student that they would be taught these topics. As a promise of what the course will be, a syllabus resembles a contract between students and teachers, and also to an extent between students and the university/institution. Teachers are bound to cover the topic as in syllabus and the students have to learn those for their evaluation and completion of the course.

RESEARCH METHODOLOGY, ANALYSIS, DISCUSSIONS AND FINDINGS

Methodology

This research was conducted by using descriptive qualitative research. Cohen, et.al, (2007:205) state that descriptive research looks at individuals, groups, institutions, methods, and materials in order to describe, compare,

classify, analyze, and interpret the entities and the events that constitute their various fields of inquiry. The purposes of this descriptive qualitative data were to create a description, illustration; facts, characteristics, and the relationship between investigate phenomena related to the study. The qualitative data was used to analyze the content validity of a summative test for the tenth grade students.

The data of this research were the English summative test items for the odd semester of the tenth grade students made by the English teacher at SMK Imelda Medan in academic year of 2018/2019.

Analysis

The techniques of collecting data was documentary, it meant that the researcher asked for the document, such as: English summative test for the odd semester of the tenth grade students at SMK Imelda Medan in academic year of 2018/2019 and the syllabus in order to get the data. The techniques of data analysis were: (1) Identifying and classifying the summative test items. It meant that the researcher identified which to the syllabus basic competences and materials. (2) Classifying the English summative test item based on the content validity into five categorizations; interpersonal text, transactional text, short functional text, genre, and grammar. (3) Describing the findings of English summative test items related to the content validity and interpreting the data by calculating the total percentage of the conformity. Then, the percentage was compared to the criteria adopted from Arikunto (2003). (4) Describing the research and making the conclusion from the analysis that was conducted.

Findings

The data were identified and classified into 5 categorizations, they were interpersonal text, transactional text, short functional text, genre, and grammar. The data analysis can be seen in Table 1:

Table 1 The Analysis of English Materials based on Syllabus

Type of Material	Number of Questions	Frequencies	Percentage
Interpersonal Text	3, 4, 5, 6, 17, 18, 19, 20	8	27%
Transactional Text	1, 2, 12, 13, 14, 15, 16, 21, 22, 1 essay	10	33%
Short Functional Text	-	-	-
Genre	-	-	-
Grammar	24, 25, 2 essay, 3 essay	4	13%
Total		22	73%

There were 2 interpersonal materials that should be studied for the tenth grade students for the odd semester: congratulating and complimenting others. From the summative test, 8 items were found suitable to the syllabus with 5 items discussed about expressing and responding congratulating and 3 items discussed about expressing and responding complimenting.

Based on the syllabus, transactional text discussed about self-introduction and intention. It was found that there were 10 questions of the English

summative test were suitable to the syllabus. There were 5 items discussed about self-introduction and 5 items discussed about intention.

Questions that discussed about grammar were also found. Based on the syllabus, the materials for grammar were pronoun (such as subject and possessive pronoun), article (such as: a, an, the, this, those), modals (such as: be going to and would like to), adjective and noun (generic structures of descriptive text), and adverb of degree (such as: quite, very, and extremely). It was found that there 4 items were suitable to the syllabus. There were 2 items discussed about adjectives and nouns and 2 items discussed about pronoun.

Whereas, there was no item categorized as short functional text and genre. Based on the syllabus, short functional text discussed the material about announcement while genre discussed about descriptive text.

After analyzing the content validity of the English summative test items for the tenth grade students at SMK Imelda Medan, the researcher found two results.

Table 2 Item Analysis based on the Syllabus Material

	Result	Frequencies	Percentage
Suitable to the Syllabus	Interpersonal Text	8	27%
	Transactional Text	10	33%
	Grammar	4	13%
Not Suitable to the Syllabus		8	27%
Total		30	100%

Table 2 shows the percentage of items analysis based on the syllabus material. There were 30 items analyzed. 22 out of 30 items (73%) were suitable to the syllabus material. The percentage obviously fell into the level

of 56 – 75% based on Arikunto's (2003) opinion which meant sufficient. Sufficient meant that the summative test was able to fulfill the requirements of the content validity. The most dominant type of material from syllabus is transactional text (33%) under the material of self-introduction and intention. While, there were 8 items (27%) were not suitable to the syllabus material.

So the findings were presented as follows:

1. The English summative test items for the odd semester of the tenth grade students at SMK Imelda Medan fell into the level of sufficient in content validity. There were three types of material that occurred in the English summative test items for the odd semester of the tenth grade students at SMK Imelda Medan. They were interpersonal text, transactional text, and grammar. The total of transactional text was 10 items (33%), interpersonal text was 8 items (27%), and grammar was 4 items (13%). The most dominant type of material was about transactional text. There was no material about short functional text and genre.
2. There were 8 out of 30 items not suitable to the syllabus materials. Those items were irrelevant to the syllabus because the items discussed about the materials that were not included to the syllabus. Moreover, there were some items that should be tested for the upper grade.

Discussions

Based on the data analysis of the English summative test of the tenth grade students at SMK Imelda Medan in academic year of 2018/2019, the summative test was sufficient in the content validity. It could be seen from Table 4.5. that showed 22 of 30 items (73%) were suitable to the syllabus, while 8 items (27%) were not suitable to the syllabus materials. Based on the types of material, transactional text was the most dominant types with 10 items (33%).

That percentage meant that the summative test was able to fulfill the requirements of the content validity, but still it is suggested for the test-maker to be paid attention to the syllabus materials and basic competences. There were found some items which were not suitable to the syllabus. It happened because there are some items that should be tested for the upper grade. Therefore, those items should be revised in order to make the test has an excellent level of content validity.

Kristiana (2014) also analyzed the content validity of a summative test. The data were the summative test for the eighth students which consisted of 46 items. She found that the content validity of summative test for the eighth grade students was sufficient. There were 15 indicators from the syllabus and it was found that 8 indicators represented by the test items. The eight indicators were covered in 27 questions (58,7%).

Different from the previous study result, Adha (2014) analyzed the content validity of English summative test for the tenth grade students by comparing the summative test to the syllabus indicators. The test consisted of 50 items, but the researcher only analyzed 35 items. Those items tested the reading and writing skills. There were 21 indicators suggested in the syllabus, and they had to be represented by 35 items number. Based on the distribution of the indicators, the researcher found that only 38% of the indicators represented the test items.

In addition, Lestari (2015) also analyzed the content validity of English summative test for the seventh grade students by comparing the summative test to the syllabus indicators. The test consisted of 35 items, 30 multiple choice and 5 essay questions. The researcher categorized the items into two skills: reading skill and writing skill. There were 4 indicators of reading skill suitable to the syllabus which represented in 13 test items. And there were 6 indicators of writing skill suitable to the syllabus which represented in 12 items. Based on the distribution of the indicators, the researcher found there were 25 items (71%) suitable to the syllabus. It can be assumed that the test was sufficient in content validity.

CONCLUSIONS AND SUGGESTIONS

Conclusions

After analyzing the data, the researcher concluded that the content validity level of the English summative test for the odd semester of the tenth grade students at SMK Imelda Medan made by the English teacher at SMK Imelda Medan in academic year of 2018/2019 was sufficient. There were 22 items or 73% that were suitable to the syllabus materials. The most dominant type of materials was transactional text with 10 items (33%). And also, it was found that 8 items (27%) were not based on the syllabus materials. The unsuitability can be caused because those items were not included as the topics that should be tested for the odd semester of the tenth grade. Those topics should be tested for the upper grade. Therefore those items should be revised, in order to make the test has an excellent content validity.

Suggestions

In relation to the conclusion, some suggestions were presented as follows:

1. The findings of the study later can inspire the teacher to increase the content validity level in summative test.
2. This research can be a guidance to evaluate the English teacher's summative test, whether the test has excellent level of content validity or not to measure the student's English achievement.

REFERENCES

- Adha, Khairul. 2014. *An Analysis on the Content Validity of Summative Test for the Tenth Year Students at SMA Dua Mei*. Unpublished Thesis. Jakarta: Syarif Hidayatullah State Islamic University.
- Arends, Richard. 2012. *Learning to Teach Ninth Edition*. New York: McGraw Hill.
- Arikunto, Suharsimi. 2003. *Prosedur Penelitian: Suatu Pendekatan Praktik*. Jakarta: Rineka Cipta.
- Arikunto, Suharsimi. 2016. *Dasar-Dasar Evaluasi Pendidikan Edisi 2*. Jakarta: PT Bumi Aksara.
- Brown, Douglas. 2003. *Language Assessment: Principle and Classroom Practices*. San Francisco: State University Longman.
- Cohen, Manion, and Morrison. 2007. *Research Methods in Education*. New York: Routledge.
- Dada, Esther Morayo and Ohia. 2014. Teacher –Made Language Test Planning, Construction, Administration and Scoring in Secondary Schools in Ekiti State. *Journal of Education and Practice*. 5(18).
- Das, Jasita. 2013. Role of Syllabus in Higher Education: A Critical Study. *Global Research Methodology Journal*. 2(7).
- Desheng, and Varghese. 2013. Testing and Evaluation of Language Skills. *Journal of Research & Method in Education*. 1(2).
- Gareis, Christopher and Grant, Leslie. 2015. *Teacher-Made Assessments: How to Connect Curriculum, Instruction, and Student Learning Second Edition*. New York: Routledge.
- Gay, Mills, Airasian. 2012. *Educational Research Competencies for Analysis and Applications 10th Edition*. New Jersey: Pearson.
- Ing, Musah, Al-Hudawi, Tahir, and Kamil. 2015. Validity of Teacher-Made Assessment: A Table of Specification Approach. *Asian Social Science*. 11(5).
- Ismail, Nyak Mutia and Marisa Yoestara. 2017. Teachers' Behalf on Language Test Construction. *Language and Language Teaching Journal*. 20(2).
- Kristiana. 2014. *An Analysis on the Content Validity of Summative Test for the Second Grade Students of Junior High School*. Unpublished Thesis. Jakarta: Syarif Hidayatullah State Islamic University.

- Lestari, Ari Puji. 2015. *An Analysis on the Content Validity of the English Summative Test (A Case Study at Second Year Students of SMP PGRI 2 Ciputat)*. Unpublished Thesis. Jakarta: Syarif Hidayatullah State Islamic University.
- Nurfiqah, Supardi, and Novita. 2015. The Analysis on the Items of the English Test Made by the Teacher. *Jurnal Pendidikan dan Pembelajaran*. 4(9).
- Qu, Wenjie and Chunling Zhang. 2013. The Analysis of Summative Assessment and Formative Assessment and Their Roles in College English Assessment System. *Journal of Language Teaching and Research*. 4(2).
- Setiyana, Rusma. 2016. Analysis of Summative Test for English. *English Education Journal*. 7(4).
- Yuksel, Hidayet and Nevin Gunduz. 2017. Formative and Summative Assessment in Higher Education: Opinions and Practices of Instructre. *European Journal of Education Studies*. 3(8).