

CONTENT VALIDITY ANALYSIS OF ENGLISH SUMMATIVE TEST BASED ON THE JUNIOR HIGH SCHOOL SYLLABUS

Kartini

Universitas Negeri Makassar
kartini719@yahoo.co.id

Annisa Shofa Tsuraya

Universitas Islam Negeri Alauddin Makassar
annisa.tsuraya@uin-alauddin.ac.id

Syahrullah

Universitas Negeri Makassar
syahrullah.s@gmail.com

ABSTRACT

This research was aimed at finding the empirical evidence of the English summative test administered to the students for the seventh grade at one of Junior High School in Makassar that is appropriate with the indicators of reading and writing skills in the English syllabus and has good content validity. The method employed in this research was descriptive research. The data was collected by observing and analyzing the English summative test. To know the content validity of the test, this research used the comparative analytic method, that is, by comparing the contents of the test with the English syllabus focusing on reading and writing indicators. The finding of the research was that the test items in the English summative test in the odd semester were appropriate with the recommended indicators in reading and writing skills in the English syllabus. They covered $11/18 \times 100\% = 61.11\%$. This percentage obviously fell into the level of 56%-75%, which means sufficient. This research proved that items of the English summative test for the odd semester of one of the Junior High Schools in Makassar have sufficient content validity. It showed that most of the materials of the English summative test are appropriate to the recommended English syllabus referring to reading and writing standard competence.

Keywords: Content validity, summative test, junior high school of Makassar

INTRODUCTION

English, as one of the main subjects in junior high school, has functioned as a developing instrument for the students in knowledge, technology, and culture. So, they are ready to take part in national development. The purpose of teaching English in Indonesia is to develop communication skills, especially in oral and written skills (listening, speaking, reading, and writing). To measure the student's skills in the teaching-learning process, teachers need to hold an evaluation.

Evaluation is a very important aspect of teaching and learning activities. Evaluation plays an important role in some activities, especially in terms of education. Through evaluation, a teacher will be able to know his or her students' achievement on the materials that have been taught in a certain period of time, so the teacher can measure his or her effectiveness in teaching that has been applied in the classroom. And the information gained through the evaluation will be very useful for making improvements in the future.

The importance of evaluation in teaching and learning activities cannot be overstated. In various tasks, particularly in the realm of education, evaluation is crucial. A teacher will be able to know his or her students' achievement on the materials that have been taught over a period of time through assessment, allowing the teacher to assess his or her efficacy in the classroom. And the data gathered during the evaluation will be extremely important in making future changes. Gronlund (1990:5), "a test is an instrument or systematic procedure for measuring a sample of behavior (answering the question "How well does the individual perform-either in comparison with others or in comparison with a domain of performance tasks". Sax in Sidiek (2010), a test is a task or series of tasks used to obtain systematic observations presumed to be representative of educational or psychological traits or attributes.

One of the ways to do the evaluation is through a test. The test can measure the students' ability. A test is a short examination of knowledge or ability, consisting of questions that must be answered. To get an accurate evaluation, a test must have the criterion of a good test; that is, validity, reliability, and usability must be arranged well. However, at this point, the writer is only concerned with the content's validity. The test, which is made by the test designer, needs to be analyzed to determine whether the test items have good content validity or not. The teacher should not design the test items carelessly and leave some of their instructional objectives in the English syllabus. In fact, the test can be called a "good test" when the test items are in line with the syllabus to measure students' abilities at the end of the lesson.

Bachman (1990:23) states that a test is one type of measurement instrument and thus necessarily quantifies characteristics of individuals according to explicit procedures. Based on the opinion above, the writer can conclude that a test is a procedure or an appliance to know or measure the students' comprehension or knowledge of the subject area more accurately. Besides, tests are the way to determine the level of education. A test is a tool or device, either in the form of a question, order, or direction, used to measure and value the students' ability to master the material given by the teacher in accordance with the curriculum.

Siddiek (2010:138) states that there are three types of tests: achievement tests, diagnostic tests, and proficiency tests. While Tinambunan (1988:17) says that there are two types of tests used in determining a person's abilities: aptitude tests and achievement tests. Gronlund (1990:12) classifies a test into four types. These are placement tests, formative tests, diagnostic tests, and summative tests. The classification of tests done by some experts above indicates that, generally, there is no too deep difference. In other words, they differ in terms of the scope of each type of test. Therefore, the writer will discuss achievement tests.

Achievement Test

According to Popham in McLachlan (2008:13), an achievement test is used to assess students' performance on a given syllabus or program. Achievement tests are designed to evaluate a student's knowledge and abilities in a particular subject as a result of schooling. An achievement test is a tool

used to assess various facets of a student's developed abilities and learning process. As a result, achievement tests are frequently administered at the conclusion of a project, a course, or a semester. Achievement tests may be developed and administered by educational ministries, official examining boards, or faculty members of educational institutions.

McNamara (2000:11) asserts that achievement tests are intrinsically linked to the instructional process. End-of-course tests, portfolio assessments, or observational procedures for tracking progress based on classroom work and participation are all examples. Achievement tests accumulate evidence throughout or at the conclusion of a study in order to determine whether and where progress has been made toward learning objectives. Achievement tests should enhance the instruction to which they are connected.

The purpose of the achievement test is to determine the degree to which individual students, groups of students, or entire courses have succeeded in fulfilling language course objectives. (2003) (Hughes, p. 13).

Tinambunan (1988:7-9) notes that numerous types of tests are utilized to assess students' achievement. There are four sorts of accomplishment tests that teachers frequently employ in the classroom, according to their function: placement exam, formative test, diagnostic test, and summative test.

Placement Test

The purpose of the placement test is to determine the student's performance prior to the start of teaching. As Gronlund (1990:12) states, the purpose of placement evaluation is to evaluate learner performance from the beginning of instruction.

Formative Test

A formative test is utilized at the end of a unit in the course book or following the design of a lesson. The purpose of formative assessment is to monitor student progress throughout instruction. Its objective is to provide continual feedback on both the student's and the teacher's progress and failures in the classroom. Formative evaluation is mainly reliant on specially designed examinations for each instructional section, Gronlund (1990:12)

Diagnostic Test

The primary goal of diagnostic evaluation is to determine the underlying causes of persistent learning difficulties and to develop a plan for remedial work. According to Gronlund (1990:12), diagnostic evaluation is used to identify learning issues that occur during instruction.

Summative Test

Summative test is used to determine student achievement at the end of instruction. Summative evaluation, according to Gronlund (1990:13), occurs at the conclusion of a course (or unit) of instruction. It is used to determine the extent to which instructional objectives have been met and is primarily used to assign course grades or to certify that students have mastered the intended learning outcomes. While summative evaluation is primarily used for grading or certifying pupil achievement,

it also provides information for determining the appropriateness of course objectives and the effectiveness of instruction.

Numerous types of tests are used to assess students' achievement. While there are a variety of tests that teachers typically administer in the classroom, they can be used in a variety of ways depending on their function. The author discusses only one of them in detail because it is directly related to the analysis presented in this paper. It is a summative examination.

Furthermore, the summative test provides teachers with not only a final report on the program's success, but also a comparison of their students' ability and achievement in the instructional objectives of teaching learning activities.

A valid test requires that the measuring instrument meets the test requirements, specifically validity, reliability, and usability. According to Gronlund (1990:47), validity refers to the adequacy and appropriateness of the interpretations drawn from tests for a particular purpose. A piece of information can be said to be valid if it corresponds to actual circumstances. The writer will elaborate on the validity in the following subchapter.

The second necessary characteristic of a good test is reliability. As a measuring instrument, a test should be reliable. The consistency of assessment results is referred to as reliability. According to Gronlund (1990:48), reliability is defined as the consistency of test results.

If teachers obtain very similar scores when administering the same test procedure to the same students on two separate occasions, they can conclude that their results are highly reliable from one occasion to the next. Similarly, if multiple teachers rate student performance on the same test task independently and obtain similar ratings, they can conclude that tests are reliable if the results persist after the test has been administered to students' numerous times.

The final characteristic of a good test is its usability in developing new tests. Thus, the term "usability" refers solely to the procedure's practicality and makes no reference to the other qualities present. The teacher must bear several very practical considerations in mind, including economy, ease of administration, scoring, and result interpretation. How long it will take to administer and score the test, and whether a shorter test is preferable to a longer one. (1990:49) (Gronlund).

In writer opinion, the practicality of a test is critical in order for test materials to be administered properly. It must be determined how much material, time, and effort it will take. The writer will not discuss all of the characteristics of a good test; therefore, the focus will be on validity, as implied by the title of this paper.

METHOD

This study was descriptive research, as Arikunto (1998: 245) states that if a researcher wants to investigate what, how, how many, level of thing are descriptive research.

The data was gathered by witnessing and evaluating the English summative test paper in order to determine the test's content validity using the Comparative Analytic Method, which involved

comparing the test's content to the English syllabus and curriculum for junior high school's seventh grade. The writer also conducted an interview with an English teacher who worked as a test designer in order to learn more about the test, how to develop a decent English test, and to determine why some aspects of the English test were still incorrect.

Variables in this research were content validity of reading test and content validity of writing test. Indicators were indicators of reading skill and indicators of writing skill of English base syllabus.

The subject of this research consisted a class of seventh-grade students of one of Junior High Schools in Makassar and one English teachers as test designer of English summative test.

The instruments that were used in collecting the data were English summative test sheets of odd semester for the seventh-grade students and English syllabus for Junior High School which is used in that school. The sources of data were reading and writing test of English summative test. The instruments of the data were question sheets, English syllabus, English curriculum, and the English teacher as a test designer to be further analyze. Besides, the writer also did the interview with the English teacher.

The collected data were analyzed quantitatively. The data was calculated by using simple percentage formula. It is the formula:

$$P = \frac{f}{n} \times 100\%$$

For more detailed information, the writer gives the explanation of the formula above:

P = percentage, is the final value of the total number of analysis that has been carried out. It is used to see how many percent the test covers indicators in the English syllabus.

f = frequency, is the total number of indicators that are sampled in determining the appropriateness of the test.

n = number or amount, is the whole total indicators available in the English syllabus.

It was used to see how many percent the test covers the instructions of the curriculum. Qualitatively, the test items were studied in terms of their appropriateness to the English syllabus and the themes.

In addition, the writer also compared the percentage with the criteria adopted from Arikunto (1998:133):

Table 1. Criteria of Test Result Percentage

76 – 100 %	Good
56 – 75 %	Sufficient
40 – 55 %	Fair
< 40 %	Bad

To increase our insight, the writer did the interview with the English teacher. The interview was done to know more information how to make a good English test.

FINDINGS AND DISCUSSIONS

The writer examined the test items to see if they corresponded to indicators of reading and writing in the English syllabus taught at one of Junior High Schools in Makassar. Question sheets, blue prints of questions, English syllabus, and the English instructor as a test designer for the seventh-grade pupils which be analyzed further. The data that the writer used in this study is the English summative test for odd semester, which is called as Ulangan Semester Ganjil, 2021/2022. Based on the regulation about the standard of educational evaluation established by minister of National Education no. 20, 2007, the summative test of particular semester must be constructed based on some indicators which should be achieved at that semester. Therefore, the summative test of the seventh-grade students of one of the Junior High Schools in Makassar was constructed based on particular indicators.

A total number of the test items are 45 items which consist of 40 multiple choice and 5 essay test items. The test was held on Monday, January 10th 2022 with the given time 80 minutes. The test was developed by the English teacher.

Table 2. The Appropriateness of English Summative Test with the Indicators of Reading Competence Standard in the English syllabus at Odd Semester of the Seventh Grade Students

Skill	Indicators	Item Number	Total Item
Reading	1. Identifying some information in short functional text : <i>greeting card</i>	28	1
	2. Identifying some information in short functional text: <i>short description</i>	1, 31,40	3
	3. Identifying some information in short functional text: <i>short message</i>	19	1
	4. Identifying some information in short functional text: <i>advertisement</i>	0	0
	5. Denoting linguistic features in short functional text	9, 12,26,30	4
	6. Identifying communicative purpose in functional text: <i>greeting card</i>	27	1
	7. Identifying communicative purpose in functional text: <i>short description</i>	0	0
	8. Identifying communicative purpose in functional text: <i>short message</i>	18	1
	9. Identifying communicative purpose in functional text: <i>advertisement</i>	0	0
Total items			11

From the table above, there are only 6 indicators which are developed by the test designers in constructing the test based on the indicators in reading. On the contrary, there are 3 indicators which are not developed. The percentage of the test validity of indicators in reading was 66.66%.

Table 3. The Appropriateness of English Summative Test with the Indicators of Writing Competence Standard in the English syllabus at Odd Semester of the Seventh Grade Students

Skill	Indicators	Item Number	Total
Writing	1. Completing short functional text	20,21,22,35, 36,37	6
	2. Jumbling text/sentence/word into good arrangement	38,39,	2
	3. Making short functional text: <i>greeting card</i>	0	0
	4. Writing short functional text: <i>short description</i>	0	0
	5. Writing short functional text: <i>short message</i>	42	1
	6. Making short functional text: <i>advertisement</i>	0	0
	7. Using rhetorical steps in writing short functional text	0	0
	8. Denoting linguistic features in functional text	7,15,16,23,29, 33,34	7
	9. Using language gambits	3,4,5,6,8,11,13, 14,17,24,32,	11
Total items			27

From the table of indicators in writing above, there are 5 indicators, which are developed in constructing test, and there are 4 indicators which are not developed. The percentages of the test validity of indicators were 55.55%.

Table 4. The Inappropriateness of English Summative Test with the Indicators of Reading and Writing Competence Standard in the English syllabus of Odd Semester of the Seventh Grade Students

Skill	Indicators	Item Number	Total
Reading	1. Identifying information from short dialog: <i>introducing</i>	2	1
	2. Identifying communicative purpose of dialog: <i>asking and giving information form</i>	10	1
	3. Identifying information/meaning from the dialog: <i>giving prohibition</i>	25	1
	Total		3

Writing	4. Making dialog: <i>introducing</i>	41	1
	5. Answering question about <i>introducing</i>	43	1
	6. writing expression: <i>asking to do something</i>	44	1
	7. Writing example of <i>giving prohibition</i>	45	1
	Total		4
Total items			7

From the table above, there were 7 indicators which were not appropriate to the English syllabus of odd semester for seventh grade student. They cover 7 items number from 45 items number. The percentage of the inappropriate items was 15.55%.

Based on the data, it showed that the highest frequency of indicators in reading which was appropriate with the English syllabus was only 66.66% from 9 indicators that was developed. This percentage obviously falls into the level of 56% - 75% which means the test validity of reading test in summative test was sufficient.

Then, the highest frequency of indicators in writing which was appropriate with the English syllabus was 55.55% from 9 indicators that was developed. This percentage falls into the level of 56% - 75% which means sufficient.

Next, the total indicators of reading and writing that are appropriate with the English syllabus are 61.11% from 18 indicators that was developed. So, the percentage falls into the level 56% - 75% which means sufficient.

According to the item analysis result above, the writer concludes that the English summative test which was administrated for the seventh-grade students of one of Junior High Schools in Makassar has sufficient content validity refer to writing and reading skill.

In an interview with the English instructor who designed the test, she stated that she created the test on her own, rather than using the teacher's group work. She drew up a blueprint of the questions before creating the test.

She devised the test using signs from the odd-numbered semesters' English course. "Not all of the elements in the indicator have to be entered in the exam," she claims. The crucial thing to remember is that we might not be able to escape the competence norm and fundamental competency. Therefore, it can be understood that a test designer has already understood how to make a good test. They recognized that the summative test must be in line with the recommended curriculum and also with the English syllabus.

CONCLUSIONS

Inferential the followings based on the findings and discussion in the preceding above:

1. The test items in the English summative test for seventh-grade students correspond to the required indicators in reading and writing proficiency in the English syllabus. They cover 11/18 x

100% = 61.11 percent of the area. This number clearly falls between 56 and 75 percent, which is considered adequate.

2. The English summative test given to seventh-grade students in the odd semester was considered a valid test since it contains sufficient content validity.
3. Before developing the test, the teacher should conduct thorough research and consider the principles of excellent test item construction.
4. The test designers should keep in mind that the test items must be appropriate for all indicators in the English curriculum, including reading and writing, and that any things that do not meet the indicators should be rejected. Because it has sufficient content validity, it is considered a valid test.
5. The next researchers should discuss content validity on the test, with a particular focus on listening and speaking standard competency. Other types of validity can be discussed in future study.

REFERENCES

- Anderson, M. 2003. *Text Types in English 1-2*. Australia: Macmillan Education Australia.
- Arikunto, Suharsimi. 1998. *Prosedur Penelitian: Suatu Pendekatan Praktek*, Yogyakarta: Bhineka Cipta.
- Assessment, Evaluation and Final Marks In About. Retrieved February 30, 2014, <http://specaled.about.com/od/assessment/a/AandE.htm>
- Bachman, Lyle. F. 1990. *Fundamental Considerations in Language Testing*. Oxford: Oxford University Press. [Read Online] available on: http://books.google.co.id/books?id=5_KJCfkWgqcC
- Badan Standar Nasional Pendidikan. 2006. *Standar Isi untuk Satuan Pendidikan Dasar dan Menenga*. Jakarta: BSNP.
- Bowers, Roger, Christopher Brumfit. 1991. *Applied Linguistics and English Language Teaching*. London: Macmillan Publishers.
- Celce-Murcia, M. 1995. *On The Need for Discourse Analysis in Curriculum Development*. San Fransisco: McGraw Hill.
- Dean Brown, James. 2005. *Testing in Language Programs: A Comprehensive Guide to English Language Assessment*, New Edition, New York: McGraw-Hill Companies, Inc.
- Departemen Pendidikan Nasional. 2006. *Panduan Pengembangan Silabus Mata Pelajaran Bahasa Inggris*. Jakarta: Departemen Pendidikan Nasional.
- Fitzpatrick, Anne R. 1980. *The Meaning of Content Validity*. Buston: American Educational Research Association.
- Fulcher, Glenn. 2010. *What Is Language Testing*. Retrieved from <http://languagetesting.info/whatis/lt.html>
- Gronlund, Norman E. and Robert L. Linn. 1990. *Measurement and Evaluation in Teaching, Sixth Edition*. New York: Macmillan Publishing Company.
- Hornby, A. S.1995. *Oxford Advanced Learner's Dictionary*, Oxford: Oxford University Press.

- Hughes, Arthur. 2003. *Testing for Language Teachers. Second Edition*. Cambridge: Cambridge University Press.
- Maheshwari, V.K. 2011. *Achievement Test: Characteristics and Construction Procedure*. India: K.L.D.A.V. (P.G) College. Posted on September 22, 2011 at: <http://www.vkmaheshwari.com/WP/?p=181>.
- McLachlan, R.Silve & Emma Lindström. 2008. *A Comparative Study of How the English A Syllabus Is Interpreted in Achievement Tests Focusing on Writing Ability*. Engelska.
- McNamara, Tim. 2000. *Language Testing*, Oxford: Oxford University Press.
- Pusat Kurikulum Badan Penelitian dan Pengembangan Departemen Pendidikan Nasional, Kurikulum 2004. 2003. *Standar Kompetensi Mata Pelajaran Bahasa Inggris Sekolah Menengah Pertama dan Madrasah Tsanawiyah*, Jakarta: Pusat Kurikulum Balitbang Diknas.
- PUSKURNAS. 2006. *Standar Isi dan Standar Kompetensi Lulusan untuk Satuan Pendidikan Sekolah Menengah Pertama (SMP)/Madrasah Tsanawiyah (MTs) Beserta Peraturan Pelaksananya (Peraturan Menteri Pendidikan Nasional Republik Indonesia No 22, 23, dan 24 Tahun 2006)*. Jakarta: Depdiknas.
- Sax, Gilber. 1980. *Principles of Educational and Psychological Measurement and Evaluation. Second Edition*. California: Wadsworth Publishing Company.
- Siddiek, Ahmed Gumaa. 2010. *The Impact of Test Content Validity on Language Teaching and Learning*. Canadian Center of Science and Education.
- Sudaryanto. 1993. *Metode dan Aneka Teknik Analisis Bahasa*. Duta Wacana University Press.
- Tinambunan, Wilmar. 1988. *Evaluation of Student Achievement*. Jakarta: Departemen Pendidikan dan Kebudayaan.
- Wiersma, William, Stephen G. Jurs. 1990. *Educational Measurement and Testing*, second edition, Boston: Allyn and Bacon.
- Worthen, B.R. Borg and K.R. White. 1993. *Measurement and Evaluation in the School*, New York: Longman.