

AN ANALYSIS OF CONTENT VALIDITY OF THE TEST

Erlik Widiyani Styati

Abstract

Analyzing the result of the learners in a foreign language is needed. Besides, many teachers still find difficulty in analyzing the test item. It is used to know the achievement of the learners. Otherwise, the teachers also still find difficulty in giving the mark. It is meant that they lack to recognize whether the learners are in a good achievement. Furthermore, the paper aims to describe how to analyze the content validity of the test item. Analyzing content validity is the way to know whether the test is appropriate with the curriculum or not.

Key Words: *Test, Content Validity*

I. Introduction

English is the worlds' lingua franca. It is very important that the students should master it well. Based on Depdiknas (2004: 3-4) the purposes of English subject are as follows: (1) to develop the ability of both oral and written communication (listening, speaking, reading and writing), (2) to develop the awareness of the importance of English as one of the foreign languages as the main learning instrument, and (3) to develop the understanding of relationship between language and culture and also to develop the students' cultural understanding. The government has realized the importance of teaching English because it is expected to enable students to carry out self development in the field of knowledge, art, and culture.

Based on the curriculum, teaching and learning English should cover the competency of four skills namely listening, speaking, reading, and writing. Agustien (2004: 1) states that the Department of National Education has made it very clear that English Education is aimed at providing school graduates with life skills in the sense that they are expected to achieve the competence required to obtain communication skills. The latest curriculum is 2006 English standard competence. This curriculum still focuses on communicative competence. Agustien (2004: 1) mentions that English is a language used as a means of communication, the competence required for this level of communication needs to be clearly defined. Theoretically, the competence required for language communication is called communicative competence. In this case, the teachers'

control towards the students is needed. Teachers do not only teach but also evaluate the students' achievement.

To measure the students' achievement, the teachers give kinds of test. Yunian (1996: 16) says that evaluation is an effective means of measuring the teaching and learning performance in language program and improving the teaching process. Heaton (1975: 1) defines that both testing and teaching are so closely interrelated that it is virtually impossible to work either field without being constantly concerned with the other. Teaching is related to the learning process and testing is an evaluation process. Besides, testing is used to determine the achievement and progress of instruction. Moreover, giving a test is a necessity in the teaching learning process.

A test is measuring the students' ability after finishing the study program. Brown (2004: 3) states that test is a method of measuring a person's ability, knowledge, or performance in a given domain. There are many purposes of doing a test. Bachman (1990:58) says that the test may be made to measure the students' ability whether or not they should enter the program or the test is to determine whether or not the students are ready to continue the program. Arikunto (2003: 10-11) states that the purpose of testing is to select the students that can be accepted in certain

school and that can pass to next grade or even graduate from school. Heaton (1975: 2) says that the purpose of a test is to enable the teacher to ascertain which parts of the language programs have been found difficult by the students, and to give students' opportunity to show their abilities to use the correct forms of the language for the language test.

A good test has characteristics. Arikunto (2003: 57-58) states that there are five criteria that are used to know whether or not the test can be considered qualified namely reliability, validity, objectivity, practicality, and economy. Grounlund (1981: 130) states that both validity and reliability are the important things related to constructing such kind of test. Nurgiyantoro (1995: 151) defines that a good test should have characteristic of appropriateness, validity, reliability, effectively, and practicality. Bachman (1990: 24) says that reliability and validity are those two essentials to the interpretation and the use of language ability. Here, the content validity is emphasized to be described in this article.

II. Definition of Test

According to Johnson and Johnson (2002: 1) test is collecting information about the quality or quantity of change students, group, teacher or administrator. As we know that in teaching learning process contains three aspects namely

education goal, teaching learning-process, and teaching-learning results. Carroll (1986: 46) as cited in Bachman (1990: 20) states that a psychological or educational test is a procedure designed to elicit certain behavior from which one can make inference about certain characteristic of individual. Brown (2004: 3) states that test is a method of measuring a person's ability, knowledge, or performance in a given domain.

Test is important not only for the students but also for the teacher. The students can measure how far their understanding about the lesson. Other than that, the teachers will know whether the objectives can be achieved or not by the students so that the students can continue the next program. Amari (1991: 1) states test is a systematic and objective instrument which is applied to get some data or some information needed about someone in form of spoken or written accurately. Besides, Arikunto (1993: 51) points out that test is a procedure to know or to determine something by certain rules. Since test maker knows that test is important, the test maker should know how to construct the test well. Test is a tool, an instrument, or a device to determine the students' ability and achievement and also it decides something which is dealing with the students or teaching learning process.

III. Objective Test

Objective test is the test that the answers are certain without asking the testee's opinion. Nurgiyantoro (1995: 75) defines objective test as short answer test. Objective test demands to give a short answer even only choosing certain codes representatives of the answer available.

Arikunto (2003: 164) defines that the testee has to weigh up carefully all the alternative and select the best one. On the other hand, it makes the teacher easy to score the test because the answer keys are provided. Johnson and Johnson (2002: 67) mention the advantages and disadvantages as follows:

Furthermore, there is a characteristic of an objective test. The characteristic includes the strength and the weaknesses. The strengths of objective tests are: (1) it can be answered in shorter period of time, (2) it has high objectivity, (3) it is easy to score. In contrary, the weaknesses of the objective test are: (1) it is much more difficult to construct objective test item than easy test item, (2) it tends to measure the cognitive aspect only, (3) it enables the students to speculate in choosing the correct answer, (4) it makes the students to cheat in doing the test.

IV. Teacher-Made Test

Teacher made test is generally arranged and scored by the teacher who

will use the test in his or her class. Johnson and Johnson (2002:62) state that teacher made tests is written or oral assessments of students' achievement that are (a) designed specifically for the teacher students and (b) not commercially produced or standardized. Arikunto (2003:149) states that he doesn't know whether or not the test makes adequate index or difficulty and discriminating index. Teacher made test is applied to measure the student's achievement based on the objectives in teaching learning process. Usually this test is given in the mid term test and also final test in first and second semester.

Teacher-made test is generally prepared, administrated, and scored by one teacher, the classroom teacher him/herself. It is also made by a group of teacher. Teacher made test is not tried out before the test is given to the students. Nurgiyantoro (1995:60) states that there are three objectives or goals of teacher made test: *“(1) Kadar pencapaian tujuan. (2) Tingkat penguasaan bahan siswa, dan (3) untuk memberikan nilai kepada siswa sebagai laporan hasil belajar disekolah itu”*. *“(1) The level of achievement, (2) students' mastery toward the material, and (3) to give grade to the students as the result of their study in a school”*. Arikunto (2003: 149) also states that he does not know whether or not the test has adequate index or difficulty and discriminating

index. Teacher made test is applied when the teacher only emphasizes on the limited objectives. Teacher made test is not tried out first because the test maker thinks that there are so many factors including efficiency, opportunity and economy, also the teachers' capability in analyzing it.

V. Standardized Test

Standardized tests are prepared for the wide nation. It is different with the teacher made test or the group of people who make it. It provides accurate and meaningful information for the students. Johnson and Johnson (2002: 53) state that there are two types of standardized tests that are: achievement test and aptitudes test. In achievement test, it focuses on the knowledge and skills learned in school and may be form of achievement batteries, diagnostic test, or subject-specific test. Aptitudes tests focuses on the potentially maximum achievement of students and may measure general intellectual aptitudes, aptitude to do well in college or certain vocational training programs, reading aptitude, mechanical aptitude or perceptual aptitude.

Arikunto (2003: 147) states that the procedure to get standardized test are constructing, pre-testing, analyzing, revising, and editing. Standardized test is usually tried out. The test maker in this case is a professional in their field in constructing the test. The use of

standardized test is to get information about the students' ability in the wide nation. It is usually held in UAN or SPMB that is made by professional test maker.

Ebel as cited by Nurgiyantoro (1995: 64) states that the other function of the standardized test is to place the students based on their capability, arrange the individual instruction and arrange the remedial teaching if the test is given early. Standardized test is also analyzed statistically and claimed its validity to be used widely. Allen and Davies (1997:49-50) states an obvious opinion that a standardized test is objective, which means that it has been tried out on a proper sample or the population from whom it is intended and that on this sample it has shown to work. Most of tests of these types are made up of items each of which have characteristics in themselves and have been shown to contribute toward the total performance of the test.

VI. Content Validity

Validity is the extent to which a test measurement claims to measure. Johnson and Johnson (2002:54) state that validity means that the test actually measure what it was designed to measure, all of what it was designed to measure and nothing but what it was designed to measure. Test has a high validity if it is able to measure what its objectives are. Sudirman dkk (1991:294) mentions that

there are three types of validity: content validity, empirical validity, and construct validity. In the line with Sudirman dkk, Weir (1993:19) states that test validity presupposes that the writer can be explicit about what is to be tested and takes steps to ensure that the test reflects realistic use of the particular ability to be measured. According to Jack and Norman (1993:139) validity is the most important idea to be considered when preparing or selecting an instrument to use.

Content validity is the extent to which a test measures a representative sample of the subject content. Content validity here should cover the content of the test based on the curriculum that is used. If the test materials are suitable with the subject matter or curriculum, it can be concluded that the test has a content validity. In other words, if the test is not suitable with the subject matter or curriculum, it can be said that the test has no content validity. So, if the test does not have content validity it is not considered as a good test anymore to be given to the students. Content validity, when a test has content validity, the items on the test represent the entire range of possible items that the test should cover. (http://en.wikipedia.org/wiki/content_validity).

Content validity is related to face validity. Face validity is a property of a

test intended to measure something. The test is said to have face validity if it “looks like” it is going to measure what it is supposed to measure. Some people use the term face validity only refers to the validity of observers who are not expert in testing methodologies

(http://en.wikipedia.org/wiki/face_validity)

. Weir (1990:26) points out that the face validity is not validity in technical sense; it refers, not to what the actually measures, but to what it appears superficially to measure. Brown (2004: 26) states that face validity refers to the degree to which a test looks right, and appears to measure the knowledge or abilities it claims to measure, based on subjective judgment of the examines who take it, the administrative personal who decide on its use, and other psychometrically unsophisticated observers.

Construct validity is a judgment based on the accumulation of correlations from the numerous studies using the instrument being evaluated. Heaton (1975:154) states that if the test has construct validity, it is capable of measuring certain specific characteristics in accordance with a theory of language behavior and learning. Construct validity seeks agreement between a theoretical concept and a specific measuring device or procedure. Construct validity can be classified into two sub-categories:

convergent validity and discriminate validity. (Mousavi 2002:244 cited in Brown 2004:25) states that construct validity is a major issue in validating large – scale standardized tests of proficiency. It is so because such tests must, for economic reasons, adhere to the principle of practicality, and because they must sample a limited number of language domain, they may not be able to contain all the contents of a particular field or skills.

VII. CONCLUSION

It is very important to make the test which has good quality. Test makers should know the aim what they are going to measure. The test makers should make the suitable instrument based on the curriculum. To know whether the test is appropriate or not, the test maker should try out the test first before giving it to the students. In constructing the test items, the test makers should understand what kinds of objective should be achieved. They should make the test based on the basic competence in the curriculum. To make the test item, there are many reference books on how to make the test which contains good index of difficulty and index of discrimination. The test makers should construct the test based on the principles of constructing a good test so that the test will be acceptable and the objectives will be achieved;

VIII. REFERENCES

- Agustien, Helena I.R. 2004. *The 2004 English Curriculum in a Nutshell*. Malang: Universitas Negeri Malang.
- Amari. 1997. *Penilaian Pencapaian Hasil Belajar Mengajar Siswa di Sekolah*. Yogyakarta: Kanisius.
- Arikunto, Suharsimi. 2002. *Prosedur Penelitian*. Jakarta: PT. Rineka Cipta.
- Arikunto, Suharsimi. 2003. *Dasar-dasar Evaluasi Pendidikan*. Jakarta: Bumi Aksara.
- Azwar, Saifudin. 2001. *Reliabilitas dan Validitas* 3rd. Yogyakarta: Pustaka Bogor.
- Bacham, Lyle F. 1990. *Fundamental Consideration in Language Testing*. New York: Oxford University Press.
- Brown, J. Dean. 2005. *Testing in Language Program: A Comprehensive Guide to English Language Assessment*. New York: McGraw-Hill Companies, Inc.
- Cohen, D. Andrew. 1994. *Assessing Language Ability in The Classroom*. USA: Heinle and Heinle Publisher.
- Fraenkel, R. Jack and Norman. 1993. *How to Design and Evaluate Research in Education*. Singapore. McGraw-Hill, Inc.
- Gronlund, Norman. E. 1981. *Measurement and Evaluation in Testing*. New York: Macmillin Publishing Co. Inc.
- Johnson, David W. and Johnson, Roger T. 2002. *Meaningful Assesment. A Manageable and Cooperative Process*. USA: Allyn and Bacon.
- Kurrikulum 2006 Mata Pelajaran Bahasa Inggris Sekolah. 2006. *Menengah Atas. Standard Isi* 2006. Jakarta: Departemen Pendidikan Nasional.
- McMillan, James H. 1992. *Educational Research. Fundamental for the Consumer*. USA: HarperCollins Publishers.
- Nurdiyanto, Burman. 1995. *Penelitian Dalam Pengajaran Bahasa dan Sastra*. Yogyakarta: BPFE.
- Nurkencana, Wayan dan Sumartana. 1986. *Evaluasi Pendidikan*. Surabaya: Usaha Nasional.
- Reid, M.Joy. 1993. *Teaching ESL Writing*. New Jersey: Prentice Hall.
- Weir, Cyrill. 1993. *Understanding and Developing Language Tests*. UK: Prentice Hall International Ltd.

