THE CONTENT VALIDITY OF LOCAL ENGLISH TEST ITEMS FOR STUDENTS OF SMP NEGERI 2 KEBAKKRAMAT KARANGANYAR

RESEARCH PAPER

Submitted as a Partial Fulfillment of the Requirements
for Getting the Bachelor Degree of Education
in English Department

by:

ASRI NURHAYATI
A 320060125

SCHOOL OF TEACHER TRAINING AND EDUCATION

MUHAMMADIYAH UNIVERSITY OF SURAKARTA

2010

CHAPTER I

INTRODUCTION

A. Background of the Study

In teaching learning process, teacher always gives materials to the students. He/she teaches the materials until the end of teaching learning period. In the end of teaching learning process, teacher usually makes evaluation because every subject needs it to measure the students' achievements of materials that have been given.

When a teacher evaluates the subject materials of the students, the teacher needs a test. According to Brown (1994: 252) test is a method of measuring a person's ability or knowledge in a given area. In other words, we can say that a test is a method of measurement. To obtain the accurate information concerning the students' achievement, the test which is given to the students must have the qualities or characteristics. Grounlund (1985: 56) states that the main goals in classroom testing are to measure and to define precisely that the test items should be able to measure what is intended to measure. It also means specifying the achievement domain in such manner that the sample of the test tasks will represent total domain of the achievement.

There are two types of test; they are teacher made-test (informal) and standardized test (formal). First, teacher made-test is a test made by a teacher implemented in his/her classroom. Teacher made-test is usually called classroom test. It is generally prepared, administered, and scored by one teacher.

Second, standardized tests are tests which are prepared by professional testing to serve the assistance of institutions in selection, placement, and evaluation of students. Standardized test is designed to be used with thousands and sometimes hundreds of thousands of subjects throughout the nation or the world, and prepared by team of testing specialists with no personal knowledge of the examinees and no opportunity to check on the consistency of individual performance. For example, local test; it is a test made by a team of teachers in certain subject in one area or regency. So, the researcher can say that the local test between one regency to other regencies is different. Local test is usually used in summative test to measure the achievement of the students and primarily for assigning course grades of certifying student's mastery of the intended learning outcomes.

Test can be called good as measurement of the student's achievement having the qualities. All good tests have three qualities; they are practicality, reliability, and validity.

Practicality refers to the relationship between the resources that will be required in the design, development, and use of the test and the resource that will be available for these activities. In another word, it is a matter of the

extent to which the demands of the particular test specifications can be met within the limits of existing resources. A test can be called practical if a test is implemented within the means of financial limitations, time constrains, ease administration, scoring and interpretation.

Reliability concerns with the consistency of the test scores. Test can be called reliable if a same test is used to measure the students' achievement implemented in same students and in different situation, the result of the test is consistent. Most simply put, a test is reliable if it is consistent within itself and across time.

By far the most complex criterion of a good test is validity. It refers to the degree to which the test actually measures what it claims to measure. Test validity is also the extent to which inferences, conclusions, and decisions made on the basis of test scores are appropriate and meaningful. There are three major types of validity: empirical validity, content validity, and face validity.

First is empirical validity. It is a concern for tests that are designed to predict student's status on an external criterion measure. A test has empirical validity if it is useful for predicting a person's behavior in a specified situation. For example, reading readiness test scores might be used to predict pupil's future achievement in reading, or a test of dictionary skills might be used to estimate pupil's current skill in the actual use of dictionary.

Second, content validity concerns the adequacy with which the test items adequately and representatively sample the content area to be measured. For example, based on School Based Curriculum, test will measure the students' achievement in identifying kind of meaning in recount text. The examples of items test measuring this ability are as follow:

> Text for No. 24 & 26
> Students and teachers from SMP Satya Karya went to the zoo during the last holiday. The students were happy to see different kind of animals and some animals doing a trick. They also saw some animals from foreign countries such as camels, bears, flamingos, and piranhas. Then, they saw elephants sitting and a monkey riding a bike, it was funny.
> 24. When did the students of SMP Satya Karya have a tour?
>   a. last week   c. next month
>   b. last holiday   d. on Saturday
> 26. These animals were from foreign countries, EXCEPT. . .
>   a. bears   c. elephant
>   b. camels   d. flamingos
> Taken from English Test in 8th level student in Karanganyar

From the indicator of School Based Curriculum, the test items above show that the test items represent the sample of the content area of students' achievement to be measured. So, by comparing the test items and indicator in School Based Curriculum, the test items are called valid.

The third is face validity. It refers to the degree to which a test looks right and appears to measure the knowledge or abilities based on the subjective judgment of the examinees who take it. Face validity is very important from the learner's perspective because a learner needs to be convinced that the test is indeed testing what it claims to test. Face validity is almost always perceived in terms of content: if the test samples the actual

content of what the learner has achieved or expected to achieve, then face validity will be perceived.

Teaching materials which are used to make the test items should be appropriate with the curriculum in the school. In elementary and middle school, the curriculum which is implemented is School Based Curriculum (*Kurikulum Tingkat Satuan Pendidikan/KTSP*). It is the newest curriculum in Indonesia to replace the Competency Based Curriculum with some change. This change occurs in every lesson that must be taught in every educational level. The implementation of School Based Curriculum in Indonesia is based on *Undang-Undang No.20 Tahun 2003*. The implementation of it between one school to other schools is different because the development of it depends on the capability of its school, the potential and characteristics of region, and also the social cultural background of the students.

The teaching materials should also cover all of learning domains proposed by Benjamin Bloom. He makes them into taxonomy of learning objectives. It is an attempt to classify forms and levels of learning. It identifies three domains of learning, each of which is organized as a series of levels; the domains are cognitive domain, affective domain, and psychomotor domain.

Cognitive domain refers to knowledge structures. It can be viewed as a sequence of progressive contextualization of the material. Cognitive domain in language testing is important because it can measure the capability or

competence of the students. It consists of six categories; they are knowledge, comprehension, application, analysis, synthesis, and evaluation.

Affective domain is concerned with values, or more precisely perhaps with perception of value issues, and ranges from awareness or receiving, through being able to distinguish implicit values through analysis. The five major categories are receiving phenomena, responding to phenomena, valuing, organization, and characterization.

The psychomotor domain includes physical movement, coordination, and use of the motor-skill areas. Psychomotor objectives usually focus on change and/or development in behavior and/or skills. Development of these skills requires practice and is measured in terms of speed, precision, distance, procedures, or techniques in execution. The six major categories are reflex movements, basic fundamental movements, perceptual abilities, physical abilities, skilled movements, and non-discursive communication.

Nowadays, the researcher finds that there are some test items which do not fulfill the requirement for a good test. The test items only reflect the understanding about the fact of the lesson. Ebel in Wallace and Larson (1978: 7) states that the tests themselves are used unwisely, misinterpretation, over interpreted, or handled as weapons rather than as tools. They only reflect particular perceptions of the goals of education which are not shared by all educators. So, the researcher is interested in conducted research about the test item that will be seen by using two indicators.

Based on the description above, the researcher will discuss about content validity of test to measure whether or not the test which is done in junior high school at eight level of SMP Negeri 2 Kebakkramat Karanganyar corresponds or not. The researcher takes Local English test items in SMP Negeri 2 Kebakkramat because its school has a good rank in educational system as a school that includes School of National Standard (*Sekolah Standart Nasional*). It will be discussed in my research paper entitled "THE CONTENT VALIDITY OF LOCAL ENGLISH TEST ITEMS FOR STUDENTS OF SMP NEGERI 2 KEBAKKRAMAT KARANGANYAR".

B. Problem Statements

In this study the researcher wants to discuss the problem of content validity of Local English test items. "Whether or not the content validity of Local English test items is valid?"

To answer the question above, the researcher raises some subsidiary research questions based on two indicators as follow:

1. Does the content validity of Local English test items correspond with School Based Curriculum?

2. Does the content validity of Local English test items correspond with the distribution of their cognitive level in Bloom's Taxonomy?

C. Limitation of the Study

To make this research specific, the researcher will only focus on analyzing the content validity of the Local English test items in SMP Negeri 2 Kebakkramat Karanganyar for students in eighth level in 2008/2009 academic years based on the syllabus in School Based Curriculum which is implemented on it and distribution of cognitive level in Bloom's Taxonomy. The test items consist of first and second semester of students in eighth level. In the first semester test contains of 45 items multiple-choice and second semester contains of 50 items multiple-choice. So, the total number of test item is 95 items. The important parts of the curriculum that will be used for the study are the indicator list on reading, speaking, listening, and writing. But in analyzing this test, the researcher only focuses on the reading and writing skill because the test items in the eight level students only measure both skills.

D. Objectives of the Study

The objectives of the studies are as follows:

1. to describe the content validity of Local English test items based on School Based Curriculum which is implemented on SMP Negeri 2 Kebakkramat, especially in reading and writing skill for the first and second semester,

2. to describe the content validity of Local English test items based on distribution of cognitive level in Bloom's Taxonomy including knowledge,

comprehension, application, analysis, synthesis, and evaluation for the first and second semester.

E. Benefits of the Study

By presenting this research, the researcher hopes that there will be some benefits as follows:

1. Theoretical Benefit

This study gives contribution to the larger body of knowledge and additional information to teaching research especially those dealing with the validity of test.

2. Practical Benefit

a. This study can deepen the understanding in literary field as the reference and develop writer's skill and ability in conducting other researches.

b. The finding of the research can improve the developing of a good quality of test.

c. The finding of the research can help the teachers to organize, develop, or select the good test for their students.

F. Research Paper Organization

In writing this research, the researcher organizes her research paper as follows:

Chapter I discusses Background of the Study, Problem Statement, Limitation of the Study, Objectives of the Study, Benefits of the Study, and Research Paper Organization. Chapter II discusses Previous Studies, Language Testing, Curriculum, and Bloom Taxonomy. Chapter III consists of Type of the Study, Object of the Study, Data and Data Source, Technique of Collecting Data, and Technique for Analyzing Data. Chapter IV consists of Research Findings and Discussion. Then, the last chapter V consists of Conclusion and Suggestion.