

CHAPTER I

INTRODUCTION

1.1 BACKGROUND OF THE STUDY

Testing is a very important part of most classroom teaching-learning activities. Heaton stresses this point saying,

"Both testing and teaching are so closely interrelated that it is virtually impossible to work in either field without being constantly concerned with the other. Test may be constructed primarily as devices to reinforce learning and to motivate the student, or primarily as a means of assessing the student's performance in language."¹

To Tuckman², a test is a tool for evaluation that must itself be evaluated or "tested" for its suitability as a prerequisite to using it. As such, a test has three main functions as follows:

1. to determine whether and to what degree each student is experiencing learning and development
2. to identify those students who are not learning and growing and in particular the area of their deficiency.
3. to determine whether instructional inputs are in general and effective.³

1 J.B. Heaton, Writing English Language Test, Longman Group Limited, 1975, p.1.

2 Bruce W.Tuckman, Measuring Educational Outcomes Fundamentals of Testing, Harcourt Brace Javanovich, Inc. New York, 1975, p.16.

3 ibid, p.43-44

Functionally, there are six kinds of tests: placement tests, progress tests, proficiency tests, aptitude tests, diagnostic tests and achievement tests. According to Gronlund, among these kinds of tests, the one that is most commonly used in schools is the achievement test of which the main purpose is to measure the success of the teaching-learning process which has been done during a period of time. The achievement test given at the end of a period of instruction for the purpose of certifying mastery or assigning grades is called a summative test.⁴

The English summative test which is held at the end of each semester has been administered by the Indonesian Junior and Senior High Schools for many years. This test is generally prepared, administered and scored by the classroom teacher himself. However, since 1990 the head of Depdikbud Office of East Java District considered it essential to hold the even semester English UUB (Ulangan Umum Bersama) as an effort to improve the quality of high school education in East Java. The even semester English UUB results were used to replace the grades of the Final English Summative Test.⁵

4 Norman E. Gronlund, Constructing Achievement Tests, Prentice-Hall, Inc., Englewood Cliffs, N.J. 07632, 1982., p.5.

5 Petunjuk Pelaksanaan Ulangan Umum Bersama (UUB) semester 2 dan 4 SMP/SMA Propinsi Jawa Timur Tahun Ajaran 1989-1990, p.1.

The 1989-1990 even semester English UUB which was firstly administered to the first and second year students of Junior and Senior High Schools in East Java, has the following aims:

1. meningkatkan dan memeratakan mutu pendidikan SMP/SMA di Jawa Timur.
2. membimbing kearah tercapainya pendidikan melalui kegiatan penilaian.
3. mendorong agar guru-guru Bidang Studi/Mata Pelajaran meningkatkan kegiatan belajar mengajar dengan berpedoman kepada Kurikulum, GBPP, Buku acuan dan alat pelajaran serta program yang telah ditetapkan.
4. memberi kepercayaan kepada guru (MGBS/MGMP) untuk meningkatkan ketrampilan menyusun perangkat penilaian.
5. meningkatkan kesiapan mental guru (MGBS/MGMP) terutama dalam menjaga kerahasiaan.

These aims are freely translated as follows:

1. to improve and to equalize the quality of high school education in East Java.
2. to lead to the achievement of education through evaluation.
3. to urge the teachers to improve teaching-learning activity by orientating to curriculum, references and teaching-learning medias as well as the programmes which have been determined.
4. to give confidence to the teachers to improve their skills in constructing evaluation items.
5. to teach character building especially in keeping professional secrets.

8 ibid, p.1

Based on the writer's observation, after the English UUB test was administered, there were a lot of students and teachers complaining about the English UUB items. According to the students, as well as some teachers' opinions, there were some ambiguous options. Besides, many teachers complained about the English UUB test results which were not quite satisfactory. Based on these facts, it is clear that the aims of UUB have not been fulfilled the way they should .

Under all these considerations, the writer is interested in analyzing the English UUB test items to find out whether or not the 1989-1990 English UUB as a formal measuring instrument has fulfilled the conditions of a test construction.

1.2 THE STATEMENTS OF THE PROBLEM

Based on the points presented as the background of the study, there were two major problems investigated by this study. They were formulated as follows:

1. Are the English UUB test items valid?
2. Are the English UUB test items well-formulated?

Specifically, the first major problem is stated in the following minor problems:

- 1.1 Are the structure test items constructed according to the syllabus?
- 1.2 Are the vocabulary test items constructed according to the syllabus?

1.3 Are the reading comprehension test items constructed according to the syllabus?

1.4 Are the conversation test items constructed according to the syllabus?

The second major problem is stated in the following minor problems:

2.1 Do the objective test items meet the principles of constructing multiple-choice items?

2.2 Do the subjective test items meet the principles of constructing essay items?

1.3 THE OBJECTIVES OF THE STUDY

In line with the statements of the problem above-mentioned, the main objectives will be achieved in this study are:

1.3.1 to find out whether the English UUB test items are valid.

1.3.2 to find out whether the English UUB test items are well-formulated

In order to achieve the main objectives, the writer specifies them into the following minor objectives:

1.3.1.1 to find whether the structure test items are constructed according to the syllabus.

1.3.1.2 to find whether the vocabulary test items are constructed according to the syllabus.

1.3.1.3 to find whether the reading comprehension

test items are constructed according to the syllabus.

1.3.1.4 to find whether the conversation test items are constructed according to the syllabus.

1.3.2.1 to analyze whether the objective test items meet the principles of constructing multiple-choice items.

1.3.2.2 to analyze whether the subjective test items meet the principles of constructing essay items.

1.4 THE SIGNIFICANCE OF THE STUDY

The information obtained from this study will describe about the validity of English UUB test items and whether the English UUB test items have fulfilled the principles of constructing multiple-choice and essay items. It will give a clearer picture about the flaws of the test items, and at last, the results of this study are expected to give the consideration to the test-item constructors which items can be maintained if the test will be administered again next time, and which items should be improved. Hopefully, the results will stimulate the SMA English test-item constructors to give more consideration to the criteria of a good test-item construction.

1.5 SCOPE AND LIMITATION

This analytical study is limited to the 1989-1990

English UUB for the first grade senior high school of the second semester in Region I North Surabaya since the writer is a teacher of the first grade senior high school in Region I Surabaya.

The test items analyzed consist of the multiple-choice items on structure, vocabulary, reading comprehension and conversation, and the essay items on reading comprehension.

It is also limited to the analysis of the contents as they match the objectives. The study of the learning-outcome and content analysis is mainly based on the syllabus.

Since the limitation of the study here is to analyze the items construction based mainly on the guidelines or principles of constructing multiple-choice items and essay items, the item analysis concerning with the analysis of the level of difficulty, discriminating power is excluded due to the time constraint.

1.6 ASSUMPTION

It was assumed that the UUB test constructors have tried hard to apply the principles of constructing test items. The test was also assumed to have been validated and the subject matters covered in the syllabus have been completed by the classroom teachers.

1.7 THEORITICAL FRAMEWORK

This study was based on the principles:

1.7.1 The Characteristics of a Good Test.

According to J.B. Heaton, a good test possesses four characteristics, i.e.: validity, reliability, discrimination and backwash. A good test, especially an achievement test, requires content validity as the most important characteristic.

1.7.2 The Steps of Planning a Test.

A careful planning is necessary to obtain the greatest extent of content validity. By planning the test carefully, relevant learning outcomes through representative samples and dependable information can be obtained.

1.7.3 The Principles of Constructing Multiple-choice Items.

1.7.4 The Principles of Constructing Essay Items.

1.8. DEFINITION OF KEY TERMS

To avoid misinterpretation or misunderstanding, it is necessary to define the key-terms used in this study:

1.8.1 UUB

UUB stands for Ulangan Umum Bersama, which is an English achievement test administered at the end of the semester. It is meant to measure the students' achievement of particular teaching learning materials.

1.8.2 Achievement Test

It is a test to measure the extent to which an

individual student has mastered the specific learning outcomes and content which are described in the syllabus of a teaching program and have been presented in formal classroom situation. e.g. end-of-year test.

1.8.3 Objective Test

It is a test of which questions are asked in such a way that there is only one predetermined correct answer. It is objective in terms of scoring.

1.8.4 Subjective Test

It is a test which the scoring depends on the subjective judgement. It means different scorers may result different scores.

1.8.5 Multiple-choice Item

It is an objective item which contains a stem, one correct option and three or more distractors and the students should choose only the best answer.

1.8.6 Validity

It is the extent to which a test measures what it is supposed to measure.

1.8.7 Content Validity

It is the validity of a test which is brought about by appropriate, adequate and representative samplings of larger tests of activities and learning outcomes to be tested and content to test these learning outcomes.

1.8.8 Content Analysis

It is an analysis of test items to determine whether each of the item really tests the intended learning outcome, whether the subject matter topic is appropriately selected to test the learning outcome, and whether the items as a whole represent adequate sampling of learning outcomes and contents.

1.8.9 Learning Outcomes

Intellectual outcomes in the cognitive domain which are divided into six classes; they are knowledge, comprehension, application, analysis, synthesis and evaluation.

1.8.10 Construction

A group of words which are combined into one sentence based on a certain manner/ way.