

CHAPTER I

INTRODUCTION

1.1 Background of the Problem

In order to increase the quality of education, the Indonesian Ministry of Education has administered a certain standardized test that is called Ulangan Umum Bersama (UUB) to the Senior High School and Junior High School students, at the beginning or in the middle of Desember.

The purpose of the English UUB, however in the Senior High School, is to know the students' ability according to their achievement in the English subject at the high school in a semester. The test scores they get can be useful for them to know the effectiveness of their learning process. In terms of interpretation of test scores, the English UUB is a norm-referenced test.

The 1989-1990 English UUB for the fifth semester students of Senior High School consists of 45 items that is 40 items of multiple choice questions with five options and 5 items of subjective questions based on the reading passage.

The English UUB is used as standard and based on the syllabus at the Senior High School. The English can, therefore, be considered as a measure of the extent of the students' achievement in their English subject at the Senior High School. Therefore, this study is undertaken to know whether the UUB as a formal, large scale, "standardized" instrument has fulfilled the conditions of a test construction. While to be considered as a good test, it should possess many characteristics, some of them, is that the test should fulfil the criteria of item difficulty, discriminating power, and the effectiveness of the distractors. So, this study conducted to analyze these three criteria.

1.2 Statement of the problem

In line with the reason for choosing the topic, this study is intended to see whether the 1989 English UUB as a measuring instrument of the students' achievement has fulfilled the standard of a test construction. The major problem is : Are the 1989-1990 English UUB items of A3 program for the third year Senior High School well-constructed ? This study, therefore, is undertaken to analyse and evaluate the items of the English UUB, through the Item difficulty analysis, Item Discriminating Power and the effectiveness of the distractors.

The statement of the problem under discussion is formulated as follows:

1. Do the questions in the English UUB fulfil the level of item difficulty ?
2. Do the questions in the English UUB fulfil the level of item discriminating power ?
3. Are the distractors effective enough ?

The answers to these questions will together answer the major problem.

1.3 The objectives of the study

Derived directly from the above formulated problems, the major objective of this study is to investigate to what extent the English UUB has fulfilled the conditions of a test construction for the third year students (A3) of SMAK St. Stanislaus Surabaya during the school year of 1989-1990, so it would be useful for the teacher to prepare all his pupils to be successful in their UUB and the minor objectives are to know:

1. How difficult each item was
2. How well each item discriminated between high-and low-scoring students.
3. How effectively the alternatives in each item functioned.

1.4 Significance of the study

Since English UUB became one of the standardized instruments to test the students achievement, the information obtained from this study will give the description about the quality of the UUB test items, whether they have fulfilled the condition of good test items. In the sense of the items are too easy or difficult for the students, whether the test items differentiate the good students from the bad ones, or the distractors are effective enough to attract the poor students away from the correct answer.

The data and information obtained from this study are expected to be able to be used as a teacher's consideration to maintain the items if the test will be administered again next time or which items should be improved. Hopefully, it will encourage the item constructors to make a better improvement in constructing a good test item.

1.5 Limitations of the study

Realizing how broad the topic of this study would be, I would like to limit it. This analytical study will be concerned only with the multiple choice questions of the 1989 - 1990 English UUB for the fifth semester students (A3) of SMAK St. Stanislaus Surabaya. The essay questions

won't be included in this study as the scoring of the essay test depends on the subjective judgement of the scorer, so different scorers may result different scores. The objective test, on the other hand, will produce the same result no matter how the examiner marks the testee's paper as it always provides one predetermined correct answer for each item. Therefore, I only deal with the multiple choice questions for the analysis as they match the objectives.

I also limit the analysis here according to the analysis of the level of difficulty, discriminating power, and the effectiveness of the distractors of the items. The analysis is based on the norm-referenced interpretation, which describes how a student's performance compares with that of others.

1.6 Assumption

It is assumed that the subject given by the third year English teacher of group A3 of SMAK St. Stanislaus are in accordance with the test material and the teacher's technique is right. Yet, it is also assumed all the students who have different level of intelligence did the analysis test items themselves.

Besides there has always been an assumption that all English instructors are capable of constructing language tests the test items met the requirement of content

validity and these language tests are considered as good measuring instruments. This may not always be the case. Very few have seriously questioned the quality of this English UUB.

1.7. Definition of key Terms

Before proceeding to further discussions, it is better to make clear some key terms used in this study.

UUB; it is the short for Ulangan Umum Bersama which is an English achievement test administered at the end of every semester by the Indonesian Ministry of Education, to measure the high school students' achievement in the English subject.

Multiple Choice Item; it is a question of a test which is cast in the form of a problem and a list of suggested solutions. The problem can be in the form of a question, a complete sentence or an incomplete sentence.

Options; they are the suggested solutions to a problem which consist of the best or correct, solution called distractors.

Essay Questions; it is a question of a test which is cast in the form of a problem. The problem can be in the form of an informative question, a yes / no question, or an inferential question.

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.

A Norm-Referenced Test; it is a test administered with a purpose to compare the level of achievement or performance which is shown by the total group to which he belongs. The test scores are used to rank the students, to compare the behavior of an individual student with the behavior of others.

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.

Subjective test; it is a test that requires an essay type answer and of which the scoring depends on the subjective judgement. It means different scorers may result different scores.

Items Analysis; it is a study of the test items to see the item difficulty, item discrimination, and effectiveness of the distractors.

Stem; it is a question or an incomplete statement in the multiple choice items which presents a problem situation.

Distractors; they are the plausible wrong answer / the incorrect solutions in the options.

Index of item difficulty analysis; it is an analysis of test items according to the level of difficulty to determine whether each of the items really test the intended learning outcome, whether the subject matter topic is appropriately selected to test the learning outcome and contents. So It can determine the percentage of the sample who answers each item correctly, implying how easy or difficult the item is. The formula for computing the item difficulty is as follows:

$$P = \frac{R}{T} \times 100$$

where: 1)

P = the percentage who answers the item correctly

R = the number of students who answer correctly

T = the total number who tries the item

Index of Item Discriminating Power Analysis; it is an analysis of test items by discriminating between high and low level examinees in order to determine how well each item by calculating the

1) Norman E. Gronlund "Constructing Achievement Tests", Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1982, p.102

difference between the proportion passing the item in the upper or lower group. Customarily the papers of the upper 27 percent and the papers of the lower 27 percent of the class are used for item analysis as Kelley points out " The selection of criterion groups based upon the upper 27 percent and lower 27 percent of the papers provides the greatest confidence that the upper group is superior in the trait measured by the test as compared to the lower group " 2) In this case, the larger the group of omitted papers (46%), the greater the probability that the upper and lower groups are different. To compute the difference between the proportion passing an item in the upper or lower group, the following formula used :

$$D = \frac{R_u - R_l}{\frac{1}{2} T}$$

Where : 3)

D = the index of discriminating power

R_u = the number in the upper group who answers the item correctly

R_l = the number in the lower group who answers the item correctly

$\frac{1}{2} T$ = one half of the total number of students included in the item analysis

2) Truman L. Kelley, " The selection of Upper and Lower Group for The Validation of Test Items", Journal of Educational Psychology 30 (1939) : P.17-24

3) Gronlund, op.cit., p.103

The Effectiveness of the distractors; it is determined by comparing the number of testees in the upper and lower group who selected each incorrect alternatives.

1.8. Theoretical framework

This analysis will be done based on these principles

1.8.1 The principle of finding the level of difficulty

The level of difficulty of each item is expressed by the difficulty index or facility value. By knowing the difficulty index of each item, the teacher will know which items are too difficult and which ones are too easy for the students.

1.8.2 The principle of finding the discriminating power of each item.

The discrimination index of each item shows the difference between the proportion passing an item in the upper and lower group of testees. From the discrimination index, the teacher will know whether the items can differentiate the able students from the poor ones.

1.8.3 The theory of the effectiveness of the distractors.

Good distractors should attract a sufficient number of testees, and they should attract more testees from the

lower group than from the upper group. If a distractor attracts none of the testees, it is categorized as an ineffective distractor and when it attracts more testees from the upper group than from the lower one, it is called a poor one. The effectiveness of the distractors is determined by comparing the number of testees in the upper and lower group who selected each of the incorrect alternatives.

1.9 Organization of the Thesis.

This thesis which analyze the Items of 1989-1990 English UUB for the third year students of A3 program of SMAK St. Stanislaus Surabaya consists of five chapters with the following organization

The first chapter is the Introduction. It deals with the background of the study, statement of the problem, the objectives, the significance, limitations, assumption, definition of key terms, theoretical framework, methodology of the study, and organization of the thesis.

The second chapter presents the review of related literature. In this chapter, the writer explains the theoretical background and review of related studies.

The third chapter discusses the methodology of research that deals with the research design, the subjects, instruments, procedures of collecting data and techniques for data analysis.

The analysis of the data and discussion are presented in chapter IV.

The findings and their interpretation will be in chapter V.

The last chapter (chapter VI) is the conclusion and suggestion.