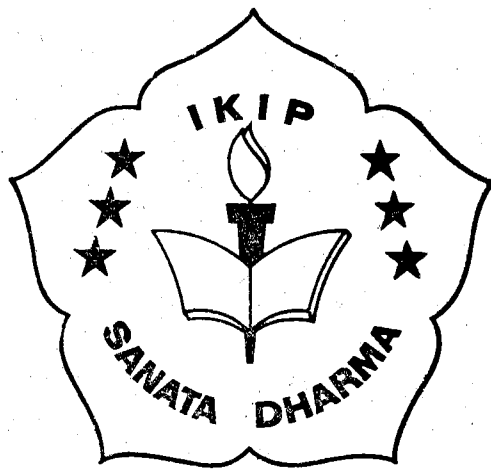


AN OUTLOOK ON LANGUAGE TESTS
(WITH THE EXPLORATORY STUDY ON ESSAY TESTS)



by

Hendro Susanto

I. 1415



THE DEPARTMENT OF ENGLISH LANGUAGE
AND LITERATURE
FACULTY ARTS AND LETTERS
IKIP SANATA DHARMA
YOGYAKARTA

1981

A C K N O W L E D G E M E N T

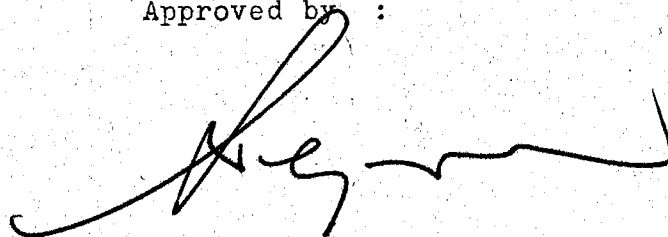
This thesis is written as partial fulfilment of the requirements for the degree of Sarjana in the English Department of IKIP Sanata Dharma Yogyakarta.

Having completed his thesis, the writer wants to express her deepest gratitude to Mr. Soepomo Poedjosoedarmo Ph.D. as his adviser, Mrs. Gloria Poedjosoedarmo Ph.D. and Drs. J. Bismoko as his readers who have patiently given suggestions and guidance.

Genuine thanks are also due to my friend, Miss Chendra who has continuously encouraged me to finish this thesis and lent me some books, to my parents who have supported my study, to the librarians who have helped me choose suitable books for reference, and in the end, to many others who have directly or indirectly helped me.

Approved by :

I.



(Soepomo Poedjosoedarmo Ph.D.)

First Adviser

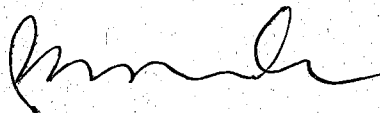
II.



(Gloria Poedjosoedarmo Ph.D.)

Second Adviser

III.

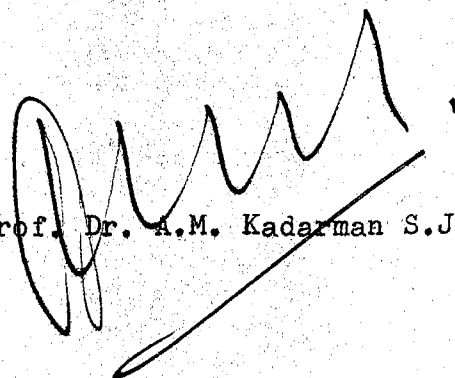


(Drs. J. Bismoko, Dip. TESL)

Third Adviser

Rector of IKIP

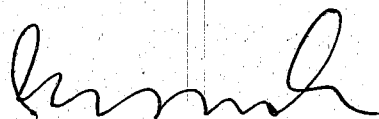
Sanata Dharma



(Prof. Dr. A.M. Kadarman S.J.)

Head of the English Department

IKIP Sanata Dharma



(Drs. J. Bismoko, Dip. TESL)

TABLE OF CONTENTS

	<u>page</u>
ACKNOWLEDGEMENTS	i
TABLE OF CONTENTS	11
INTRODUCTION	1
CHAPTER I : TEACHING AND TESTING	3
I. 1. Language Teaching	3
I. 2. The Functions of Testing	8
CHAPTER II : LANGUAGE TESTING IN GENERAL	12
II. 1. The Process of Testing	12
II. 2. The Variables of Language To Be Tested	18
II. 3. Types of Tests	21
II. 3. 1. Based on The Aims	23
II. 3. 2. Based on Who Makes Them	30
II. 3. 3. Based on How They Are Administered	33
II. 3. 4. Based on The Scoring	36
II. 3. 5. Based on The Language Variables Tested	49
CHAPTER III : ESSAY TESTS	66
III. 1. The Nature of Essay Tests	66
III. 2. The Advantages And The Disadvantages of Essay Tests	69
III. 3. The Stages in Essay Tests	73
CHAPTER IV : EVALUATIONS OF ESSAY TESTS	82
IV. 1. Validity	82
IV. 2. Reliability	87



IV. 3. Scorability	<u>page</u> 88
IV. 4. Administrability	92
IV. 5. Economy	93
IV. 6. Preparability	94
CONCLUSION	96
BIBLIOGRAPHY	99

I N T R O D U C T I O N

Language testing is of the same importance as language teaching and therefore, these two activities must be given equal attention. Up to this time, however, many educators as well as teachers only pay attention to language teaching. This can be seen from the available textbooks on language instruction. They deal mostly with language teaching and only a few deal with language testing. The reason for this unbalanced proportion between language teaching and testing is that language testing used to be considered unimportant in the past because there were only a few students in the class. The class teacher already knew their achievement from the daily activities done in class. Nowadays, since the number of students is getting bigger and bigger, it is almost impossible to determine the individual achievement from the class activities alone. Thus, testing is a must for assessing the students' achievement. Consequently, the urge on the part of the teacher to know about testing, as well as to construct tests becomes greater. Since there have been only a few textbooks on language testing, this thesis is meant as a contribution to this field. By reading this thesis, it is hoped that the class teachers will be able to use the most effective type of test for measuring their students' mastery of the language they are learning.

Although this thesis will discuss tests in general, special emphasis will be given to essay tests. The choice of essay tests is based on the important usage of this type of tests. Thus, various language tests will be discussed first. This is followed by a further and detailed discussion of essay tests. From the discussion, it will be con-

cluded whether this type of test is worth using.

The materials in this thesis were based partly on the "Seminar on Language Testing" held by the doctoral students in the English Department of Sanata Dharma Teachers' Training Institute in which the writer took part with the topic on essay tests. Other sources for this thesis were various articles and books concerning language testing as well as language teaching.

The order of presentation for this thesis will be as follows : Chapter I will discuss the relationship between language testing and language teaching; the first part of Chapter II will discuss the process of testing. Subsequently, it will discuss the various types of tests looked at from different points of view. This discussion is meant to give a clear picture of the place of essay tests among various language tests; Chapter III will discuss essay tests in connection with the term 'essay', the advantages and the disadvantages and the stages found in this type of test; the last chapter will discuss the evaluation of essay tests. From this discussion, conclusions will be drawn; this will be shown in the last part of this thesis. In this discussion, whether essay tests are worth using will be decided, followed by some suggestions on how to get the best results in using the tests.

CHAPTER I

TEACHING AND TESTING

Language testing is concerned with language and language learning besides the testing or the evaluation itself. Therefore, it involves both linguistics and psychology. However, this section will only deal with the testing and the linguistics and it will not talk and discuss about the psychology part separately.

Language testing is closely related to language learning, and thus, to language teaching and vice versa. The approach, the method and the techniques of teaching are made in such a way that they will make it easier for the students to take tests. This means that the students are given tests in the same way as when they do the exercises during their learning. If a test uses completely new procedures, the learners will become confused and thus, the test will not be valid.

This chapter is divided into two parts. The first part deals with language teaching, including some methods used in teaching, while the second part concerns language testing in connection with language teaching.

I.1. Language Teaching.

Teaching and learning are two activities that are considered as one unity; this means that teaching and learning are actually one process, the process of teaching and learning. In this process, it is not only important to give attention to what the teachers are doing but it is also important to know and improve the methods and techniques used so that the teaching will reach its maximum results.

Before discussing specific approaches, methods or techniques, the meaning of these terms will be given so that they

will not be confusing. The relationship is that techniques carry out a method which is consistent with an approach.

An approach can be described as a set of correlative assumptions concerning the relationship between the nature of language and the nature of language teaching and learning.¹ It describes the nature of subject matter to be taught. It states a point of view, a philosophy, an article of faith; something which one believes, but cannot necessarily prove. In short, an approach is axiomatic.

There are several approaches in language teaching. One of these is the aural-oral approach. Some assumptions can be found within this approach. The most important of them is that the aural-oral aspects should be taught before reading and writing aspects. Understanding the spoken language is taught more efficiently before oral production, and is needed as a first step toward production; reading is a first step in learning to write.

Another important assumption in this approach is that languages are habits and habits are established by repetition; languages must therefore be taught through repetition of some sort. There is no need to explain grammatical items.

Method is an overall plan for the orderly presentation of language material, no part of which contradicts any other part. All of this material is based upon the selected approach. From this definition, it can be concluded that there can be many methods just within one approach.

No matter how important the selection of one method is, it does not mean that it is the only determining factor in language teaching. The nature of the student's language as

1. Edward M. Anthony. Approach, Method, and Technique.

compared to English (the target language) will influence the orderly presentation of language to the students. The age of the student, his cultural background, and his previous experience with the target language modify the method employed.

There are various methods that are used in language teaching. The most common types in use are (1) the Direct Method, (2) the Grammar-Translation Method, (3) the Pattern-Practice Method, (4) the Eclectic Method, and (5) the Mimicry Memorization Method.

The grammar-translation method was commonly used in the past. This method was so popular among the teachers that it seemed there was no other method beside this one. There was no attempt to search for new methods since everything that was done within this method was always considered right. There was no question concerning its results.

Using this method, the teachers' task was quite easy. They just gave grammatical items as they were and some translation exercises as well. Thus, under this method, the teaching began with rules, isolated vocabulary items, paradigms, and translation. Pronunciation was either not taught, or was limited to a few introductory notes. Grammar rules were memorized as units, which often included illustrative sentences. In this way, it can be said that the students learn about the language and not learn the language itself because using this method, the students are required to know the rules of the language and they need not be able to speak it.

Furtunately, this situation has changed. Nowadays, the teachers as well as the educators realize the important need to improve language teaching as they begin to realize the urge to communicate using the language learned. Among the most developed methods are the ones mentioned above.

Two of the methods above are based on one approach. They are the mimicry memorization and the pattern practice methods. Both share the factor of goal; they aim at automatic oral production coupled with skill in understanding the stream of speech. They each function best under intensive course conditions. However, the order of presentation differs.

The mimicry memorization (mim-mem) method begins with a situation such as getting a room at a hotel and at the theatre. The learner must mimic a native speaker, real or recorded, and remember a rather large number of useful sentences within the situation. From the memorized sentences, certain structures, expressions and certain phonemes are drawn for particular emphasis and drill. There is nothing in the mim-mem method which contradicts the assumptions which make up the aural-oral approach.

On the other hand, the pattern practice method ideally uses bilingual comparison at the very beginning, and starts with grammatical and phonological structures chosen. These structures are drilled and built up into a situation through the addition of lexical items. Again, there is nothing here which contradicts the aural-oral approach. Both methods have been used with success. Both lie with the same approach, and yet, each has distinctive features.

The last method to be discussed is the eclectic method. Before talking about this method, however, the method that becomes the basis of the eclectic method will be discussed; it is called the direct method. There are several characteristics of the direct method. The main ones are: (1) The use of everyday vocabulary and structure. (2) Grammar taught by situation. (3) Oral teaching of grammar and vocabulary. (4) Extensive listening and imi-

tation until forms become automatic.

The eclectic method is a compromise between the direct method, with its many demands on the teacher coupled with the more formal methods based on grammar rules and translation. So in this method, a certain amount of translation is introduced, with some deductive grammar, and some audio-visual aids. Thus, it can be said that this method is a combination of many methods. It takes the good points of these methods, and combines them.

A technique is an activity that takes place in the classroom. It is a particular trick, strategem, or contrivance used to accomplish an immediate objective. Techniques must be consistent with a method and therefore, in harmony with an approach as well.

Techniques depend on the teacher, his individual artistry, and on the composition of the class. It can be said that teaching is an art because each teacher has his own techniques in tackling particular problems. For example, in teaching the difference between the pronunciation of English /l/ and /r/ to some Indonesian students, some teachers get good results only by imitation. Some other teachers may use another technique that requires the use of pencil in the mouth to prevent the student's tongue from touching the alveolar ridge, hence inhibiting the pronunciation of /l/.

Laboratory tape-recorder, slides, teaching machines, phonographs, closed circuit television and other audio visual aids are techniques. The teacher may use them in his teaching and the effectiveness of these devices depend partly on how the teachers use them. Different techniques are needed to teach different elements and skills of a language.

Understanding the terms 'approach', 'method' and 'tech-

nique', we can see that teaching in the class will be much influenced by the approach and method used in the class and the techniques presented cleverly by the teacher. The more the teacher knows about teaching techniques and is able to use them, the more interesting his teaching will be. Consequently, no matter how great his store of teaching techniques is, it will be useless if he does not know how to use them in effective ways. What effect does the teaching have on testing? This question will be answered in the next part of this chapter.

1.2. The Functions of Testing.

For many years textbooks writers and educators have been concerned with language teaching only; only a little attention is given to language testing whereas actually language testing is just as important as language teaching. It is not only important for teachers but it is also useful for the students. Language testing can even be said to be the ultimate activity in the whole process of teaching-learning.

Language testing is a tool to measure the ability, knowledge, intelligence, and capacity of an individual or group. The test has a function in controlling the learning process both for the teacher and for the students.

After giving a test, a teacher can reflect whether his teaching has been successful or quite a failure. If the results of the test are too good, there may be two possibilities, either the test is too easy or the students have mastered the materials completely. If the teacher considers that the test was too easy, he may give a more difficult test on the next occasion. On the contrary, if he considers

that the students have really mastered the materials, he may then continue his previous procedures in his next lessons. However, if the results of the test are too bad, he may then reflect that either his test is too difficult or he has given insufficient explanation on the materials tested. If the class considers that the test is too difficult the teacher may give a repeat test which is somewhat easier. But if the failure is due to the lack of understanding on the part of the students, the teacher may then change his future teaching with different techniques or he may explain in such a way that the students can really master the materials taught.

Seeing the results of his teaching is just one function of giving a test from the teacher's point of view. There are several other functions of giving a test. These functions are : (1) as guidance in giving grades, (2) as preparation for further tests, (3) as a means to stimulate himself to prepare and study his lesson plans better, (4) as a way to find out the students' difficulties, (5) as a means of evaluating textbooks and other materials, especially the ones used in the class, and (6) to determine which students can be promoted and which can possibly be given scholarships.

In past years, when the classes were small, meaning to say that the number of students was relatively not large or to be more exact, there were less than ten students in each class, it might be understood if testing was considered unimportant as guidance in giving grades since the teacher might already know the students' achievement in the daily exercises. This, however, is not the case nowadays. The class is big now so that the urge to give the test to the students cannot be delayed any longer.

If the test proves to be successful, then the teacher

might repeat his teaching procedures in his next lessons and thus, the students are supposed to be able to do the with the same success. On the contrary, if the test proves to be a failure, he might prepare his teaching procedures in his next lessons better so that the students will be able to do the next tests better.

Furthermore, if he finds out in the test that the majority of his students cannot do it, then he perhaps realizes that this is caused by his shortcomings in preparing his lesson plan so that in his next class he might be able to do better.

By having the test, the teacher might increase his own effectiveness by making adjustments in his teaching to enable certain groups of students or individuals in the class to benefit more. The teacher might give certain attention to the the students who need special help.

By giving tests the teacher might estimate whether the materials used are quite representative or not. If not, he might search the other materials for his reference in teaching.

From the test, the teacher might know the students' real achievement so that the is able to determine in the end of the school year which students should stay in the same class and which of them can be promoted.

Those are the functions of a test from the point of view of the teacher. However, it is not only the teacher who might receive good effects from a test. The students might be able to use the test for their own benefits as well. For them, the main function of doing a test is to know their own capacities if they compare the result of that test with the

earlier tests and even with the results of their classmates. If their work is below standard, that is, their achievement is worse than the achievement of their friends, they might think that they should work harder before the next tests. In this way the test functions as a teaching device, as far as it sets out to measure the performance of the students as fairly as possible without in any way setting traps for them. In other words, a good test can be effectively used to motivate the students.²

To conclude the description above, it can be said that although language testing mainly is concerned with the performance or achievement of the students, the types of tests or examinations that evaluate their knowledge of the language often determine the methods and techniques which are being used by the teachers.

Some textbook writers sometimes mention the term wash-back or backwash. This term refers to the effects of a test on teaching. If a test has good backwash effects, it will exert a good influence on the learning and teaching that takes place before the test.³

On the other hand, if a test has bad backwash effects, this means that either the students do not understand the materials or they have not prepared adequately.

2. J.B. Heaton. Writing English Language Tests. Longman, p.2.

3. Ibid. p. x.

CHAPTER II

LANGUAGE TESTING IN GENERAL

Testing is an important activity in the whole process of teaching-learning. Besides giving feedback to the teacher of what he has taught during a certain period of time, it proves to give benefits to the students as well since by taking the test, they may have feedback concerning what they have been learning.

Since testing is important, tests should be properly prepared. This activity might occupy the teacher's time considerably although the test itself only takes a short period of time. In his preparation, the teacher has to know what tests are made of and how they differ. In other words, the following question is to be answered in his preparation : What language variables are to be tested ? The answer to this question is in the first part of this chapter, following the discussion on the process of testing.

How the test differs will be discussed in the next or last section of this chapter under the headings : Types of Tests. In this section, different types of tests will be looked at from different points of view. This section is further subdivided into the points of view that determine the types of tests.

II.1. The Process of Testing.

The process of testing can be divided into three parts: preparation, administration and marking. The portion of these three activities is different from one to another. A teacher may need a greater amount of time for the preparation than for the other activities in one test. This is the case when he uses an objective test in testing his students. He

has to prepare the items of which there are many. However, in the administration as well as in marking, he only needs a short period of time. He need not read the questions again in the marking. What he needs is just to read the one-word answer for each question in the answer sheets of his students. This marking can even be done mechanically, which means that the correction of his students' work can be done using a machine. An example of a test using this procedure, hard preparation and easy marking, is an objective test.

On the other hand, there are tests which need little preparation and are easy to administer but which take a great portion of the teacher's time in marking. An example of such a test is a composition test. Using this kind of test, the teacher's task in composing the question is much easier than when he uses an objective test, for example. He just thinks of a topic and perhaps a little guidance for his students and then lets the students do what is being asked. Nevertheless, in marking, it is quite the reverse. The teacher has to read long answers, hence it takes a lot of his time if he uses a composition test while if he uses an objective test, this does not happen.

Still, there are other types of tests which need little preparation and which are easy to mark as well but which take a large portion of the teacher's time in the administration. An oral test is just one of the examples. If a test is done orally the teacher should examine one or two students at a time so that for the whole class he needs several hours.

Thus, in the preparation it should be decided what kinds of tests are to be used. Of course, this choice is

influenced by the decision concerning what language variables are to be tested. Suppose, a teacher wants to test his students' speaking skill. The best way to do this is by giving his students an oral test.

After deciding the kind of test to be used, the teacher may prepare the questions. These questions are better graded from the easy to the more difficult ones or from the previous to the later lessons. The number of questions should also be considered. It has to be in accordance with the time available. Those questions should also be representative. This means that all of the materials should be covered in the limited number of questions. The directions in the test should also be clear so that the students know what they should do and there is no misunderstanding of what to do.

In preparing tests, several criteria of a good test should also be fulfilled besides the things that have been mentioned above. These criteria are as follow :

1. Validity : A test is valid if it measures what it intended to measure. This criterion sometimes is also called the appropriateness or relevancy of the questions in a test. Suppose, a teacher wants to test the students' knowledge in composing answers to English questions. This test should be composed in such a way that their knowledge of geography, for example, does not influence the result of the test. In other words, it can be said that the test possesses validity if it has questions related to the objectives of the test.

According to this criterion, then, the teacher already has the answers to the questions asked in the test as early as in the preparation stage, so that the test he planned will not fluctuate very much.

2. Reliability : A test is reliable if it gives a consistent measurement. For example, if a student does two tests of equal difficulty on the same material, the result must be approximately the same unless other factors intervene, such as illness or the fact that he has not worked for his next test. Or, if the test is given to two students who are of the same level, the result must be approximately the same. Thus, to know the reliability of the test, the tester can give two tests of equal difficulty and length to the students or he may give the same test twice provided that neither teaching nor learning has taken place in the interval.

In other words, based on this principle, the scores should not be subject to fluctuations, and hence, the persons obtaining the scores should maintain their standings relative to one another.

3. Scorability : A test is scorable if it can be scored with ease so that the users are able to handle it. Based on this criterion, perhaps the most scorable test is an objective test since this kind of test can be scored mechanically. However, the other tests can be made more scorable if in the preparation, the teacher already prepares the answers of the given test.

This criterion is sometimes referred to as objectivity. The test is objective if there is not much fluctuation in the scores given by different examiners on the same test. For example, there are two examiners who are giving an oral test. If these two examiners give more or less the same scores to each student, we may then call this test objective. It may also happen that the same teacher scores the same student's test twice. In this case, if the scores are about the same, then it can be called objective.

4. **Administrability** : A test is considered to be administrable if it can be given under the prevailing conditions by the person available. In the preparation, the teacher has to consider this principle so that he is fully aware whether the test he made can be given to the students or not. If not, then he should not give it or it is better to change it. For example, he prepares to give an oral test whereas he knows that he only has an hour for almost fifty students and there are no other teachers who can help him administer it. In this case, the test only has low administrability since it is impossible to have enough time to test each student. If he knows this principle, he must change the planned test with another kind of test, namely, a written test.

5. **Economy** : A test is economical if it measures what it is intended to measure in a reasonable time considering the testing situations such as setting and facility. An objective test with a large number of questions will be uneconomical if there is no printing equipment in schools since it will be very time-consuming. In this case, an essay test is better used because the teacher is able to write the questions within a short time on the blackboard.

Those are the criteria of a good language test. These criteria are mainly based on Lado's Language Testing. To these criteria, adequacy and preparability are sometimes added. A test is adequate if the questions in it are representative. This means that this test includes all the materials given to the students in proportion to their importance in the course. Some of the questions must be easy, but some others must be average and some are even dif-

ficult. It is important to keep the ratio of the proportion constant or else the results will be distorted.

With regard to the preparability of a test, it can be said that a test is preparable if there is not much time needed for composing its questions.

If these criteria have been followed, the administration as well as the marking will be easy since these two activities have also been thought of in the preparation. If, however, these criteria of preparing a good test are not carried out consistently, the administration and the marking will be very time-consuming. In the preparation the answers have been thought of so that when the time for marking the students' work comes, there will be no questions as to what should be scored.

Concerning the time allotted in the administration of the test, it is wise to give enough time so that the majority of the class can finish the test without being in a hurry. Thus, if around 85% of the testees can finish the test, the time given is sufficient.

In conclusion to the discussion of the process of language testing, it can be said that the preparation or to be more exact, the composing of the blueprint is generally more difficult and hence, needs more attention from the examiners than the other two stages. The blueprint is a plan in which all the details concerning the test, including the objectives and the number of questions are thought of. In the administration stage, the task of the examiner is concerned mostly with keeping the students from looking at each other's work or from cheating during the examination or the test. In marking, the difficulty usually arises in trying to be objective. If this difficulty can be overcome then this stage

becomes easy for the examiner since everything connected with the answers has already been thought of in the preparation stage.

II. 2. The Variables of Language To Be Tested.

The discussion of language testing so far has been somewhat general. That is we have discussed testing the ability of students either in using a language or in their mastery of the language learned. Language, however, can be analysed into a number of variables. So the question is : Which one of these variables is going to be tested ?

There are two kinds of variables in a language. They are the language elements and the language skills. The language elements themselves can be subdivided into phonemes, morphemes, syntax, discourse and vocabulary. Each of these elements is studied separately in the branches of language study. On the other hand, language skills are also subdivided into five skills, namely, reading, writing, speaking, listening comprehension and translation.

Since the elements and skills of language are studied in different branches of language study, they can be tested separately. Phonemes are studied in phonology and they are tested using pronunciation tests, either orally or in written form. However, since it is a test of pronunciation which in turn becomes the basis in speaking, it may best be tested orally. In written tests, the testees can be asked to describe some problematic and difficult sounds such as the difference between the sounds /l/ and /r/ to some Indonesian students. This can be asked

in written form, for example, using this question : How do the positions of the tongue when we produce the sounds /l/ and /r/ differ ? It can also be asked using pictures as the answers, for example using this question : Draw a picture of the positions of the tongue when we produce the sounds /l/ and /r/.

Morphemes and words are studied in the branch of language study called morphology while syntax is the study of the word order, word group, clause and sentence. Both morphology and syntax are part of grammar.

The units in grammar form a rank scale. The highest unit on the scale in grammar is called the sentence; while the lowest unit is the morpheme. Between these two units are words, groups or phrases, and clauses. Perhaps most English lessons in the past and even today comprise and concern grammar since grammar is the easiest thing to teach and to test. The students can be asked to make sentences using certain constructions such as : Make a sentence using : so that, such that, the same as, different from, more than. The teacher can even ask the students to change the sentences, either from a positive sentence into negative and interrogative sentences or from present tense into past tense or present continuous tense. The test on grammar can be carried out using substitution drills or transformational drills as well.

A paragraph, a higher unit above the sentence, is outside grammar since its unity is mainly based on categories of logic, not of structure. This unit and also a discourse can be taught in the composition class. Consequently, the best test that can be used for paragraph

and discourse structure is a composition test. In this test, the learner are required to use their whole mastery of the language learned in the production of their compositions.

The last language element to be discussed is the lexicon. This mainly concerns vocabulary building. In teaching vocabulary many techniques can be used, such as :

1. demonstration by using an object, a picture or by performing an action.
2. pictures which can be found in photographs or which can easily be drawn on the blackboard.
3. explanations using words which usually describe the object or by giving synonyms, words that mean the same or nearly the same as another word, or by giving the translation in the students' native language.

Vocabulary tests are not much different from the techniques of teaching above. The students are sometimes required to give the antonyms, words that are contrary or opposite to another. The teacher may also use jumbled letters to form a word, for example, the word diamond is written as

nddmoai

okob for the word book

eiplnc for the word pencil.

Of course, this technique must be given to the students in the exercises first before it is used for testing.

Those are the elements of a language which must be mastered by the students to make them able to use the language. Knowing these elements alone is certainly not the purpose of language learning because only knowing the elements is the same as only knowing about the language. To know the language itself, one has to master as well as practise the skills of the language. Thus, it can be said that the elements of

language serve as a means to know and be able to use the language. One should not stop with knowing the elements of the language because in that case he would only know about the language and when asked to use it, he would certainly be unable to do so.

As mentioned earlier, there are five language skills. To know the ability of the students in using these skills, different tests are used. Thus, there are five kinds of tests for examining the students' skills. These tests are oral tests for examining the students' ability in speaking, listening comprehension tests for listening comprehension, reading tests, both reading aloud and reading comprehension, for reading skill, composition tests for writing skill and translation tests for examining translation skill.

In the past, when the grammar translation was largely used, testing was mainly concerned with language elements. Today, however, more and more tests are concerned with language skills. These tests are well developed and they are very popular among teachers. A further description concerning these tests will be given in the next section of this chapter, under the heading types of tests.

II. 3. Types of Tests.

In preparing the test, the teacher must already have in mind what his purpose is in giving the test to the students. As mentioned earlier, there are various aims in having the tests either from the teacher's point of view or from the students' point of view. Based on these aims, one kind of test can be drawn. Different tests are used for different aims, but different aims are not always a

guarantee of making a difference in the forms of tests.

Beside the type of tests based on the aims of the testing, there is also a classification of tests based on who makes the tests. A test is usually made by the class teacher. Some of the tests are, however, composed by experts. The difference between these two types of tests will be explained in this part.

A test can be given orally or in written form. It can even be in the form of an action or a demonstration. This difference will be described as types of tests based on how they are carried out. Of course, each test, either oral or written, has its own advantages and disadvantages. In preparing the test, it is the situation that mostly determines which one of these two is going to be used.

A test must be scored. The scoring that is done using the judgement of the scorer is called an essay test or subjective test. On the other hand, a test that does not require any judgement on the part of the examiner is called an objective test. Thus, based on the scoring, a test can be differentiated as either subjective or objective. These two tests will be discussed in the fourth section of this part.

The last discussion of types of tests will deal with various tests based on the emphasis of the language skills and elements being tested. Different variables, either the skills or the elements, require different tests although it may happen that some of these tests overlap. Unless the learners are studying a language in an intensive course of language such as in the English Department, these skills and elements will not be separated too distinctly. In high schools, for example, the test is usually referred to as a language test.

Tests of skills are quite different from ones for measuring the elements. Tests on skills usually also require the students' knowledge of the elements of the language. Thus, an oral test which measures the speaking skill of the examinees will also involve phonology since whenever they speak, they have to know the basic or fundamental sounds of the language tested. This test will also involve vocabulary indirectly. This means that whenever they speak, they have to use the words or morphemes in the proper way. Hence, they should master the vocabulary or the lexicon. A test of the elements of the language, on the other hand, will only require little skill on the part of the examinees.

II. 3. 1. Types of Tests Based on The Aims.

As mentioned earlier, there are various aims or functions of having a test from the point of view of either the teacher or the learners. Based on these different aims, tests can be divided into four. This classification is by no means the only one since some textbook writers classify the tests into only three groups although some others group them into five. The four tests in this classification will be described in detail one at a time.

II. 3. 1. 1. Achievement Tests.

As indicated by the term, this type of test is mainly concerned with the achievement of the students in a course. These tests determine how much of the material in a course has been mastered. Some writers refer to these tests as attainment tests.

"Attainment tests estimate the present achievement of the learner, what he knows and can do now irrespective

of how he learnt it or what his learning ability is."⁴

There are two kinds of achievement tests; achievement tests as a measurement of the students' daily progress which are known as progress tests and achievement tests which serve to measure what the student is presumed to have learnt. For convenience, these two tests will be differentiated as progress tests and achievement tests. These two tests will be discussed one by one.

Progress tests are designed to measure the extent to which the students have mastered the material which is given in the classroom. These tests are usually constructed by the class teacher. Each progress test is different from the others since in each test the teacher will evaluate its situation which is usually different from other situations.

Moreover, each progress test is constructed based on the language programme which the class has been following. In this way, each test will be unique since the class will always make progress and hence, the next test will be different from the previous one. This kind of test is much influenced by the syllabus and the teacher's particular aims or goals.

A progress test may also serve as feedback either for the teacher or the students. Results obtained from the progress test enable the teacher to have feedback on his teaching. Whether his teaching is successful or not can be seen from the results of the test. If the majority of the students cannot do the test, it is likely that they have not understood the material being tested. This kind of test also enables the teacher to become more familiar with the work of each of his students and hence to know the real capacity of his students. The progress of the class in general can also be deduced from such a test.

4. Alan Davies. English Testing Symposium. Oxford, p. 71.

Although the progress test is mainly concerned with measuring what the student has mastered, it can be used to evaluate the effectiveness of the instruction as well. This means that a progress test can function as a teaching device. A good progress test enables the learners to have a positive manner over the target language and to gain additional confidence. In short, it can be said that the aim of the progress test is to stimulate learning and to reinforce what has been learnt.

On the other hand, achievement tests are made to measure achievement on a larger scale and to show mastery of a particular syllabus. Public exams which used to be given to the high school students in the past are examples of achievement or attainment tests in this sense. These tests are based on what the student is presumed to have learnt.

Since such tests are given to the schools which have the same syllabus but different materials and techniques to achieve them, they can be used to evaluate which materials and techniques are best used in those schools. For example, the final exams of some junior high schools are the same although different teachers are used. From the results of the exams, the best school can be determined. This school can be made into a model for the other schools. Thus, the materials used in that school will be used in other schools in the coming year. In this way, the whole instruction will become more effective.

An achievement test which is used for various schools is usually constructed by experts who rarely teach any of the students being tested.

In short, it can be said that achievement tests, either progress tests or the more general kind, are used to find



out how much of a course a learner has actually learned. The scores that the teacher gets from such tests are used to determine which students can be promoted to the higher class.

II. 3. 1. 2. Proficiency Tests.

Proficiency tests indicate what an individual is capable of doing now. This ability is achieved as a result of his cumulative learning experiences.⁵ A proficiency test looks forward which means that it does not measure what the learner has studied in a course but it defines a student's language proficiency with reference to a particular task which he will be expected to perform in the future.⁶

In connection with the aim of testing, proficiency tests are used to separate the learners who are capable of following an academic or training program from those who are not. In these tests, the degree of success or failure may not be considered important since such tests only measure the readiness of the students to follow instructional programs.

Proficiency tests are also used to put the learners into categories according to their knowledge of the foreign or second language they have learnt. Every individual will be placed or classified in an appropriate class so that there are no pass-fail distinctions in these tests.

The proficiency test may be used for the purpose of finding the level at which a learner should start or continue his study of the language.

Those are the functions of proficiency tests in connection with the aims of language testing. The level of proficiency for all the language skills may not be the same. A reading proficiency test may place a learner at one level ,

5. David P. Harris. Testing English as a Second Language. p.3.

6. J.B. Heaton. Writing English Language Tests. Longman, p.164.

while a speech proficiency test may classify him at another.

There are many examples of good proficiency tests. Among those tests, three tests will be mentioned including the purpose for each test. The first test is the Michigan Proficiency Test of English as a Foreign Language.⁷ This test is used to measure whether the individual's mastery of English is sufficient to study in universities such as the University of Michigan. If it is proved that he is still insufficient in his mastery of English, the score of his test will show how much teaching he should receive before he enters a certain university.

The second example of such a test is the Test of English as a Foreign Language (TOEFL) which is published by the Educational Testing Service. This test is probably the most widespread proficiency test used in Indonesia. For certain purposes such as studying in North America or trying to get a scholarship for continuing one's study in North America, this test is usually used. In other words, this test measures how much English the candidate actually knows compared to what he needs to know for a certain purpose. Thus, if he wants to get a scholarship from a certain institution, a higher score may be required than for another institution. Also, a certain university may accept a candidate only if he is able to get a certain score. This passing score may be higher than the scores determined by other universities.

The last example of proficiency tests is the Davies Test which is administered by the British Council. This test is designed to assess whether a candidate's command of English is adequate for higher education against a British background.⁸

7. Alan Davies. Language Testing Symposium. Oxford, p. 71.

8. Ibid. p. 72.

Those are the examples of proficiency tests. There are, however, several tests which are partly achievement and partly proficiency tests such as the tests of modern foreign languages which are constructed and run by the College Entrance Education Board (CEEBO) of the USA. These tests are based on what is common to a number of courses in some schools. This is the achievement part of the tests. However, since these tests are used as entrance qualification for high school graduates by certain colleges, they can be considered to be proficiency tests.

To differentiate between achievement and proficiency tests, the preliminary work undertaken by the test constructor must be observed. If the test is constructed based on a certain syllabus, it is usually an achievement test. If not, it is likely to be a proficiency test.

II. 3. 1. 3. Aptitude Tests.

Aptitude tests are referred to as prognostic tests by some writers. In this discussion, only the term aptitude is used.

This kind of test is used to predict the future performance of the learners. At the time of the testing, the examinee may have little or even no knowledge of the language to be studied, and hence the test is employed to assess their potentials. By giving this test, the examiner may select the learners who are likely to profit from language instruction.

Aptitude tests assess ability to learn a particular subject such as a foreign language. Some specialists, however, suggest that it is neither possible nor desirable to take an overall measurement of language aptitude. Thus, aptitude

tests assess the ability to learn the specific tasks in which the person is being trained such as listening comprehension , translating and reading comprehension. In the first assumption, the whole subject is considered particular and specific, while in the second assumption a subject is still a general ability. Therefore, this general ability should be divided into several aspects.

Learning involves other factors besides intelligence . Therefore, language learning aptitude, and consequently language aptitude testing, deals with such factors as intelligence, age, motivation, memory span, and sensitivity to grammatical patterning. Thus, based on this opinion, different individuals will have different aptitude tests or if the tests should be the same at all, different weighting must be given to those factors.

There are several methods of estimating individual language aptitude. The most common ones are the estimations which are based on the child's IQ or on his grades in his native language. Based on this kind of test, it is hoped that the pupils will be put into classes for which they are well suited. The better the test, the less likely it will be that a pupil will be put into a class in which he will not succeed or the more likely it will be that a pupil will not be excluded from a class in which he will succeed.

II. 3. 1. 4. Diagnostic Tests.

The purpose of diagnostic tests is to diagnose the weaknesses and strengths of an individual. By doing so, the teacher may know how to proceed in his teaching and what part of his teachings should be repeated.

Since this kind of test is used to determine the strengths and weaknesses of the students, it is very useful for the teacher. He may discover the learners' previous accomplishments and then, determine what remains to be taught. However, this does not mean that he should always teach new materials. The learners do not always show their capability over the materials that have been taught. Thus, from this test, it will be found out whether they have mastered the previous materials or not. If not, they can be taught the same materials again. In this case, it may be useful to use different techniques so that the learners will not be bored.

In connection with this purpose, other kinds of tests can be used. For example, the teacher may use the achievement test or the progress test to determine the students' or the learners' strengths and weaknesses and then, based on this test, he is able to continue his teaching. He may teach the new material or else he has to repeat his previous teaching if it is found that his students are still incapable of doing the test and hence they have not mastered the materials given. Thus, this kind of test may make use of other kinds of tests.

II. 3. 2. Types of Tests Based on Who Makes Them.

Based on who makes the tests, one may distinguish two different kinds of tests. The first kind of test is made by the class teacher, while the second is that which is made by an expert or a group of experts.

Tests that are made by the class teacher are usually called classroom tests. There are several characteris-

tics of this kind of test. The obvious one is that this test is made by the class teacher and is used for his own class only. Thus, the preparation, the administration as well as the scoring only involve one person, viz. the teacher. Since this test is intended for classroom use only, it has a rather informal characteristic. This means that the test objectives should be in accordance with the course objectives. Thus, this test is used for measuring the students' ability in class in connection with the objectives of the instruction.

Moreover, since a classroom test has the aim of measuring the daily progress of the students, it is very important to the teacher. The content of this kind of test should be derived from a specific course content which has been given to the students. Thus, each test is unique which means that one test is different from every other one.

The classroom test is used in class only. This means that the number of students involved is relatively small and hence, the teacher may use an essay test as well as an objective one. The main problem with this kind of test is the scoring. Since the teacher is in contact with the students everyday, he cannot avoid scoring less objectively than if he had no contact with them. The students' interpretations of the questions and their answers are less fixed than the ones made by a group of experts. To overcome this difficulty, the teacher concerned may set up a certain standard which is reasonably consistent. This means that the scoring does not fluctuate too much from the standard made beforehand. Also, to make the scoring more objective, it is better not to rely on one test only. Thus,

several tests are used and from these given tests, the teacher may draw a conclusion concerning the general ability of his students.

The second kind of test under this classification is known as a standardized test. This test is made by a group of people working for the professional testing services. In this group, there may be test specialists, linguists as well as psychologists. The test objectives need not be the same as the course objectives. Thus, students with different educational backgrounds can have the same test objectives. The test objectives themselves may range from selection and placement to evaluation.

In this testing situation, the tester has no personal or direct contact with the testees. Thus, the scoring is more objective since the tester is not much influenced by the personality of the testee. Moreover, the questions in this test require fixed answers and the interpretation of the test scores is also fixed so that the scoring is more objective than the classroom tests. Since this test involves a large number of students, objective tests are used. An essay test cannot be used because the scoring would be troublesome. Those who give this test need not be the same as those who compose and score it. Thus, if this test has an essay form, it is impossible to be objective since each tester will interpret the students' answers in his own way.

From the teacher's view, it is quite obvious that the classroom test is somewhat more important than the standardized one since the classroom test has something to do with the class progress. Therefore, the tests referred to in this thesis are classroom tests.

II. 3. 3. Types of Tests Based on How They Are Administered.

Based on this criterion, tests can be divided into several groups. However, from these various kinds, two major kinds are used here. They are verbal-verbal tests and verbal-non verbal tests.

II. 3. 3. 1. Verbal-Verbal Tests.

There are four varieties in this classification. They are given directly to the students either in the oral or in the written forms. The first variety is what is usually called a written test. In this variety, both the questions and the answers are written. The written questions can be given either on question papers or on the blackboard. On the other hand, the students may write their answers either on blank sheets of paper or on the answer sheet specially prepared for a particular test. The choice of the questions as well as the answers depends largely on the facilities that the school has. For example, in the school with no printing equipment, it is better to write the questions on the blackboard and the students must write their answers on their own paper. In this case, an essay test is preferable since the teacher only needs a short period of time if he uses this kind of test.

The second variety in this classification is what is usually called an oral test. In this test, both the questions and the answers are done orally. The teacher asks his students one or two at a time. Thus, this kind of test is time-consuming.

The third variety as well as the fourth one are a mixture of the first two tests above. In the third variety, the

teacher asks his students orally while his students write their answers on a piece of paper. This kind of test is called a semi-written test. The main point in this test is to get the students to write the target language as much as possible. On the other hand, the fourth or the last type of test under this classification is usually called a semi-oral test. The questions are written while the answers are given orally. Instead of getting the students to write the target language, this test is aimed at getting them to speak it as much as possible. This kind of test has a clear weakness. The student hardly pays attention to the answers which other students are giving. He is only concerned with the questions which are going to be asked. Thus, the first student gets less time to think than the later ones, since the later students already know the questions while their friend is giving his answers. Therefore, an oral test is usually preferable to a semi-oral test.

II. 3. 3. 2. Verbal-Nonverbal Tests.

In this classification, one of the activities in the test, either the questions or the answers, is in the form of pictures, actions, or activities. There are also four varieties under this classification. The first variety is when the teacher or the examiner uses one or more pictures to ask questions while the students give their answers either orally or in a written form. This type of test may make use the same or similar techniques as the substitution drills. For example, in the test the teacher shows the students a picture of a boy who is making a kite; if the test

is oral, the students have to say "He is making a kite.", if the test is written, then the students have to write the answer.

The second variety is the test which requires the students to draw the pictures to the questions which are written or given orally. This kind of test is rarely used since not all students can draw pictures.

The third variety is the test with the written or oral questions while the answers are in the form of activities. In this test, the teacher asks his students to perform something as is used in the oral-aural approach. For example, the teacher says: "Please close the door.", then the student who is asked should walk to the door and close it.

The last variety of the test in this classification is the one with the written or oral answers while the questions are in the form of activities. This is just the reverse of the third kind above. If in the third variety, the main aim is to test the students' understanding of the questions, the last one aims at the students' speaking skill. For example, the teacher opens a box, then one of his students should say: "You are opening the box," or in a better way, "You're opening a box."

Those are the varieties found in the two major divisions based on the way the test is administered. The choice for those varieties, whether the teacher wants to use both the questions and the answers orally or written or he prefers to use the written or oral questions with answers in the form of activities depends largely on what he aims at in the test. Suppose, he wants to measure the students' skill in speaking, then he has to choose the kind of test which makes the students speak in giving their answers. He may use either oral

or written questions or activities to elicit responses. However, the best of these varieties is the one which uses oral questions although with this kind of test, it cannot be avoided that the students' comprehension of the questions will interfere their answers. If the teacher uses activities in his questions, the answers are not so fixed. For example, if he opens a box, the student's answer may vary such as, "He opens a box." or "He is opening a box." Also, the questions will be limited since not all questions can be performed.

II. 3. 4. Types of Tests Based on The Scoring.

There are three processes in all tests, the preparation, the administration and the scoring. Based on the scoring, the tests can be differentiated according to whether they are objective tests or subjective ones. The subjective tests are also called essay tests. The characteristic that mostly differentiates these two kinds of tests is that in essay tests, the judgement of the examiner in scoring is required, while in the objective ones, it is not required. Actually, there is one more characteristic that differentiates these two types of tests. In essay tests, the questions require long answers from the examinees, while in the objective ones, the responses are relatively short. These two types of tests will be discussed one at a time although the essay tests will just be looked at in a glance since they will be discussed in detail in the next two chapters.

II. 3. 4. 1. Objective Tests.

An objective test is an assessment that needs relatively easy scoring, because the questions in this test only

require short, usually one-word, answers. This scoring can even be done mechanically. The questions in this test are commonly called the items while the responses are used to indicate the answers.

Within the objective tests themselves, there can be various types. They can be in the form of multiple choice, matching, true-false, and completion. The cloze test can be included as a kind of completion test. This test, the completion one, is on the border between an objective and an essay test. This will be discussed in detail under the sub-heading, Completion Tests.

Since an objective test is concerned with the procedures of testing or to be more exact, the procedures of scoring, it can be used for measuring all language variables.

There are advantages as well as disadvantages of using this type of test. The advantages of having an objective test if compared to an essay test are especially with regard to scoring. In the students' answers, only one answer is made possible for each question. Thus, there will be no ambiguity as to the right answers for the questions given. This in turn makes the scoring more objective. Moreover, since the answers are very short, the teacher only needs a short time in scoring the students' work. Also, since an objective test gives a large number of questions, the coverage of the materials is adequate. The last advantage that can be mentioned is that with this type of test, the students need not demonstrate their ability in using the language. Thus, the test is more valid since the students only answer what is being asked without interference of the medium.

Those are the advantages of having objective tests. On

the other hand, there are some disadvantages of having this kind of test as well. The most prominent weakness of objective tests as a means of measuring the students' ability and mastery over the target language is that they may encourage the students to guess wildly. Thus, without real understanding of the problems asked, the students may give correct answers. If this happens, the teacher may have a wrong impression of the students' real capability. An experienced teacher may reduce this wild guessing using certain techniques such as a correction technique. This means that the number of correct answers is reduced by the number of wrong ones.

Since objective tests only require short or even one-word answers, they are not good to be used for measuring productive skills such as speaking and writing skills. Moreover, since the number of questions in this type of test is large, it requires a lot of preparation on the part of the teacher. An objective test is a good one if it consists of at least forty questions so that it can cover all the materials tested. A school with no printing equipment will not use this test for measuring the daily progress of the students because it is almost impossible to write the questions on the blackboard.

Those are the advantages and disadvantages of objective tests. As mentioned before, there are various kinds of objective tests. These tests will be discussed one by one.

II. 3. 4. 1. 1. Multiple Choice.

As the term indicates, this kind of test offers several alternatives for each item. The task of the examinees is quite easy; they must choose the most suitable alternative

from among the given alternatives. In preparing such tests, it should be remembered that only one possible correct alternative is provided for each item. The other alternatives serve as distractors. Each distractor, or incorrect option should be reasonably attractive and plausible so that the students who do not really understand the problems tested may possibly choose these alternatives as well. Thus, in choosing the correct alternative, the students do not begin with eliminating the ones which are very obvious as incorrect alternatives. On the contrary, they have to choose the correct one by direct selection.⁹ An example of this will be given. Suppose there is an item like this :

The present tax reforms have benefited poor.

A. that B. the C. a D. an

If this item is given to the students, they will immediately exclude D since this alternative is grammatically incorrect and impossible. Thus, for most purposes, each distractor should be grammatically correct when it stands by itself.

The alternatives themselves may be in the form of a single word, short sentences, sets of pictures, or punctuation marks. The choice for these forms depends on the purpose of the test. Suppose multiple choice is used as a means to test vocabulary. The examiner may use either definitions or sets of pictures for each item. Let's take this item as an example : "They came here by plane."

The plane in the sentence above refers to: A. a means of transportation via sea. B. a means of transportation via air. or C. a means of transportation on land.

9. J.B. Heaton. Writing English Language Tests. p. 19.

However, for this item, the examiner may draw a set of pictures. A. a picture of a ship. B. a picture of a bus. C. a picture of a train. D. a picture of an aeroplane. The students' task is easy. They just choose which picture is a plane.

In recent years, there has been a tendency to use objective tests with more than one correct alternative. An example of this is an entry test used by some public universities. Thus, the instructions will be like this :¹⁰

For items number 1 to 20, choose :

- (A) if (1), (2), and (3) are correct
- (B) if (1), and (3) are correct
- (C) if (2) and (4) are correct
- (D) if only (4) is correct
- (E) if all answers are correct

1. "He was told to return John's camera on Monday."

This means :

- (1) "John could borrow his camera until Monday."
- (2) "John allowed him to borrow the camera until Monday."
- (3) "John should return his camera on Monday."
- (4) "He had to return John's camera on Monday."

From the alternatives above, the candidates have to choose (C) because numbers (2) and (4) are correct. Another item is this :

2. "John would not have been punished if he had admitted his mistake."

- (1) "John admitted his mistake and was not punished."
- (2) "John didn't admit his mistake so he was punished."

10. Questions from the entry test in "Perintis I" 1980.

- (3) "John was punished because he had admitted his mistake."
- (4) "John was punished because he didn't admit his mistake."

Again, for this item the correct answer is (C) since the correct alternatives are (2) and (4). The last example is :

- 3. (1) Reading a newspaper, the sad news startled me.
- (2) Reading a newspaper, the sad news was startled by me.
- (3) Reading a newspaper, I startled the sad news.
- (4) Reading the newspaper, I was startled by the sad news.

For this item, the correct answer is (D) since only alternative (4) is correct.

That is only one variety of multiple choice test. It can reduce guessing. On the other hand, this test is more difficult to compose. Thus, for the hard work done by the tester, there is a reward in the results of the test. The test will reflect the capacity of the students in a better way since guessing can be reduced to a minimum.

There is still another variety of multiple choice as a means for measuring the capacity of the students. This is what is usually called completion multiple choice. The items in this kind of multiple choice are very similar to the items in the completion test. The difference lies on how to answer the items asked. If in the completion test the students have to answer them using their own words, in the completion multiple choice one, they only choose the most suitable alternative from the given alternatives. Examples of completion multiple choice items are :

- I am sorry I cannot go. You know, I if I could.
 - A. had gone C. would have gone
 - B. would go D. went

- "Is John coming to the party tonight ?"

"Yes, he asked me"

 - A. if he could go with us
 - B. can he go with us
 - C. he went with us
 - D. going with us
 - E. whether he goes with us

Those are the varieties found within multiple choice tests. This type of test is suitable for measuring factual knowledge, vocabulary, ability to apply principles, and ability to interpret and to compare two things. Thus, it is no wonder if this type of test is widely used. This test, however, is difficult to compose. It is difficult to find alternatives which are equally attractive, and plausible, and are of the same length. For each item, the examiner should have at least three alternatives; and a test consisting of multiple choice items only is said to be reliable if it has fifty items. To help the examiner compose a good multiple choice test, several hints below can be followed.

- a. The materials or problems to be tested should be clear so that the examiner will not deviate from them.
- b. The examiner makes a blue print. A blue print is a scheme in which the examiner considers how many items will be asked for each problem and whether these items have something to do with memory, knowledge, understanding or application. After deciding this, he may begin to compose the items. The time needed

to do this test is stated here. This provided time should be enough for the students to complete the test on time. At least, 80% of the students should be able to do the test calmly.

- c. In composing the items, the principles of grading should be stressed. This means that the items should be in order from the previous to the later lessons and from the easiest to the more difficult ones.
- d. If the test consists of multiple choice only, it is better to prepare an answer sheet so that it can be used on another occasion.
- e. There must be clear instructions, either orally or written so that the students will not get confused.

To conclude the discussion on multiple choice tests, it can be said that although they are of great usefulness they are difficult to compose. Concerning language testing, this type of test is very effective for measuring skill in grammar, vocabulary, reading comprehension, and listening skill. It can also be used for testing writing skill and pronunciation although other types of tests are preferable for testing these two skills. A multiple choice test is an unlikely choice for measuring speaking skill. To test this skill, another type of test, viz. the oral test should be used.

II. 3. 4. 1. 2. Matching.

As indicated by the term, the learners' task is to match two things so that the whole thing can make sense. The two parts in the matching item are called the premise and the response. Thus, if the examiner intends to use this

type of test he has to provide a set of premises and a set of responses and the examinees' task is to match these premises to the responses.

A matching test is suitable for measuring the factual knowledge of the students but it is not good for measuring the understanding that needs application. In terms of language testing, this type of test is best used for measuring vocabulary and reading skill since these two variables only need a little application. As a means for measuring other elements and skills of a language, a matching test is not so effective.

There are two types of matching tests as far as the number of premises and responses is concerned. The first type is the even or perfect test; the second is the uneven or imperfect one. In the even matching test, the number of premises is the same as the number of responses. On the other hand, in the uneven matching type, the number of premises is unbalanced with the number of responses. With regard to language testing, it is suggested to use the uneven one because this type reduces the chance of guessing on the part of the students. Also, with this type, the examinees feel challenged to do the test.

The difficulty faced by the examiner in composing the matching test lies in how to present the premises and the responses that are of the same category. If one of the premises is a noun or a name of an author, for example, then the other ones should also be nouns or names of authors. Then, in the responses, a list of books are given. Thus, the form is like this :

- | | |
|-------------------------|------------------------------------------------|
| 1. Shakespeare | A. At Castle Boterel |
| 2. John Donne | B. Quiet Work |
| 3. William Butler Yeats | C. Let Me Not to The
Marriage of True Minds |

4. etc.

D. The Canonization

E. etc.

Both the premises and the responses can be arranged in alphabetical or chronological orders. The instructions have to be simple but clear. It should be made clear whether one response can be used more than once or not.

Due to the difficulty in composing this type of test, it is rarely used in itself. It is usually accompanied by other types of objective tests. Also, due to the limitation of the elements and skills that can be tested, this type of test is rarely used. Thus, this test is good if it is only a part of an objective test.

II. 3. 4. 1. 3. Completion.

This type of test is the most difficult one among the other types of objective tests. The items in this test offers the least possibility of wild guessing. Thus, the examinees who do not master the things being tested usually fail to produce the correct answer, since in this type of test they have to give their answers using their store of knowledge.

The items in the completion test should be well prepared so that only one correct answer is possible for each question or blank. If not, this test would tend to be subjective test since more than one answer is possible and the scoring depends on the examiner. For this reason, this type of test is often considered as the test which is on the border between an objective and a subjective test.

Among other objective tests, this type is the best one

for measuring the productive skills of a language. This type of test is quite satisfactory for measuring vocabulary, spelling, grammatical structure and writing skill. Thus, although this type of test can be used for measuring all elements and skills of a language, it is advisable to use it only for the elements and skills mentioned above.

In composing a completion test, it should be remembered that each item should only require a restricted short answer. Also, the instructions should be clear. If possible, the dotted lines to be filled in are not at the beginning of the items because such a position will confuse the students.

In scoring, this type is the most time-consuming among the objective tests. For this reason, and because of the difficulty in composing it, this type is very rarely used alone. It can be used for supplementing other types of objective tests.

There are several varieties in presenting a completion test. It can be accompanied by pictures which serve as stimuli for the examinees. An example of this is :

- A large animal with a trunk, and ivory tusks is an ...



- My often barks at night.



- A kind of fruit from which people make wine is



A cloze test can also be considered as a variety of completion test. In its simple way, a cloze test is described as a technique or method of deleting every n^{th} word, which should be less than the tenth word and more than every fifth word. This deletion is done systematically and mechanically. If it

is a test of reading comprehension, the examiner may delete every fifth word in the passage and the examinees' task is to supply the missing parts, and to make the passage whole again by using the clues provided. The deletion can be either every -nth word or a certain class of words. Since the examinees' task in the cloze test is the same as in the completion test, it can be said that a cloze test is a variety of completion test. Only in recent years has the cloze test been sufficiently developed.

II. 3. 4. 1. 4. True - False.

This type of test is the easiest to compose among the objective tests. However, it is also the type that mostly encourages the examinees to guess wildly. The reason for this is that there are only two alternatives in each item. The examinees' task is just to choose between true and false for each item. Thus, without knowing the problems tested, they could make a guess easily with a 50% chance of being right. To reduce such a bad effect, this type is usually combined with other kinds of objective tests.

In composing this test, it should be remembered that the items tested should be perfectly right or wrong so that there will not be any confusion. These items are best arranged in the same order as they were taught. The right and the wrong items should be balanced, with the tendency for the right items to be more in number than the wrong ones. If a test consists of true-false items only, it should have at least 80 items or else it cannot be considered a reliable test.

With regard to testing a language, this type of test

at its best may be used for measuring recognition of the language elements and skills. If this type of test is used by itself at all, the selection is made for practical reasons such as the fact that this test is easy to score, easy to administer, and is economical as well. However, if the validity and the reliability have to be considered, this test should not be used by itself. It must be accompanied by other types of objective tests.

Those are the types of objective tests that can be used for language testing. Except for multiple choice, other types are rarely used by themselves. They are often used in combination. Thus, a test may consist of 20 items of the true-false type, 20 of the matching type, 10 items of multiple choice and 10 items of the completion type.

II. 3. 4. 2. Essay Test.

As mentioned earlier, this kind of test will be discussed in detail in the next two chapters since it is the main topic for this thesis. Thus, in this section, this test will be viewed at a glance.

An essay test is usually defined as a test in which the examinees are required to answer the questions by giving explanations and descriptions. From this definition, it is clear that this type of test is very useful in language testing. It can be used for measuring the productive skills of a language. However, that is not the only use of an essay test. This test may be used for measuring all the elements and skills of a language except pronunciation. Actually, this skill, pronunciation, can also be tested using this type although it may be less effective

than if other types of tests such as objective tests are used. It is for this reason, the great use of an essay test, that this thesis is made.

II. 3. 5. Types of Tests Based on The Language Variables Tested.

As have been discussed in the previous section, the language variables consist of two parts, the elements and the skills. Different elements result in different types of tests; and so do different language skills. Thus, there are pronunciation tests for measuring phonology, grammar tests for syntax, vocabulary tests for measuring the students' vocabulary, as well as different tests of skills such as oral and composition tests for measuring the language skills.

Although there are various types of tests based on the language variables as mentioned above, in practice, especially for beginners, there is usually only one test, a language test. Thus, unless the course has special purposes such as a conversation class, the test used is usually a mixture of those variables. Moreover, unless a learner is in the intermediate or advanced class, the test used is just a language test. Nowadays, more and more people learn a foreign language either for special purposes or for general purposes so that it is worthwhile to know in detail the types of tests mentioned above.

II. 3. 5. 1. Pronunciation Tests.

There are two opinions from the teacher's point of view concerning pronunciation tests. The first opinion is that there is no need to test pronunciation at all. The reason for this is either that the teacher does not under -

stand the role of pronunciation in language or because it is difficult to find an efficient and simple pronunciation test. In the second view, the teacher realizes the important role of pronunciation and thus, believes that it must be tested. The problem faced here is what the standard should be. Some teachers test pronunciation requiring the accuracy of a native accent; others take the phonemic criterion. The second view is preferable since a pronunciation test with a native accent criterion is extremely difficult to give and to score objectively. Thus, in this section only the second view will be discussed.

The thing to be tested in the pronunciation test is the phonological system. This system involves the phonemic system of the language with its chief phonetic features and distribution of phonemes.¹¹ This system is operated through firmly set habits and transfer to the foreign language when the student is learning it.

The test itself may make use of other types of tests, such as objective tests or to be more specific, multiple choice tests. It may be used for measuring sound contrasts. For this purpose, multiple choice may be presented in the following forms; sound to digits, sound to letters, sound to phonemic symbols, minimal pairs, triplets, quadruplets, comparing sound to a model, and sound to pictures.¹²

In the form of sound to digits, the examiner reads sets of words such as :

1. a. sheep ; b. sheep ; c. ship ; d. sheep.
2. a. bad ; b. bet ; c. bad ; d. bad.
3. a. foot ; b. food ; c. food ; d. food.
4. a. she ; b. she ; c. she ; d. see.

11. Robert Lado PhD. Language Testing. Longman, London, 1965, p. 42.

12. Ibid. pp. 46 - 77.



In reading the sets of words above, it should be remembered that there must be a pause for a few seconds after each word so that the examinees may recognize it. The task of the examinees in this test is to identify which of the words read by the examiner is different from the others in each set. Thus, the answers for the sets above are c, b, a, d.

In the sound to letters form, several sets of words containing the troublesome sounds are given such as :

1. a. shall ; b. shell ; c. sell ; d. cell.
2. a. foot ; b. food ; c. put ; d. flute.
3. a. sink ; b. thing ; c. sing ; d. think.

In this test, the examiner just reads the alternative containing the troublesome sound and the examinees have to choose which of the words is read by him. Suppose, for number one the examiner reads the word which contains the sound / æ / , then the examinees have to cross or to write the letter a on their answer sheet.

If the sound to phonemic symbols form is used as a test, the examinees have to be made familiar with the phonemic symbols first. This form of test is only suitable for learners who are studying the language intensively or for special purposes. It is not suitable for high school students in Indonesia. Thus, unless phonemic symbols have been taught, it is not proper to give this kind of test to the students.

Pronunciation tests which use minimal pairs, triplets , or quadruplets are all basically the same. If minimal pairs are used, there are only two alternatives given. The minimal pair is a pair of words in which the difference lies only in one sound. The sounds in the two words given are phonemically different. Thus, in the test, the examiner may read several

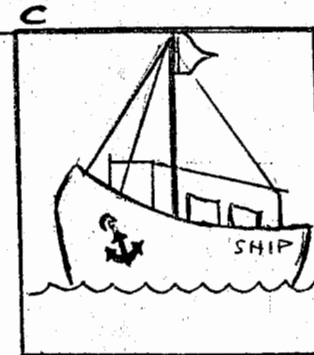
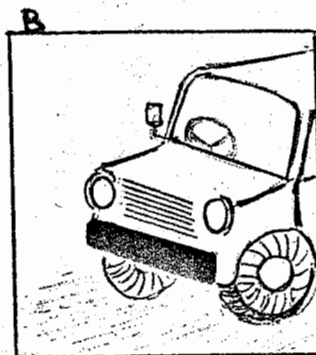
sets of words in which each set consists of two words only. The sets of words read consist of the same or different sounds interchangeably. For example,

- | | |
|-------------------|-----------------------|
| 1. sleep - sleep. | 4. rid - read. |
| 2. grape - grip. | 5. thirsty - thirsty. |
| 3. beat - bit. | 6. etc. |

The triplets and quadruplets have three and four alternatives respectively. From these alternatives, the examinees have to choose the one that is different from the rest.

In the comparing sound to a model form, the examiner gives a model, followed by saying several words containing different sounds. The task of the examinees is to identify the alternative which is in accordance with the model. With this kind of test, the examinees tend to complain of the heavy burden of remembering the model, so that it is seldom used.

The last technique to be discussed for pronunciation tests is the sound to picture form. This type has limited usage; only the words containing the trouble sounds which can be drawn are the ones that can be used. In this type, some pictures are supplied and the examiner says a word or a sentence containing the word that has the troublesome sound. The examinees choose which picture corresponds to the word said by the examiner. This is the technique which is least used since it is limited and time-consuming. For example :



The pictures above are used to test the sound contrast between the vowels / i / and / i: /, and the sound contrast between the consonants / dʒ / and / ʃ /. Suppose the examiner says: "sheep" or "It's a sheep." If the examinees choose B, it means that they are troubled by the consonant contrast; if they choose C, it means that they are troubled by the vowel contrast.

These techniques of testing pronunciation are only valuable for recognition. If the examinees can do the test, it does not guarantee that they can pronounce the problematic or troublesome sounds correctly. To test their pronunciation, an oral test must be applied.

There is still another disadvantage to use such techniques. The examinees' achievement is greatly influenced by their ability in hearing. They may know the problem sounds tested and yet, they may not be able to answer the questions since they mishear what the examiner says. This kind of test is in part the same as the dictation test since in this test, the examiner reads several words or sentences and the examinees' task is to write the right sound. This is specially true for the sound to phonemic symbols form.

Besides multiple choice techniques and oral tests, there are still various ways of measuring pronunciation. Some of them are by using translation, reading, and completion items. However, these techniques will not be discussed here since by using them, other skills are involved. Thus, if translation is used, then the test will not be so valid for the skill of translation is involved here.

II. 3. 5. 2. Grammar Tests.

Grammar tests are commonly carried out using objective

tests, translation or they are done orally. Thus, there is no specific means for testing grammar. Essay tests were commonly used in past years and they are maintained for measuring grammar up to this time by some teachers.

Nowadays, the most commonly used method for measuring the students' ability in grammar is what is usually called communicative grammar tests. Using this method, the teacher trains his students to be able to use the language, not just know about the language. Thus, grammar is taught deductively, not separately. It is not taught in itself but as a part of communication which becomes the natural aim of learning a language. Language is for communication and hence, learning it, the students should be able to communicate with it. Thus, if a teacher follows this line, there will be no testing of grammar.

The techniques for testing communicative grammar itself range from pattern practice to situational grammar. Using these techniques, the test is usually done orally while the students' grammar is usually deduced from it. Actually, they are not so valid since in the test, the speaking skill of the students will influence their performance. Yet, they are better than other ways of testing grammar which only measure whether the students know about the grammar or not.

II. 3. 5. 3. Vocabulary Tests.

Vocabulary can be learned through reading materials or from a list of words. This list is usually selected either on the basis of frequency or coverage of a maximum range of situation; and although it is rarely done, it can be made on the basis of power to define other meanings.¹³

13. Ibid. pp. 181 - 203.

To test vocabulary, many types of tests can be used. They range from objective tests, essay tests or translation to using a reading passage. Using translation, vocabulary can be tested either using loose words, sentences or a reading passage. In using this technique, however, the examiner must be careful since translation is a separate skill and hence, it needs a special training. This will be discussed further in the section deals with translation tests.

There are two kinds of testing vocabulary, either for recognition or for production. Words for recognition are more difficult than for production since a learner usually hears or reads oftener than he speaks or writes the language. Moreover, there is a difference both in size and in the nature of knowledge between vocabulary for recognition and for reproduction. Knowing the words for speaking, however, means that a unit can be recalled and used correctly whenever the learners wants to.

A list of words, either lexical or grammatical units usually constitute vocabulary. This does not mean, however, that all lexical units are words. Some idiomatic expressions such as merry-go-round, call up, and Jack hit the sack have meanings in themselves and hence, they are lexical units. In testing vocabulary, therefore, this must be paid attention too.

Pictures can also be used for testing vocabulary. Using this technique, some pictures are distributed to the students and then, the teacher pronounces one word that indicates the chosen picture, for instance 'a cat' while the students give a mark to the picture of a cat.

Those are the types of tests that can be used for

testing the language elements. The most importance in giving such tests is that they should not be artificial. They should be put in context. The examples of such tests are :

- John needs a pencil.
He can use one
a. of me. c. mine.
b. my. d. of mine.
- Your ideas may be good, but you must them.
a. fulminate c. equivocate
b. discrete d. substantiate

The trouble sound to be tested is /s/ in the word thousand; the question is :

- Ten hundred is one

If the students can answer these questions, it shows that they are able to use them in communication. Thus, it is assumed that the ability to do the test will show the ability to communicate, using the words or patterns tested.

With regard to language skills, however, the problem is quite different. Each skill should be trained and learned and to know whether the learner has mastered it, a test should be given. Thus, there are tests for oral skill, listening skill, reading skill as well as writing and translation skills. These tests will be discussed in the following sections. Before that, however, it should be remembered that those skills are the integrated skills which mean that they involve all the language elements at once.

II. 3. 5. 4. Oral Tests.

This type of test can be used for measuring almost every

aspect of language learning. Thus, it can be used for testing grammar, pronunciation and vocabulary simultaneously although it is best used for measuring the speaking skill of the examinees. Since the nature of language is for communication, this type of test can be used as the final device to judge whether the examinee is qualified or not.

By having this type of test, the examinees get the chance to use the language tested as total experience since in this test, they will be tested on their ability to understand the spoken language at a normal speed and to express themselves orally. During the test, the situation should be made in such a way that the examinees will feel at ease in executing the test, because it may happen that some examinees who actually know the answers to the questions are incapable to express them orally.

Although an oral test is the most suitable form of language testing, it has not been widely carried out since it is difficult to score and is time-consuming. In scoring, the examiner may be influenced by many irrelevant aspects such as the examinee's appearance. Also, it is difficult to score many things at once. In giving this type of test, it should be decided beforehand what things to be scored, so that this test can be valid. Concerning the time, this test will be time-consuming either in preparing the questions or in the administration. The number of questions should be as much as possible and if possible, a set of questions are only given to two students or groups successively so that they will not get the information of the questions before the test.

II. 3. 5. 5. Listening or Auditory Comprehension Tests.

Communication system involves incoding and decoding. It means that in communication, information is sent and received by means of a code. Incoding deals with choosing the signals to express information, while decoding is getting the information from the encoded signals. This type of test is meant for measuring the examinees' ability to comprehend or understand the spoken utterances in the language being tested. Actually, this test is important and should be given before the examinees execute an oral test since in the oral test, the ability to understand the questions is very decisive.

This type of test is not meant for measuring a specific language element but for the integrated of the language elements. It does not mean, however, that things like technical term or technical vocabulary and technical matters of literary interpretation can be included in such a test. On the other hand, they should be excluded. This test should only deal with matters in which native speakers, without any specific or particular training, can understand them. Thus, this test is used as a means for revealing how well the examinees grasp the whole meaning of utterances which are not technical.

The utterances used as the stimuli for this test may be varied in the length or difficulty. They may be short, long, simple or complicated questions and statements, requests and imperatives. They may also be a brief dialogue or a short conversation. The choice for this depends on the capability of the examinees, whether they are beginners, or in the advanced level. An example of items asked in this test :

- There were 40 seats in the bus, but only half of them were taken. How many passengers were there in the bus?

- John dropped the letter in the mailbox.
 - a. John sent the letter.
 - b. John opened the letter.
 - c. John lost the letter.
 - d. John destroyed the letter.

- Please, open that window !

If the student comprehends this question, he will certainly go to the window and open it.

II. 3. 5. 6. Reading Tests.

There are two kinds of reading that can be learned and hence, tested. They are reading aloud and reading comprehension. Reading aloud should not be encouraged because by reading aloud, the reader will get no or less information than if he reads silently. A reading test in which the examinee should read the passage aloud can only be used for measuring the pronunciation of the examinee and there should be no questions following the passage.

Reading comprehension, on the other hand, is an integrated skill. It involves graphic symbols, arrangement of words, the use of punctuation, and the message or the content of the passage. However, the main aim in teaching reading comprehension is to get the learners as quickly as possible in reading and yet, still understand what they are reading. Thus, a reading comprehension test is given for measuring the language elements without having an emphasis on a special language element, and for measuring how they understand the relationship between the graphic symbols and the idea conveyed through them. Such a test may make use of essay or objective tests.

The passage used in the reading comprehension test should not be too long. Usually a passage of 200-300 words is sufficient because if it is longer, it will be difficult for the examinees to remember all of its content. For beginners, the passage can even be shorter and the words used are those found in the thousand word list.

To determine the time limit, a kind of pre-test can be given. This test is used to predict the students' achievement and the time needed for doing it so that in the test, sufficient time can be given. This means that almost all of the examinees can finish the test calmly.

Reading tests which use passages are called intensive reading tests. There is, however, still another kind of reading comprehension test which is called extensive reading test. In this test, instead of having a passage, the examinee reads or prepares his reading at home. The material is usually a short story or a book. This kind of test is usually used for advanced students only since it is very difficult for the beginners to read such a long story.

Those are the kinds of reading comprehension tests that can be given to the students. In the intensive reading comprehension test, the passage may contain scientific things. Also, the vocabulary found in the reading materials is of a far greater range than the vocabulary ordinarily used in speaking. The length of utterances that are found in reading, especially from books, is greater than the length of utterances heard in conversation.¹⁴ As a conclusion, it can be said that reading comprehension is an important skill that should be developed if the learner wants to master the language well.

¹⁴. Ibid. p. 224.

II. 3. 5. 7. Composition Tests.

Writing skill is usually considered the most difficult skill in language learning and this skill is only given after the learners have learned the language for some time. However, learning a language without having the ability to write is incomplete. Since this skill has to be learned, it must also be tested.

Writing a piece of composition needs talent as well as exercises. Since it is a very difficult skill to acquire, it is divided into three categories : guided composition in which the students are given frames to do the composition , partly guided or partly free composition in which the given guide is reduced, and completely free composition where the students are given freedom in expressing their ideas. Based on these categories, the composition tests are divided into three as well. Which one of these three is the best ?

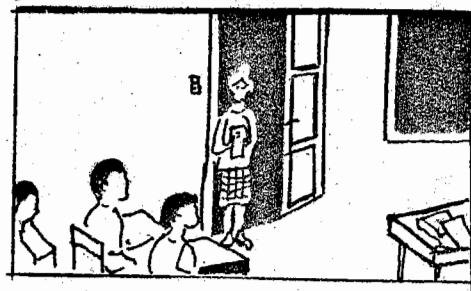
In terms of scoring, obviously the guided composition test is the best since in such a test, the score is only based on the examinee's ability to arrange the composition from the frame given. For the partly guided, and moreover for free composition, the scoring is difficult since in this kind of composition, the creativity, the completeness, the balance of the content and the relevancy are scored. Thus, the score includes many things. In practice, two scores can be given for each composition. The first score is for the content of that composition. This includes creativity, completeness, the balance and the relevancy. The other score is for the expression which includes grammatical mistakes, spelling, punctuation, paragraphing and the diction.

In giving a guided composition test, many techniques

such as cloze test, dictation, sentence completion and error recognition can be used. In the cloze test, for example, the examiner gives a passage of 100 words. The students are supposed to read. After that, another passage with the same topic but some words are left out is given. The students are required to fill in the blanks based on the previous passage. This is to test whether their ability in mastering the 100 word list is workable or not.

In the partly guided composition test, some pictures can be used as stimuli. For example, the examiner shows a set of pictures of a school situation and asks the examinees to write down what they can find out from those pictures. These pictures are given numbers so that the examinees are not confused in starting their work. The pictures are like this :

1



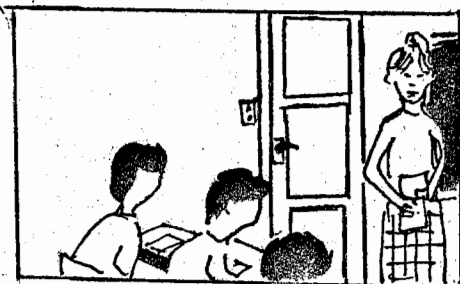
2



3



4



For free composition, the examiner just prepares some topics and the examinees may choose one of them and then write down their composition. The topics may vary from the general to the particular ones. Some of the general topics are :

- Write about your daily routine.
- Write about your school and your class.
- Write about your family, etc.

The examples of the particular topics are:

- What do you want to do after graduating from this school ?
- What would you do if you got lost in an unknown town ?

As a conclusion, it can be said that the free composition test is the aim of language teaching-learning although to come to it, some guided and partly guided composition tests should be given. Free composition tests are only suitable for advanced learners and those whose knowledge is adequate.

II. 3. 5. 8. Translation Tests.

The last type of test to be discussed in this section is concerned with skill, the skill of translating which is also an integrated skill. Since it is a separate skill, it needs a special training and hence, a special test. For example, a learner who possesses excellent speaking and writing abilities may not be able to translate well. However, the reverse is not true. A learner who can translate well must be able either to read or write well. If the translation is done orally, the examinee should be able to comprehend or speak well. If the translation is from the mother tongue into the target language the examinee has to be able to express his translation either orally or written. On the contrary, if it is from the target language into the mother tongue, the examinee should mainly be able to comprehend the translated passage either by reading or by listening.

In giving the translation test, it is better for the ex-

aminer to tell the examinees what kind of translation is expected. There are three kinds of translation based on the style. Literal translation or word-to-word translation which is only suitable for measuring vocabulary and not the skill of translation is the first kind. The second kind is formal translation. This is better for measuring the translation skill for it does not stick too much to the original language, including its word order and sentence patterns. However, the best test in term of its use for measuring the skill of translation is the psychological translation. In this type of translation, the psychological effect is the most important factor. Thus, the examinee may translate in his own way and it will be considered correct as far as the whole translation has the same or similar effect as the original one. Of course, the last type is the most difficult and it is usually given to the intermediate or advanced students only. From the teacher's point of view, this kind of translation is the most difficult to score.

If a translation test is used for measuring vocabulary, there are four possibilities. The translation of the two words may be identical, overlapping, included in the meaning of the translated word or no exact translation for a word in one language into another.

If the meaning of two words is overlapping it means that each word has a feature which the other does not. The example of such words is the English word 'sister' and the word 'adik' in Indonesian. The word 'sister' refers to the older or younger female sibling while the word 'adik' refers to the younger male or female sibling. In translation, it may happen that 'sister' is translated into 'adik' whereas the meaning of 'sister' is not always 'adik'.

The meaning of a word is included into the translated word if the translated word has a wider meaning. An example of this is the translation of the Indonesian word 'kucing' into the English 'cat'. Actually, 'a cat' is not referring to 'kucing' only but it also includes animals which have tiger-shape, such as leopard, panther and tiger. On the other hand, 'kucing' cannot refer to tiger.

Words which have no exact translation into another language are usually words which have specific cultural backgrounds. The example of such words is the Indonesian 'lauk'. This word cannot be translated into English since there is no such thing as 'lauk' in English. The word 'lauk' refers to the thing or things that go on or accompany the main food (rice) while in English, the main food is meat.

Those are the problems faced in translation. Beside that, a translation test is time-consuming either in the administration or in the scoring. Thus, although it is a good means for measuring the students' achievement, it must be used with care or else, it will become an ineffective test.

As a conclusion on the discussion of the types of tests based on language variables, it should be remembered that the language element tests should be contextual so that the learners do not only learn about the language but they learn how to use the language itself. Of course, tests on the integrated skills are better than on the elements.

CHAPTER III

ESSAY TESTS

In the previous chapter, the place of an essay test among various types of language tests has been revealed. In this chapter, the essay test will be observed in detail.

In the first section, the nature of essay tests will be observed. This will answer the questions : What is an essay test ? Why is an essay test sometimes considered as the same as a subjective test ? Is there any difference between an essay test and subjective tests ?

Every test must have its own advantages as well as its disadvantages. The second section in this chapter will deal with the advantages as well as the disadvantages of essay tests, either from the point of view of the teacher or of the examinees.

The last section in this chapter will be concerned with the stages in the essay test. An essay test may require answers which range from one-sentence-answers to a paragraph or an essay if it is concerned with composition, for example.

III. 1. The Nature of Essay Tests.

In the testing situation, there must be an examiner and one or more examinees. The examiner gives a number of questions; while the examinees answer those questions. The answer for each question can be just one word or it can be a sentence or even an essay. The test which requires the examinees to answer fixedly without giving freedom is an objective test; while the test in which the examinees are

able to answer in their own words in the form of a description, explanation, or the like is called an essay test.

An essay test is often also called a subjective test. This term is often used to contrast it with the objective test. These two terms refer to the processes of the test and the examination. There are three processes in the test: the preparation which involves the compilation or construction of the questions, the administration or the stage in which the examinee answers the questions, and the scoring or marking of the examinees' answers to these questions.¹⁵

The first process is certainly subjective in the sense that it is the examiner who chooses the questions. Each examiner may compile or construct different questions even if the materials and the type of test used are the same. It is difficult to include all the materials in the test so these materials must be selected. It is in this selection that different constructors will frame their questions differently, depending on their likings. Thus, one constructor may emphasize or give priorities to the first part of the materials, while another to the last part; still, other constructors will do that on the points which they like most.

The second process which refers to the activity in which the examinees answer the questions is even more subjective than the first. The examinees may answer the questions according to their likings. They may choose which of the questions they want to answer first and if it is an essay test, how much they will answer the questions. Moreover, if there is a reserved question which means that the students may choose to answer five out of six questions, they may choose differently. Thus, in answering the questions, they may base their answers on their

15. Alan Davies. Language Testing Symposium. Oxford, p. 20.

memory which is different from one to another and they may choose which question will be answered first. For example, in answering a multiple choice question, the examinee may choose the wrong alternative although a sufficient clue to the correct answer has been given. Therefore, this choice is subjective although the question itself is considered objective.

The third or the last process of testing is the scoring. It is in this process that a test can be called subjective or objective. The first two processes are of the same kind and no difference exists. In the scoring, if the examiner has to judge, if he has to decide whether the answer is adequate or not and if he has to choose between awarding the answer a high or low mark, then the scoring process is called subjective. In the objective process, the examiner's task in scoring can even be replaced by a machine.¹⁶

If an essay test is compared to a subjective test, the difference lies on the points in the process of testing. The 'essay' refers to the relatively long answer given by the student while the 'subjective' refers to the process of scoring. It can be seen from the discussion above that it is the long answers of an essay test that require judgement in scoring. Hence, it can be understood if these two terms are often used interchangeably.

An assessment of the students' achievement is sometimes differentiated into an examination and a test. To do this, the terms 'subjective' and 'objective' are often used. A subjective test is often used to indicate that the assessment is an examination. An objective test, on the other hand, is often

16. Ibid. p. 21.

referred to as a test. This difference is actually unnecessary if it is realized that either subjective or objective kind can be used both in the test or in the examination. This problem arises because in the examination or the final achievement, the subjective kind is often used.

If a test and an examination should be differentiated at all, the difference usually lies in the formality of the assessment, and not in the use of the types of tests. The daily progress assessment is usually called a test. This test can be in the form of an objective or subjective type. At the end of a course, the final achievement of the learners is measured. In this case, the assessment is often referred to as an examination. The situation in the daily tests is less formal than in the final examination. In the examination, the global knowledge of the examinees is measured. So, it includes all the materials taught during the course period, while in the daily progress test, the material is limited to, for example one chapter for each test. In conclusion, it can be said that the different terms 'test' and 'examination' are due to the formality of the assessment and they have no concern with the kind of test, whether it is of the subjective or objective type.

In summary, it can be said that an essay or subjective test is an assessment that requires the examinees to give relatively long answers to the questions asked and it requires the examiner's judgement in scoring the examinees' work.

III. 2. The Advantages And The Disadvantages of Essay Tests.

From the discussion of essay and objective tests, it

can be concluded that these two kinds of tests actually only serve as channels for measuring or testing every subject, including language. The essay type, thus, includes all kinds of tests in which the examinees are required to answer in their own words and the examiner uses judgement in his scoring.

The two types of tests above, the essay and objective, can be used for measuring all language variables although the degree of effectiveness is different. The essay type is at its best if it is used for measuring productive skills since such a test gives the examinees a chance to show their mastery of the skills as long as possible. Applying essay tests to the language variables will be discussed in the last part of this thesis.

As for the advantages of applying essay tests, the most obvious one is that it gives a chance to the examinees to describe their answers in their own words and styles. They are able to express what they have in mind. This is very important in language learning since after passing their course, they will be forced to express their ideas and not just say 'yes' or 'no' to the questions asked to them.

The second advantage of an essay test is that it can be used for measuring the integrated skills which means that the achievement of the examinee is observed. Thus, it can be revealed which elements or parts are still unclear to the students so that they can be given more emphasis on the next occasion. After all, it is the improvement of teaching as well as learning that will be achieved as the result of a test. If this feedback is not taken into account, the whole process of testing which takes a lot of time and energy will be useless.

If an examinee has to choose or to answer the questions

with one word without explanation of his choice or response, he may guess with the chance of being correct more than 25% if it is a multiple choice test with four alternatives for each question. In an essay test, such a thing does not occur since the examinee is required to explain his answer. This is the third advantage of applying an essay test. Thus, in essay tests, it is hoped that guessing will be avoided.

On the part of the constructor, the advantage is that this type of test does not require much preparation. The number of questions is limited to ten at the most. In the preparation, it should be mentioned whether the examinees have to analyze, compare or tell the relationship of two things, give reasons or examples or illustrations, or to evaluate. Thus, in the preparation, the examiner just chooses the most important points within the materials and composes them using words such as describe, explain, compare, illustrate and evaluate. It is relatively easy and it does not take much time since the number of questions is relatively small. The hardest task in this preparation is to choose a small number of representative questions from the whole of the materials.

If an essay test is easy to prepare, the scoring is hard and time-consuming. It is hard since the examiner has to make judgement concerning the various answers given by the examinees for the same questions. He should judge whether the answers are complete, how complete they are, what score he will give them, and what standard or criterion is required for the examinee to pass. In short, it can be concluded that the variations of the examinees' answers make the scoring difficult and in this case, it is difficult to employ perfect

objectivity in the scoring. With regard to the time required by the examiner for scoring, this type of test will consume a great deal of the examiner's time. The scoring should be done by the teacher himself since the judgement made by the examiner will be different from that made by another person and this scorer will not have the same point as the teacher on the weight of the answers. Scoring is thus the weakest point of essay test and it is the first disadvantage of applying essay tests in the examination.

Since this type of test has only a small number of questions, the coverage of the materials is also limited. This disadvantage or shortage can only be hindered by the careful choice or selection of the questions being asked. Unless the selection is carefully done, the examinees may speculate in their preparation. Thus, they may speculate to study only the last part of the materials since this part is usually the part asked in the test if selection is not applied.

This type of test requires the examinee to express his ideas in his own words and style. This kind of test may thus be useful for measuring language variables. On the other hand, if this test is used for measuring the examinee's comprehension, it may have a bad effect. In expressing their ideas, the examinees may have difficulties. Thus, although they know the answers to the comprehensive questions, they are considered wrong since their answers are quite unintelligible. To overcome this shortage, it is advisable to give two scores on the comprehensive tests. The first score is for the comprehension, whether their ideas are good or not; while the second is for the expression, whether they express their ideas well or badly. In this case, the validity of the test can be maintain

tained, and the tests can reach their aims, viz. measuring the examinees' comprehension.

III. 3. The Stages in Essay Tests.

In learning a language, there are three domains that become the aims. These three domains are the cognitive, the psychomotoric and the affective domains.

The cognitive domain is concerned with the students' understanding of the subject they are learning. This domain is measured by means of a tool called a test. However, not all kinds of tests can be used for measuring this domain effectively. In terms of an essay test, it serves as a good tool whereby the examiner can measure the knowledge of the examinee as freely, broadly, deeply, and accurately as possible by asking explanations, descriptions, illustrations, and comparisons in the questions set for the examinee. Consequently, the quality of the questions given is very important in determining the kinds of knowledge demanded from the examinees.

The psychomotoric domain deals more with skills rather than knowledge. If knowledge needs memorization, the skills demand more practice. As far as essay tests are used for measuring language skills, they serve well to encourage this domain. For example, if an essay test is used for measuring the comprehensive reading of the examinees, it will help the examinees to have the skill in speed reading. This will, in turn, encourage the psychomotoric domain of the examinees.

The last domain in language learning, the affective domain, has to do with the attitude of the students towards

what they are learning. In terms of this domain, an essay test can increase it indirectly. If the examinees can answer correctly all the questions given in the essay test, they will be happy and hence, they will build a positive attitude toward the language tested. Even if they fail to answer the questions, they may still build a positive attitude as far as the examiner is able to give a wise explanation to them so that they are challenged to get better results in the next tests.

Those are the domains that will be achieved in every instruction. Since it is the cognitive domain that is directly related to the essay test, which is revealed in the answer given, it is only this domain that will be observed in this section.

It has been mentioned that an essay test requires relatively long answers. The length of the answers is, however, not the same from one question to another. It ranges from only two words up to an essay. Thus, there are several stages in the essay test in terms of the cognitive domain. It is these stages that will be discussed in this section. As a matter of fact, there are five stages of an essay test. These stages are differentiated according to the questions given and they will be discussed one at a time with the examples of each stage for the questions in the test.

III. 3. 1. The Mere Knowledge Stage.

This is the easiest stage in the essay test. In this stage, the questions in the test have something to do with mere knowledge. The examinees are able to answer the questions by memorization only or if it is a reading comprehension

sion test, they are able to deduce the answers easily from the passage or the text given. The questions themselves often begin with question-words such as who, what, where and when. On certain occasion, the questions may also be in the form of yes-no or one-word questions.

An example of an essay test on reading comprehension and its questions which are included in this stage is this :

Read This Passage And Answer The Questions : 17

Peter's uncle lived in the country. Once Peter went to stay with him for a few weeks. Whenever they went for a walk or for a drive in the car and they passed somebody, his uncle waved. Peter was surprised, and said, 'Uncle George, you know everybody here. Where did you meet them all ?'

'I don't know all these people,' said his uncle.

'Then why do you wave to them ?' asked Peter.

'Well, Peter,' answered his uncle, 'when I wave to someone and he knows me, he is pleased. He continues his journey with a happier heart. But when I wave to someone and he doesn't know me, he is surprised and says to himself, "Who is that man ? Why did he wave to me ?" So he has something to think about during the rest of his journey, and that makes his journey seem shorter. So I make everybody happy.'

The Questions :

1. Where did Peter's uncle live ?
2. What did Peter once do ?
3. What happened ?

17. L.A. Hill. Elementary Stories for Reproduction. Oxford, 1965,

4. What did Peter feel about this ?
5. What did he say ?
6. What did his uncle answer ?
7. What did Peter ask him then ?
8. What did his uncle's answer ?

In answering these questions, there is no need to use the power to analyze, to apply certain things or to create anything. What the examinee has to do is just read the passage carefully and take the answers from the passage.

III. 3. 2. The Comprehension Stage.

As indicated by its name, this stage deals with comprehension questions. To be able to answer these questions, the examinees are required to understand the subject fully, not just memorize it. In case of a reading comprehension test, the examinees cannot just take the answers from the passage but they have to use logic.

The questions in this stage are seldom used alone. Instead, they often mix with questions in the first stage, with various proportion. Thus, in an essay test, half the questions may be taken from the first stage while the other half are from the second. On another occasion, it may be that the questions from the comprehension stage are only three while the rest are taken from the first stage. The proportion of the questions depends on the level of the examinees and the aim of the test. The questions in this stage often begin with the question-words why and how. If the reading passage used in the first stage above is used here, the questions will be :

1. Why was Peter surprised ?

2. Why did Peter's uncle wave to everybody ?
3. How was Peter's uncle whenever he waved somebody ?

III. 3. 3. The Application Stage.

An essay test which uses questions from this stage will require the ability to compare or relate two or more things. Thus, in answering the questions, the examinees have to apply what they have learned by comparing or relating things. In foreign language testing, for example, the examinee is required to compare the vowel system in the target language and in the mother tongue. In this case, the examinee should, of course, already have enough knowledge of the vowel system in the target language and that question is meant as an assessment to this knowledge.

In testing writing skill, the examiner may make use of pictures as stimuli for the examinees to write. This kind of test is included in this stage since in such a test, the examinees should apply their knowledge after having a look at the pictures given. Another way of measuring their writing ability is by giving several questions. Instead of answering these questions, the examinees are required to make a composition based on the questions given. For example :

Do people visit your house ? Who are they ? Do they sometimes stay at your house for several days? Do you sometimes stay at other people's houses ? Do you find that you enjoy having some guests, but then you do not enjoy having certain others ? What sort of people do you like as your guests ? What sort of people do you dislike as your guests ? What sorts of things make a person a good guest ? What ones make a person a bad guest ?

In testing grammar, the questions of the essay type in this stage can be used effectively. These questions are relevant to the aim of the test, that is, measuring the examinees' knowledge of grammar. The question can be like this:

Make a sentence for each word or phrase below :

- therefore
- so that
- such ...that
- different from
- the same as

III. 3. 4. The Analysis Stage.

This stage is suitable for intermediate level examinees, at least, and it is too difficult for beginners. In this stage, the examinee should analyze the subject through observation or research. In the test, the examiner gives a set of data which should be observed by the examinee and the examinee should draw conclusions from the data and he gives reasons to support his conclusions. The support can be various. It depends on how reasonable he is and on how wide his knowledge is. The more reasonable the examinee is and the wider his knowledge of the subject given is, the better the support will be. It may happen that the data are not given to the examinee. Instead, the examinee should use his knowledge from the lessons or lectures given and use them as the data for his analysis. Suppose, the examiner has once given a lesson on the advantages and the disadvantages of comprehensive reading if it is compared to reading aloud. In the test, he may ask a question like this :

"Make an analysis of comprehensive reading and reading aloud !"

This stage is also very suitable for measuring or testing writing skill. For example :

Write an article on the conclusions that you draw from the data below. Support your conclusions with reasons. Your article should be about 250 words.¹⁸

	1950	1970
Travelling abroad	4	17
Going to seaside	38	31
Camping	8	31
Visiting friends/relatives in another town	11	10
Going to another town (but not to visit friends/relatives)	16	3
Staying at home	23	8
TOTAL	100	100

In the examinee's work, he should mention why people who are travelling abroad and who are camping increase in number while those who do other activities decrease. It should even be explained why the decrease is so sharp for some activities. In this way, the examinee exercises the way to analyse the data. This is very important, especially for those who are leaving schools since in real life, one time or another, they must be forced to analyse in such a way.

III. 3. 5. The Creative Stage.

This last stage is also the most complex one. The questions in this stage require the examinee to create some ideas

18. J.B. Heaton. Writing English Tests. Longman, 1975, pp.131-132.

of his own and then express these ideas clearly to the examiner. Thus, in this stage, besides the ability to describe and explain, his ability to create ideas is also required. It is, therefore, only appropriate for advanced students. For the purpose of language testing, this is very effective. Instead of the questions that should be answered in the class, the examiner gives an assignment to the examinee in the form of a paper or a report. This paper may be done at home. If it is a kind of report, the examinee has to insert his own ideas or opinion on the problem being discussed. For example, in the teaching methods subject, the examinee can be asked to make a paper on the ways to teach a grammatical pattern so that this pattern is taught on a larger scale, namely, in communication. Another examinee may be asked to make a paper on communicative grammar without special emphasis on certain patterns.

In terms of language skills, this stage of essay tests is appropriate for measuring productive skills. In the test, the examiner just gives a certain topic and lets the examinees develop their ideas as broadly and completely as possible within the time given. The examiner may give, for example, a topic like :

- What are your plans after you have graduated from this school ?

This topic may be given in the oral productive test or in the composition test.

Those are the stages that can be found within the essay test. Each stage is applicable for certain levels of the examinees. The more complex the stage is, the more judgement is required from the examiner and the more the test approaches the language learning aim and the nature of language learning

since in such a test, all language variables are involved, which means that it is an integrated test. In society, the examinee is often required to practise what he has learned. Thus, he has to employ the integrated skills. In short, it can be said that although an essay test which involves integrated skills is difficult to score, such a test has a high value for the examinee.

CHAPTER IV

EVALUATIONS OF ESSAY TESTS

In the previous discussion, the criteria for a good test have been given. Those criteria are validity, reliability, scorability, administrability, economy, preparability and some times adequacy. These criteria are mainly based on Lado's Language Testing : The Construction and Use of Foreign Language Tests. The scorability, economy and preparability are often referred to as practicality.

In this chapter, these criteria will be applied to essay tests so that it will be revealed whether the essay or the subjective tests are worth using. Each criterion will be discussed and in the discussion, it will be observed which language variables are suitable to be tested by means of the essay tests.

IV. 1. Validity.

A test is valid if it measures what it is intended to measure. An essay test can be valid if it is used consistently. This means that if it is used for measuring the comprehension of the examinee in reading, as far as the examiner sticks to this point and no other things, this test will be valid.

All language variables can be measured by using essay tests. These tests are valid since they measure what is going to be measured. In practice, however, it is sometimes difficult or even unreasonable to neglect the mistakes made by the examinees although these mistakes are concerned with expression, and not the content. To overcome this problem, it is advisable to give two marks, one is for the content and the other is for the expression, which includes the grammatical mistakes, the

spelling, the punctuation, and every aspect of the language usage.

For measuring the language elements, the essay test is valid. For testing pronunciation, the questions can be in this form :

- Mention the voiceless as well as the voiced consonants in English !
- What is the difference between the voiced and voiceless ones ?

For testing grammar, an essay test is often used. This test can be done orally or in written form. The questions asked range from just giving one clue and ask the examinees to make a sentence using that clue word, to the form of questions which require the examinees to make a paper on specific grammatical patterns. Thus, the questions can be like this:

- Make a sentence or two for each clue word below :
 - on the other hand
 - on the contrary
 - unless
 - in spite of
- Make a brief paper on the grammatical items 'some' and 'any' so that the difference in the usage is clear.

Make a brief paper on the teaching of grammatical items 'some' and 'any' using certain techniques so that they can be used in a clear context.

For testing vocabulary, an essay test is rather ineffective but it can still be used with quite high validity. The example is this question :

- Mention words which have meanings related to the parts of the body.

To conclude, an essay test is valid if it is used for measuring language elements although such a test should be reduced since passing such a test does not guarantee that the students are able to use the language. It only reflects the ability of the examinee over the language and hence, he will only know about the language. To know the language itself, the tests on the language skills should be employed.

For measuring the speaking skill, an essay test is an effective means if the test is done orally. In such a test, however, it should be remembered that the test is mostly concerned with the fluency of the examinee. Consequently, the content of the test should be kept easy so that the examinee is able to concentrate on his speech. If this requirement is maintained, the test will show its validity. The topic in this test is like :

- Tell me about your first day in school.
- Describe one of your relatives or friends.
- Write about your daily activities.

For testing listening or auditory comprehension, an essay test may have quite high validity, as far as the examinee is not asked to respond orally in writing. If the questions in the essay test for measuring comprehension require the examinee to speak or write in his response, the response may be influenced by his skills of speaking or writing. Thus, it will not be a valid test since it does not measure comprehension only. To overcome this, questions which require the examinee to do something can be employed. Thus, the questions will be :

- Please, open that door !
- Please, show me your bag !

This kind of question is, of course, unable to be used as broadly as possible since the vocabulary of the examinee has to be taken into account as well. Suppose, the examinee does not do anything after the examiner gives his question, it may be that the examinee does not understand the question or he simply does not know the meaning of the word used in the question. In summary, it can be said that this skill is difficult to test or to measure without overlapping with other skills.

For measuring reading comprehension, the essay test has high validity. In such a test, the examiner gives a reading passage followed by several questions of the essay type. If the examinee is able to answer the questions, it means that his comprehension of the passage is good. An example is this :

Read this passage and then answer the questions based on it !

Outdoor activities are those which are carried out in open spaces. These include different forms of exercises, sports and games like running, jumping, skipping and swimming. All these help to keep the person fit physically.

If every person were to take an active part in games and sports out-of-doors, then the population of that country could be said to be physically fit and active. Of course, this one hundred per cent rugged populace can never be successfully obtained. The reasons are obvious : the populace included very young children whose ages range from a few months to four years, old people, the physically handicapped and the sick or bed-ridden people, as well as housewives.

Outdoor activities, which comprise strenuous games in

main, teach the spirit of teamwork, of working hard to win and to do better, and of losing gracefully. These qualities, when added to a healthy body, give a sound mind.

The questions are :

1. What are outdoor activities ?
2. What is the use of outdoor activities ?
3. How can the population of a country become active and fit ?
4. Can a hundred per cent rugged populace ever be obtained ?
If not, why ?
5. How can a person be said to have a sound mind ?

These questions should ask for the information found in the passage so that the test will be valid for measuring the examinee's comprehension of the passage.

For measuring writing skill, the essay test is at its best, since an essay test requires the examinee to write. An example of a valid essay test for measuring the writing ability is :

- Write a reply to this letter using the same format.¹⁹

176 Wood Lane

London, N.W. 2

15th May, 1974.

Dear Mr. Johnson,

I wish to complain about the noise which has come from your home late every night this week. While I realise that you must practise your trumpet some time, I feel you ought to do it at a more suitable time. Ten o'clock in the evening is rather late to start playing. Even if you could play well, the noise would still be unbearable at that time.

19. Ibid. pp. 129 - 130.

I hope that in future you will be a little more considerate of the feelings of others.

Yours sincerely,

W. Robinson.

The examiner has to tell the examinees that the reply should not be too long. It can be as long as the given letter. In the reply, the examinees should deny the accusation. Instead, they would tell about or complain of the loud noise from Mr. Robinson's television; this complaint must be reasonably polite since they do not want to be enemies.

For measuring the translation skill, an essay test is also valid. In the test, the examiner may give some sentences or a short passage for the examinees to translate either from the target language into the mother tongue or vice versa. Since translation is an integrated skill, all the mistakes made by the examinee should be taken into account.

IV. 2. Reliability.

Reliability is necessary for any test to be good. In terms of an essay test, it is reliable since the examinees who are able to answer the questions in the test will also be able to answer those questions again in the next occasion unless he is not prepared for the test or he is in a bad condition. Even if the questions are changed, they may still be able to do the test as far as the questions are of equal difficulty between the two tests. For example, if the examiner gives a topic for a composition, the examinee may make a one-page essay. Next time when another topic of equal kind is given to him, he will make approximately the same essay as the previous one. In the first

occasion, the examinee is asked to describe his school, for instance, then on another occasion, he may be asked to describe one of his relatives or friends. These two essays will approximately be the same and hence, the scores will relatively be the same if the examiner is able to maintain his reliability in scoring.

The other method of measuring the reliability of a test is by readministering the test after a certain period of time. During this period of time, the examinee is assumed to have the same treatment as the others. They may either have certain instruction or have no added instruction at all. Then, the results of the two tests are compared to get how great the reliability of the test is. This method is, in any case, inefficient since certain students will benefit more than the others in the familiarity with the type and format of the test. Moreover, in addition to the changes in performance resulting from the memory factor and instruction, personal factors such as motivation and differential maturation will also account for differences in the performance of certain students.

IV. 3. Scorability.

This criterion is, in fact, the weakest point of an essay test since the scoring should be done by the examiner himself and it depends a lot on the examiner's judgement. In his judgement, the examiner should try to be critical and careful. Even if the examiner has tried hard to be careful and critical, his judgement will never be a hundred per cent objective because he has his own taste, background, logic, emotion, motivation and knowledge. It is very probable that the examiner will give a good mark for he is in a good mood. Perhaps, he

will consider that same work to be bad when he is feeling very annoyed or tired.

Two different examiners may give different scores for the same work so that the reliability of the scoring of an essay test will tend to be low. To overcome this problem, several general steps can be used to help the examiner to be more objective in scoring an essay test. If these steps are followed by different examiners who are scoring the same work, the scores will not be much different. These steps are:

IV. 3. 1. The examiner prepares the key-answers which should be the ideal ones. An ideal answer should have the features of a good answer, namely, correct, relevant, complete, systematic and clear.

The state of being correct, relevant, complete and systematic deals more with the idea of the answer, while the state of being clear deals more with the language used or the expression.

It is not difficult to see whether an answer is correct. It is correct if it answers what is being asked. Thus, if the examiner asks the examinee to mention the vowel system in English and the examinee mentions i, a, u, o, e, ai, and oi, the answer is correct although it is incomplete.

The relevancy of the answer is not easy to trace since an answer may be irrelevant although it is not wrong. For example, there are two persons, say Ali and Amat, who are talking to each other concerning the results of their exams. Ali asks : "Why did you fail in the last exam ?". Then Amat answers :

"I was sick, you know, I had seen a doctor. The doctor said that I had been too tired. Therefore, I had to and so on ." In this case, Amat's answer is irrelevant because he did not answer what Ali asked. As a matter of fact, this answer is true in terms of the content. Amat was sick and had seen a doctor. Yet, this true answer is not relevant with the question. If Amat is to give a relevant answer, he should say : "Because I was unable to answer the questions well.", or other similar answers. This is a relevant answer because it is his inability to answer well that makes him fail, and not the fact of his being sick. Even if he was sick, he might have passed if he had been able to answer the questions well.

A good answer should also be complete. This means that in answering the questions, the examinee should include all points that are relevant to the questions. For instance, in the test, the examinee is asked to mention the English tenses. In his answer, he only mentions the simple present tense and the past continuous tense. This means that his answer is still incomplete and hence, his score should be reduced.

The examinee's answer is said to be systematical if it uses a logical order. For example, an examinee is asked to explain A, B, and C. If in his answer he explains A partly, then C, then A again, B, and C again, it means that his explanation is unsystematical.

The last feature of a good answer is its being clear. This deals more with the expression, which means that an answer is clear if the sentences in it are grammatically

correct and they are not complicated, the diction and the idioms are properly used, and the spelling or the pronunciation is correct.

- IV.3.2. There are two kinds of correction methods, the general and the specific ones. In correcting the daily progress tests, the examiner may use the general method, that is, he gives a score or a mark directly to the whole work. In the final exams, however, it is advisable to employ specific method. Using this method, the examiner scores or examines the work part by part. For example, the examiner gives scores to the answers of all examinees for question number one. Then he gives scores to question number two, and so on. This scoring is done successively for each number. The reason for choosing this method is to help the examiner avoid the subjective factors such as sympathy or antipathy towards certain examinees.
- IV.3.3. The examiner tells the examinees to put their students' numbers only, and not their names on their answer sheets so that the examiner does not see directly whose work he is examining. This is meant to prevent the examiner from being influenced by the previous performance and achievement of the examinees.
- IV.3.4. After the correction, the work is better given back to the examinees and is discussed in class so that the examinees know why they get good or bad marks. This is important since the idea of giving scores is to improve both the teaching and the learning.

Those are several good steps in scoring. An example of an essay type used for measuring the writing skill with its ideal answer is :

Give an account, in not more than 70 words, of how you opened a bottle without a corkscrew ! ²⁰

Idea list :

Tried banging bottle underneath. Hot, cold, water. Pushing cork down. Breaking neck ? Danger glass. Picking cork out. Had to push down remains.

The ideal answer should be like this :

Unsuccessfully I banged the bottle underneath with my hand to loosen the cork. Next I ran hot and cold water over it alternately, and banged it again unsuccessfully. I could not push the cork and would have broken the neck off, but glass might have got in. Finally I picked at the cork with a knife. Halfway, the other half began slipping down, so I pushed it in.

IV. 4. Administrability.

In terms of administrability, an essay test is usually very administrable. This means that a test which uses an essay type can be handled by the persons available under various conditions, except if the test is done orally. In the oral essay test, the presence of the examiner is required since in such a test, the score is usually given right after the administration of the test.

On the contrary, in the written essay test, the examiner need not be there during the administration. Thus, if an examiner is sick at the time of the test or if he gets trouble

20. Richard and Long, Breakthrough 3, p. 40.

to be in the class during the administration of the test, the test can still be carried out as long as there is another person, who is willing to replace him in administering the test; of course, assuming that the questions have been prepared before.

A test of the essay type is also administrable in terms of the number of the required personnel. For one class, one examiner is enough for administering a composition test, for example. This is possible because the examinees are unlikely cheating each other during the test.

IV. 5. Economy.

This criterion is very well suited to an essay test. An essay test is economical because it can be used for measuring what it is intended to measure in a reasonable time considering the testing situations such as setting and facility.

For Indonesian schools where the majority of them have no printing equipment, a written essay test is better used at least for the daily progress tests. This is economical since the questions in the test are relatively small, and therefore, they can be written down on the blackboard. Thus, the examiner need not pay for the composing of the questions since he does not rent a printing equipment.

The Indonesian classes are usually large. For the administration of the test to be economical, an essay test is better used. In the essay test, only one examiner is required for administering the test since it is usually difficult to cheat the answers to the questions asked in such a test.

IV. 6. Preparability.

An essay test is preparable since there is not much time needed for composing its questions. The number of questions is relatively small so that it is easier to prepare such a test if it is compared to the tests which use a big number of questions such as in the objective tests.

The problem faced in the preparation of the test which uses an essay type is to choose questions which are representative and adequate. The questions are representative if all the materials can be included in them. To choose a small number of questions for the whole materials can be difficult. On the contrary, if the materials consist of only one or two chapters, an essay test can be used effectively since it is easy to choose representative questions from those limited materials. Thus, if after each chapter, a test is given, an essay type is recommendable.

The questions in the test should be adequate as well. This means that the questions include all materials in proportion to their importance in the course. Some of the questions must be easy, but some others must be average and some are even difficult.

Those are the criteria for a good test. Sometimes, another criterion, discrimination, is also included. A test discriminates positively if the examinees who are bright answer the questions better than the less bright ones. If in the test, the poor examinee can answer, while the good ones cannot, it means that the test discriminates negatively. The idea of this criterion is to separate the good and the poor learners.

From the six criteria above, the last three can be referred to as one, practicality. A test which is economical, administrable and preparable is a practical test. The essay test is practical since it is economical, administrable, and preparable. Moreover, an administrable and preparable test is usually also economical, but the reverse is not always true. An economical test may be neither administrable nor preparable. For example, if there is printing equipment in the school, the teacher may print the questions of an objective test. This means that the test is economical since the examiner can use the equipment in the school effectively. However, in terms of preparability and administrability, this test is not effective. The examiner needs a relatively long time to prepare the questions in such a test. Furthermore, if the class is large as usually exists in Indonesian schools, more than one examiner is needed for the administration of the test. If not, the chance for the examinees to cheat becomes greater.

C O N C L U S I O N

After a long discussion of language tests, it can be concluded that language testing is just as important as language teaching. However, due to economic conditions in Indonesia, teachers usually experience difficulty in devoting sufficient time to their profession, especially in constructing suitable tests for their students. Some teachers may think that giving a test is not a pleasant job at all since testing involves a great deal of work either in the preparation, administration or in scoring. Some others might feel compelled to give a test or tests for they have to hand in the marks for the report-cards.

Based on this fact this thesis has attempted to discuss the various types and techniques of language, especially English language, testing and the importance of testing in the whole of language instruction. Hence, after reading this thesis it is hoped that language teachers may have a positive attitude towards testing. They may realize that the aim of testing their students is mainly to control the students' learning process, not just to fill in the marks in the report-cards. Moreover, since there is a tendency to use objective tests for all purposes, this thesis tries to place subjective tests in the proper perspective. From the discussion, it can be concluded that the subjective or essay type of test is very important in the whole of language instruction.

The essay type can be used for measuring all language variables, from the easiest stage to the more difficult ones. The problem with this type of test only lies in the scoring.

Scoring, however, is not everything for determining whether a test is good or not. This is especially true for language testing. Language is quite complex and it is very difficult just to measure one part of the whole of language variables. A language test is at its best if it measures the general mastery of the students. Therefore, an essay type which measures the general mastery of the examinee is the best one in terms of language testing.

At last, some hints that will help language teachers improve their teaching and testing will be given. These hints are by no means the only ones but if they are followed, they will contribute to the improvement of the whole of language instruction.

First, to get the chance with language usage, a weekly meeting with a native speaker is suggested. If there is no native speaker, a senior teacher can be used. With this meeting, the class teacher is assumed to have better insights so that he may give new things in his teachings.

The school provides some useful works of reference such as encyclopaedias, dictionaries, and books on teaching methods as well as on testing. This reference is very useful since one time or another, the class teacher may find difficulties and hence, he may consult the reference concerning those difficulties.

If possible, the school provides a subscription to newspapers and magazines using the target language. The idea of this subscription is to keep the teachers in touch with modern ideas and with the English way of expressing them, if the target language is English.

The last thing to be suggested is that the class teach

er is given a chance to follow up-grading and refreshing courses regularly, either in the field of teaching or testing. The idea of this activity is to help the teacher know the newest methods both in teaching and in testing. After mastering those new methods and techniques, he is supposed to apply them to his students.

Those are the things that will likely help the teachers in improving both their teaching and testing methods and techniques. It is the earnest desire of the writer to make a contribution, through this thesis, to the progress in the field of testing especially and to the progress of the whole language instruction generally.

B I B L I O G R A P H Y

- Anthony, Edward M., "Approach, Method and Techniques" in Teaching English as a Second Language, edited by Allen, M.B., New York: McGraw-Hill, 1965.
- Bismoko, J., An Introductory Reader to Methods of Teaching English in Indonesia. Yogyakarta: IKIP Sanata Dharma, 1976.
- Byrne, Donn, English Teaching Extracts, London: Longmans Group Ltd., 1969.
- Davies, Alan, Language Testing Symposium. A Psycholinguistic Approach, Oxford University Press, 1968.
- Departemen Pendidikan dan Kebudayaan, Kurikulum 1975, buku III C, Pedoman Evaluasi, 1978.
- Departemen Pendidikan dan Kebudayaan, Pedoman Pelaksanaan Penilaian Pendidikan dan Penyusunan Program Pengajaran, 1978.
- Harris, David P., Testing English as a Second Language. Tata McGraw-Hill, New Delhi, 1974.
- Heaton, J.B., Writing English Language Tests, London: Longmans Group Ltd., 1975.
- Hill, L.A., Elementary Stories for Reproduction. Oxford University Press, 1965.
- Lado, Robert, Language Teaching, New York: McGraw Hill, Inc., 1964.
- Lado, Robert, Language Testing, London: Longman, 1975.
- Nation, I.S.P., Language Teaching Techniques. Yogya: IKIP Negeri, 1974.
- Poedjosoedarmo, Gloria R., Ph.D., Lectures on Semantics IKIP Sanata Dharma, Yogyakarta.

Poedjosoedarmo, Soepomo, Ph.D., Lectures on Teaching Methods. IKIP Sanata Dharma, Yogyakarta.

Richard & Long, Breakthrough 2 & 3, Oxford University Press, 1977.

Winkel, W.S., SJ, M.Sc., Pengantar Psikologi Pendidikan. FIP, IKIP Sanata Dharma, Yogyakarta, 1974.

