

CHAPTER ONE

BASIC ASSESSMENT CONCEPTS

Assessment is a topic of concern to language teachers. Although teachers are primarily concerned with teaching, rather, than assessing, they must periodically measure their students' performance. This chapter introduces the students to the basic assessment concepts as raised by the following questions:

Why do we assess?

How do we assess?

What do we assess?

When do we assess?

1.1 What is assessment?

Any means of checking what students can do with the language. It also includes what they cannot do, but proper assessment gives due weight to positive side of their achievement.

Assessment may be carried out before, during or after a course, or it may not even be connected with a course. Assessment may be of individual students or it may be to check the capability of a whole class. This means that assessment is concerned with teaching as well as learning.

1.2 Why do we assess?

There is no universally good way of assessing. What and how you assess should obviously depend on your purpose. We can distinguish at least seven important purposes of assessment which indicate seven different emphases in measuring student ability and potential

The purposes of assessment

(1) **Placement:** to place or classify students on a suitable course.

(2) **Diagnosis:** to specify the student's particular strengths and weaknesses.

(3) **Selection:** to select as fairly as possible the best candidates for a limited number of places or jobs or for higher education.

(4) **Evaluation:** to evaluate the effectiveness of the syllabus as well as the method of teaching so as to make adjustments when needed.

(5) **Progress:** to assess what the students have learned of the specific syllabus and what progress they have made.

(6) **Prediction:** to discover potential abilities and aptitudes and to predict probable future successes whether in school or outside.

(7) **Accreditation:** to provide evidence that the student has undergone a course of training or study and is qualified to practise his profession.

Types of tests:

Consequently, these different purposes of assessment can be realized by four major types of language tests: **achievement/attainment tests, proficiency tests, aptitude tests and diagnostic tests.**

(1) Achievement/Attainment Tests:

These tests can be further subdivided into class progress tests and achievement tests

Class progress tests:

The progress test is designed to measure the extent to which the students have mastered the material taught in the classroom. It is based on the language programme which the class has been following. Its aim is to stimulate learning and to reinforce what has been taught. Its backwash effect on learning and teaching is valuable. In addition, the progress test enables the teacher to become more familiar with the work of each of his students and with the progress of the class in general.

Achievement tests:

Achievement tests, though similar to progress tests in some ways, are far more formal tests and are intended to measure achievement on a larger scale. Most annual school examinations, and all public examinations which are intended to show mastery of a particular syllabus are of this type.

(2) Proficiency tests:

The proficiency tests also measure what testees have learnt but they are in no way related to any syllabus or teaching programme. These tests are concerned simply with measuring the student's control of the language in the light of what he will be expected to do with it in his future performance of a particular task. In other words, language tests of this type aim to assess global ability. Typically proficiency tests are used in the placing of individuals in learning groups appropriate to their level of knowledge-beginner, intermediate, advanced. There are several internationally or proficiency tests in English. Among the best known are the TOEFL and the IELTS.

(3) Aptitude tests:

A language aptitude test (or prognostic test) is designed to measure the student's probable performance in a foreign language which he has not started to learn. Aptitude tests generally seek to predict the student probable strengths and weaknesses in learning foreign languages by measuring his performance in an artificial language. Language learning aptitude consists of many factors including: intelligence, motivation, age, memory, etc.

(4) Diagnostic tests:

Although the term 'diagnostic tests' is widely used few tests are constructed solely as diagnostic tests. Achievements and proficiency tests are frequently used for diagnostic purposes. The aim is to find out the strong and weak points of the class as a whole and individual student so that appropriate remedial action can be taken.

1.3 How do we assess?

The question about how to assess may be answered either in terms of describing the variety of testing techniques or through the consideration of a series of conflicting modes of assessment. A considerable range of testing techniques are discussed thoroughly in chapter four. Modes of assessment, on the other hand, require the use of differing techniques. We may identify some of the major ones as follows:

Modes of assessment

Formal vs Informal

Formative vs Summative

Continuous vs Terminal

Internal vs External

Coursework vs Examination

Although the distinction made are sometimes too fine, and there is often considerable overlap between the modes; this contrasting presentation is significant for the construction of assessment.

Formal assessment is standardized, large-scale tests which are prepared by professional testing services to assist institutions in the selection, placement and evaluation of students. **Informal assessment** is often used diagnostically by the teacher to obtain reliable insights about the students' ability and state of development.

Formative assessment helps to show which aspects the student has mastered and where remedial work is necessary. The tests are to be given during the course of instruction. **The summative assessment** is usually given at the end of a course to measure the 'sum total of the material covered.

Continuous assessment means continuous updating of judgements about students' performance in relation to specific subject. **Terminal** means assessment that takes place at the end of the course. The two modes could be put together to form some final assessment.

Internal assessment is usually carried out by the teacher of the subject once a month or a year to check students' progress. **External assessment** is usually constructed by a team of specialists or an organization outside the school and applied on thousands of students for the purpose of selection or placement.

Examinations are either school examinations constructed by the teacher or teachers of the subject in a given school, or Ministry examinations constructed and administered by the Ministry of Education. **Coursework** is work completed by the students under the supervision of their teachers, for example, assignment or fieldwork

1.4 What do we assess?

Clearly the answer to this question is language. When analysing the language being tested we have to consider the whole system of language skills and components. Indeed language skills are complex and so related to each other that it may seem impossible to separate them for the purpose of any kind of assessment.

Four major skills in communicating through language are often broadly defined as listening speaking reading and writing. Ways of assessing performance in these skills may take the form of tests of:

Listening comprehension: including phoneme discrimination, stress and intonation as well as recorded texts followed by questions of varying techniques.

Speaking ability: in the form of an interview, picture-description or roleplaying.

Reading comprehension: in which questions are given to test students' ability to understand a text by applying certain cognitive skills as scanning, skimming, and drawing conclusions.

Writing ability: usually in the form of letters, instructions, accounts of past events. The linguistic components of the foregoing skills are: grammar, vocabulary and phonology.

Test of phonology:

Which vowel sound is different?

- A. go B. got C. close D. smoke

Test of grammar

You would get better soon if you ----- a doctor.

A. see B. saw C. are seeing D. had seen

Test of vocabulary

To ----- someone means to save him from danger.

A. praise B. injure C. rescue D. murder

However, at all levels, it is generally advisable to include test items which measure the ability to communicate in the target language in speech and in writing as well as to understand what they hear and read.

1.5 When do we assess?

It is important to determine when to assess, as it is important to determine how and what to assess.

The three terms terminal, periodic and continuous assessments are often used. Assessment can be introduced at different stages of a course of study:

(1) **Terminal of final**: simply means assessment that takes place at the end of a predetermined period or unit, as, for example, a school year or a course of study. It ought, therefore, to concern itself with the whole of what is being taught.

(2) **Periodic assessment**: measures levels of achievement reached at predetermined intervals throughout the course.

(3) **Continuous assessment**: is given to have a continuous updating of judgement about a student's performance.

CHAPTER TWO

APPROACHES TO LANGUAGE ASSESSMENT

Modern theories about the nature of language seem to influence theories about language learning, which in turn influence ways and methods of teaching and testing. The direction of influence is usually from linguistic theory to learning theory to teaching methods and eventually to testing.

This chapter will touch lightly on the major approaches to language testing. These are roughly classified according to four main ones: (1) the essay-translation approach: (2) the structuralist approach (3) the integrative approach: (4) the communicative approach.

2.1 The essay-translation approach

This approach is referred to as the pre-scientific stage of language testing because no special skill in testing is required and also there is a lack of concern for objectivity and for test reliability.

The tests are mainly written exercises, such as: translation, composition, grammatical analysis, and sometimes, aural/oral component.

2.2 The structural approach

This approach is mainly characterized by the view that testing can be objective, precise, reliable and scientific. Objective tests have been developed the most widely used being the multiple-choice items.

Statistical methods have been improved to measure the reliability and validity of tests. Contrastive analysis is also used to identify problems faced by second or foreign learners. The tests measure the learners' mastery of the separate elements of the language: phonology, vocabulary and grammar. The skills of listening speaking, reading and writing are also tested separately. The best example of these tests is the TOEFL (Test of English as a Foreign Language).

According to the foregoing view, the analytical assessment or (discrete-point approach) is followed. Analytical assessment seeks to measure the

knowledge of students on individual distinct points of separate sounds, special grammatical features or vocabulary items.

These tests often used long lists of unrelated sentences that were incomplete or that contained errors in grammar or usage. Students have to complete or correct these sentences by selecting or supplying the correct answers as in the following examples of structure and vocabulary:

1. I want ----- home now.

A. to go B. go C. going D. went

2. She needs to get up earlier so she's buying ----- clock

A. a time B. a bell C. a watch D. an alarm

2.3 The integrative approach

This approach involves the testing of language in context and is concerned with meaning and the total communicative effect of discourse. Thus, it is the exact reverse of analytical assessment since it attempts to assess several elements and perhaps skills at the same time. Integrative tests are best characterized by the use of cloze testing, which reflects the best feature of the two approaches: analytical and integrative.

2.4 The communicative approach

This approach is sometimes linked to the integrative approach as they both emphasize the importance of the meaning of utterance. However the communicative tests are concerned primarily with assessing real communication. The tests should indicate how well a student can function in the target language. Consequently the tasks should relate to real-life situations with emphasis on authentic materials.

Language use is often emphasized to the exclusion of language usage. Use is concerned with how people use language for a multitude of different purposes while usage concerns the formal patterns of language.

Finally, communicative testing has introduced the concept of criterion-referenced assessment in preference to norm-referenced assessment. Criterion assessment rates students against certain standards, but norm-assessment compares each student with his classmates.

CHAPTER THREE

CHARACTERISTICS OF ASSESSMENT

3.1 VALIDITY

Validity is the degree to which a test measure what it is supposed to measure or can be used successfully for the purpose for which it is intended.

Two questions must always be considered:

1. What precisely does the test measure?
2. How well does it do it?

There are, in fact, four kinds of validity, the first is Content validity; and the remaining three are Construct, Empirical and Face validity.

(a) Content validity: is the extent to which a test adequately covers the syllabus to be tested. A valid test must be based upon a careful analysis of subject or skill we are testing. It must also be constructed as to represent adequately each portion of this analysis.

(b) Empirical validity: If we are to check the effectiveness of a test and to determine how well the test measures, we should relate the test scores to some independent outside criteria such as scores given at the end of the course or the teacher's judgement of his students or an external examination.

(c) Face validity: This simply means the way the test looks to the examinees, supervisors or in general to the people concerned with the education of the students. Obviously face validity can never replace content validity or empirical validity, yet its importance should not be underestimated.

(d) Construct validity: A test is said to possess construct validity if it is designed in accordance with a theory of language behaviour and learning.

3.2 RELIABILITY

Reliability means the stability of test scores. Presumably, if the same test is given twice to the same group of students, under the same

conditions, it would give the same results. The requisites of a dependable test are the following:

(1) Multiple samples: The more samples of students' performance we take the more reliable will be our assessment of their knowledge and ability. Consequently, the test must be long enough to provide a generous sampling of the areas tested. And it should also contain a wide variety of levels of difficulty.

(2) Standard conditions: The reliability of the test scores can be assured only if all students take the examination under identical conditions. In a listening test, for example, all students must be able to hear the items clearly.

(3) Standard tasks: All students must be given the same items or items of equal difficulty. In other words, the tests must be identical so as the format

(4) Standard scoring: All tests must be scored in an identical manner. The scorer should give the same or nearly the same score repeatedly for the same test performance.

Objective tests tend to be more reliable than free-response tests like composition, where individual judgement must be made.

3.3 PRACTICALITY

A third characteristic of assessment is practicality or usability. The test which is appropriate in terms of our objective, provides dependable evidence and can do this with ease and economy is practical. A test which is valid and reliable but difficult to administer or score or which is expensive may fail to gain acceptance. Two parameters appear to be involved:

(1) Economy: The cost in time, money and personnel of administering a particular test.

(2) Ease: The degree of difficulty experienced in the administering and scoring of the test, for example, an oral test that demands the use of a tape recorder is not practical if it has to be administered to thousands of students.

CHAPTER FOUR

TECHNIQUES OF ASSESSMENT

1) Objective Items (or Closed response)

An objective item (item is the term more commonly used than question) is one which is asked in such a way that there is only one predetermined correct answer.

Example:

The country to the north of Iraq is -----.

A. Syria

B. Kuwait

Options C. Iran

D. Turkey

4.3 CHARACTERISTICS OF OBJECTIVE ITEMS

1. Each objective test item has a precisely defined problem that can be solved in a short time.
2. An objective test is usually highly reliable.
3. An objective test usually has high content validity.
4. The test provides wide coverage of subject content.
5. It can be speedily and easily marked
6. Marking is not influenced by the marker's prejudices.
7. It has the advantage of being easy to examine statistically.

2) Semi-objective Questions

Semi-objective items are designed to elicit answers which may be marked more or less objectively not so objectively as with multiple choice tests and not so subjectively as with open-ended or essay questions. Semi objective questions are fundamentally short questions that require more than just recognition as with objective items. Actually they are considered the most promising forms for measuring understanding, reasoning, creative thinking and problem

solving. Semi-objective questions combine some of the virtues of both objective, and subjective questions. The problems are short and highly structured; nevertheless, they provide the examinee with the opportunity to compose his own answers.

Examples:

(1) Fill the blank space with a word derived from the word between brackets.

To be ----- it is generally necessary to be well educated. (success)

(2) Define the term validity?

4.17 ADVANTAGES AND DISADVANTAGES OF SEMI OBJECTIVE QUESTIONS

Advantages:

- (1) The questions are quick and easy to write.
- (2) They are less time consuming to construct than multiple choice items.
- (3) They require much less scoring time than would a composition.
- (4) They embrace a greater subject area than a composition question.
- (5) They demand a certain amount of coherence in the answer.
- (6) They require more than just recognition and recall.
- (7) The opportunity for guessing the correct answer is much less than that of a multiple-choice item.
- (8) The relatively high objectivity of marking makes this test reliable.

Disadvantages

- (1) They take longer to score than objective tests.
- (2) Quite frequently there are a number of possible right answers, some of which might not have been considered by the tester when preparing the test.
- (3) They are not suitable for certain such as answers to questions on literary appreciation.

3) Subjective (or Free-response) Questions

The term free-response or sometimes open-ended implies two things: first, the student has to decide what to include and what to exclude in his answer and secondly, there is normally no precise limits in the questions themselves as to the length or style of answer. Essay questions and composition are the most appropriate forms for assessing command of language. Marking these questions is subjective, since it is affected by the judgement the tester.

Techniques of testing:

<u>OBJECTIVE</u>	-----	<u>SEMI OBJECTIVE</u>	-----	<u>SURJECTIVE</u>
Multiple-choice		Transformation		Composition
True/false		Completion		Essay
Matching		Cloze test		Precis
Supplying		Short-answer questions		
Labelling				

CHAPTER FIVE

TESTING USAGE AND USE

The testing of usage and use assesses the learner's active control of grammatical and sentence structure. The test should include the full range of structures that were taught in the course. These include (verb form, verb tense, prepositions, pronouns, modal verbs auxiliaries, questions words, word order, conjunctions, determiners, active and passive, direct and indirect, and the like.

TECHNIQUES FOR TESTING, USAGE AND USE

5.2 MULTIPLE CHOICE COMPLETION ITEMS

The following are some examples of how to construct this format.

1. ----- are you going? To school.
A. When "B. Where C. What D. How
2. This is the man ----- son is a mechanic.
A. who B. whose C. what D. where
3. ----- he tries, he will not succeed.
A. Although B.If C. Before D. Until
4. I'm upstairs. I ----- a bath.
A. have B. am having C. was having D. had

5.3 SUPPLY-TYPE ITE

Complete the following sentences by supplying the correct form of the verb between brackets.

1. I haven't seen him since he (leave) school.
2. I (write) to him last summer.
3. Peter (punish) by his father for coming late.

5.4 COMPLETION (WORD FORMATION OR INFLECTION)

From the word in capital form a word that fits suitably in the blank

space.

1. Thousands of people have been made ----- by the war. (HOME)
2. I am not saying he's ----- but he's not very good with money. (HONEST)
3. He said good morning in a ----- way. (FRIEND)

5.5 COMPLETION

This format is more of a semantic test than a structural one. It tests the student's active control of sentence structure.

Complete the following sentences with suitable words:

1. If I had a camera, ----- of this beautiful sight.
2. He must leave early because -----.
3. They are so tired ----- this evening.

5.6 TRANSFORMATION

Transformation technique is preferable to multiple choice items because it tests a wide range of structures with no need for supplying alternatives, For example: Finish each of the following in such a way that it means exactly the same as the sentence before it.

1. "Where is the post office? John asked. John asked -----
2. Someone must meet him at the station. He -----
3. My brother and I both went to that school. I went to that school and so -----
4. John can dance better than Jill. Jill cannot -----
5. This suitcase is too heavy to carry.
6. This suitcase is so heavy that -----

Transformation format may require testees to substitute a given word in the sentence. The new word needs a change in the structural pattern of the sentence.

For each of the sentences below, write a new sentence similar in meaning to the original sentence.

1. She couldn't type very fast. (Use: able)
2. He won't listen to a word I say. (Use: refuse)
3. They accused me of causing the accident. (Use: Blame)

5.7 ARRANGEMENT

The testees are given sets of words to form correct sentences complete each sentence by putting the words in the right order.

1. went, yesterday, film, John, see, to, and, a, evening
2. the, he, situation, by, was, worried
3. Of, can, which, louder, you, shout?

5.8 ERROR RECOGNITION

Sometimes students are given incorrect sentences. Then they are required to recognize the source of error by selecting the part that contains the error. They may also be asked to produce the correct forms. This type of test is closely related to those skills required when students check, proof-read an insider, article or composition they have just written.

Each sentence contains three words or phrases underlined and marked A, B and C. Write the letter of the incorrect word or phrase

1. Susan likes tennis so I do.
 A B C
2. I did not see Ali since he went to Mosul.
 A B C
3. I do hope you are pleased in my work.
 A B C
4. Tom cannot sing as good as you do.
 A B C

Another variant of this technique is to give the students sentences. They are, then, required to decide whether they right or wrong, and correct the wrong ones. Another be

1. I'm usually going to work by car.
2. Look! That man is climbing up the tree.
3. No, he hasn't wrote to me recently.
4. Three men have been arrested by the police.

CHAPTER SIX

VOCABULARY TEST

6.1. It is considered important to test knowledge of individual vocabulary items for three main reasons:

First, they represent one of the principle components of language. **Second**, the testing of individual vocabulary items enables the tester to include a large number of items in a limited time available for any test. **Third**, the testing of individual items allows for greater objectivity marking which provides a balance for the more subjective assessment involved in marking pieces of connected writing. A distinction should be made between the active and passive vocabulary. **Active vocabulary** is tested as part of integrated skills in the connected writing or speaking questions; whereas **passive vocabulary** is tested in the comprehension questions. The vocabulary items are usually drawn directly from the textbook used in class.

TECHNIQUES FOR TESTING VOCABULARY ITEMS

6.2. (1) Multiple-choice items

Multiple-choice vocabulary tests include the following types:

(A) Multiple-choice completion

Write the number of item and the letter of the word which best completes each sentence.

1. If you walk along this road you will see the signpost ----- to the beach.

A. pointing B. showing C. directing D. indicating

2. They ----- me for sending them postcards

A. think B. thank C. believe D. help

(B) Multiple-choice synonyms

Write the number of item and the letter of the option that is closest in meaning to the underlined word (or words) in each sentence:

1. A man was murdered last night.

A. moved B. killed C. stopped D. called

2. He spoke rudely in class yesterday.

A. unkindly B. impolitel C. heautifully D. respectfully

(C) Multiple-choice Antonyms

Write the number of the sentence and the letter of the word that gives the opposite meaning to the underlined word.

1. She lives in a modern house.

A. old B. beautiful C. nice D. new

2. They have won the match.

A. found B. got C. finished D. lost

6.3. Matching items

This technique is used for testing synonyms, antonyms, definitions, collocations and the like.

Write the number in Column A and the letter in Column B which matches it.

Objects and Their Uses

A

1. A thermometer
2. A vacuum cleaner
3. A camera
4. A calculator
5. A key
6. A match

B

- a. for locking doors
- b. for riding
- c. for cleaning carpets
- d. for making holes
- e. for taking pictures
- f. for doing arithmetics
- h. for lighting a cigarette
- i. for taking temperature

The words in column A can be replaced by pictures of the objects.

6.4 Word sets

This format is used for measuring student's familiarity with a range of associations:

A) Put the following words in the correct columns.

weightlifting, tie, journalist, Cheese, jeans, manager, T-shirt, climbing, employee, swimming, lettuce, tennis, sweets, pajamas, biscuits.

Clothes	Food	Athletics	Profession

B) In the following list of words there are 4 professions and 4 words associated with each profession. Put the words in the right column.

- bandage - technology - oven – cook
- headline - to interview - X-ray – programme
- to roast - newspaper - spoon – journalist
- injection - thermometer - pan - collect data
- nurse -break into systems - article - computer expert

6.5 Completion items (gap filling)

The lack of context of the former type is felt to be a drawback as a large part of lexical meaning comes from the context. Gap-filling format for testing vocabulary is generally preferable particularly when based on connected text rather than single sentences.

A) Write the number of the blank spaces and the words you choose from below to fill in the gaps in the following passage:

watched saw went visited
rested ate got up returned

This morning we (1) ----- Kuala Lumpur the capital of Malaysia. We (2) ----- at 8.00. After breakfast we (3) ----- to the National Museum. At the museum we (4) ----- some beautiful Malay handicraft. After visiting the museum (5) ----- to the hotel for lunch. We (6) ----- a lovely meal of fish, rice and fruit. We (7) ----- in the afternoon. In the evening, we (8) ----- a beautiful cultural show.

B) The following is a cloze test with the omission of lexical items only. Fill each of the numbered blanks in the following passage with ONE suitable word.

Dear editor,

I read your newspaper every week and usually, I (1) ----- with what you say but last week's (2) ----- 'Our Town' made me very (3) ----- . You said, our (4) ----- was crowded, noisy and (5) ----- . But you said nothing about how we could make things (6) ----- . It is easy to say that there is too much (7) ----- in the streets. But what do you (8) ----- we should do to make our town a better (9) ----- to live? We know the (10) ----- . Give us your answer.

An angry reader

C)

D)

Last Saturday a friend of mine and I were spending a pleasant afternoon in the garden, when some (1) ----- started making an (2) ----- noise in the street outside. I went to ask them what was (3) ----- on, and they explained that they were digging up the road to (4) ----- the gas pipes. I fully understand that the gas (5) ----- cannot send workmen only at times (6) ----- to me but I feel very strongly that (7) ----- is the wrong day to choose.

Yours faithfully,
(Miss) T. Brendon

- | | | |
|----------------|-------------|---------------|
| 1. a. men | b. workmen | c. hoys |
| 2. a. dreadful | b. awful | c. terrible |
| 3. a. going | b. making | c. happening |
| 4. a. make | b. do | c. repair |
| 5. a. company | b. service | c. workshop |
| 6. a. right | b. suitable | c. convenient |
| 7. a. today | b. Saturday | c. Sunday |

Animals that are unable to control their body temperature are said to be cold-blooded and their temperature rises and falls with changes in the external temperature. In freezing conditions the temperature of

such animals might be as low as 36 F. Warm-blooded animals, on the other hand, are not dependent on external temperature. Their body temperature is constant usually in the region of 100 F. Only birds and mammals are warm-blooded.

A. Find words or phrases in the passage above that can be replaced by the following:

cannot	-----
fluctuates	-----
however	-----
remains the same	-----
about	-----

B. What words in the story mean the opposite of:

warm-blooded	-----
internal	-----
high	-----
fall	-----
constant	-----