Chapter Six

Evaluation

Of all the constituents of English language teaching programme at the tertiary level English teaching / learning, examination systems have drawn severe criticism for the past few decades. Various attempts at examination reforms at tertiary level yielded only partial success. “Inefficient examination system” is almost a cliché now days. Those who are associated with tertiary education are very much concerned with the drawbacks of the examination systems.

Since 1950, because of the advances made in the field of psychology and linguistics, there is a flux of literature on testing and educational assessment. These literatures have tried to illustrate the distinctions among the various terms in the domain of the judgment on the nature and quantum of learning, - evaluation, assessment, testing, examination etc.

Quite often, these terms are used interchangeably, mostly because of the fact that the definitions of these terms are seldom consistent, and differ according to author, or context of publication, and thereby confusing the users. It is also
noteworthy that testing has been found serving different functions. For example, the context of evaluation serves improvement of learning (formative function-), for selection and accountability (summative function), to motivate and gain public support (socio-political function) and to exercise authority (administrative function). Based on the functions and purposes, mainly four types of language tests have been identified-proficiency tests, placement tests, achievement tests and diagnostic tests.

Based on the information obtained, the end of the course examinations of English as a second language can be considered an achievement test. The following are the chief features of the examinations in practice at the tertiary level

1. It has already been mentioned that curriculum/syllabus, which should clearly define the terminal behaviour of the learners in terms of vocabulary, structures and other linguistic skills to be developed during the course, is silent about them. The textbooks, which should be a guide for developing the language proficiency of the learners, as required by the curriculum/ syllabus, turn out helpless in these matters. Those who prepare the question papers at the tertiary level are often provided with a model question paper / old question paper, along with the prescribed text books which provide no guide lines, which will help the paper setters to make devices to measure the proficiency, expected to be developed, on the part of the testees during the course. From these resources, the paper setters are expected to determine the terminal behaviour of the learners. An examiner, who prepares only one question paper, has no occasion to know about the textbooks concerned with that. If the test / examination is to be effective the examiner, just like the class room teacher, the text book writer, and
the curriculum planner, should have a thorough knowledge of all the constituents connected with the course. The paper setter, has to do his work limiting to the narrow domain of the supplied materials, causing a serious set back to the entire system. The examiner, in the modern educational context, is one who has sufficient knowledge of the total educational programme. The existing examination system demands from him, only a fragment of the system.

2. The examination paper setters devise an evaluation scheme which may be easier for every one associated with it, but not at all desirable in accordance with the norms, device as a result of the recent developments in the techniques of evaluation. Most of the questions remain stereotypes and decades old. The system provides no scope for making use of any modern developments in the domain of educational assessment and evaluation.

3. The end of the course examination (now it is annual) constitutes the only measure testifying the work done over an academic year. This single occasion of the expression of the attainment of the learner whose attempt to learn in a span of about a year, decides the achievement of the learner. This will never satisfy the requirements of a good test.

4. It is an undisputed fact that the proficiency of a learner, attained during a period of one year or two years, cannot be gauged by means of a test, the duration of which is three hours, by means of only a few stereotyped questions which the learners can easily guess.
Two or three essays, with which a learner can secure minimum marks required far a pass constitute a major part of the test item. A question as this will cover only a small percentage of the syllabus. Most learners who are prone to resort to guess papers / ready made answers learnt by heart, may get higher marks if fortunately they get the ‘expected questions.’ Really answering questions like these, is supposed to demand from the testees, an integrative view of the details of the text, but in practice those examinees who haven’t read the text, answer the questions as expected by the examiners, defeating the entire purpose of the system.

Essays and annotations, if approached in the true spirit of learning / testing demand from the student a proficiency in the language in terms of vocabulary and structures required to be taught in accordance with the syllabus / curriculum. In addition to this, sufficient knowledge of the content of the text is also a condition for answering this. If properly implemented, these questions are effective tools to evaluate the learner’s proficiency in the language as well as the content of the text. Therefore, as an item of test, essays and annotations have a multifaceted objective.

In a three hour test consisting of three essays, five annotations and a few paragraph questions / comprehension questions, there is an attempt at testing the same skill repeatedly. All these questions test knowledge and linguistic skills all at once. Because of this overlapping, a systematic and thorough coverage of the knowledge dimensions and the skills becomes difficult.
6.5 All these formats test only the content relegating linguistic skill expected to be developed by the course, to the background.

1. The experience of college teachers including the investigator, who evaluate the answer scripts of the tertiary examinations, shows that the marking of the essay type questions is largely subjective, arbitrary and inconsistent depending upon the individual examiner’s notion of what a good essay should be like. There cannot be, and has not been any uniformity in assessing the essay. Each examiner uses a different criterion for marking the answer.

The essay is an open-ended test item. The scoring key cannot be exactly specified. Usually only certain ineffective guidelines are given. Quite often, scoring is dependent on the examiner’s judgment, which is sometimes subjective, hasty, and in certain cases indifferent. Therefore, essay type questions cannot be depended on fully to measure the achievement of a learner, who has learnt English for a longer duration. So, essays and annotations can be only a part of the components of evaluation.

The system of evaluation in practice today has been considered only tests of memory rather than constructive ability. The kind of ability called ‘pure parrot memory’ is coming to be less and less esteemed. The learner who knows the readymade answers by heart unintelligently is not sure of the utility of what he has been taught in the class.
It is also obvious that written answers to a series of specific questions do not provide any basis for a rational judgment of the linguistic proficiency of the learners.

2. The scheme of valuation and instructions to the examiners are vague and inadequate, giving little indication of specific answers, expected of the examinees.

3. The minimum marks required for a pass is 35%, which encourages slip-shod work. This system, which has been prevalent in previous courses, also in effect result in a cumulative language deficit, when the learner is at the tertiary level. Every course, passed by the student, securing only the minimum marks required for a pass, will have on passing the tertiary examination, only a proficiency level, much below the expectation of the course designers. The experience of many teachers shows that the students who make elementary mistakes in English, are allowed to pass the examination.

4. Since the evaluation systems have an inseparable link with, both the educational objectives and learning experiences, the activities associated with evaluation may help clarify objectives and learning experiences. The recent findings have made it clear that a well-designed course and evaluation system may yield guidelines for planning and improved course and evaluation. Thus evaluation in the
broad sense of the term refers to multifaceted activity which serves a variety of purposes.

Researches, in the field of testing and evaluation in the past few decades have yielded a variety of formats in language testing, at various levels of teaching / learning. However, for evaluation purposes at the tertiary level, in the universities of Kerala, only a limited number of traditional formats are still in use.

5. The examinations are expected to determine the efficiency of a course of study, and curriculum, providing sufficient feedback. Nevertheless, the present examinations cannot be considered an integral part of the curriculum and course.

A condition as this makes examination reforms a very difficult task. However, the following efforts will bring about a considerable change.

The defect of the tertiary education is that it is a system in tolerant and rigid. No allowance is made for the idiosyncrasies of the individual student, who is often sacrificed for the average of the class.

1. As teaching / learning aims at the development of linguistic skills and examinations measure these skills, it is inevitable to state clearly the instructional objectives the course, precisely and realistically. Definitive statements of the objectives of the course and the examinations have to be a sure guide to making purposeful efforts to achieve the goals.
A description of the objectives of the course will also define the degree of excellence with respect to various language skills. These descriptions will consist of the advances made by the learners from entry-level behaviour to the terminal level behaviour as indicated as the objective of the course.

Each instructional objective has to be listed based on its priority, and an examination system has to evaluate the skills in accordance with the weight age given each item. The system of marking each item of the test is to be devised accordingly.

A test as this would ensure learning on the part of the learner as they prepare for the examination.

The tests will have to provide opportunities for the learner to understand what he knows and what he does not know.

The feedback the learners get after the test, both from their own reflection and from teachers / evaluations will be instrumental in further learning.

2. While preparing the question papers adequate coverage of the syllabus is to be made inevitable, if the students are to give proper attention to all the parts of the prescribed course. What Jack C. Richards says about this is very relevant. He says: “In recent years, there has been growing, interest in the application of assessment procedures that are radically different from traditional forms of assessment”(335). More authentic methods of assessment which are student centred have been evolved recently. These methods, in addition to being an assessment tool, provide students with a
tool to be more involved in their learning and give them a better sense of control for their own learning. A proper assessment tool provides the teachers with useful information that can form the basis for improving their instructional plans, and practices.

Interest in the use of non-traditional form of assessment in the classroom reflects the changing paradigm in education in general and in second language teaching in particular. The old paradigms are slowly giving way to a new one, as exemplified below.

<table>
<thead>
<tr>
<th>Old Paradigm</th>
<th>New paradigm</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Focus of language</td>
<td>Focus on communication</td>
</tr>
<tr>
<td>2. Teacher centred</td>
<td>Learner centred</td>
</tr>
<tr>
<td>3. Isolated skills</td>
<td>integrated skills</td>
</tr>
<tr>
<td>4. Emphasis on product</td>
<td>Emphasis on process.</td>
</tr>
<tr>
<td>5. One answer, one-way</td>
<td>Open – ended multiple solutions</td>
</tr>
<tr>
<td>6. Tests that test</td>
<td>Tests that also teaches.</td>
</tr>
</tbody>
</table>

How can we assess our students in a way that consistently reflects abilities in the second language? Jack C.Richards further observes:

Although traditional forms of assessment can provide psychometrically valid measures of students’ performance they often fail to provide the kind of information that the typical classroom teachers are interested in, namely, what
the students can do in their second idiom. Because of this, alternative to the traditional forms of assessment has been proposed in recent years. (335-36).

Assessment is a process of finding out who the students are, what their abilities are, what they need to know, and how they perceive the learning will affect them. It enables the students to identify their own strengths and weaknesses, which, in turn will make them know how to go about improving themselves and become more effective users of language. Brian Tomlinson observes:

If the test, either formal or informal, is of the learners explicit knowledge of the language (either declarative or procedural) than it is both fair and useful to announce the test in advance, and to spend time helping the students to revise and extend what they know. (43)

Today ‘testing and evaluation’ as a part of educational assessment is a highly developed and sophisticated field, with numerous facets. One of these facets is the increased emphasis on ongoing assessment of the students’ performance as a course progresses or what has been called formative evaluation. An ideal examination should consist of a series of tests designed systematically to gauge the language skills. The results of such a testing would serve as the basis for an accurate judgment of each individual examiner. About the recent developments in testing Jack C. Richards Observes:

With the advent of techniques for performance based assessment, portfolio development, oral production inventories, co-operative student –techniques,
and other authentic testing rubrics, we are quickly developing the capacity to produce an ongoing programme of assessment throughout student’s course of study. With formative process of assessment in place, the teacher can make appropriate mid-course pedagogical changes to more effectively reach goals. The notion that evaluation must be confined to summative, end of term, or end of unit tests alone is vanishing. However, it is important to note that summative evaluation is also important component of a language programme. The difference between current summative testing philosophy, and the presupposition behind methods – that “one size fits all” – can be seen in a wide variety of assessment batteries that cover both production and comprehension skills, a range of assessment tasks, individualized (including computer adaptive) test, and increased attention to the communicative properties of tests.(17)

Teachers’ long experience in marking written work has made them familiar with the level of written competence people need, in order to succeed at a specific level; but they are less secure when dealing with tests which measure speaking and listening, even though these skills are more important for learners of second language in the contemporary situation. It is true that the second language English learners come from a written culture rather than an oral culture; the correct level of oral skills in English becomes more challenging and important.

Recently a mode of testing known as ‘alternative assessment ‘ has been in use as an alternative to standardized testing and all of the problems found with such
testing. Alternative assessment consists of all those “effort that do not adhere to the traditional criteria of standardization, efficiency, cost effectiveness, objectivity and machine scorability” (Richards 357). It aims at gathering “evidence about how students are approaching processing and completing ‘real-life’ tasks in a particular domain” (Richards 357). Alternative assessment as the name suggests, “provides alternatives to traditional testing in that it (a) does not intrude on regular classroom activities (b) reflects curriculum that is actually being implemented in the classroom (c) provides information on the strengths and weaknesses of the individual student, (d) provides multiple indices that can be used to gauge student progress, and (e) is more multiculturally sensitive and free of non-linguistic and cultural biases found in traditional testing.”(Richards 457)

Various commissions appointed at the national level and state level since 1947 to better the quality of education at the school level, tertiary and post graduate level have insisted on reforming the examination system in the first instance. There has been a unanimous opinion that the examination systems are mainly responsible for deteriorating the standard of education. Many institutions at higher level of learning have implemented a series of reforms to make educational standards cope with the objectives envisaged. These evaluation systems are at once an assessment of the whole educational system as well as the educational achievement of each learner.

At present, the main concern of teachers and students is a meritorious success in the examination, the main index of which is a high percentage of marks. Quite naturally the teachers and students alike subordinate teaching / learning to
examinations. More over the evaluation of what the learners have achieved over a period in the course of their study is made by a single external examination. This has resulted in neglecting learning for most of the duration of the course by many learners. Often, majority of students, towards the end of the course, by the time the examinations commence, resort to learning things on the basis of guess papers or probable questions. This condition must go. Proper devices for internal assessment have to be introduced. Internal assessment must be an integral part of teaching/learning/evaluation. A number of formats have been successfully tried for internal assessment.

Examinations on the one side look at what the examinees can do and at another side at what they need to do. It is to be particularly noted that no test can even be wholly valid or wholly reliable. In this chapter, the investigator outlined and provided examples of some of the main approaches that have been adopted for the evaluation process of the English learners under consideration. It has also been found that developing valid and reliable language tests is a complex process.