

TESTING AND EVALUATION: AN EFFECTIVE PROCESS OF LEARNING LANGUAGE

Eshita Das

M.A. Forth Semester

Department of English Rishi Bankim Chandra College, Naihati

Abstract:

The procedure of testing and evaluation is an indispensable whack of our educational system. The aim of this paper is to display how this process is used to learn and to teach a language. Not only that, this paper also projects the mode of testing and evaluation in Indian context. It makes an expurgated contrast between the testing and the evaluation. Evaluators and teachers are not tantamount, specialisation in profession, student's intuition and admissibility of this process that is also comprised. The qualities of a good and prosperous test and the techniques of testing and the typologies of the test are necessitating by the teaching learning ecology to enhance their inclined and consolatory debenture.

Keywords: testing, evaluation, learners, exam, techniques

Introduction:

Testing and evaluation, these two are extremely very important concepts related to English Language Teaching (ELT) and learning because everyone as a learner as well as a teacher has to face some kind of major experiences regarding the testing and evaluation. Like, a learner, from his/her nursery level to the level of his/her higher studies; he/she has to face some kind of practical observations which are associated with testing and evaluation. There are different kinds of feelings that arise within the learners whenever he/she thinks of giving an exam. Some of those feelings are positive and some of those are negatives. Confidence level, own feedback, faster accumulation under pressure, sense of achievement, post exam relaxation all these are positive feelings that occur in a mind of student or learner and the negative feelings are anxiety, lack of confidence level, parental pressure, the problems of time management, parental pressure etc. So learner's attitude towards exam is very complicated.

Testing

Now what is testing? A teacher always believes that the testing is an integral part of our educational system as well as a device of measurement. Testing gives us (learners) the sense of achievement and our (learners) own feedback. And finally a testing always should be learner-centred or in a learner friendly manner. We never accept an educational system without testing. According to Geetha Nagaraj, "A test is a procedure of measuring ability, knowledge or performance. A test is used not only in an academic course, but also everywhere; it is used in situations where the performance of people or things needs to be measured or discriminated."

Comparison between a thermometer and a testing

If any kind of testing, which is a related term with our examination, is compared to a thermometer there will be some kinds of valid similarities and dissimilarities between a testing and a thermometer. The similarities are: firstly, both are measuring device. A Testing helps to measure the merits and demerits of a learner and a thermometer measures the temperature of our body. The thermometer is also known as a diagnostic device. Secondly, a thermometer has a minimum

and a maximum level. In case of testing the same process is maintained. Thirdly, both are scientific processes. Though in case of testing, it depends on the tester. And finally, testing and thermometer both has a minimum passing level. In case of testing, if the grace mark is crossed the student or the learner will be declared as failed. In the same way, in case of thermometer, if the minimum level is crossed, anyone is considered as sick. So, both have a minimal requirement. And if we focus on the dissimilarities between a testing and a thermometer, we will find these following points. Firstly, thermometer measures accurately but in case of testing, it varies. Irrespective of the user is become subjective, in case of testing. Secondly, testing is an academic process while thermometer measuring is a clinical process. Thirdly, the range or the scope or the capacity of a thermometer is limited, but in case of testing the scope or the range is unlimited. And finally, testing varies contextually, but the thermometer measuring does not differ from context to context.

Evaluation

Next is evaluation. According to Mary Thrope, “Evaluation is the collection of, analysis and interpretation of information about any aspect of a programme of education or training as a part of a recognised process of judging its effectiveness, its efficiency and any other outcomes it may have.” Regarding the above statements, Rowntree comments about evaluation that it is a process which should not be ignored, it does not equal assessment, and evaluation is a planned, systematic and open endeavour.

Difference between testing and evaluation

Although the terms testing and evaluation are interrelated with each other but there is a slight distinction between them. Testing is a term which is related to examine someone’s knowledge of something to determine what he or she knows or has learned. Testing estimates the level of skill or knowledge that has been reached. Evaluation can be regarded as a process of making judgements based on criteria and evidence.

Testing and evaluation, both these terms are interconnected with the profession of an ideal teacher, which is very unique because the teachers are just like the medical practitioners or doctors. As a doctor examines his/her patients, a teacher also deals with his/her students, not with the toys. Learners are just like the patients as they have also the emotions and feelings. Teachers deal with the students with flesh and blood. A good teacher has to be very sensitive to the requirements of the learners. Requirements do not mean the mechanical requirements. An ideal teacher should careful about the emotions and the feelings of the students. So, in any kind of teaching-learning scenario, there should be well maintenances of the ecology of teaching. And the testing and evaluation helps to balance this ecology. In Marxian terms, if testing and evaluation is the “base”, the ecology of teaching is the “superstructure”.

How just are the teachers to evaluate their students?

Whenever a teacher starts to evaluate any answer script that belongs to any particular board or council or university examination, there has a standard norm of minimum pass marks. Suppose, the pass marks is forty. And a student got thirty-nine. Then there is a psychological conflict occurs in the mind of the examiner. And the teacher is being pulled into two different directions. Teacher’s humanitarian self tells to give one more mark but his/her administrative self tells not to give that grace mark. Assume that this paper is the script of a JNU student, then why should the teacher be liberal as an examiner? Liberal mind of a teacher can dilute the whole system. If

the teacher will give one more mark to this fellow, then why should the teacher give one more mark to that fellow who got fifty nine? Then the teacher is questioning himself/herself as an evaluator. The teacher should evaluate any script in a balanced way. The teacher as an evaluator should perform in a subjective way. If the teacher gives the priority to his/her humanitarian self, the whole system will be questionable.

How qualified are the teachers to evaluate their students?

If this question can be asked to the teaching community of India, their answers will not be at the edge of their mouth. In India, the teachers take it for granted that once they are teachers, they consider themselves as evaluators. Do they really have the knowledge of evaluation? Do they have proper training in evaluation? How qualified are they? Maximum teachers are not properly trained to evaluate their students. These kind of untrained teachers evaluate their students from their own perspective, like how they were treated by their teachers. They are biased kinds of teachers. An ideal teacher should be properly trained in the process of testing and evaluation. As early it is discussed and drawn similarities between a teacher and a doctor. Whenever we are sick, we go to a doctor, who suggests for the medicines and diagnostic tests, according to the requirements. For any kind of diagnostic test, we have to go to a diagnostic centre. And there is a pathologist who diagnoses the patient. So we can see how specialisation is maintained in the medical profession. If the teachers are like doctors, they should be trained in their specialisation. But unfortunately, the teachers are not trained as an evaluator in our India.

How effective and realistic are the process of testing and evaluation?

Suppose, in any course of English as a L2, one basic purpose should be communication. Most of the students are unable to communicate. Then, what is the justification of setting questions like write a short note of a particular term or analyse or write a critical comment on a particular thing. Definitely these kinds of questions are needed to judge some skills and sub skills of a student but do this kind of question help the achievement of the basic purpose of communication? Absolutely that is not true. After passing the two years degree course in M.A., the learners have to face lots of interview, group discussions. But our education system never teaches us properly to face all these things. So, it is clear that the process of our testing and evaluation are not realistic and effective.

From all these perspectives, it is clear that there should be some qualities of a good test. Now what are they? A good test should be based on conceptualisation as well as information. Sometimes both are needed in a balanced way. Objective pattern should be included in a question paper and of course that should be in a simple language. Next, the time management: a proper time should be given, recess time is necessary for the long time process examination. Time management is not only a factor for the students but also for the learners. After that, a proper classroom ambiance is required, question pattern should not be lengthy and the questions should cover the whole syllabus. After analysing all these points, they should be arranged in a systematic order. Then we find the qualities of a good test. Any good test should have validity, reliability, practicality and instructional value.

Validity

It means the appropriateness of testing. Geetha Nagaraj said, "A test is valid when it measures what it sets out to measure." There are five types of validity- Face validity, content validity, criterion related validity, predictive validity and constructive validity.

Face validity- It is a very simple one but it ‘is not based on any scientific criterion, but it is important nevertheless.’ Face validity deals with what the teachers and learners think of a test.

Content validity- it deals with what goes into a test. ‘The content of a test should be finalised after deciding what is to be assessed and then specifying the content.’ It clarifies the exact areas which are going to be tested. In other words, content validity discusses that the good questions should be uniformly distributive.

Criterion related validity- It is related to the backwardness of a testing.

Predictive validity- It is related to the forwardness of a testing.

For example, suppose a learner who is in a course of semester system, he/she scores 50% or 60% in his/her end semester exam. So, there can be an expectation from him/her to get around 50% or 60% in the mid semester exam, or vice versa. In criterion related validity, the performance of the examinee should be tuned with his/her previous records; while in predictive validity, the performance of the examinee should be tuned with his/her future records.

Constructive validity- this is true for the language papers. The question paper should be based on good established theory of language that means the testing of Behaviourism or Grammar Translation Method or something else.

Reliability

It means the objectivity of the test. A good test should be reliable when it is only objective. If the testing and evaluation process will be subjective, there will be some kinds of tendency of being biased and prejudiced. Suppose someone measures his/her waistline with an elastic tape, the measurement will not be appropriate. So, reliability means objectivity as well as dependability. Because, testing system should not be elastic. If it will be elastic, the testing system will be unreliable, undependable and subjective. Reliability is mainly two types: Mark-Remark Reliability and Test-Retest Reliability.

Mark-Remark Reliability- It is also known as Score Reliability. In that case, whenever the learners are not satisfied with our results, we place it for a review. Because, the learner or the examinee may not satisfied after marking. If the marks are not satisfied, that means there is a loss of reliability. So, there should have been Mark-Remark Reliability. Whether examiner A or B examines the paper, there should be some kind of clarity and consistency.

Test-Retest Reliability- It is a very critical concept. It is a measure of reliability which is obtained by the administration. The same test twice over a period of time to a group of individuals. The scores from first time to second time can be correlated in order to evaluate the test for stability over time. So, any good test should have reliability. And more importantly, we should give prominence both, validity as well as reliability, though it is difficult to maintain these in a balanced way for an examination.

Practicality

This issue is related to practical issues. A good test is not only concerned with the question paper, not only concerned with the awardees but also concerned with the management issues. The students will have to answer the questions in a proper time, and this proper time will be also maintained by the evaluator to check the script. So, the testing should be practical in terms of time management, in terms of answer the question, in terms of evaluation, in terms of manpower,

in terms of scoring, in terms of processing, in terms of administration and in terms of interpretation of the results.

For example, in a 10th standard board examination, if the oral based examination starts to be valid, then it will need lots of manpower, time and money. Not only that the results will not be published in a particular time, there will be some kinds of serious problems. So this is related to the practicality, it is related to the idealistic issues. Practically, though this kind of examination is possible but it is not visible. So, practicality refers to the acceptability of the test.

Instructional value

In this term, “instruction” means teaching. Instructional value refers to the two way relationship between teaching and testing. Whenever the teacher is teaching, there are two possibilities. One is the question that the teacher will expect from the answer of the test of the learners. In that case the learner depends on the teaching of the teacher. So the teaching will determine the question paper. In other way, there is a possibility that the question paper will determine the teaching. Teaching should determine the value and this is called backwash effect. Backwash effect is two types: beneficial backwash effect and harmful backwash effect.

Beneficial backwash effect In that case, teaching regulates governs and conditions the testing. In any good test there should be beneficial backwash effect. **Harmful backwash effect** is whenever the testing regulates the teaching.

Regarding all these valid, reliable, practical and instructional issues reserved in mind, in terms of English language, there are different types of test as well as different types of testing.

Typology of test

According to performance, there are seven types of test. And they are- Proficiency test, Achievement Test, Aptitude test, Diagnostic test, Qualifying test, Competitive test and the Placement test.

Proficiency test- It is not based on a fixed or particular syllabus. Its main target is whether the learner can apply and handle the real life communicative situation outside of the classroom. Every time the learners have to communicate with new people, with new situation. So, this test justifies whether the learner is proficient in different situations, or not. Because sometimes the learners have to communicate with different people, like vegetable seller, railway ticket checker, bus contractor etc. The communication has to be done not only with different people but also in different ambiance, society, culture. We know that in India English is used as lingua franca but there is cultural and linguistic difference between the English of a British and an American. A British says “truck”, an American says “lorry”, and a British says “biscuit”, “lift” and “taxi” while an American says “wafer”, “elevator”, and “cab”. Suppose, a student wants to take admission in a university of USA and the university wants that the student should qualify the TOEFL exam. In USA, TOEFL exam is mandatory because they want to check whether the learner is able to communicate in their country, or not. TOEFEL test falls under the category of Proficiency test. Because it justifies that the learner is proficient to continue the purpose of communicative skill, or not. Proficiency test helps to develop the eligibility of a learner. It is a non-exhausted test where performances are predictable.

Achievement test- It is based on a particular syllabus but it cannot ensure the learners’ future performances. This kind of test is just a gateway of eligibility to get a particular job or admission. For example, a student has taken admission in a M.A. course and he/she knows that there is a syllabus to follow during his/her course. At the end of the period, the student has to sit in exam.

And finally at the end of the course the student will get a degree. So, this is a kind of Achievement test because there is a sense of achievement but there is no guarantee.

Aptitude test- this kind of test refers to the potentiality of the learner. For example, if anyone wants to be a pilot then he/she should have a good physical dexterity, sensibility as well as a stout physic and obviously good eyesight. In language learning, whenever the learner starts to learn or adopt a new language, he/she should have the potentiality to learn new phonetic sounds, pronunciation etc. In this test, natural inclination of the learner is tested.

Diagnostic test- It tests the shortcomings, negative qualities and positive qualities of a learner. In an educational ambience, it is important for both the teacher and the learner comes to know where they stand. The learner should give more time to work on his/her shortcomings, and the teacher should also think that his/her students may need some remedial classes for the betterment of their learning. So, it can be said that the Diagnostic test is a kind of improvement test.

Qualifying test- in that case, there is no individual distinction. It only makes the distinctions between the two groups of students, one is qualified and another is non-qualified. Suppose in Net exam, there is a cut off mark, if the learner is able to qualify the cut off mark, he/she will qualify the NET exam.

Competitive test- This kind of test compares the merits of the learners. There is a individual distinctions. Because, through comparative study, a judgement is cast, regarding the competitive merits of the learners. There is a scope for an individual. Competitive test publishes the merit list of the learner, according to their brilliance.

Placement test- in the term “placement”, ‘place’ word clarifies that the learners are placed according to their tests. The placement is determined by the judgement of the merits of the learners. Though it is not so much popular in India but placement test is vehemently use by the countries, like Singapore, USA. In case of this test, the performance of the learner is never static. The judgement is done in a very short time. There is a tendency of taking challenge among the learners as they have to improve their level of performance. So, placement of the learners according to their merits is a very much scientific process.

So these are all the different types of test, according to performance; and which are base on an exact process of testing and evaluation.

Techniques of Testing

Now what is techniques of testing and how it is related to the process of testing and evaluation? Whenever the teachers take our (learners) exams, we find that our question paper includes different items of questions, like paragraphs, letters, translations etc. and that is a valid question paper. A good question paper should not only be consisted of any particular item or one type of variety. The question paper which includes different items, that is regarded as a beneficial question paper for the learners. The exam will not be monotonous if there will be multiple items in a question paper. If the student has to answer the same type of question, the student may be very bored. So diversity is needed in a question paper which will be interesting towards the students. Not only that, the student will also know how to answer and write an informative as well as a descriptive type of question. In this way it increases the validity as well as reliability of the test and also the different types of techniques in the process of testing and evaluation is helpful for the skill development of a learner.

Techniques of testing are mainly three types: Objective test, Subjective Test and Pragmatic test.

Objective test- It is non-exhaustive in nature. There is hardly any scope of biased scoring. It tests one skill at a time and also one particular level at a time (that means it may test phonology, grammar or vocabulary in a particular time). Objective test is also known as discrete point test that means the test is conducted discretely one item by another item.

Some major types of objective tests are – True or False with supportive statements, MCQ, Matching, Completion and Rearrangement or Transformation.

Subjective test- IT is non impressionistic in nature. There is a lot of scope of being biased scoring. It tests multiple skills at a time as well as multiple level of language at a time. Subjective test is also known as an integrative test because it tests so many skill of a learner at a time.

Some major types of subjective tests are- Short-answer type, Essay type, Translation and interview.

Pragmatic test- it is practical in terms and but not always scientific in nature. It evaluates the real life communication of a learner. Pragmatic tests are two types. One is cloze test and another is Dictation.

Cloze test- ‘This is a technique used to measure reading comprehension and general language ability.’ This is originated to assess the difficulty level of reading passage for the native speakers. The cloze test is derived from the ‘gestalt theory’. Though it looks like an objective test but that is not true. It cannot create any answer key. Under this type, two types scoring procedure are used; the exact word method and the appropriate or acceptable word method.

Dictation- It should be done in a general communicative space. Though it is unscientific in nature but it increases the sub skills like receptivity and rote learning ability of a learner. Among the four basic skills it passionately develops the listening skill of a learner. This kind of test has a validity as well as reliability.

From the above mentioned features and characteristics, techniques and study of experiments, it becomes clear to us that the process of testing and evaluation are very much effective in our educational system. Not only that, how testing and evaluation are also helpful for the learners as well as teachers to learn and to teach a new language in a justified way.

End notes:

1. L2 – it means the second language as well as the foreign language. And in Indian context, the second language refers to the English language.
2. Native speakers- someone who learns to speak a language as a part of his/her childhood development. Someone’s first language is regarded as his/her native language.
3. Lingua franca- the common language which is shared by the most of the people in the community.
4. TOEFEL- Test of English as a Foreign Language
5. Beneficial backwash effect- it is related to when teaching regulates the testing.
6. Harmful backwash effect- it is related to when testing regulates the teaching.

7. Gestalt theory- this theory of learning was originated in Germany by the theorists like, Graf Christian von, Ehrenfels, Wertheimer, Kohler, Koffka. Gestalt means the pattern or form. The main tenet of this theory is that the whole is greater than the sum of its parts. Learning is more than just invoking mechanical responses from learners.
8. Exact word method- in this case, the exact word used in the original text must be supplied.
9. Appropriate or acceptable word method- in this case; the testee can supply any word which is acceptable or appropriate.

References:

Hudson, R.A. 2015. *Sociolinguistics*. Second Edition, Cambridge University Press, Delhi.
Nagaraj, Geetha. 2010. *English Language Teaching: Approaches, Methods, Techniques*. Second Edition, Orient Black Swan, Delhi.

Lyons, Johns. 2012. *Language and Linguistics: An Introduction*. Cambridge University Press, United Kingdom

Harrison, A. 1983. *A language Testing Handbook*. Macmillan, Oxford

Larsen-Freeman, Daine. 1986. *Techniques and Principles in Language Teaching*. Oxford University Press, Oxford

Ellington, Percival, Race. 1988. *Handbook of Educational Technology*. Third Edition