

MODULE 5 TESTING AND EVALUATION

INTRODUCTION

The fifth Module talks about evaluation and assessment in ESP. Evaluation is a process which begins with determining what information to gather, and which ends with learners and courses. It makes use of quantitative methods (e.g. tests) and qualitative methods (e.g. interview and questionnaire). It can be formative (on-going) or summative (end-of-course).

Unit 1 Evaluation and Testing: Meaning, Effects and Purpose
Unit 2 Types of Evaluation

UNIT 1 EVALUATION AND TESTING: MEANING, EFFECTS AND PURPOSE

CONTENTS

- 1.0 Introduction
- 2.0 Objectives
- 3.0 Main Content
 - 3.1 What is Evaluation/Testing?
 - 3.2 Purpose of Testing
- 4.0 Conclusion
- 5.0 Summary
- 6.0 Tutor-Marked Assignment
- 7.0 References/Further Reading

1.0 INTRODUCTION

As we have tried to establish in the whole course work, ESP is of utilitarian value to the learners who are undergoing a course for a purpose; the course sponsors, who also know why they are sponsoring the course. As much is expected of ESP programme, there must be a way of ascertaining when the learners have learnt and whether the course is effective, achieving its goal.

This unit will examine the terms evaluation/assessment/test (often used interchangeably). You should, however, know that they don't exactly mean the same thing. The benefits of testing will also be examined.

OBJECTIVES

By the end of the unit, you should be able to;

- define and explain the terms evaluation and testing;
- state some of the benefits of testing;
- state some of the effects of testing and assessment; and
- describe communicative testing, stating some of the features.

HOW TO STUDY THE UNIT

- a. Read this unit as diligently as possible.
- b. Find meaning of unfamiliar words in the unit using your dictionary.
- c. As you read, put major points down in a piece of paper or jotter.
- d. Do not go to the next section until you have fully understood the section you are reading now.
- e. Do all the Self-Assessment exercises in the unit as honestly as you can. In some areas where it is not feasible to provide answers to Self-Assessment exercises, go to the relevant sections of the unit to derive the answers.

3.0 MAIN CONTENT

As a distance learner under the National Teacher's Institute, Kaduna, you are subjected to different kind of tests and evaluation. Try to see the relationship between the way you are tested and evaluated in this course and the way this unit said you should be tested and evaluated. Bring out areas of similarity and difference.

3.1 What is Testing/Evaluation/Assessment?

Evaluation is a process beginning with determining what information to gather and ends with bringing about changes in current activities or future ones (Dudley-Evans & St. John, 1998). It applies to both learners and courses. It makes use of quantitative methods (e.g. tests) and qualitative methods (e.g. interview and questionnaire). It can be formative (on-going) or summative (end-of-course). Evaluating students' work, teachers' work or course evaluation are the necessary parts of each learning and teaching process. Evaluation is not only a motivating factor but also shows students' progress or effectiveness in the course, or on the other hand it can disclose possible inadequacies that are not successfully covered. It helps teacher to provide information whether he or she does a good job or not. There exist many types of tests, questionnaires, tasks or the evaluation can be done in form of talk (discussions, interviews). The test is, perhaps, the best way for learner assessment. Teacher finds if the content of course meets learner's

expectation and whether the learner is able to dispose with the new information and employs learnt skills in a particular situation. “This assessment takes on a greater importance in ESP, because ESP is concerned with the ability to perform particular communicative tasks” (Hutchinson and Waters, 1992:144).

Testing could be defined as a pause at reasonable intervals to look back at how well the students are performing. It is a kind of assessment. Other methods of assessing a child’s work are by asking the child some questions or taking a look at certain works performed by the child at his extra time. Testing or assessment is a useful working instrument at the teacher’s tool box.

Language testing is a complex activity tasking the teacher’s ingenuity. It is the duty of the teacher to know the appropriate test materials for different language skills.

The technique for assessing speech work and the type of assessment instrument may not be identical with the technique and type of assessment instrument you will require for assessing reading comprehension. In testing, the teachers should not rely on only the questions set by the textbook writers. You need to draw up your own questions and it must be well framed. You should utilize different questioning techniques and make your instructions as clear and unambiguous as possible.

Hughes (1989:1) refers to the effect of testing on teaching and learning as “backwash”. Backwash is something which can be harmful or beneficial depending on the handling. It is harmful if the preparation dominates all teaching and learning activities and if the “test content and testing techniques are at variance with the objectives of the course”. A test should test the language skill it is intended to test. Multiple choice items should not be used to test writing skill. Situations like this result to “harmful backwash”. Where the design of a test brings about beneficial changes in the syllabus and higher standard in English for students, it is a “beneficial backwash”.

For Davis (1968:5), “the good test is an obedient servant since it follows and apes the teaching”. Hughes (1989:2) disagrees. He sees the relationship between teaching and testing as that of ‘partnership’. Testing is not always a servant to teaching because there are occasions where teaching is good and appropriate and testing is not; and equally, there may be occasions when teaching is poor or inappropriate and testing is able to exert a beneficial influence. Testing should not always follow teaching rather it should be supportive of good teaching and where necessary exert a corrective influence on bad teaching.

There are many effects which testing can have on teaching and learning. It has significant influence on how a teacher works with the learners and also influences how learners learn.

SELF-ASSESSMENT EXERCISE 1

“Testing or assessment is a useful working instrument at the teacher’s tool box”. Discuss.

3.2 Purpose of Testing

Information about people’s language ability is often very useful and sometimes necessary. The type of test determines why it is needed (Proficiency test, placement test, etc.). In the teaching systems test measures the students’ achievement in the second or foreign language. Tests also provide information about the achievement of groups of learners without which it is difficult to see how rational educational decisions can be made.

Language testers should first of all be clear about the purpose of testing in a particular situation. This is because different purposes will usually require different kinds of tests. Hughes (1989, p.7) identifies the different purposes of testing to include the capacity to:

- measure language proficiency, regardless of any language courses that candidates may have followed;
- discover how far students have achieved the objectives of a course of study;
- diagnose students’ strengths and weaknesses, to identify what they have learnt and what they have not learnt; and
- assist placement of students by identifying the stage or part of a teaching programme most appropriate to their ability.

It could be summarized that tests and testing are of great benefit to the child or learner, the teacher and to educational practices.

3.2.1 Benefits to the Learner

Testing is a source of help and encouragement in your work, which may help you to progress in your work/learning. The child sees at a glance where s/he stands and decides on his/her own the amount of effort to put in future.

Secondly, when the assessment is done in a loving, honest manner, the students can assess their works themselves. The teacher can tell the students what to look for in the assessment in advance. This enables them to mark each other’s work, though subject to the teacher’s cross-checking. A child who could assess another’s work could assess himself.

In addition, an assessment in which the teachers visit the student at home gives the child the feeling that the teachers care about his/her progress. To achieve all these, the teacher should be fair and honest in his assessment, he should avoid being an extreme assessor, that is, one who says 'very good' or 'excellent' to one student and to another 'very poor'. The teacher should operate an open door policy and win the confidence of the children. Finally, the teacher must not show preference or dislike for some students, either in attitude or in the way their scripts are marked.

3.2.2 Benefits to the Teacher

Testing is the best way of ensuring maximum effectiveness in the teaching programme. A teacher will be able to adapt the teaching process to the needs of the students.

It enables the teacher to identify the areas of weaknesses, either of an individual or members of the entire class. The teacher does this by writing down all the errors the students can possibly make and ticking against each time the error is committed. It is diagnostic in nature. Testing enables the teacher to determine the pupils' readiness for the learning task that has been set for them. It also helps in determining the sequence of programme tasks to be followed. After assessment, one may have to re-arrange what one intended teaching after realizing through the administered test that the class needs more practice in the previous lesson.

In addition, test helps the teacher to evaluate himself/herself. Mass failure of the students is an indication that the lesson was not well taught. A teacher should always take a critical look at his/her role as a teacher.

Finally, testing helps the teacher to place the child well. Positioning, promotion, admissions into certain schools and jobs are all dependent on test scores and analysis of results.

3.2.3 Benefits to Educational Practices

The Nigerian National Policy on Education advocates continuous assessment and summative examination for certification. Assessment is important for record keeping. The students should be tested severally and the aggregate recorded before the final assessment. It is a device for clarifying objectives. It discovers strengths and weaknesses of pupils and programmes.

SELF-ASSESSMENT EXERCISE 2

- i. What is evaluation?
- ii. Of what benefit is ESP testing?

3.3 Communicative Language Testing

You should note that if ESP involves communicative methodology, the testing should also be communicative in nature. This is measuring students' ability to take part in acts of communication. A communicative language testing is bound to concern itself with 'capacity' (Widdowson, 1983) or 'communicative ability' (Bachman, 1990). It tests communicative performance or students' language ability in one isolated situation or specific context of use. Alderson and Hughes (1981) accept that to follow the communicative paradigm one needs to define what it is that students have to do with language in a specific situation or series of situations and recognize that by specifying performance in this manner, "one might end up describing an impossible variety of situations which one cannot encompass for testing purposes" (p.59). Weir (1990) identifies the following as some of the distinguishing features of communicative tests:

- Test constructions must closely identify those skills and performance conditions that are the most important components of language use in particular context and incorporate them where appropriate. This will indicate the degree to which the test task reflects the attributes of the activity in real life that it is meant to replicate.
- The sample of communicative language ability in our tests should be as representative as possible. Tests should meet the performance conditions of the context as fully as possible.
- Integrative approach to assessment is strongly recommended as against a decontextualised approach. Language devoid of context (linguistic, discoursal and socio-cultural) is meaningless.
- Authenticity of tasks and the genuineness of texts in tests should be pursued. Different tests need to be constructed to match different purposes.
- Test of oral interaction should reflect the interactive nature of normal spoken discourse, conducted under normal time constraint, paying attention to the element of unpredictability in oral interaction.
- In the area of marking, the holistic and qualitative assessment of productive skills and the implications of this for test reliability need to be taken on board. There is this demand for criterion-referenced approach to testing communicative language ability.

- Testers under communicative paradigm have the greater pressure to validate tests because of an expressed desire to make the tests as direct as possible, both in terms of tasks and criteria.
- Communicative testing requires a high degree of explicitness both at the test design stage, where one is concerned with the required result and at the evaluation stage where one is estimating the acquired result. It should have a beneficial backwash effect in encouraging the development of communicative capacity in the classroom (p. 10).

3.4 Authenticity in Language Testing

If you remember that ESP supports the use of authentic materials, you will now also understand that there should also be authenticity in language testing. Bachman (1990), defines authenticity as a quality of the relationship between features of the test and those of the non-test target-use context. There are two approaches on authenticity; the real-life approach and the interactional ability approach. 'Real-life (RL) approach' tries to develop tests that mirror the 'reality' of non-test language use. This approach has been considered as naive because the test setting itself does not exactly resemble its real-life setting, also, "this approach does not distinguish between language ability and the context in which this ability is observed since non-test language performance constitutes the criterion for authenticity and the definition of proficiency" (p.302).

In the second approach, the authenticity of language tests arises from their 'situational' and their 'interactional' authenticity. 'Situational authenticity' refers to the relationship of features of the test method to particular features of the target-use situation. 'Interactional authenticity' mentions the extent to which an examinee's language ability is engaged in the test task. Thus, the emphasis in this model shifts from "attempting to sample actual instances of non-test language use to that of determining what combination of test method facets is likely to promote an appropriate interaction of a particular group of test takers with the testing context" (Bachman, 1990, p. 317).

Assessment can be used to improve instruction and help students take control of their learning (Bostwick & Gakuen, 1995). Accordingly, it is also necessary to briefly examine 'backwash effect' as a concept.

4.0 CONCLUSION

Evaluation or testing is a wrap up of teach-learn activity. The benefits cannot be over-emphasized. Evaluation of the learners reflects not just the learners' performance but to some extent the effectiveness or otherwise of the course too. A successful ESP course is the one that has

enabled particular learners to do particular things with language. Where this is not so, it is an indication of the fact that something is wrong with the course design or that the objectives are ambiguous.

5.0 SUMMARY

It is the role of evaluation to give feedback for the onerous job of teaching and learning. In this unit, we have examined what is meant by evaluation and testing. Because ESP is a communicative language teaching, we have also looked at communicative testing and the features that make it so. The benefits of testing were also highlighted.

6.0 TUTOR-MARKED ASSIGNMENT

What do you understand by communicative testing?

7.0 REFERENCES/FURTHER READING

- Alderson, J. C. & Hughes, A. (eds.) (1981). *Issues in language testing. ELT Documents 111*, London: The British Council.
- Bachman, L. F. (1990). *Fundamental considerations in testing*. London: Oxford University Press.
- Davis, A. (ed.) (1968). *Language testing symposium: A psycholinguistic perspective*. Oxford: Oxford University Press.
- Dudley-Evans, T., & St. John, M. J. (1998). *Developments in English for specific purposes: A multi-disciplinary approach*. Cambridge, Cambridge University Press.
- Hughes, Arthur. (1989). *Testing for language teachers*. UK: Cambridge University Press.
- Hutchinson, T. & Waters, A. (1987). *English for specific purposes: A learning-centred approach*. Cambridge, Cambridge University Press.
- Hutchinson, T., and Waters, A., (1984). How communicative is ESP?, *ELT Journal*, Vol.38/2, p.109.
- Olaofe I. A (2003) *Teaching English in a Second Language Adverse Situations*. Zaria: Applied Linguistics and Language Education Centre.
- Weir, Cyril J. (1990). *Communicative language testing*. UK: Prentice Hall International (UK) Ltd.
- Widdowson, H. G. (1983). *Learning purpose and language use*. Oxford: Oxford University Press.

UNIT 2 TYPES AND FUNCTIONS OF EVALUATION AND TESTING

CONTENTS

- 1.0 Introduction
- 2.0 Objectives
- 2.0 Main Content
 - 3.1 Types of Evaluation
 - 3.1.1 Formative and Summative Evaluation
 - 3.1.2 Product and Process Evaluation
 - 3.2 Functions of Evaluation
 - 3.2.1 Learner Assessment
 - 3.2.2 Course Assessment
 - 3.3 Difference between a Summative Evaluation and Learner Assessment?
 - 3.4 Advantages of Evaluation and Testing
- 4.0 Conclusion
- 5.0 Summary
- 6.0 Tutor Marked Assignment
- 7.0 References/Further Reading

1.0 INTRODUCTION

Now that you have learnt the meaning of evaluation and its great importance to ESP, this unit will examine further the types and functions of evaluation. The aim of ESP is helping learners achieve communicative competence in the target language, that is, satisfying the learners target and learning needs. Evaluation performs many functions in ESP. The learners as well as the course are assessed to find out if the goals and objectives of the course have been achieved. The classification of evaluation in ESP is dependent on the purpose it is meant to achieve. This unit will look into these classifications. See if the types and functions of evaluation and testing described in the unit are reflected by NTI in your Degree Programme.

2.0 OBJECTIVES

By the end of this unit, you should be able to:

- define and explain each of the types of ESP evaluation and when they are used;
- state some of the functions of Evaluation; and
- distinguish between formative and summative evaluation.

HOW TO STUDY THE UNIT

- a. Read this unit as diligently as possible.
- b. Find meaning of unfamiliar words in the unit using your dictionary.
- c. As you read, put major points down in a piece of paper or jotter.
- d. Do not go to the next section until you have fully understood the section you are reading now.
- e. Do all the Self-Assessment exercises in the unit as honestly as you can. In some areas where it is not feasible to provide answers to Self-Assessment exercises, go to the relevant sections of the unit to derive the answers.

3.0 MAIN CONTENT

3.1 Types of Evaluation

Evaluation usually deals with the learners' performance in terms of mastery or non-mastery of language programme objectives, when current performance is measured versus desired performance by means of testing, using both quantitative and qualitative criteria. There exist different approaches to evaluation; it can be product-oriented (summative evaluation) and/or process-oriented (formative evaluation) (Weir & Roberts, 1994; Robinson, 1991). It can be quantitative (based on quantitative criteria) and/or qualitative (based on qualitative criteria). All these types of evaluation are complementary and not mutually exclusive, their interdependence being of great importance to obtain valid findings.

Evaluation may serve two, complementary functions. In one context, the aim is prospective, or *formative* -- to improve, to understand strengths in order to amplify them, or to isolate weaknesses to mend. The other context is retrospective, or *summative* -- to assess concrete achievement, perhaps as part of a process of acknowledgement or giving awards. Here are some ways to think about the distinction further:

3.1.1 Formative Evaluation

Formative evaluation (sometimes referred to as internal) is typically conducted during the development or improvement of a program or product (or person, and so on). It is conducted, often more than once, for in-house staff of the program with the intent to improve. The reports normally remain in-house. Serious formative evaluation may be done by an internal or an external evaluator or preferably, a combination. Of course, many program staff are, in an informal sense, constantly doing formative evaluation. Formative evaluation is conducted to provide program staff evaluative information useful in improving the program.

The purpose of formative evaluation is to validate or ensure that the goals of the instruction are being achieved. It is also to improve the instruction, if necessary, by means of identification and subsequent remediation of problematic aspects.

Formative assessment helps teachers determine the next steps during the learning process as the instruction approaches the summative assessment of student learning. A good analogy for this is the road test that is required to receive a driver's license. What if, before getting your driver's license, you received a grade every time you sat behind the wheel to practise driving? What if your final grade for the driving test was the average of all of the grades you received while practicing? Because of the initial low grades you received during the process of learning to drive, your final grade would not accurately reflect your ability to drive a car. In the beginning of learning to drive, how confident or motivated to learn would you feel? Would any of the grades you received provide you with guidance on what you needed to do next to improve your driving skills? Your final driving test, or summative assessment, would be the accountability measure that establishes whether or not you have the driving skills necessary for a driver's license. It is not a reflection of all the driving practice that leads to it. The same holds true for classroom instruction, learning, and assessment.

Another distinction that underpins formative assessment is student involvement. If students are not involved in the assessment process, formative assessment is not practiced or implemented to its full effectiveness. Students need to be involved both as assessors of their own learning and as resources to other students. There are numerous strategies teachers can implement to engage students. In fact, research shows that the involvement in and ownership of their work increases students' motivation to learn. This does not mean the absence of teacher involvement. To the contrary, teachers are critical in identifying learning goals, setting clear criteria for success, and designing assessment tasks that provide evidence of student learning.

One of the key components of engaging students in the assessment of their own learning is providing them with descriptive feedback as they learn. In fact, research shows descriptive feedback to be the most significant instructional strategy to move students forward in their learning. Descriptive feedback provides students with an understanding of what they are doing well, links to classroom learning, and gives specific input on how to reach the next step in the learning progression. In other words, descriptive feedback is not a grade, a sticker, or "good job!" A significant body of research indicates that such limited feedback does not lead to improved student learning.

There are many classroom instructional strategies that are part of the repertoire of good teaching. When teachers use sound instructional practice for the purpose of gathering information on student learning, they are applying this information in a formative way. In this sense, formative assessment is pedagogy and clearly cannot be separated from the instruction. It is what good teachers do. The distinction lies in what teachers actually do with the information they gather. How is it being used to inform instruction? How is it being shared with and engaging students? It's not teachers just collecting information/data on student learning; it's what they do with the information they collect.

Some of the instructional strategies that can be used formatively include the following:

- **Criteria and goal setting** with students engage them in instruction and the learning process by creating clear expectations. In order to be successful, students need to understand and know the learning target/goal and the criteria for reaching it. Establishing and defining quality work together, asking students to participate in establishing norm behaviours for classroom culture, and determining what should be included in criteria for success are all examples of this strategy. Using student work, classroom tests, or exemplars of what is expected helps students understand where they are, where they need to be, and an effective process for getting there.
- **Observations** go beyond walking around the room to see if students are on task or need clarification. Observations assist teachers in gathering evidence of student learning to inform instructional planning. This evidence can be recorded and used as feedback for students about their learning or as anecdotal data shared with them during conferences.
- **Questioning strategies** should be embedded in lesson/unit planning. Asking better questions allows an opportunity for deeper thinking and provides teachers with significant insight into the degree and depth of understanding. Questions of this nature engage students in classroom dialogue that both uncovers and expands learning. An "exit slip" at the end of a class period to determine students' understanding of the day's lesson or quick checks during instruction such as "thumbs up/down" or "red/green" (stop/go) cards are also examples of questioning strategies that elicit immediate information about student learning. Helping students ask better questions is another aspect of this formative assessment strategy.
- **Self and peer assessment** helps to create a learning community within a classroom. Students who can reflect while engaged in metacognitive thinking are involved in their learning. When students have been involved in criteria and goal setting, self-evaluation is a logical step in the learning process. With peer

evaluation, students see each other as resources for understanding and checking for quality work against previously established criteria.

- **Student record keeping** helps students better understand their own learning as evidenced by their classroom work. This process of students keeping ongoing records of their work not only engages students, it also helps them, beyond a "grade," to see where they started and the progress they are making toward the learning goal.

All of these strategies are integral to the formative assessment process, and they have been suggested by models of effective middle school instruction.

3.1.2 Summative Evaluation

Summative evaluation (sometime referred to as external) is popularly called "assessment", "grading", "marking", or "testing". The purpose, according to Hamp-Lyons and Heasley is "to inform the teacher, the learner and, often, others, as precisely as possible, how far the learner has progressed towards control over the written language." Summative evaluation provides information on the product's efficacy (its ability to do what it was designed to do). For example, did the learners learn what they were supposed to learn after using the instructional module? In a sense, it lets the learner know "how they did," but more importantly, by looking at how the learner's did, it helps you know whether the product teaches what it is supposed to teach.

Summative assessment at the classroom level is an accountability measure that is generally used as part of the grading process. The list is long, but here are some examples of summative assessments:

- State assessments
- Interim assessments
- End-of-unit or chapter tests
- End-of-term or semester exams
- Scores that are used for accountability for schools (AYP) and students (report card grades) (Saddler, 1998).

The key is to think of summative assessment as a means to gauge, at a particular point in time, student learning relative to content standards. Although the information that is gleaned from this type of assessment is important, it can only help in evaluating certain aspects of the learning process. Because they are spread out and occur *after* instruction every few weeks, months, or once a year, summative assessments are tools to help evaluate the effectiveness of programs, school improvement goals, alignment of curriculum, or student placement in specific programs.

Summative assessments happen too far down the learning path to provide information at the classroom level and to make instructional adjustments and interventions *during* the learning process. It takes formative assessment to accomplish this.

The judgment is recorded for consultation by the learner's parents, head teacher, for admission purposes or by other authorities. Once recorded, the learner cannot do anything to improve on that. Summative evaluation is typically quantitative, using numeric scores or letter grades to assess learner achievement, such as 8/10, 12/20, 65/100 or grades such as A-; C+; F.

We can also distinguish between formative and summative evaluation. This is as seen in the table below:

Formative Vs Summative Evaluation

Formative	Summative
Primarily prospective	Primarily retrospective
Analyzes strengths and weaknesses towards improving	Documents achievement
Develops habits	Documents habits
Shape direction of professional development	Show results of such forays
Opportunity to reflect on meaning of past achievements.	Evidence of regular formative evaluation
Feedback	Evidence

Source: CETaL:

<http://sunconference.utep.edu/CETaL/resources/portfolios/form-sum.htm>

Ideally, the two modes are complementary. Also, as noted in the table above, the *process* of formative evaluation may be an important component in summative evaluation. What questions do you hope to answer? You may wish to turn the programme components that you have identified into questions assessing:

- Was the component completed as indicated?
- What were the strengths in implementation?
- What were the barriers or challenges in implementation?
- What were the apparent strengths and weaknesses of each step of the intervention?
- Did the recipient understand the intervention?
- Were resources available to sustain project activities?
- What were staff perceptions?
- What were community perceptions?
- What was the nature of the interaction between staff and clients?

What questions do you hope to answer? You may wish to turn the program components that you have just identified into questions assessing:

- Was the component completed as indicated?
- What were the strengths in implementation?
- What were the barriers or challenges in implementation?
- What were the apparent strengths and weaknesses of each step of the intervention?
- Did the recipient understand the intervention?
- Were resources available to sustain project activities?
- What were staff perceptions?
- What were community perceptions?
- What was the nature of the interaction between staff and clients?

3.1.3 Process Evaluation

Process evaluation addresses how a project was conducted, in terms of consistency with the stated plan of action and the effectiveness of the various activities within the plan.

Why is Process Evaluation Important?

- (1) To determine the extent to which the program is being implemented according to plan.
- (2) To assess and document the degree of fidelity and variability in program implementation, expected or unexpected, planned or unplanned.
- (3) To compare multiple sites with respect to fidelity.
- (4) To provide validity for the relationship between the intervention and the outcomes.
- (5) To provide information on what components of the intervention are responsible for outcomes.
- (6) To understand the relationship between program context (i.e., setting characteristics) and program processes (i.e., levels of implementation).
- (7) To provide managers feedback on the quality of implementation
- (8) To refine delivery components.
- (9) To provide program accountability to sponsors, the public, clients, and funders.
- (10) To improve the quality of the program, as the act of evaluating is an intervention.

3.1.4 Product Evaluation

This is the evaluation of the outcome of the program to decide to accept, amend, or terminate the program, using criteria directly related to the goals and objectives (i.e. put desired student outcomes into question form and survey pre- and post-). Judging training outcome and the costs incurred for a program offering. This also involves relating the outcomes to pre-specified objectives and considering both positive and unintended outcomes.

3.2 Functions of Evaluation

Evaluation can perform two functions: assessment and feedback. The two can provide important input to the content and methods for future work. Hutchinson and Waters (1992:144) stress two prominent levels of evaluation based on assessment and feedback: ‘learner assessment’ and ‘course evaluation’.

3.2.1 Learner Assessment

This is the assessment of students performance at strategic points in the course, e.g., at the beginning or at the end of the course. In ESP, what is tested is the communicative competence of the learners as ESP is concerned with the learners’ abilities to perform communicative tasks. This forms a basis for decisions to be made by sponsors, teachers and learners. A number of tests are available in EAP/EOP for evaluating learner performance. Placement tests, achievement tests and proficiency tests are three basic types of assessment. Although tests are here to stay, there is a pervasive prejudice against testing (Hutchinson & Waters, 1987). As a matter of fact, tests play an important role in the teaching-learning process. Tests provide feedback to inform teachers and learners about what and how they might improve their future work. Therefore, we need to develop a positive attitude to tests (Hutchinson & Waters, 1987). Learner assessment is made up of the following:

(a) Placement Tests

These are tests given to learners at the beginning of a new course. This is to determine what the learners know in the second language. The teacher uses the result to place them into suitable classes or groups. Placement tests are constructed for particular situations. No one placement test will work for every institution. The key features at different levels of teaching in institutions must be identified before the construction of the test. If well constructed, it makes for accurate placement. This kind of test can serve as needs assessment instrument.

(b) Achievement Tests

Achievement tests are directly related to language courses. They are given to learners at the end of the course to determine how successful individual students, group of students or the courses themselves have been able to achieve the objectives. What are the learners able to do at the end of the course which they could not do at the beginning? For example, being able to pronounce the dental sound ‘th’, make simple requests; write good essay introduction, etc.

Hughes (1989, p. 10) divides achievement tests into two:

- Final achievement tests, and
- Progress achievement tests.

Final achievement tests are those administered by Ministries of Education (Junior Secondary Exam in Nigeria); official examining boards like West African Examination Council that organizes Senior Secondary School Certificate exams and G. C. E. Ordinary Level. The test is based on detailed course syllabus or on books and materials used during the course of study. The advantage is that students are tested on what they have studied. Success in the exam indicates successful achievement of the course objectives.

Progress Achievement tests measure the progress the students are making. It is given during a course to see how far their language ability has developed. It enables the teacher to assess himself to see how far he is achieving his objectives, what needs to be re-taught and what to do next. It makes for progression towards the final achievement test based on course objectives. For Hughes (1989), if the syllabus and teaching are appropriate to these objectives, progress tests based on short-term objectives will fit well with what has been taught. If not, there will be pressure to create a better fit. If it is the syllabus that is at fault, it is the tester’s responsibility to make clear that it is *there* that change is needed not in the tests (p. 12).

It should be recalled that evaluation in ESP is an on-going activity.

(c) Proficiency Tests

Proficiency tests are designed to measure people’s ability in a language regardless of training they may have had in the language (Hughes 1989, p. 9). He defines ‘proficient’ as ‘having sufficient command of the language for a particular purpose.’ For example, the new aptitude tests by Nigerian Universities for those in the faculties of Arts and Social sciences is a proficiency test designed to determine whether a student’s English is good enough to follow a course of study in the universities in

those faculties. The content of the test does not follow the content or objectives of the language courses which people taking the test may have studied.

Lindsay and Knight (2006) identify a sub-category of proficiency test known as “external proficiency examinations” (p. 122). They describe it to mean that which “may be produced by the Ministry of Education in a particular country, or by an organization which sets language examination internationally.” Learners from diverse institutions and countries may take the examination. The result from the test may be used for job placement or for admission purposes. Examples are TOFEL (Test of English as a Foreign Language), IELTS (International English Language Testing System) and ESOL offered by Cambridge.

3.2.2 Course Evaluation

The course evaluation helps to assess whether the characteristic features of designing the course were met. It involves all those who share the learning process in making up the ESP course. In course evaluation, factors such as materials, classroom activities, out-of-class support, course design, methodology and assessment should be evaluated. Where everything relevant cannot be evaluated, priorities should be set and the type and timing of data collection should be planned. In assessing a course, questions as to what, how, who, when and how often of your evaluation should be asked.

- What should be evaluated? This involves the teacher’s ability to collect information and use them; ability to satisfy the learners’ needs as language learners and language users. The idea of what should be evaluated, according to Rea (1983), is that “different areas of evaluation are important to different people at different times and for different reasons” (p. 90).
- How can it be evaluated? It could be through tests, questionnaires, discussions, interview, comments, etc.
- Who should evaluate? The ESP teacher, learners, sponsors or course designers? It depends on who mounted the course and why it was mounted.
- When and how often it could be done? This should not be too often because it is time consuming, complex and at times frustrating.

That is, after prioritizing what should be evaluated, techniques such as tests, questionnaire, discussion and interview will be used to collect data. Then, the information is discussed and conclusions drawn.

3.3 Difference between a Summative Evaluation and Learner Assessment?

Although both might look at the same data, a Learner Assessment generally looks at how an individual learner performed on a learning task. It **assesses** a student's learning - hence the name Learner Assessment. For example, you may assess an entire class of students, but you are assessing them individually to see how each performs.

A Summative Evaluation, on the other hand, looks at more than one learner's performance to see how well a group did on a learning task that utilized specific learning materials and methods. By looking at the group, the instructional designer can **evaluate the learning materials and learning process** -- hence the name Summative Evaluation. For example, here you may find that, as a group, all of the students did well on Section A of some instructional materials, but didn't do so well on Section B. That would indicate that the designer should go back and look at the design or delivery of Section B.

3.4 Advantages of Evaluation and Testing

- Evaluation is the process of examining a program or process to determine what's working, what's not, and why.
- Evaluation determines the value of programs and acts as blueprints for judgment and improvement (Rossett & Sheldon, 2001).
- It provides data as an input to possible change.
- As a result of evaluation, programmes may be changed for the better, thus it is used as part of quality control.
- Evaluation is a source of information and experience. From information obtained, the teacher assesses himself, the students and the material.
- It ensures that money invested into the programme is not wasted.
- It highlights real problems and areas of success in classroom teaching.
- Summative evaluation may bring about a systematic programme of in-service teacher training as the programme being evaluated will have already been completed (Robinson, 1991).
- Testing is a tool for needs analysis (Umera-Okeke, 2005, p. 80).

SELF-ASSESSMENT EXERCISE

- i. What is the difference between process and product evaluation?
- ii. What is the difference between formative and summative evaluation?

4.0 CONCLUSION

A final consideration in ESP concerns the role of formative and summative evaluation in developing the programme. Systematic formative and summative tests of the students' achievements on the parts and the whole of the EFL/ESL programme (ESP and GE components) help not only in making decisions concerning the continuation and/or discontinuation of the programme and the placement of students, but they can also "...indicate weaknesses in specific components of the programme as a whole, and therefore provide information for making decisions about revising the programme itself" (Mackay & Palmer (Ed.), 1981).

5.0 SUMMARY

Evaluation has been said to deal with the learners' performance in terms of mastery or non-mastery of language programme objectives, when current performance is measured versus desired performance by means of testing, using both quantitative and qualitative criteria. It can be summative or formative; product or process. Learners can be assessed through tests which include placement, achievement or proficiency tests; each is used depending on the function you would want the evaluation to achieve. Finally, the ESP course we have designed can also be evaluated.

6.0 TUTOR-MARKED ASSIGNMENT

What is learner assessment? Briefly define about three tests you can use to assess a learner.

7.0 REFERENCES/FURTHER READING

- Davis, A. (ed.) (1968). *Language testing symposium: a psycholinguistic perspective*. Oxford: Oxford University Press.
- Dudley-Evans, T. & M. J. St. John (1998). *Development in English for specific purposes: a multi-disciplinary approach* [M]. Cambridge: Cambridge University Press.
- Hamp-Lyons, Liz and Heasley, Ben (1987). *Study Writing*. New York: Cambridge University Press.
- Hedge, T. (2000). *Teaching and learning in the language classroom*. Oxford: Oxford University Press.

- Hughes, Arthur (1989). *Testing for language teachers*. UK: Cambridge University Press.
- Hutchinson, T. & Waters, A. (1987). *English for Specific Purposes* [M]. Cambridge: Cambridge University Press.
- Lindsay, Cora and Paul Knight (2006). *Learning and Teaching English, A Course for Teachers*. Oxford: Oxford University Press.
- Mackay and Palmer (Ed.), (1981). *Languages for specific purposes, program design and evaluation*, NY: Newbury House Publishers, Inc., p.108.
- Robinson, P.C. (1991). *ESP today: a practitioner's guide*. New York: Prentice Hall.
- Rossett, Allison & Sheldon, Kendra (2001). [*Beyond the podium: delivering training and performance to a digital world.*](#) San Francisco: Jossey-Bass/Pfeiffer
- Saddler, D. R. (1998). Formative assessment: revisiting the territory. *Assessment in Education*, 5 (1), 77-84.
- Weir C. and Roberts, J. (1991). *Evaluation in ELT*. Oxford: OUP
- Widdowson, H. G. (1983). *Learning purpose and language use*. Oxford: Oxford University Press.