TYPES OF EVALUATION AND TESTING TECHNIQUES

Oana-Iuliana ŞTEFĂNESCU[*]

Introduction

Before being a pedagogical concept, evaluation has been an essential tool used in the management of companies, in order to reach a correlation between the targeted objectives and the methods used, on the one hand and the results obtained on the other.

In carrying on an evaluation in any area, including that of foreign language teaching, we have to pay full attention to the following four steps:

(1) *the intention,* which specifies the purpose and the modality of evaluation;

(2) *the measurement,* which involves: the gathering of information closely related to what we want to measure; the organization of information in order to study and interpret it; the interpretation (analysis) of information;

(3) *the appreciation (judgement)*, based on the information received and analyzed;

(4) *the decision*, which will take into consideration the purpose of the evaluation.

So, in education evaluation represents a process of obtaining pieces of information on the student, on the teacher, on the program, or on the educational system in its totality, helped by an evaluative instrument, ending with the elaboration of some appreciation on which a series of decisions regarding the content, the methods, the strategies, and so on, will be taken.

Formative and summative evaluation

These are the two basic types of evaluation which differ primarily in terms of aim, but also as regards the time and method of evaluating.

Formative evaluation is oriented towards the improvement of learning conditions at the level of the student, of the group, of the faculty, and so on. Its purpose is that of arrangement of learning activities and the checking of the acquisitions at different moments. It is put into practice at the beginning, during or at the end of a more or less extensive learning sequence. This function is formative, because it represents an instrument of formation, by adapting the learning activity to different types of intelligence. Formative evaluation cannot be defined without the association of the notion of learning objective.

Summative evaluation is the testing of the student at the end of a course, of a cycle or of a program of study in view of classification, evaluation of the academic progress or in order to check the efficiency of the course and teaching methods.

This evaluation refers to terminal objectives and tries to check the knowledge that each student has obtained during the year or up to the moment when the evaluation is carried out. It has a summative function because it gathers all existing knowledge at the end of the year/course in order to check the attainment of all targeted objectives.

When and why do we evaluate our students?

Pedagogic evaluation is a continuous process that may be realized:

a) at the beginning of the training program, in order to check students' already existing knowledge and their ability to start a new sequence of learning: *initial or predictive evaluation*;

b) during the course, meant to evaluate students' progress, to identify the gaps and the reason of their existence and to make all the corrections needed for the adjustment of the teaching /learning process: *continuous or formative evaluation*;

c) after a shorter or longer instruction period, in order to check the punctual result, to check whether the objectives planned for a certain sequence have been fulfilled, in view of continuation of the instruction or of resuming it in order to make the right corrections: *punctual formative evaluation*;

d) at the end of the training program in order to check the level of realization of the objectives of a course or of a terminal part of it and to decide the promotion, or the obtaining of a diploma, a certificate, and so on: *summative or cumulative evaluation.*

---

[*] *Junior Assistant Lecturer, The University of Medicine and Pharmacy, Craiova*

No matter what the function of evaluation is, formative or summative, the action presents the same stages: *intention, measurement, appreciation, decision.* **Measurement** is the stage of the evaluative process that allows the collecting of the quantitative and qualitative information necessary for making the appreciation and for decision-making.

In order to collect the desired information, different procedures are used. Thus, in the field of foreign languages, informal procedures (informal evaluation) such as observation, oral questioning, the analysis of answers as well as other instruments projected by the teacher are a current practice. During formal procedures (formal evaluation) the teacher uses standardized instruments (written tests, tests of oral evaluation) for the established control as well as for the exams.

The **interpretation** of the information (results) obtained through standard procedures or not, is possible either by reference to a norm (*normative evaluation*), or by reference to a criterion (*criterion evaluation*).

a) *Normative evaluation* allows the comparison of the performance of each student to the performances of the group he/she is a part of, by reference to the standard norms. Normative evaluation is done through a test that should contain easy, intermediate and difficult questions, for each category of students. The results are expressed in percentages, on a scale from 0 to 100. From a total of grades a mean of the entire group will be obtained.

b) *Criterial evaluation (interpretation)* consists in assessing each student in relation to the objectives of instruction.

The results of this evaluation may be expressed in percentages (from 0 to 100%), through letters (A,B,C,D,E) or by using a five-scale appreciation (excellent, very good, good, satisfactory, poor).

Evaluation techniques

The techniques for the evaluation of knowledge represent those modalities structured and identified as such, through which the evaluator orients the obtaining of some answers from the students, in conformity with the objectives or specifications of the test.

In order to choose the most appropriate techniques, the evaluation of *communicative competence* should take into account the fact that the competence of communication involves various other subcompetences:

a) *the linguistic or grammatical one*, manifested through the capacity of using linguistic forms (phonology, morphology, syntax and so on);

b) *the socio-linguistic and socio-cultural ones,* which are concerned with the capacity of acting and of using social rules in the physical and socio-cultural context of the situation of communication;

c) *the competence of discourse,* which manifests itself through cohesion (the ability of bringing together phrases, paragraphs, messages, the usage of conjunctions, prepositions and so on) and coherence (the ability to organize ideas in a logical way).

d) *the strategic competence,* manifested by the ability of using various compensatory ways (verbal or non-verbal) in order to maintain communication.

But, as communicative competence is manifested by means of linguistic abilities, their evaluation will give us pieces of information about the students' competence.

The performance criteria used for the evaluation of the competence of communication are related to the sub-competences mentioned above (linguistic, socio-cultural, discursive, strategic). They have been grouped into two categories:

of linguistic nature (linguistic resources)
general linguistic knowledge;
the richness of the vocabulary;
grammatical correctness;
the mastery of vocabulary;
the mastery of the phonological system;
the mastery of orthography;
a) *of pragmatic nature (the usage of language)*
-   cohesion;
-   coherence;
thematic development;
precision.

The first category above comprises criteria of linguistic nature: grammatical correctness; lexical richness and variety, the orthographic correctness and the correct usage of the punctuation signs. The second category refers to the internal structure of the text: *content* (the way in which the work task has been observed), *coherence* (clear, well-structured organization), *cohesion* (the usage of articulators), *style and tone* (that come out from the stylistic quality of the paper).

Whenever we start a linguistic evaluation, we will have to pay careful attention to the capacity's content, *identification,* and *production*. And, in what the content is concerned: the signification, the relationship and all communicative aspects. This will be done

according to the table below:

Capacity → Content ↓	Identification	Production
Signification -the sense of the sentence; -the main idea of the paragraph; -the main idea of the text/ of parts of the text; Relationships -reference; -relations among elements; -structure; Communicative Aspects -(author's) opinions; -(reader's) opinions; -stylistic ways; -situation.		

So, in order to test the students' writing skills, for example, we should test them not simply by means of grammatical or lexical exercises, but by asking them to:
1) identify the intentions, or the topics expressed by a document;
2) answer a letter, an invitation, a proposal, and so on.

In the literature, the techniques of evaluation are grouped into *objective* and *subjective* techniques. If the assessment of productive skills (speaking and writing) is by definition subjective, receptive skills can be marked objectively, without involving the teacher's judgement. Objective testing techniques are used for the evaluation of mental processes such as understanding and application. Items (an item being the smallest noticeable component of the test) can be based on definitions, differences, generalization, cause-effect relationships.

Objective techniques may include: multiple choice; dual choice (True/False); association; ordination; classification. They can be done in a relatively short period of time, as students do not have to write much, yet have enough time to read the different options and to choose the right one(s).

Multiple choice items are formed of an initial part (corpus) that may have the form of a question, of an order or of an incomplete definition followed by several options from which the student chooses:
- only one good answer;
- several correct answers;
- the best answer;
- the only false answer.

With this technique different kinds of learning, excepting oral communication, may be measured. For example, multiple choice items are mostly used to test the students' knowledge of the vocabulary and grammar. However, they can also be used to assess the students' listening or reading comprehension, by asking them to choose the right answers (from among a series of alternatives) to a number of questions referring to the text they have read or listened to.

When designing multiple choice items, some requirements should be met. Thus, the correct and the wrong answers should have the same length, the same abundance of details and the same degree of abstractization. They have to be homogenous, from the point of view of the content and of the grammatical structure. The distractors should be plausible or logical. Each one of them, through its form and structure, should seem to be related to the corpus of the item.

Dual items contain 2 answers from which the student should choose the 'right' one (yes/no; correct/incorrect; true/false). In this case:
- each item should contain only one idea;
- there should be an equal number of items from each type;
- items of each type should be of similar length;
- usage of some words should be avoided (e.g. ever/never; sometimes/generally);
- items should be presented in random order.

Ordinative items propose placing a series of elements or statements given in random order in a certain order, logical or chronological. This technique allows the evaluation of the students' discursive performance, i.e. of coherence (the logical order of ideas) and cohesion (the adequate usage of the linking words).

Finally, an item type that is of great help for the evaluation of our students' progress is *classification*. The item consists in the classification by categories of a number of elements presented in random order.

Conclusion

The assessment of the students' performance in language learning can be done for a variety of purposes, and by using a wide range of techniques. Each of these techniques should be adapted to the students' different levels of knowledge (i.e. beginner, intermediate or advanced), to the skills being tested and, above all, to the purpose of the evaluation. And we need to remember that communicative competence involves, besides a real mastery of language as a code

(lexical-semantic-grammatical), an understanding of the socio-cultural and situational implications of language in use, which require a careful analysis of the lexico-grammatical units used in various discourse types.

BIBLIOGRAPHY

1. BLOOM, B.S., HASTINGS, J.T., MADAUS, G.S. - *Handbook on Formative and Summative Evaluation of Student Learning*. New York, McGraw-Hill, 1971
2. CARROL, B.J., WEST, R, - *Performance Scales for English Language Examinations*. ESU, Framework, London, Longman, 1989
3. ELLIOT, G. - *The Discourse of Medicine*, Nordwood New Jersey, Aslex Publishing Corporation
4. HARRISON, A. **-***A Language Testing Handbook,* ELTS, MacMillan, 1986
5. NEACȘU I., STOICA A. **-** *Ghid general de evaluare și comunicare*, Editura Aramis, 1996
6. NICOLA, I, - *Pedagogie*, București, EDP, 1992
7. NORTH, B., ALDERSON, J.C, - *Language Testing in the 1990's. Modern English Publications*, The British Council, MacMillan, 1991
8. NOVEANU, E., PANA, L.I. - *Didactica limbilor moderne,* Bucuresti, EDP, 1981
9. RADU, I.T. - *Teorie și practică în evaluarea eficienței învățământului*, București, EDP, 1981
10. STANTON, N. - *Comunicarea,* Soc. de Știință și Tehnică S.A., 1995