

EVALUATION AND ASSESSMENT IN ELT

Supriusman

Fakultas Keguruan dan Ilmu Pendidikan
Universitas Riau

Abstract:

Confusion in understanding the terms of 'evaluation' and 'assessment' by the majority of teachers and learners is inevitable. Therefore, a clear explanation needs to be given to overcome the problem, since these two terms are the main duties of every teacher as stated in the teachers and lecturers law No. 14, 2005. It is stated in the law that a teacher is a professional teaching staff whose main duties are to educate, to teach, to guide, to direct, to train, to evaluate, and to assess learners of formal education at the levels of primary education and secondary education, including early childhood education. Besides, assessment is one of the contextual teaching and learning (CTL) components that a teacher does in teaching learning process. Some other terms that need to be described in order to possess good understanding on evaluation and assessment are 'measurement' and 'test'. Some differences of traditional assessment and alternative assessment are also discussed in this paper. In addition, the characteristics of assessment are also introduced to get clear description about this term. It is also mentioned the importance of evaluation carried out in teaching and learning language skills which include productive and receptive skills.

Key words: evaluation, assessment, measurement, test, teaching, and learning.

INTRODUCTION

It is generally accepted that the majority of people especially teachers and learners still get confused to understand and make distinction between the terms of 'evaluation' and 'assessment'. Other important terms such as 'measurement' and 'test' are also described in this article. Therefore, this article tries to describe them in details in order to get full understanding about them.

Literature review

Experts especially educators, and curriculum designers, define the key terms found in this paper such as 'evaluation', 'assessment', 'measurement', 'test', 'teaching', and 'learning' in a number different ways. However, the writer tries to select some definitions as shown in the following.

Evaluation in this paper refers to a process which consists of the subprocesses of measurement and assessment providing adequate and useful data upon which the final judgments are made. Puhl (1997:3) defines evaluation as 'the process of making a judgment of a product, a response, or performance based on criteria.' Evaluation in our schools is essentially concerned with two major approaches to making judgments: *Firstly*, product evaluation which means an evaluation of student performance in a specific learning context. Such an evaluation essentially seeks to determine how well the student has achieved the stated objectives of the learning situation. In this sense the student's performance is seen as a product of the educational experience. A school report is an example of product evaluation. The *second approach* is known as process evaluation which includes the experiences and activities involved in the learning process i.e. making judgments about the process by which students required learning or examining the learning experience before it has been concluded. In most cases, process



evaluation is used when making judgments about school effectiveness, classroom interactions, the curriculum and about the effectiveness of specific programs. For example, process evaluation may be carried out upon the nature of student-teacher interaction, instruction methods, school curricula, and so forth.

The Purposes of Writing of This Article

This article is intended to answer the following questions:

1. What is the difference between evaluation and assessment?
2. What are the purposes of assessment done by language teachers?
3. What is the difference between traditional assessment and alternative assessment?
4. How do language teachers assess speaking and writing skills?
5. How do language teachers assess speaking and writing skills?
6. What should the language teacher do to carry out evaluation and assessment?

From the above questions, the purposes of writing this article are to be able to make distinction between evaluation and assessment, to know the ways and the indicators used to assess the productive skills, and to indicate the correct steps before evaluation and assessment are carried out.

METHODOLOGY

The writer applies library research to collect the information needed. This means that all relevant sources are used to discuss and describe about evaluation, assessment, measurement, and test as mentioned in the following:

DESCRIPTION AND DISCUSSION

The following are described the definition, the purpose, the difference between evaluation and assessment, the difference between assessment and measurement, the difference between teaching and learning, assessing productive skills: Speaking and Writing skills, and other important description about traditional and alternative assessment.

According **Print** (1993:190) Evaluation has several functions: to provide feedback to learners, to determine how well learners have achieved the stated objectives, to provide information to improve curricula, and to clarify the stated objectives by the curriculum developers. From the above the definitions, it can be concluded that evaluation refers to a process of collecting information or data about anything (student, school, teacher and program) to make decision about it based on value judgment.

Assessment as described by **Puhl** (1997:3) is the process of collecting and integrating information about learners from various sources to help us understand these students and describe them. While **Richards, Platt, and Platt** (1992:23), in the one hand, define assessment as the measurement of the ability of a person or the quality or success of a teaching course, etc. Then, assessment according to **Print** (1993:215) involves the interpretation of measurement data. It makes sense of the data collected on student performance. It can be seen from the definitions that to know the information about what students can do or cannot do needs measurement by



using the appropriate instrument such as test, interview, questionnaire, observation, etc. The general purpose of the assessment as described by **Cohen** (1994:23) can be categorized into administrative, instructional, and research functions as shown in the following Table.

Table 1: The Purpose of the Assessment

| General purpose of the assessment | Specific reason for the assessment |
|------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <ul style="list-style-type: none"> Administrative | <ol style="list-style-type: none"> general assessment placement exemption certification promotion |
| <ul style="list-style-type: none"> Instructional | <ol style="list-style-type: none"> diagnosis evidence of progress feedback to respondent evaluation of teaching or curriculum |
| <ul style="list-style-type: none"> Research | <ol style="list-style-type: none"> evaluation experimentation knowledge about language learning and language use |

Measurement, on the other hand is defined as the collection of data or information, usually in quantitative terms, of student performance. It can be also described that to measure means to convert information into numbers or figures. It can be seen from this definition that measurement needs figure or a certain scales such as nominal, ordinal, interval and ratio scales.

The differences between assessment and measurement are as follows: Assessment is *interpretative* which means that assessment gives information about something through measurement, for example, people whose height are 165 centimeters are ideal heights for Indonesians. Assessment is also *qualitative*. This means that the information of the assessment determines the quality of something such as excellent, sweet, tall, etc. The last characteristic of assessment is *subjective*. The information of the assessment is dependent on situation and condition of assessor, for example, a person with 165 centimeters tall is categorized short for western people. Measurement, on the other hand, is *descriptive*, *quantitative*, and *objective*. Descriptive measurement refers to information shows the real fact of the object or person. For example, his brother is 150 centimeters tall. The second characteristic means that measurement provides information in figure or number. Measurement is objective because it presents information as it is without changing the information or data.

A test can be defined as an instrument which is used to determine a learner's ability, knowledge or performance in language skills or language aspects. **Brown** (2004:3) defines a test as 'a method of a person's ability, knowledge, or skill, or performance in a given domain'. In order to know learner's ability, knowledge, and skills, s/he is asked some questions. In other words, a test may refer to a number of questions asked to a learner or student to be responded or answered based on the instruction given.

Teaching in traditional understanding is transferring knowledge, skill, and experience from a teacher to his/her students usually in the classroom. From this definition it can be concluded that when teaching learning process takes place in or outside the classroom there must be a teacher, and students who discuss a topic to teach and learn. Teaching according to **Brown** (1987:7) refers to guiding and facilitating learning, enabling the learner to learn, setting the condition for learning. It is implied from this definition that a teacher has function to guide, facilitate, and enable students to learn in order to get knowledge, skill, and experience.

Learning in general understanding can be stated as the change of behavior because of reinforced practice. Learning can be done by students with or without a teacher. In other words, students can apply his learning style how to acquire or get knowledge of a subject or skill by study, experience, or instruction. This is in line with the definition stated by **Knight et al** (1992:430). They describe learning 'as gaining knowledge of or acquiring skill in (something) by study, practice and teaching.'

Assessing Productive skills

Four macro language skills in English language teaching include listening, speaking, reading and writing. These skills can be categorized into *receptive skills* (listening and reading) and *productive skills* (speaking and writing). Based on the writer's knowledge, experience, and observation, English teachers have difficulty in assessing the second category namely speaking and writing. In other words, English teachers need clear criteria to assess these two skills. To help teacher assess speaking and writing skills, the writer proposes to use criteria as shown in the appendix attached. It is clear in the appendix that there are some aspects to be considered in assessing speaking such as **pronunciation, grammar accuracy, vocabulary, fluency, and interactive communication**. The scoring guide consists of two parts: (1) a description of the aspects of speaking ability to be assessed, and (2) the scale criteria based on the rater makes a judgment.

Then, the criteria used in assessing writing are **content, organization, vocabulary, grammar, and mechanics**. This analytic scoring can help English teachers assess students' writing skill. These criteria are offered by **Cohen** (1994:328-329). He proposes three alternatives in assessing writing skill; they are holistic scoring, analytic scoring, and primary trait scoring. However, the second type scoring seems very easy to apply by English teachers. Therefore, the writer only shows the analytic scoring in assessing writing. The aspects to be assessed in this skill are also in line with **Heaton** (1991:146). He also includes content, organization, vocabulary, language use, and mechanics. The only difference is the term of 'grammar' is substituted with 'language use.'

Some Differences of Traditional and Alternative Assessment

Brown (2004:13) describes some differences of traditional and alternative assessment as shown in the following table.

Table 2: Traditional and Alternative Assessment according to Brown (2004:13)

| Traditional Assessment | Alternative Assessment |
|-------------------------------|--------------------------------------|
| One-shot, standardized exams | Continuous long-term assessment |
| Timed, multiple-choice format | Untimed, free-response format |
| Decontextualized test items | Contextualized communicative tasks |
| Score suffice for feedback | Individualized feedback and washback |
| Norm-referenced scores | Criterion-reference scores |
| Focus on the 'right' answer | Open-ended, creative answers |
| Summative | Formative |
| Oriented to product | Oriented to process |
| Non-interactive performance | Interactive performance |
| Foster extrinsic motivation | Foster intrinsic motivation |

Puhl (1997:5), on the one hand, also describes some differences between traditional and alternative assessment, however, she uses the term of continuous assessment instead of alternative assessment. In addition, she describes the differences based on dimension which includes purpose, focus, feedback, test task, classroom management, frame of references, and stakeholder as can be seen in the following table.

Table 3. Traditional and Continuous Assessment according Puhl

| Dimension | Traditional Assessment | Continuous Assessment |
|-----------------|------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------|
| Purpose of test | Summative; it "sum up" what has been happening Judgmental; forces learner to study | Formative; it generates input to inform and guide teaching Developmental; diagnostic; directs instructional attention |
| Focus | Product of instruction Teacher-created activity Heavy on memorization De-contextualized | Process of instruction Learner-created activity Heavy on thinking, integration holistic |
| Feedback | A score or mark; final, no changing it | A range of comments from peers, teachers; happens the process while still time to change |
| Test task | Typically written work Medium: paper and pencil Narrow focus Exercises (for the future) Formal | Typically a range of tasks Multimedia Multidimensional Authentic (real life tasks for now) informal |

| | | |
|----------------------|---------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------|
| Classroom management | Intrusive; interrupts class process “one-shot;” only one change to show competence Results need time to be determined Books closed | Integrated; part of class routine Overtime; change to revise, improve, add Feedback comes quickly References available |
| Frame of Reference | Norm-referenced Learner compared against norms based on other test takers | Criterion-referenced Learner compared against specified criteria of achievement |
| Stakeholders | Learners, parents, principal receive results | Learners, parents, principal invited to help assess |

CONCLUSIONS AND SUGGESTIONS

Measurement and assessment have different characteristics. Three characteristics of measurement are descriptive, quantitative and objective. Three different characteristics of assessment, on the other hand, are interpretative, qualitative, and objective. Assessment can be done when measurement has been done. As one of the CTL components, assessment is known as authentic assessment which means a process of collecting information or data from various sources during and after teaching learning process. A test in the one hand is one of the assessment types which is used to know the ability of student in language skills. Evaluation in English language teaching can be described as making a decision on something through measurement and assessment. Based on the teachers and lecturers law, evaluation and assessment are the two of the seven main duties of professional educators in educational world.

Teachers and learners must be able to understand the difference between evaluation and assessment by looking at the function, the feature and the purpose of these two terms. The teachers and learners can assess language productive skills which include speaking and writing by taking into account the following components, for example, when assessing learners' speaking skill, notice their pronunciation, structure, vocabulary, fluency, and interactive communication. Then, when assessing students' writing skill, look at their *content, organization, vocabulary, grammar and mechanics*.

REFERENCES

- Brown, HD. 1987. **Principles of Language Learning and Teaching**. New Jersey: Prentice Hall Regents.
- Brown, HD. 2001. *Teaching by Principles: An Interactive Approach to Language Pedagogy*. Second Edition. San Francisco: Addison Wesley Longman, Inc.
- Brown, HD. 2004. *Language Assessment: Principles and Classroom Practices*. San Francisco: Longman.
- Cohen, Andrew D. 1994. *Assessing Language Ability in the Classroom*. Second Edition. Boston: Heinle & Heinle Publishers.

- Departemen Pendidikan Nasional. 2003. Pendekatan Kontekstual (Contextual Teaching and Learning). Jakarta: Diknas Dirjen Dikdasmen Direktorat Pendidikan Lanjutan Pertama.
- Dickins, Pauline Rea and Germaine, Kevin. 1992. *Evaluation*. Oxford: Oxford University Press.
- Ham, Liz and Heasley, Lyons Ben. 1987. *Study Writing: A Course in Written English for Academic and Professional Purposes*. Cambridge: Cambridge University Press.
- Harmer, Jeremy. 2006. *The Practice of English Language Teaching*. Malaysia: Pearson Education Limited.
- Heaton, JB. 1991. *Writing English Language Tests*. London: Longman.
- Hedge, Tricia. 2008. *Teaching and Learning in the Language Classroom*. Oxford: Oxford University Press.
- Knight, Lorna *et al.* 1992. *Collins Concise Dictionary and Thesaurus*. Great Britain: HarperCollins Publishers.
- Mukminatien, Nur. 2000. The Advantages of Using an Analytic Scoring Procedure in Speaking Assessment in *TEFLIN*. Volume XI Number 1 August 2000.
- Nicolls, A and Nicolls, SW. 1983. *Developing a Curriculum: A Practical Guide*. London: George Allen & Unwin Inc.
- Puhl, Carol A. 1997. Develop, Not Judge Continuous Assessment in the ESL Classroom in *English Teaching Forum*. April 1997.
- Print, Murray. 1993. *Curriculum Development and Design*. Australia: Allen & Unwin Pty Ltd.
- Richards, Jack C., Platt, John and Platt, Heidi 1992. *Longman Dictionary of Language Teaching & Applied Linguistics*. Longman: Longman Group UK Limited.

