Assessment in English Language Teaching : *Assessment in Practice*

Muhammad Zaki Pahrul Hadi
Faculty of Culture, Universitas Bumigora, Indonesia
*e-mail: zakipahrul@universitasbumigora.ac.id

Abstract

The terms of test, assessment, examination, and evaluation are sometimes difficult to differentiate. The purpose of this interview is to measure teachers understand the terms of test, assessment, examination, and evaluation. As educators, actively involved in evaluation practices, we would argue that assessment has an important influence on students' learning. In this study, the researcher tried to find the data from the interview with the English teachers. The interview also tried to find out the difficulties faced by teachers in organizing (preparing, administering, marking, and interpreting) the test. The English teachers teaching at SMK Pelayaran Semarang, based on their requests, so their names are in this article are replaced with relevant codes, Mrs. A and Mr.B. From the findings, the researcher can conclude that the teacher had tried to make the test valid and reliable for the students. Based on the principles, the teacher also tried to implement wash-back principles by giving feedback to the students. Moreover, the teachers also need to know the brief distinction between test, assessment, examination, and evaluation.

Keywords : assessment, language teaching, measurement, evaluation

1. INTRODUCTION

It is undeniable that teaching English is ensured not easy. There are a number of difficulties related to this model of teaching, learning and assessment. Besides teaching, there are many things to do which are considered to be very important. They have to teach, test, assess, examine, and evaluate the students' achievement. The purpose of this interview is to measure teachers understand the terms of test, assessment, examination, and evaluation. As educators, actively involved in evaluation practices, we would argue that assessment has an important influence on students' learning. We also propose that the nature of student learning in turn is very closely related to the student's approach to learning. The way in which a student thinks about learning and studying, determines the way in which he tackles assignments and evaluation tasks. Conversely, the learner's experience of evaluation and assessment determines the way in which the student approaches (future) learning. Assessment is thus logically, but also empiri- cally, one of the defining features of students' approaches to learning (see Entwistle & Entwistle, 1991; Marton & Säljö, 1997; Ramsden, 1997. In education terminology, to find out whether the educational process is successful in accordance with educational goals, it is necessary to conduct an evaluation to find out the extent of

its success. Measuring is the most common activity carried out and is an action that initiates evaluation activities in the assessment of learning outcomes. (Hayati et al. 2020)

Many teachers assume that "assessment" simply means giving students tests and assigning them grades. This conception of assessment is not only limited, but also limiting (see section below on Assessment versus grading). It fails to take into account both the utility of assessment and its importance in the teaching and learning process. In the most general sense, *assessment* is the process of making a judgment or measurement of worth of an entity (e.g., person, process, or program). *Educational assessment* involves gathering and evaluating data evolving from planned learning activities or programs. This form of assessment is often referred to as *evaluation* (see section below on Assessment versus Evaluation). *Learner assessment* represents a particular type of educational assessment normally conducted by teachers and designed to serve several related purpose (Brissenden and Slater, n.d.). These purposed include:

(1) Motivating and directing learning

(2) Providing feedback to student on their performance

(3) Providing feedback on instruction and/or the curriculum

(4) Ensuring standards of progression are met

Learner assessment is best conceived as a form of two-way communication in which feedback on the educational process or product is provided to its key stakeholders (McAlpine, 2002). Specifically, learner assessment involves communication to teachers (feedback on teaching); students (feedback on learning); curriculum designers (feedback on curriculum) and administrators (feedback on use of resources). It can be practiced anytime without any schedules and even without any prior information to students. Teacher may give score to how students responding questions, giving comments, or creating structures. Overton (2012) defines assessment as a process of gathering information to monitor progress and make educational decisions if necessary. As noted in my definition of test, an assessment may include a test, but also includes methods such as observations, interviews, behavior monitoring, etc.

According to Brown (2004: 4) is 'an ongoing process that encompasses a much wider domain'. It can be practiced anytime without any schedules and even without any prior information to students. It can be both formal and informal which can be conducted by observing, taking note, test, and etc in order to know the students' skill in a subject.

Test or testing according to Brown (2004: 3) is a subset of assessment used as device to measure the learners' ability, knowledge or performance in a given domain in identifiable times. As Brown defines assessment as device, Calderon and Gonzales (2007: 7) define it as 'measuring

instruments whose general characteristic is that it forces responses from a pupil and such responses are considered to be indicative of the pupil's skill, knowledge, attitude'. Carroll (1968: 46) in Bachman (1992: 20) as cited in Hartoyo (2011: 6) states that test is 'a procedure designed to elicit certain behavior from which one can make inferences about certain characteristics of an individual'. It is concluded that a test is a mean of measuring students' understanding in a subject and done in a certain time. On the other hand, there is examination which is done in the last semester to measure the students' competencies in all subjects. The last term, evaluation is defined as systematic gathering of information for the purpose of making decisions (Weiss, 1972 in Bachman, 1997: 22 as cited in Hartoyo, 2011: 3). Calmorin (2004: 19) also states that 'evaluation measures broad range of goals or objectives of the modern school curricula rather than limited subject matter achievement. The writer then assumes that evaluation is broader than assessment.

On the one hand, test is classified as a set of assessment. It is usually on schedule with prior announcement to students. so test is part of assessment strategy which measures someone's ability, knowledge, or performances in a given domain (Brown, 2004). In most cases, a test will be one part of a fully developed assessment plan. Tests are commonly used in association with cognitive goals in order to review student achievement with respect to a common body of knowledge associated with a discipline or program. Bachman (2004) underlines that language testing practice should cover theoretical view of language ability as consisting skills (listening, speaking, reading, writing) and components (grammar, vocabulary, pronunciation) with an approach to test isolated discrete points in a student. He also concerns on the importance of psychometric validity and reliability. The purpose of this article is to measure the teachers'' understanding in terms of test, assessment, and evaluation. In addition, this study also describes how the teachers of the teachers of SMK Pelayaran Semarang constructs the test items

2. METHOD

In this study, the researcher used descriptive study. Descriptive research means the research that describes the fact. According to Tarigan (2008:105), descriptive research is a research which is describing the phenomenon naturally without manipulation or experiment. This research had focused on activities of the teachers at SMK Pelayaran Semarang. To get the data and information, the researcher gave detail data and information as detail as possible the object that being observed. Data collected in this study were in the forms of statements and description taken during interview, observation and questionnaire. Therefore, the data in this study belong to qualitative data. Qualitative inquiry seeks to understand and interpret human and social behavior as it is lived by participant in a particular social setting . The interview also tried to find out the

difficulties faced by teachers in organizing (preparing, administering, marking, and interpreting) the test. The English teachers teaching at SMK Pelayaran Semarang. Based on their requests, so their names are in this article are replaced with relevant codes, Mrs. A and Mr.B. They graduated from reputable university in Semarang which could be assumed that their knowledge on language assessment and testing theories are still fresh yet on the real practices might have some shortcomings.

3.   FINDING AND DISCUSSION

After having observation and interview, there are some findings that the researcher could explain about the implementation of assessment theory implemented at SMK Pelayaran Semarang. SMK Pelayaran Semarang is a formal educational institution under the Ministry of National Education and the Ministry of Sea Transport.

3.1. First Teacher (Mrs.X)

Firstly, the researcher asked Mrs. X to differentiate test, assessment, evaluation based on her understanding. The researcher gave some questions about the differences between terms of test, assessment, examination, and evaluation and her answers show that she missed some parts of differences between those definitions..

The next question relates to how she constructs her test items. She answered that she does it based on SK/KD which covers 4 skills of language. She determined the ability of students to determine the level of difficulty. She also determined the amount of matter, formulated questions had some try out to determine valid or not, the account after invalid then tested. The third question asks whether she tries out her test items before she gives them to the students. She states that she tries them out to her colleague. She chose this step as a way to reach the reliability and validity because she expects some advice from her colleague whether the test is too difficult or too long for students. Mrs.A admits she always tries to apply principles of testing in constructing the test items for students.

In assessing the test, she prefers using multiple choices, because it is easy to calculate (pass judgment). Multiple choice is also considered to be easy to tested and marked. She admits that there are some difficulties that she faces when she has to organizing the test. It starts from the preparation step when she has to construct or select questions based on SK/KD, administering the items when she thinks that she has to prepare and give remedial test for the students. based on my opinion, Mrs.A misinterprets the administering step here. It is meant to the step which relates to administration of the test which may correspond to the staff of school, or even the preparation of

the copies. She also finds difficulty when she is asked to mark the test. She added that administering and preparing really need a long time to do.

The transcripts of the interview are as follows;

1. Do you think they are similar to each other? If no, what makes them different?

No.

Assessment of teacher is assessments to measure students' level of academic ability, skill, attitude.

Test is an assessment of a student's academic ability scoring

Examination is a test / a media of test

Evaluation is an activity the teacher to evaluate student learning outcomes

2. How do you construct test items?

See SK / KD.

Determining the ability of students to determine the level of difficulty.

Determining the amount of matter.

Formulate questions.

Try out to determine valid or not, the account after invalid then tested

3. Do you try out the test items before giving them to students? Why/ Why not?

Yes. To measure the validity of test.

4. Do you understand the principles of good items and tests?

Yes, I do

5. Do you apply the principles in constructing and organizing the tes? Why?

Yes, Because I think a good and valid questions to be prepared about the preparation of the basic principles.

6. How do you assess or mark and interpret students' work?

To measure per SK ,multiple choice is used, because it is easy to calculate (pass judgment). To measure the craft in following the lesson by giving tasks such as notes and assessed with memberik token reinforcement.

7. What type(s) of test do you tend to construct?

Multiple choice because it is considered to be easy to tested and marked.

8. What difficulties or problems do you encounter in organizing (preparing,   administering, marking, assessing, and interpreting) the test?

Administering and preparing because we have to prepare and determine the validity of a matter and it really takes time.

3.2. The Second Teacher (Mrs.B)

First the researcher asked Mrs.B to define test, assessment, evaluation based on her understanding and the answers show that she missed some parts of differences between three of those as Mrs. A.

The next question relates to how she construct her test items. She answered that she always considers SK/KD to make sure that the test fulfills the needs and he thinks that SK/KD already cover 4 skills of language. Compared with the theory, this step represents his will to make the test valid especially in term of content.

The third question asks whether he tries out her test items before he gives them to the students. She stated that he didn't do that. Before giving them the test, she usually gives them some exercises. Therefore, she thought that it's not a must to give the try out first.

Mrs. B admits she always trying to apply principles of testing in constructing the test items for students but sometimes he is just too busy to make sure that he has already made goo test items.

In assessing the test, she prefers using numerical scores to the students because in her opinion, it is the fastest and simplest way and the score already represents student's ability. According to washback theory, this step is not classified as an activity support the theory.

When asked about the type of test she usually gives to the students, she said that she will Consider the purpose of the test. If she wants to assess the students' basic ability, she will use multiple choice. But if the purpose of the test is to assess / measure the student's ability to analyze a problem, then she will use essay test. So, we still have to look at the purpose of giving the test.

Mrs.B admits that there are some difficulties in organizing the test. Selecting questions or test items based on SK/KD is seen as the most difficult step he has to do in preparing test. She has to select the items carefully so they represent his objectives of teaching. In administering the test, sometimes he finds difficulty when he has to report the test item to the principle and prepare the copies of the paper by himself. The transricpts of the interview are as follows;

1. Do you think they are similar to each other? If no, what makes them different?

> *No.*
>
> *assessment of the student assessment of all aspects, not just the cognitive*
>
> *Test is an instrument to assess students' cognitive abilities*
>
> *Examination was used to test the ability of the student. Value is not the end   result.*
>
> *Evaluation is an assessment at the end of the time KBM (1semester)   overall*

2. How do you construct tes items?

> Based on SK-KD, teaching and learning objectives

3. Do you try out the test items before giving them to students? Why/ Why not?

> No, I dont. Because before giving them the test, I usually give them some exercises.
>
> Therefore, I think it's not a must to give the try out first.

4. Do you understand the principles of good items and tests?

> Yes, I do

5. Do you apply the principles in constructing and organizing the test? Why?

> Yes. Because it is the teacher's task to do that in order to get the maximal   result from

students' assessment.

6. How do you assess or mark and interpret students' work?

Based on SK-KD and from the guideline. For example, when I have to assess students' work on writing item test, I will use writing guideline like coherence, punctuation, grammar, etc. And the generic structure and language features of news item text itself.

7. What type(s) of test do you tend to construct?

First, I will Consider the purpose of the test. If I want to assess students' basic ability, then I will use multiple choice. If I want to assess / measure the student's ability to analyze a problem, then I'll use essay test. So, we still have to look at the purpose of giving the test.

8. What difficulties or problems do you encounter in organizing (preparing, administering, marking, assessing, and interpreting) the test?

Preparing. Because SK-KD should look and level of difficulty to be tested on the ability of heterogeneous students. It is the most difficult.

4. CONCLUSSION

From the findings and discussion, the researcher can conclude that Mrs. A has tried to make her test valid and reliable for her students. Based on the principles, she also tried to implement washback principles by giving feedback to her students. it is better for her to make certain standards since she prefers constructing only multiple choices to assess the students.

> Based on Mrs.B's explanation, the writer suggested her to pay more attention to the activity

of organizing test, especially the need of students to get feedback from teacher. Because sometimes

students expect to get more attention from teacher and feel close to the teacher. she also has to try to construct short answer or essay test if he wants to know the real ability of the students. The researcher gave a special credit to Mrs.B on how she chooses the type of test she usually gives to the students; she said that she will consider the purpose of the test. She still have to look at the purpose of giving the test. From this findings and brief discussion, the writer recommend the teachers to broaden their knowledge on how to construct test items because as teachers, they have to know their students' quality and test is considered as one way to know it. Moreover, the teachers also need to know the brief distinction between test, assessment, examination, and evaluation.

REFERENCES

Bachman, L.F. (2004). Statistical Analyses for Language Assessment. Cambridge: Cambridge University Press.

Brown, H. (2004). Language Assessment: Principles and Classroom Practices. New York: Pearson Education, Inc.

Brissenden, G., & Slater, T. (2007). Assessment primer. In College Level One (CL-1) Team. *Fieldtested learning assessment guide*. Available at http://www.flaguide.org.

McAlpine, M. (2002). Principles of assessment*. Glasgow:* University of Glasgow, *Robert Clark Center for Technological Education. Available at:*

http://www.caacentre.ac.uk/dldocs/Bluepaper1.pdf.

Hayati, U. et al. (2020). Test Technique as a Tool for Evaluation of Learning Outcomes. *Budapest International Research and Critics Institute-Journal (BIRCI-Journal)*. Volume 3, No 2, Page: 1198-1205. DOI: https://doi.org/10.33258/birci.v3i2.961

Setyowati, Y., Heriyawati,D.F., & Kuswahono, D. (2020) The Implementation of "Test of Evaluating" and "Test of Creating" in the Assessment of Learning by EFL Lecturers in Pandemic Era. *IDEAS Journal of Language Teaching and Learning, Linguistics and Literature*. Volume 8, Number 2, December 2020 pp. 578 – 587
DOI: 10.24256/ideas.v8i2.1520

Sitepu, T. (2019). Cooperative Model in Language Reasoning: Indonesian Language Teaching Materials in Teacher Training and Education Faculty of Muhammadiyah Sumatera Utara University. Britain International of Linguistics, Arts and Education (BIoLAE) Journal, 51-59.