

## IS TEACHER-MADE TEST STILL NEEDED FOR UNIVERSITY STUDENTS? TASKS AND CONSIDERATIONS

**Hendrikus Male**

*hendrikus.male@uki.ac.id*  
Universitas Kristen Indonesia

**Erni Murniarti**

*erni.murniarti@uki.ac.id*  
Universitas Kristen Indonesia

### **Abstrak**

*Tulisan ini merupakan salah satu upaya untuk membahas tugas dan pertimbangan dari peran tes yang dibuat guru sebagai salah satu bentuk evaluasi dalam kegiatan belajar mengajar bagi mahasiswa. Hal ini juga diharapkan dapat memberikan kontribusi dan manfaat bagi para guru bahasa Inggris, pembuat tes atau para praktisi untuk mendorong mereka dalam membuat tes yang memenuhi syarat tanpa mengabaikan prosedur standar untuk membuat tes yang baik. Tujuan artikel ini adalah untuk memberikan informasi tentang mengapa tes yang baik diperlukan dan harus diperhatikan oleh pembuat tes mengenai keandalan, keajegan, dan kepraktisannya dalam merancang tes yang baik (lihat Brown, 2004).*

**Kata kunci:** *evaluasi, reliabilitas, validitas, kepraktisan, tes yang baik*

### **ABSTRACT**

This paper is an attempt to discuss the tasks and considerations of the role of teacher-made test as one type of the evaluation in teaching and learning

activities for university students. It is also expected to give contributions and benefits to the English teachers, test makers or practitioners to encourage them in making a qualified test without neglecting the standard procedure of making a good test. The aim of this article is to provide information about why good test is needed and has to be noticed by the test makers regarding its reliability, validity, and practicality of designing a good test (see Brown, 2004).

**Keywords:** evaluation, reliability, validity, practicality, good test

### Introduction

Learning is a lifelong process one should have in his/her daily life. In university level, the students may undergo many obstacles. The obstacles that may appear is usually from two aspects that is academic and non academic. In the context of academic for example the curriculum that has been changed a lot for the past few decades. The national curriculum in Indonesia for university has now changed into a KKNI-based curriculum (see Presidential Decree No. 8 of 2012); KKNI stands for *Kerangka Kualifikasi Nasional Indonesia* or Indonesian Qualification

Framework (IQF)<sup>1</sup>. The IQF curriculum requires a lot of qualifications that the students must possess and acquire in their learning process, so the result may focus more on process rather than products. Learning in the academic sense will create the learners to acquire more knowledge. Therefore, the problem that may appear during the teaching and learning activities is that when there is no evaluation or test in a process of teaching and learning activities.

At the end of teaching and learning activities, the teacher may

---

<sup>1</sup> Indonesian Qualification Framework (IQF) is the author's translation

Male & Murniarti, *Is Teacher-Made Test Still Needed for University students? Tasks and Considerations for Test Makers?*

deal with the evaluation or test. Such activity is done to know the students ability or their competence in the learning itself. Therefore, teachers need to create a good test before addressing it to them. Designing a good test may not an easy task to the test makers or teachers. A number of works should be put into account such as preparing and determining the material and creating a clear instruction.

Zohrabi (2011) stated that

testing is one of the means within the assessment procedure which only measures the students' attainment of course objectives and materials. Testing is more concerned with the mechanical ways of measuring the structural and grammatical knowledge of the students. It reveals nothing about the functional and practical use of language by the students. Testing is only fulfilled through the end-of-semester exams and is carried out via the conventional paper and-pencil means (i.e. written form).

However, a sound teaching and learning process must have an evaluation implemented in every single teaching and learning activity. This way is to see the learners' progress on the subject taking and to find out how far they have achieved the material learned. The teaching and learning process, therefore, has a close relationship to evaluation. Gronlund (1981) pointed out that, "Evaluation is the systematic process of collecting, analyzing and interpreting information to determine the extent to which pupils are achieving instructional objectives."

In short, evaluation means that it gives both much information about how far the instructional objective has been achieved by the learners and gives teachers feedback for more effective in the teaching and learning process. Therefore, teachers should provide good tests to evaluate their students' competence and performance in learning the

language. Unlike content based curriculum which focuses much on the content itself, the IQF based requires the process to reach a better result. In relation to evaluation and test, they cannot be separated from the curriculum.

### **Review of Literature**

This section briefly covers merely the evaluation and test. It does not discuss the measurement and assessment.

#### **A Brief Overview of Evaluation**

In the field of education, doing a careful evaluation is a crucial responsibility of being an educator or teacher. What an educator evaluate may be what an educator value and wish to accomplish. The evaluation will help an educator what to do next to achieve his/her teaching and learning goals. The good evaluation will give meaningful feedbacks if they have both clear procedures and criteria that are

understood by both students and teachers. In the meantime, a good evaluation also requires a constant observation and a large number of sources of information.

In addition, evaluation cannot be separated from the learning process. In the teaching and learning of a language, evaluation is essentially needed. Through evaluation teachers or lecturers can measure the students' progress on their learning and also to see how far their ability or competence. Frey, Alman, Frey and Christinger (2012) defines evaluation as the systematic process of collecting, analyzing, and interpreting information to determine the extent to which pupils are achieving instructional objectives. Nunan (1999), "Evaluation is the collection and interpretation of information about aspects of the curriculum...for decision making purposes" Sudijono (1998) explained that evaluation as attempts to obtain information such as feedback and

Male & Murniarti, *Is Teacher-Made Test Still Needed for University students? Tasks and Considerations for Test Makers?*

the systematic process which determines instructional objective for betterment in education.

To sum up, evaluation is the process of activities which is planned systematically for gathering information about students, teachers and educational programs to get feedback of education progress and to make decision.

### **Purposes for Evaluation**

Indonesian curriculum has the aims for evaluation that is obviously stated in DEPDIKNAS regulation (2003). It said that the aims of evaluations in learning are a) to see the productivity and effectiveness of the learning. b) to improve the teachers' activities. c) to improve and develop teaching and learning program. d) to know the obstacles encountered by the students during the learning activities and to find the solution. e) to place the students in teaching and learning situation based on their ability. Likewise, a number of

experts have also defined a few purposes of evaluation as follows: Weir (1993) distinguished two purposes categories i.e., general and specific purposes. He further states that general evaluation purposes may be undertaken for three principal reasons; 1) accountability. 2) curriculum development and betterment. 3) self-development: teachers and other language teaching professionals. Ahmann and Glock (1968) have defined the purpose of the evaluation as follows: 1) it helps the teacher evaluate the degree to which educational objectives have been achieved; 2), it helps the teachers know his pupils such a degree that educational experiences can be planned according to their varied interest, aptitudes and prior experiences.

To sum up, the aim of the evaluation is to know the students' competence. Apart from its purposes, in general, evaluation is

conducted using two kinds of instruments; they are test and non test. The test is carried out through the form of subjective and objective. Non-test, however is done by using form of observation, questionnaires, and checklist.

### **The Nature of Test**

The test is an instrument to obtain information from one's ability or competence. Norris (2003) affirms that, "Language tests are simply instruments or procedures for gathering particular kinds of information, typically information having to do with students' language abilities." Loannou and Pavlou (2003) have stated that, "testing is one of the procedures that can be used to assess child's performance. A test has a certain objective, for example, to see to what extent a child understand a written text."

Genesse and Upshur (1996, p. 141) define test as follows:

A test is, first of all, about something. That is, it is about intelligence, or European history, or second language proficiency. In educational term, test has subject matter or content. Second, a test is a task that elicits observable behavior from the test taker. The test may consist of only one task, such as writing composition, or a set of task, such as in a lengthy multiple – choice examination in which each question can be thought of a separate task. Different test tasks represent different methods of eliciting performance so that, taken together, test are not single methods of collecting information. Third, test yield scores that represent attributes of characteristics of individuals.

Brown (2004) states," A test, in plain words, is method of measuring a person's ability or knowledge in a given domain." He further states," a test is a method meaning that it is set of procedures that constitute an instrument of some sort that requires activity on the part of the test-taker" (p.384).

Male & Murniarti, *Is Teacher-Made Test Still Needed for University students? Tasks and Considerations for Test Makers?*

Other scholars like Hopkin and Antes (1990) have stated similar ideas regarding the test, “test is one of evaluation instrument that consists of series of task about particular point that students have to answer. The test result is used to measure students’ ability.” Hopkin and Antes (1990, p. 130) further stated, “Test is an instrument, device, or procedure that proposes a sequence of tasks to which a student is to respond- the result of which are used as measures of a specified trait”. It is believed that a good test fulfills three characteristics; they are validity, reliability, and practicality. This is in line with Weir (1993, p. 19) stating “a number of general principles that should underlie all good test design on test need to be valid, reliable and practical”. Other scholars like Griffin and Nix (1989) also described “test is any assessment that is conducted in a formal setting with specified procedures and provide comparability of results.”

In other words, it means that a test is a method consists of a set of procedures such as arranging the instruments and activities on the test – taker. The purpose of the test is measuring the learners’ ability or knowledge. As Hopkin and Antes (1990, p. 130) state “Test is an instrument, device, or procedure that proposes a sequence of tasks to which a student is to respond- the result of which are used as measures of a specified trait”

### **Types of Tests**

Teachers or test makers should understand types of test and know what kind of test they are to design before giving it to their students. There are a number of different types of tests classified by the scholars. (see Brown, 2004; Heaton, 1986)

1. Placement test

Placement test aims to place new students to be posted in the right class or level of language proficiency

based on the result of their language skills. The test also indicates how good the students are in their English or their prior knowledge of English.

2. Diagnostic test

This type of a test is to know the students' weaknesses or strength. It is also used to discover the students' problems in a course.

3. Progress/Achievement test

This type of a test is designed to measure the students' language skills and to see their progress in relation to their syllabus. The test is normally done during the course.

4. Final progress/achievement test

Final progress test is done in the end of the course. This way is to measure the students' progress of their achievement of the course.

5. Proficiency test

The test is designed based on a certain courses for example a course for getting a new job, admitting a foreign university. It aims to measure the students' knowledge and ability in language.

6. Aptitude test

The aptitude test is design to discover the students' basic talent or ability for example a test for learning a new language.

It is hoped that by understanding the types of tests, the teachers can create their own test and select what appropriate test they can provide for their students.

**Characteristics of a Good Test**

In making a test, test makers are hoped to understand how to create a good test. Moreover, they have to understand that a good test should have the following criteria i.e., reliability, validity, and practicality (see



Male & Murniarti, *Is Teacher-Made Test Still Needed for University students? Tasks and Considerations for Test Makers?*

Brown, 2004). Why it is called reliable test is due to its score stability. Gronlund and Linn (1990) “reliability refers to the consistency of measurement—that is, how consistent test scores or other evaluation results are from one measurement to other.” Therefore, no matter how many times the test is conducted, the test gives consistent result. It is considered that the test is reliable. Meanwhile if the test does not give consistent result, it may be less reliable. Thorndike, et al (1991) affirmed that there are three techniques to know whether the test is categorized as reliable or not. First, the test is reliable when it is repeating the same test or it is named as test-retest method. Second, administering a second equivalent form of the test which is known as parallel test form, and the third, subdividing the test into two or more equivalent fractions or it is more known as a single-administration method.

Next, the test is considered valid when it measures what it intends to measure. There are also a number of different types of validity (see Fulcher & Davidson, 2007; Brown, 2004).

1. Operational validity- it deals with the sufficiency of evaluation of definite activities or qualities.
2. Predictive validity-it has predictive validity if the scores can be predicted for future performance.
3. Content validity- it refers to the representative of sample items in the whole course content.
4. Construct validity- it deals with the test scores psychologically.

In addition, concerning its practicality, the test is categorized as practical if it is easy to administer, score, and be economical. This is in line with Bachman and Palmer (1996) state practicality means the relationship between the sources such as human resources, material

resources, time and many others that is required to develop, design, and use of the test. As mentioned earlier that a good test has to fulfill three characteristics; they are validity, reliability, and practicality. It is believed that teaching without evaluation is not a complete activity. Teaching has to deal with evaluation. Test, therefore, is one of the evaluation instruments which are used to collect information and to measure something. It has a close relationship to one another. Gronlund (1981, p. 5) has defined that "Evaluation is the systematic process of collecting, analyzing and interpreting information to determine the extent to which pupils are achieving instructional objectives." Weir (1993, p. 19) pointed out that "A number of general principles that should underlie all good test design on test need to be valid, reliable and practical". Therefore, teachers should consider the role of evaluation and test in their teaching and learning activities so that their

students' learning outcome may be well improved.

In brief, a good test can be considered valid when it can measure what is supposed to measure and it is reliable when it can consistently give scores every time it is conducted for many times, and it is practical when the test can be easy to administer. All of these characteristics must not be separated one another in order to have best result in the learners' level of their learning outcomes.

### **Discussion**

In the teaching and learning the language in university context, the teachers or test makers are required to be able to design good tests in order to reach his/her students' achievement or target in their learning activities process. Therefore, as part of teachers or test makers' tasks to create good tests, they should understand the importance and objectives as well as its purposes of tests or

Male & Murniarti, *Is Teacher-Made Test Still Needed for University students? Tasks and Considerations for Test Makers?*

evaluation and all the requirements of how to make good tests before creating and addressing them to the test takers. A simple question is then raised, why teacher made-test is still needed in university context? One of the reasons that can be taken into account is that the teacher-made test can help the students realize that their competence or ability in learning the language is much more developing or have the progress on educational system.

Another reason is that the test can be source of input or feedback to the teachers. This way is to see the students' learning outcome. For instance, by taking the test, the students result may not have the same result one another and the bad result can even become input for the teachers as the test makers to be more careful in designing the test next time. In other words, it means that test can give not only a lot of information about how far the instructional objective has been

achieved by students but also give teachers feedback for more effective in the teaching learning process let alone their readiness and preparation to make a better tested lessons.

Apart from the aforementioned description, the tests that are to be designed by the test makers in this case the teachers might have lost of weaknesses compared to the standardized test for example the teacher-made test might have low validity and reliability. It also has specific objectives and applies merely for a certain class or group and the items of the test are rarely tested before it is used to be part of the test itself. Unlike the standardized tests that have high validity and reliability because they are often tested. Regarding its objectives, the standardized tests have general objectives and it is designed by the experts.

Concerning the test for university students, the test should be made based on the university

curriculum criteria and its national framework qualification that has been standardized by the government throughout the nation. The teachers or test makers must not neglect the requirements of how to make a good test. Additionally, the test makers should also pay attention to the process of how to achieve better result since the IQF curriculum based (*Kurikulum berbasis pada KKNl*) focuses more on the process rather than the products. It does not mean that the products are not important, but it is more on the process to reach better products or results. So, the teachers' jobs or tasks in this stage are to facilitate their students with giving them a comprehensive and good test even though the test may have its effect-what is so called backwash/washback both good or bad (see, Hughes, 2013).

In summary, the quality of the test itself depends greatly on teachers or test makers' capabilities in designing the

questions or items of tests as well as the teacher's knowledge. In the case that teacher or test makers' lacking of knowledge of designing an appropriate evaluation or test, the test is likely to be inaccuracy in measuring the students' comprehension regarding to the topic given or subject matter. Through considering such points and combining with their knowledge, it is hoped that the tests have been fulfilled all the requirements so that one the problems faced by the students might be solved.

### **Conclusion**

It is suggested that teachers as the test makers consider some points in making the test. The test must fulfill the criteria or the requirements of good test. In doing so, they can determine a valid, reliable and practical test for students. Moreover, good quality of a test depends upon the teacher's capability and knowledge. The test makers are also hoped to provide

Male & Murniarti, *Is Teacher-Made Test Still Needed for University students? Tasks and Considerations for Test Makers?*

good tests by implementing all the requirements for making a good test. Lastly, based on the writer's observations during the teaching process, the teacher-made tests are still needed for the students to at least exercise or practice their knowledge and comprehension on the topics or material given. This way is also to help them improve their proficiency on the subject matter higher and much better result of learning outcome as well as getting feedback for a better and educational progress.

### References

- Ahmann, J. S., & Glock, M. D. (1968). *Evaluating Pupil Growth*. Boston: Allyn and Bacon, INC.
- Alman, S. W., Frey, B. A., & Christinger. (2012). Social and Cognitive Presence as Factors in Learning and Student Retention: An Investigation of the Cohort Model in an iSchool Setting. *Journal of Education for Library and Information*, 53, 290-302.
- Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice: Designing and developing useful language tests*. New York: Oxford University Press.
- Brown, D. H. (2004). *Language Assessment Principles and Classroom Practices*. New York: Pearson Education.
- Fulcher, G., & Davidson, F. (2007). *Language testing and assessment: An advanced resource book*. New York: Routledge.
- Genesse, F., & Upshur, J. A. (1996). *Classroom-Based Evaluation In Second Language Education*. New York: Cambridge University Press.
- Griffin, P., & Nix, P. (1989). *Educational assessment and reporting*. California: Harcourt Brace Jovanovich Inc.
- Grondlund, E. N. (1981). *Measurement and evaluation in teaching*. New York: Mac Millan Publishing co. Inc.
- Gronlund, E. N. (1981). *Measurement and*

- evaluation in teaching*. New York: Mac Millan Publishing co. Inc.
- Gronlund, N. E., & Linn, R. (1990). *Measurement and evaluation in teaching* (6th ed.). Singapore: Mac Millan.
- Heaton, J. (1986). *Writing English Language Tests*. London : Longman.
- Hopkins, C. D., & Antes., R. L. (1990). *Classroom measurement and evaluation*. Itasca, Illinois: Peacock Publisher, INC.
- Hughes, A. (2013). *Testing for language teachers*. Cambridge: Cambridge University Press.
- Kemendikbud. (2003). UNDANG-UNDANG SISTEM PENDIDIKAN NASIONAL.
- Loannou, S., & Pavlou, P. (2003). *Assessing young Learners*. Oxford: Oxford University Press.
- Norris, J. (2003). Purposeful language assessment: Selecting the right alternative test. *English Teaching Forum Online*.
- Nunan, D. (1999). *Second language teaching and learning* . Cambridge: Cambridge University Press.
- Sudijono, A. (1998). *Pengantar evaluasi pendidikan*. Jakarta: Raja Grafindo Persada.
- Thorndike, et al. (1991). *Measurement and evaluation in psychology and education*. Singapore: Mac Millan.
- Weir, C. (1993). *Understanding & developing language test*. Englewood Cliffs: Prentice Hall International Ltd.
- Zohrabi, M. (2011). An Investigation of curriculum elements for the enhancement of the teaching-learning process. *Higher Education Studies*, 67-78.