Celt

Celt

A Journal of Culture, English Language Teaching and Literature

ORAL TEST: A POWERFUL TOOL FOR ASSESSING STUDENTS' ACTUAL ACHIEVEMENT IN LANGUAGE LEARNING

Muhamad Ahsanu[1]

Abstract: *Teaching and testing are inseparable elements in pedagogical world irrespective of the course a teacher teaches. Phrased differently, there is no teaching without testing and vice versa. The results of testing should ideally motivate students in learning and give better perspectives to teachers on how to devise a better teaching-learning. Accordingly, a teacher needs a sort of test that can sufficiently assess students' actual achievement in learning, in their given courses. One of which is so-called "Oral Test", the test that can give a feel of confidence that the test really measures what is purported to measure and provide relatively consistent results over the time (validity and reliability respectively), which, in the end can opaquely discriminate the proficiency levels amongst the students. Thus, this paper is a humble attempt to juxtapose teaching and testing and to run a critical diagnosis on the fruitfulness of oral test, the test type worth trying.*

Key words: *oral test, assessment, learning achievement*

---

[1] **Muhamad Asanu, S.Pd., M.Sc.** <ahsanu.muham mad@yahoo.com> +62281625152 is a leturer at the Study Program of English Language and Letters, Department of Hunmanities, Faculty of Social and Political Sciences, of Jenderal Soedirman University, Purwokerto.

INTRODUCTION

There could be no science as we know it without measurement. Testing, including all forms of language testing, is one form of measurement. In language testing, there are so many types of test and one of which is *achievement test.* The achievement tests (Henning, 1987: 6) are used to measure the extent of learning in a prescribed content domain, often in accordance with explicitly stated objectives of learning program. In other words, it is to provide information about the effectiveness of program of instruction.

According to Heaton (1989: 5), both testing and teaching are so closely interrelated that it is virtually impossible to work in either field without being constantly concerned with the other. He further asserts that test may be constructed primarily as devices to reinforce learning and to motivate the student or primarily as a means of assessing the student's performance in the language.

In general, a language test seeks to find out what candidates can do with language and provides a focus for purposeful, everyday communication activities. A good communicative test of language should have a much more positive effect on learning and teaching and should generally result in improved learning habits. Not the least the fact that testing student retention of fundamental/powerful concepts is a challenge for any discipline, but this can be especially difficult for a certain course like literature, linguistics, and humanities courses.

The foregoing explanation gives partly an answer to the question: why test? Besides being used as an evaluation for the purpose of selection or screening, Heaton also mentions that the classroom test is concerned with evaluation for the purpose of enabling teachers to increase their own effectiveness by

making adjustments in their teaching to enable certain groups of students or individuals in the class to benefit more. In addition, a good classroom test will also help to locate the precise areas of difficulty encountered by the class or by the individual student. Generally speaking, a reliable method of obtaining measurements of oral production skills is that which involves the students' class teacher. In the words of Hughes (2008: 134), to have an accurate measurement of oral ability is not easy. It takes considerable time and effort, including training, to obtain valid and reliable results. Nevertheless, whereas a test is high stakes, the investment of such time and effort may be considered necessary.

THEORETICAL FRAMEWORK

Probably it entails truth that, as assumed by Hingle and Linington (2002: 354),   many language teachers have been comfortable setting pencil-and-paper tests. Years of experience marking written work have made them familiar with the level of written competence pupils need in order to succeed in a specific standard. Conversely, teachers often feel much less secure when coping with tests which measures speaking. Speaking test is perceived to be appropriate with Indonesian learners as they are considered to come from an oral rather than a written culture, and so are likely to be more proficient in this mode of communication.

The query is: How does one set a test which does not intimidate learners but encourage them to provide an accurate picture of their oral ability? According to Madsen in Hingle and Linington (ibid.), "the testing of speaking is widely regarded as the most challenging of all language tests to prepare, administer and score". The theorists suggest three

reasons why this type of test is so different from more conventional types of tests. First, the nature of speaking skill itself is difficult to define. Because of this, it is not easy to establish criteria to evaluate a speaking test. For example, is "fluency" more important than "accuracy"? Second, a set of difficulties emerges if one tries to treat an oral test like any other more conventional ones. In oral test the people involved are important, not the test, and what goes on between tester and testee may have   an existence independent of the test instrument and still remain a valid response.

A. Teaching, Learning and Testing

Teaching sets up the practice games of language learning: the opportunities for learners to listen, think, take risks, set goals, and process feedback from the teacher and then recycle through the skills that they are trying to master. What about testing and teaching? Like teaching and learning, both testing and teaching are so closely interrelated that it is virtually impossible to work in either field without being constantly concerned with the other (Heaton 1989: 5).

As a rational follow-up of teaching and learning, test may be constructed primarily as devices to reinforce learning and to motivate the students or primarily as a means of assessing the students' performance in the language (ibid.). In this respect, the test that becomes the issue is the classroom test which is concerned with evaluation for the purpose of enabling teachers to increase their own effectiveness by making adjustments in their teaching to enable certain groups of students or individuals in the class to benefit more (ibid. 1989: 6).

B. Assessment

We might think that testing and assessing are synonymous terms, but they are not. Tests are prepared administrative procedures that occur at identifiable times in a curriculum. Assessment, on the other hand, is an ongoing process that encompasses a much wider domain. Whenever a student responds to a question, offers a comment, or tries out a new word or structure, the teacher subconsciously makes an assessment of the students' performance.

In assessment, there are two mostly known types, namely formative and summative assessment. According to Hughes (2008), assessment is formative when teachers use it to check on the progress of their students, to see how far they have mastered what they should have learned, and then use this information to modify their future teaching plans. Summative assessment, on the other hand, is used at the end of the term, semester, or year in order to measure what has been achieved both by groups and individuals.

C. Achievement Test

Tests are a subset of assessment that a teacher can make. Brown (2001) defines test as a method of measuring a person's ability, knowledge, or performance in a given domain. Hence, a test measures performance, but the results imply the test-takers' ability, or competence. It is common to find tests designed to tap into a test-taker's knowledge about language. Thus, a test is a method that measures performance and competence in a given domain. A well-constructed test is an instrument that provides an accurate measure of the test-taker's ability within a particular domain. There is a number of test-types and one of

which, dealing with the issue presented in this paper, is classroom achievement test.

In line with this, Brown (2005) asserts that all language teachers are in the business of fostering achievement in the form of language learning. And the purpose of most language programs is to maximize the possibilities for students to achieve a high degree of language learning. This fact will lead language teachers to make achievement decisions. Achievement decisions are decisions about the amount of learning that students have accomplished. Such tests are typically administered at the end of the term, and such decisions make take the form of deciding which students will be advanced to the next level of study, determining which students should graduate, or simply for grading the students (cf. Brown 2004; Hughes 2003).

Thus, achievement tests should be designed with very specific reference to a particular course. This means that the achievement tests will be directly based on course objectives and will therefore be criterion-referenced. A good achievement test can tell teachers a great deal about their students' achievement and about the adequacy of the course. Achievement test can be executed in many different ways and one of which is via oral test or oral presentation.

D. Oral Test

There are five basic types of speaking/oral test. They include imitative, intensive, responsive, interactive and extensive. Brown (2004) laments that extensive oral production task includes speeches, oral presentations, and story-telling, during which the opportunity for oral interaction from listeners is either highly limited (perhaps to nonverbal responses) or

ruled out altogether (Brown ibid.) further affirms that in the academic and professional arenas it would not be uncommon to be called on to present a report, a paper, a marketing plan, a sales idea, a design of a new product, or a method.

A summary of oral assessment techniques would therefore be incomplete without some consideration of extensive speaking tasks. Once again the rules for effective assessment must be invoked: (a) specify the criterion, (b) set appropriate tasks, (c) elicit optimal output, and (d) establish practical, reliable scoring procedures. And once again scoring is the key assessment challenge.

For oral presentations, a checklist or grid is a common means of scoring or evaluation. Holistic scores are tempting to use for their apparent practicality, but they may obscure the variability of performance across several subcategories, especially the two major components of content and delivery. Following is an example of a checklist for a prepared oral presentation at the intermediate or advanced level of English (ibid.).

E. Oral presentation checklist

Evaluation of oral presentation
Assign a number to teach box according to your assessment of the various aspects of the speaker's presentation and performance.
4   Excellent
3   Good
2   Fair
1   Poor

Content:

- ☐ The purpose or the objective of the presentation was accomplished.
- ☐ The introduction was lively and got my attention
- ☐ The main idea or point was clearly stated toward the beginning.
- ☐ The supporting points were
  - Clearly expressed
  - Supported well by facts, argument
- ☐ The conclusion restated the main idea or purpose.

Delivery:

- ☐ The speaker used gestures and body language well.
- ☐ The speaker maintained eye contact with the audience.
- ☐ The speaker's language was natural and fluent.
- ☐ The speaker's volume of speech was appropriate.
- ☐ The speaker's rate of speech was appropriate.
- ☐ The speaker's pronunciation was clear and comprehensible.
- ☐ The speaker's grammar was correct and didn't prevent understanding.
- ☐ The speaker used visual aids, handouts, etc., effectively.
- ☐ The speaker showed enthusiasm and interest.
- ☐ (if appropriate) the speaker responded to audience questions well.

Such a checklist is reasonably practical. Its reliability can vary if clear standards for scoring are not maintained. Its authenticity can be supported in that all of the items on the list contribute to an effective presentation. The washback effect of such a checklist will be enhanced by written comments from the teacher, a conference with the teacher, peer evaluations using the same form, and self-assessment.

In the perspective of Harris (1969), there is no language skill which is so difficult to assess with precision as speaking ability. Like writing, speaking is a complex skill requiring the simultaneous—use of a number of different abilities which often develop at different rates. There are at least four components that are generally recognized in analyses of speech process: pronunciation, grammar, vocabulary, and fluency. Harris (ibid.) further underlies that when we refer to a student's skill in speaking a second language, our fundamental concern is with his ability to communicate informally on everyday subjects with sufficient ease and fluency to hold the attention of his listener. Thus in the test of speaking ability we are primarily concerned with the student's control of the signaling systems of English—his pronunciation, grammar, and vocabulary—and not with the idea content or formal organization of the message he conveys.

The emerging question would probably be about how 'powerful' the oral test is in terms of its practicality in designing, implementing and scoring. The standard textbooks tend to conform that it is easy in the first two phases yet highly 'subjective' in the second last part. Subjectivity in scoring will in a nutshell be easily referred to as a 'weak' tool for assessing student's actual learning. This can be either true or false. However, as language teachers who teach and test language have to be aware that in fact all kinds of evaluating tools can never identify, generate and represent an accurate measure of 100% of student's actual learning performance or language ability. There is no flawless testing. There is always blind-spot in every measure. Irrespective of debatable strengths and weaknesses, oral test can serve as a practical and 'powerful' tool for measuring student's actual learning achievement.

DISCUSSION

The discussion here is principally on the ground of the writer's actual experience in teaching some courses like speaking, writing, grammar, English culture, SLA, and language testing. However, the explanatory data are not resulted from 'field' research but generated from a continuous reflection and tinkering as a language teacher. Doing reflections has provided the 'state-of-the-art' perspectives of teaching and learning processes, things the writer has done well and things the writer needs to change, revise and improve. Thus, let's have this discussion as though the writer was reiterating an old story.

Having tried many kinds of test like matching test, transformation tests, picture-cued tests, multiple choice, essay and the like, the writer always feels that there is something missing in the test, a sort of 'unconscious doubt' that my tests have not really measured what the tests were supposed to measure. This uneasiness became more overt when it came to scoring the test results. Let's take, for instance, the case of multiple choice items. Thomas (2011), who had done one study on this, found this kind of test to be problematic. First, a cumulative final exam like this one teacher would have thousands of questions and possible answers. Students are intimidated by a test that has approximately the same number of pages as their textbooks, and teachers are intimidated by the prospect of making and grading such a test. Secondly, tests of this type encourage the "3 R's"—read, remember, regurgitate (and then forget). Retention of these concepts simply does not occur with such a testing format.

The writer has never been so sure that what students had chosen strongly represented their very understanding of a given course.   Did they answer it because they really knew the

answer, or did they answer it correctly because they had the right guess? Would it be possible for those who haven't been exposed to previous learning at all could have the possibility to get the right answers when doing it with multiple choices? Another irritating question perhaps would be, "Did language teachers, with no exception, design a multiple choice test just because they were very concerned with the 'practicality' in scoring part for he/she was unlucky to teach a big class with more or less 50 to 100 students therefore neglecting the critical moment in the selection of test items? The worst of all, when the tests items were not carefully opted based   upon the syllabuses or objectives of the course or on what the teacher was supposed to teach and the students were supposed to learn, whereas the students did smart guesses, the test didn't mean anything either for the students, the teachers or the pedagogical process itself.

This sort of test cannot be relied on. If this sounds over-exaggerated, let us put it away. However, if that once happened to us especially when we set out our first teaching practice, let us take a 'moment of silence' how the test meant to us, as a teacher, and to students, as learners. Certainly, flawed testing practice means a lot: unprofessional. Of no doubt, these are not intended to disregard the virtues and values of multiple choice test. As long as it is very carefully designed, multiple choices probably can bring about a desired testing result.

This brings us to the second option—an essay test. What is your perception about essay test?   As the name implies, Essay Test is "easy" to make yet hard to grade. Jonathan Thomas (2011) firmly believes that a literature/humanities class should be writing intensive, and a test like this one definitely satisfies that criteria. But there are problems here, too. Such a test may be difficult to complete in two hours. If a choice of

writing prompts is offered, several students may opt to write on the same question, thereby limiting the number of fundamental/powerful concepts addressed in the class.

The writer is 100% certain that all of language teachers had an experience in conducting an essay test. Many of them very probably favor or disfavor of it due to one of these reasons, which is being very easy to make and being very tiresome to score respectively. Probably, there are few of language teachers who are very idealistic in composing an essay test, for instance, to purely generate students' comprehensive ability and therefore design a comprehensive test items covering the details of the given domain. At the same time the teachers become passionate teachers who are willing to go a miles away (time and energy) to do a scoring by looking at the details of the answer, the logic of the answers, the smoothness of the sentence structure, etc., and finally make a final judgment that student X deserves an A and student Y deserves a B.

The language teachers might possess a delicate idea when doing the scoring that dragged them to a social judgment postulating that well even though the answer does not mean anything, not really answering the question, but it is fine to reward the student with a certain score for his/her effort in making 'good and long' handwriting. Then, they were scoring the handwriting, not the answer. Making unwise wisdom is never wise. If the language teachers were taken to the church and asked to confess honestly on what they have done in their testing using essay, they would very likely to admit one of the following:

First, the writer designed the test following the whole procedures and scored the test results very attentively (the

writer did it as a passion of being a teacher, paid or unpaid that is not a big deal as my job is to serve and to educate).

Second, the writer designed the test seriously and carefully, but due to certain reasons he did the scoring half-heartedly (he just wanted to finish it at once as he didn't want to do something hard that he was not paid for)

Third, as the writer had not ample time or he was busy completing his academic-related responsibilities, the writer did not really look into the details of what essential points to involve in the questions. Yet, he was very concerned with the scoring and he had to see as well other factors during the class like students' activeness, assignment, and attendance. It means that he scored the students seriously especially based on the essay test.

Finally, the writer thought that he knew who his students were, who were and who were not active in the class. The tests were just complementary aspects. He made the essay test as the way he presumed the students could answer and corrected the test when had enough time or when he deemed he had to. There is no teacher who wants to sacrifice his/her students.

In short, it can be inferred that both multiple choice and essay tests do not really promise the language teachers a haven of comfort in being both a teacher and a tester. If there are teachers who feel great doing the testing as reflected above, perhaps they are having a made-up ambience that does not last constantly. Apart from these, some language teachers might propose the third option, which is by merging the two: partly multiple and partly essay. This sounds great and better yet very tricky in practice.

The underlying assumption is perhaps acceptable in which the two will complete each other in the sense that the

weakness of each is mended by strength of the other. Bear in mind that the language teachers who prefer to marry these two run the risk of having imbalance in terms of the proportion of the test items, yet very often overlapping to one another (the items are taken randomly, the proportion is lessened, and the multiple choice outnumbered the essay, even some language teachers might take one or two essay questions used just to complement the multiple choice). These do not stop there as the scoring will be more uncertain in terms of the value for each item particularly essay items. These models of tests are said to lack of so-called 'validity and reliability' in their loose sense.

What about making a paper that has to be submitted on examination day? It sounds an 'academic' task. However, has the teacher or lecturer asked of how the students have made the paper, whether they really make it themselves, where they got the references, how much they really understood the contents, and the like. Teachers should be alert at least on two queries: 1) how was the paper made? and 2)   how much the student know about what they put on the paper? Why these two? Because today is the electronic era in which all sources can be found available in the internet. Therefore, this era is also learnt as "copy and paste" era. Students can access e-books, articles, theses, book reviews, etc. easily with similar topics given by their teacher or lecturer.

The writer holds a belief that the essence of tasking a paper to students is to see how far students can formulate and arrange his understanding in a systematically arranged paper. So, how to come to grips with the issue of ideal and practical testing in which the aspects of validity and reliability are debatably negotiable?

Assuredly, the answer is Oral Test. It is a test in which the students can express their ideas, understanding, and

perception, argument, thought verbally. It is a test, regardless of what the subject is, that can be designed relatively easily (i.e. giving students some course-related topics to prepare in a paper or slides, just in their memory), practically implemented (asking the students to present orally the given topics, the course materials right on the spot and asking them to explain some statements or questions or to verify some contrasting /conflicting ideas opaquely), and effectively scoring their very performance reflecting their competence (of their respective course).

In scoring students' oral test performance, the language teachers are no longer haunted by uncertainty as to whether the answers really represent the students' actual understanding and knowledge as the teachers can directly observe the students' linguistic behavior (eye contact, mimicry, body language, etc.), logical of reasoning, and depth of understanding (in relation to content knowledge), use of media (visual aids, handouts, etc.), and students' linguistic performance (fluency, accuracy, intonation, pronunciation, and vocabulary) in uttering their answers, ideas and feedback. With the rating scale at hands, language teachers can easily put a score on every item tested.

How about the test efficiency? Oral testing can be claimed to be more efficient than other types of test in the sense that language teachers can directly score the students, while other test types language teaches have to look at the test sheets one by one which might take extra time at home. With other test types, language teachers work on the test twice, even three times or more for very often uncertainty shadows the teachers' mind on what actual score a student should get. In oral test, language teachers might seemingly take much time to test. That is definitely true yet we do not bring the job back home as both testing and scoring are enacted simultaneously.

What about its validity and reliability? It is valid as the students are tested what they have learnt previously and they are given ample time to prepare before the oral tests. In addition, students are also required to make a 'small' paper as per given topic. In other words, all questions and materials to be orally presented have been well informed in advance so they students make themselves well prepared. The test scores are reliable in the sense that all questions and materials have been 'pre-tested' in the class with the students and the scores are also reliable since the scores are not provided on a random basis but with a certain rating scale that has been informed to the students earlier.

What about the subjectivity of the scoring as it is scored by one person? It is true that more scores will minimize the subjectivity, yet it does not mean that a single rater cannot have an object score. This rests on the answer of the question: are you sure that you are a good teacher of English who knows English well who can write and speak the language correctly and accurately which sound English? If the answer is yes that there is no doubt that you can give an objective score. It is always better to have one good Malang apple than a basket of rotten Washington apples.

Naturally, there are problems with this format. The written part of the exam (the paper) simply has to be completed at home, and students will grumble about having to speak in front of the class. But the oral part can be factored into the exam grade using a speaking rubric, and the presentation still tests a student's ability to synthesize and analyze information in a finite amount of time. More importantly, it fosters a true knowledge and retention of the subject matter that does not fade as soon as the pen cools.

With a full-fledged understanding of the ideas of how to conduct oral test and of the advantages resulted from running the test, language teachers can have a sound result of language assessment. Language teachers can also do formative and summative assessment using oral test set formally based on a fixed schedule or informally on random basis. By advocating oral test, language teachers can truly perform classroom achievement test since they can easily diagnose the strengths and weaknesses of the students as the students can directly expose what they believe they can and cannot do without hesitation.

The oral test, therefore, enables language teachers to identify how much their students have learnt their respective course and what they lack of so that the teachers can have a strong basis on what to improve in their future teaching and learning process. Best of all, oral test can help teachers in sketching the real ability of their students so that they can console and guide their students accordingly. While at the same time oral test can give language teachers direct input on their teaching styles, teaching methodologies, teaching strategies, teaching materials, teaching aids and teaching commitment to achieving the success of students' learning.

CONCLUSION

Like teaching, testing can be done in many varied ways. Multiple choices, essay and other forms are not something to leave and to forget but something to review and to renew on the way how it should be held. Oral testing as a part of language assessment is another kind that is worth practicing. Each of them has advantages and disadvantages. However, to excel in performing it we need to focus on the good points and to be

aware of its drawbacks so we are alert on which track to go through. The end does not mean the last, yet it is the angle to see where to start. Succinctly, in the conclusion we are not closing but opening possibilities or new ways in doing things: teaching and testing. Let us see what goes well and do something extra to keep it well. Too, see what does not go in tandem with our teaching blue prints and have a plan to do better things to make them better. Bear in mind that both teaching and testing are intertwined in which one cannot stand still without the other. The niche of the issue of language teaching, language learning, and language testing is that language learners should be ensured that they are not solely good at acquiring linguistic competence but more importantly they are solemnly great at performing their communicative competence which can be done through oral testing.

REFERENCES

Brown, J. D. *Testing in Language Program: A Comprehensive Guide to English Language Assessment*. Singapore: McGraw-Hill, Inc., 2005.

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

Harris, D. P. *Testing English as a Second Language*. New York: McGraw-Hill Book Company, 1969.

Heaton, J. B. *Writing English Language Test: Longman Handbooks for Language Teachers*. London: Longman Inc., 1989.

Henning, G. *A Guide to Language Testing: Development, Evaluation, Research*. Boston: Heinle & Heinle Publishers., 1987.

Hingle, I. and Linington, V. "English Proficiency Test: The Oral Component of a Primary School". In Richards, J.C. and

Renandya, W.A. (eds.). *Methodology in Language Teaching: An Anthology of Current Practice*. Cambridge: Cambridge University Press, 2002.

Hughes, A. *Testing for Language Teachers*, 2nd ed. Cambridge: Cambridge University Press, 2008.

Jonathan Thomas, J. *After the Test: An Oral Examination of Fundamental/Powerful Concepts,* 2001. In www.surry.edu/ about/critical_ thinking/thomas_afterthetest.pdf, retrieved 7 December 2012.