

TYPES OF ENGLISH LANGUAGE TEST

Nike Angraini¹⁾

¹⁾English Education Study Program, Faculty of Teacher Training and Education, Indo Global Mandiri University
Jl. Jend. Sudirman No. 629 KM.4 Palembang Kode Pos 30129
Email : nikeangraini210191@gmail.com¹⁾

ABSTRACT

In this globalization era, English is very prominent for many people all over the world starts from students in elementary school till students in university level. The students are expected to master English language both in oral and written communication. If the students are able to master English, they can use it to interact and socialize with people who are not only from Indonesia but also from other countries. Therefore, to see whether or not the students succeed in mastering English, the test must be administered. This article aims to discuss miscellaneous of tests which can be used in teaching and learning environment. Some conclusions and suggestions are made to ensure both the teacher and students will gain benefits from these types of English language test.

Keyword : *types, English language, test*

1. Introduction

The use of English as an international language plays a significant role in the globalization era particularly on classroom instruction toward EFL learners in which they have different linguistic and cultural background. The ability to verbalize English effectively is strongly demanded not only for teacher and learners but also for learner and learners in order to create good communication and understanding in the classroom. And to measure whether or not the learners achieve the instructional of study could be administered by a test. Over the past few years, researchers started to pay great attention to studies on language testing to improve the methodologies of administering, constructing and evaluating tests. Testing is the most controversial areas related to any kind of teaching and at the same time something that is necessary as a sort of completion of the teaching and learning progress (Gonzalez, 1996, p. 17). In fact, if adequately focused, it checks the effectiveness of the whole process. In line with this statement, Brown (2004, p. 219) states that test is a method in which to measure a person's ability or knowledge in a given area. And speaking about language testing, and more specifically, foreign language testing, the first problem arises because language is both the instrument and the object of measurement (Bachman, 1990, p. 2), that is to say, we use language to measure language ability. Testing is often the basis for taking an important decision. These decisions influence the learners in the first place, parents, and also other people who are responsible for education.

When test are used in making educational decisions for individual students, it is important that they accurately measure students' abilities, knowledge, skills, or needs, and that they do so in ways that do not discriminate in violation of federal law on the basis of students' race, national origin, sex, or disability (Riley and Cantu, 2000). Tests that are used in educationally appropriate ways and that are valid for the purposes used

can serve as important instruments to help educators do their job. In a similar way, tests help the educators know if his/her teaching is effective and can provide insights into ways to improve the testing process. By giving a test, we understand any formal and informal task set at a given moment for one or several purposes. Therefore, we, as a teacher of English must be aware of giving a test to the learners in order to reach the instructional goal of study.

A. Discussion

a) *Different Types of Test*

There are different kinds of test according to different criteria. Gonzalez (1996, p. 31) divides test into five categories based on the purpose as follows:

1. *Placement tests*

The function of this test is to place new students in the right class in the school. Usually based on syllabuses and materials the students will follow and use once their level has been decided on, these test grammar and vocabulary knowledge and assess students' productive and receptive skills. Some; schools ask students to assess themselves as a part of placement process, adding this self-analysis into the final placing decision. In line with that, Hughes (1989) reveals that placement tests are intended to provide information which will help to place students at the stage (or in a part) of the teaching program most appropriate to their abilities (p. 14). On the other word, this test helps the educators to place students in the learning group that is appropriate for their level of competence. For instance, when the students want to join formal schools and universities. Virtually, placement tests are also administered in some non-formal institution.

2. *Diagnostic tests*

Hughes (1989, p. 12) defines that diagnostic tests are used to identify students' strengths and weaknesses. They are intended primarily to ascertain what further

teaching is necessary. In addition, Gonzalez (1996) asserts that diagnostic tests are also called formative or progress tests and they are used to diagnose a particular aspect of a particular language or to check on students' progress in learning particular elements of a course. They help the teachers to decide what needs to be taught to students. Generally, these tests refer to short-term objectives. When you want to join "students' exchange", you must join diagnostic tests.

3. Achievement tests

These tests are used to measure the students' language and skill progress in relation to the syllabus they have been following (Harmer, 2001, p. 321). Similarly, Hughes (1989) mentions that achievement tests are directly related to language courses, their purpose being to establish how successful individual students, groups of students, or the courses themselves have been in achieving objectives (p. 10). In line with that, Hungerland (2004) asserts that these tests are associated with process of instruction. In short, these tests assess whether learners have acquired specific elements of language that they were taught in the language course they took part in. There are two types of achievement tests: final tests at the end of the course and progress tests during the course. For instance; quiz, mid-term, final examination, and others.

4. Proficiency tests

Gonzalez (1996) claims that proficiency tests are not limited to any course or curriculum and they measure global competence in a language. Some of them are external examination which contains several papers and machine scorable and sometimes they add free writing or speaking with the subsequent problem of practicality of scoring. It is supported by Hughes (1989, p. 9) who asserts that proficiency tests are designed to gauge people's ability in a language regardless of any training they may have had in that language. The examples of these tests are TOEFL, TOEIC, IELTS, Cambridge Examinations and others.

5. Aptitude tests

The function of these tests is to find out whether or not the learners will be successful in learning a foreign language. As Brown (2004) points out that these tests are rarely used in learning environment due to the prediction of the general scholarly success of the students is merely. These tests are also not to tell anything about the strategies students may use. Apart from that, there are even serious ethical objections because they bias both teacher and students. *The Modern Language Aptitude Test* and *The Pimsleur Language Aptitude Battery* are kind of aptitude tests.

Hughes (1989) mentions that there are two kinds of testing based on frame of reference as follows:

1) Norm-referenced testing

In norm-referenced tests, candidate's performance is assessed in comparison with that of the other candidates. For these reason the cut-off points (line between fail and pass) are determined after the test

results are obtained from the group of students based on the distribution of the scores. This kind of test relates one candidate's performance to that of other candidates. In line with that statement, Waugh and Grolund (2013, p. 27) state that norm-referenced testing is used to measure individual differences in achievement.

2) Criterion-referenced testing

Criterion-referenced tests compare all the testees to predetermined criterion. In such tests everybody whose achievement comes up to the pre-set criterion will receive a pass mark, while those under it will fail. The criteria are often set in terms of task that students have to be able to perform. For instance: to interact with an interlocutor with ease; to ask for information and understand instructions. Apart from that, Gonzalez (1996, p. 32) asserts that these tests create a more positive attitude and reduce the negative effect of competition. Criterion-referenced tests are intended to classify people according to whether or not they are able to perform some task or set of tasks satisfactorily (Hughes, 1989, p. 18). In line with that, Waugh and Grolund (2013) point out that these tests are used to describe tasks students can perform and to compare performance to a clearly specified achievement domain (p. 27).

6. Types of Test Item

Harmer (2001, p. 322) reveals two kinds of test item that is direct and indirect test items.

1) Direct test item

A test item is direct if it asks the candidate to perform the communicative skill which is being tested. In line with that, Hughes (1989) states that the test is direct when it requires the candidate to perform precisely the skill which we wish to measure (p. 15). This test is easier to carry out when it is intended to measure the skills of speaking and writing. Direct test items try to be as much like real-life language use as possible. The examples of this test items are; (a) an interviewer questioning a candidate about themselves, (b) role-play activities, (c) writing compositions and stories and so on.

2) Indirect test item

The purpose of this test items is to measure students' knowledge and ability by getting at what lies beneath their receptive and productive skills. Indirect items try to find out about students' language knowledge through more controlled items such as multiple choice questions, grammar transformation items, cloze procedures and others. These are often quicker to design, crucially, easier to mark, and produce greater scorer reliability.

B. Scoring Procedure

Commonly, certain skills and areas of language may be tested more effectively by one method than another (Gonzalez, 1996). Objective tests are more useful to test

recognition level, while subjective tests are better for the production level.

a) Objective testing

The scoring of a task is objective if the rater does not have to make judgment because the scoring is unambiguous. But we must admit that objective could not test the ability to communicate in the foreign language. This test could not evaluate performance. This paper discusses in some detail two standard formats for objective scoring. For example: multiple-choice questions and true-false items.

1) Multiple-choice questions (MCQs)

This kind of test requires the examinees to answer by choosing one of a number of alternatives. Waugh and Grolund (2013) mention that the multiple-choice item consists of a *stem*, which presents a problem situation, and several alternatives (options or choices), which provide possible solutions to the problem (p. 91). The stem may be a question or an incomplete statement. The alternatives include the correct answer and several plausible wrong answers called *distracters*. The function of the latter is to distract those examinees who are uncertain of the answer. Here are the examples of multiple-choice items as follows:

The journalist was _____ by enemy fires as he tried to send story by radio.

- a. wronged
- b. wounded
- c. injured
- d. damaged

He looks angry. He _____ his wallet.

- a. lost
- b. b. has lost
- c. has been losing
- d. loses

Harmer (2001) claims that multiple-choice tests are considered to be ideal test instruments for measure students' knowledge of grammar and vocabulary (p. 323). However, there are number of problems with multiple-choice questions. In the first place, they are extremely difficult to write well, especially in the design of incorrect choices. These "distracters" may virtually put ideas into students' heads that they do not have before they read them. Second, while it is possible to train students so that their multiple-choice question (MCQ) abilities are enhanced, this may not actually improve their English. According to Applied, Skills, & Knowledge, LLC (2001, p. 4), there are some advantages of using multiple-choice items as follow:

- They are easy to administer and can be answered quickly by the examinees.
- They are easy to score and can be scored electronically.
- They are easy to handle statistically.
- They have good reliability (i.e., examinees choose the same answer again if the same item is presented on two separate occasions).

- They permit the measurement of several levels of learning due to the incorrect answers can be selected to attract examinees of varying degrees of skill.
- In addition, multiple-choice items also have disadvantages, including:
- Though better than true-false items, guesses will be correct a certain percentage of time. For example, if there are four alternatives, one-fourth of the guesses will be correct.
- The development of good multiple-choice items with "attractive" incorrect answer is difficult. Multiple-choice items can measure a wide range of skill complexity including simple recall of facts to application, analysis and synthesis of information.

2) True-false items

Gonzalez (1996, p. 34) points out that this kind of test is used both for reading and listening comprehension and also for understanding of specific elements of the language. Apart from that, Waugh and Grolund (2013, p. 110) assert that true-false items are typically used to measure the ability to identify whether statements of fact are correct. The basic format is simply a declarative statement that the student must judge as true or false. Virtually, true-false items are multiple-choice tests with only two options. This means that there is a fifty percent of probability of getting the right answer and, obviously, this encourages guessing.

Here is the example of true-false items as follows:

*TF Lengthening a test by adding items like those in the test will increase its reliability.

T*F A true-false item with the correct answer being "false" provides evidence that the student knows the correct answer.

3) Subjective testing

In subjective test tasks, raters have to make a judgment when assessing candidate's performance. In spite of the scoring difficulty, they are very useful because they test production, both oral and written (Gonzalez, 1996). In addition, they are very easy to set. This paper discusses briefly two traditional tests, oral interview (speaking) and composition.

b) Oral interview (speaking)

Oral interview is a standardized, global assessment of functional speaking ability. Taking the form of a conversation between the tester and test-taker, the test measure how well a person speaks a language by assessing their performance of a range of language tasks against specified criteria. The questions could be made personal if the interviewer knows something about the students. A range of yes/ no, wh- either/

questions should be used, together with the statements (Gonzalez, 1996). It is important to include some questions or statements that require some kinds of correction or modification to make students talk. Commonly, interview test takes 11-14 minutes and is made up of three sections as follows:

- Part 1 (introduction and interview): The examiner will introduce him or herself and ask you to introduce yourself and confirm your identity. The examiner will ask you general questions on familiar topics, e.g. home, family, job, study and interest. This section should help you relax and talk naturally. This part takes 4-5 minutes.
- Part 2 (individual long turn): The examiner will give you a task card which asks you to talk about a particular topic, including points to include in your talk. You will be given one minute to prepare and make notes. You will then be asked to talk for 1-2 minutes on the topic. You will not be interrupted during these times, so it is important to keep talking. The examiner will then ask you one or two questions on the same topic. This part takes 3-4 minutes.
- Part 3 (two-way discussion): The examiner will ask you further questions which are connected to the topic of part 2. These questions are designed to give you an opportunity to discuss more abstract issues and ideas. This part takes 4-5 minutes.

c) *Composition*

The writing displays an ability to communicate in a way which gives the reader full satisfaction. It displays a completely logical organizational structure which enables the message to be followed effortlessly. Gonzalez (1996, p. 35) points out that composition is one of the easiest tasks to set both as classroom homework and as a testing exercise. In these tasks, the students are demanded to think critically to compose paragraph or essay. We, as a teacher of English can provide the title of the tasks so that the students will be able to do it well.

C. *Principles and Characteristics of Communicative Test Design*

a) *Principles*

Communicative language tests are used with the goal of measuring language learners' ability to take part in acts of communication or to use language in real life situations. Communicative tests, which cover the four language skills of listening, speaking, reading, and writing, are designed on the basis of communicative competence. According to Canale and Swain (1980), communicative competence involves linguistic competence (knowledge of linguistic forms), sociolinguistic competence (the ability to use language appropriately in contexts), discourse competence (coherence and cohesion), and strategic competence (knowledge of verbal and non-verbal communicative strategies) (p. 4).

Understanding the model of communicative competence is necessary and helpful for developing communicative language tests, which involves formulating test objectives and considering the effects of the test on teaching and learning. Following the model of communicative competence, a team at the Ontario Institute for Studies in Education (OISE) has established four principles of communicative test design (Bailey, 1998). The first principle is to "start from some where." The study of Katsumasa (1997) and Wesche (1983) showed an agreement with this statement by confirming that when designing the test, test makers should state carefully what they expect test takers to perform when they use the target language in a specific context, which means that test writers must know what they want to test. After that, test makers can establish scales and criteria in assessment procedures to measure exactly the stated features of testees' performance. Bailey (1998) provided an example to illustrate this principle, which is that even though conveying and capturing meaning while maintaining accuracy are two important elements in communication, if the tester focuses on checking the test takers' ability to convey meaning, then accuracy will not be put into the scoring criteria.

Therefore, it is unfair for students if test writers take into account other unstated objectives when grading tests, which also negatively affects the test's validity. "Concentration on content" is the second principle in designing a communicative language test. The content here refers to not only topics but also tasks that will be implemented. Appropriate content matches 4 or fits learners' or test takers' age, proficiency level, interests, and goals/needs. According to Carroll (1983), "the language tasks our learners are expected to perform in their future jobs will guide us with the tasks we will set them in our tests" (p. 37). In other words, tasks should be constructed based on the testee's relevant needs. For instance, some tasks for children at the end of primary school may include introducing themselves and others, reading traffic signs such as "Danger," "No Entry," and so forth.

The tasks should aim to be authentic and have clear reference in reality. These tasks match children's proficiency level and children's age. Children are supposed to be able to do the tasks because what they do is what society (at least based on opinions of experienced teachers and advisors) expects of them in real life. The third principle of communicative test design is "bias for best." This means that test makers should bear in mind that they should create a test that can exploit test takers' performance at their best. The work of Brown (2003) also supports this principle in designing the test.

According to Brown (2003), "biased for best" is "a term that goes little beyond how the student views the test to a degree of strategic involvement on the part of student and teacher in preparing for, setting up, and following the test itself" (p. 34). To illustrate this third principle, Bailey (1998) said that when she gave a test that requires students to do a dictocomp, she noticed that the text might be higher than students' proficiency level;

therefore, sheread the text three times and also encouragedthem to ask for new words involved inthe text. Swain (1984) declared that in order to create an assessment procedure which is biased for best, test makers and teachers should provide students or test takers appropriate review to help them to be well-prepareda nd ready for the test, suggest strategies that will be beneficial, and constructthe test in a way that it is modestly challenging to the best students/testees and yet the weaker will not be over whelmed (ascited in Brown, 2003, p. 34).

“Working on washback” is the fourth principle of communicative language testing. In order to obtain positive wash back, test writers should create clear scoring criteria that will be provided to *both* teachers and test takers. Course objectives and test contentare also put into consideration in thehope of promoting positive washback. Accordingto Wesche (1983), when formulating objectives, it is important to clarify (a) the purpose of the interaction concerningtopics and the language functions that thelearner will need, (b) the context that mayimpact language use, and (c) appropriatetypes of discourse, and the degree of skill attestees’ levels. Clarification for these factorsnot only promotes positive washback but also makes it easier for test makers tochoose good stimulus material that willprovide necessary language forms such asstructures and vocabulary.The principles constructed by the teamat the OISE are not unique. Katsumasa(1997), citing Morrow’s (1981) study, discussed similar principles in the context of communicative language testing. The seprinciples are (a) knowing what you are measuring, (b) measuring the ability to deal with discourse, (c) focusing on communication processes (d) setting up real situations,and (e) understanding the fact that mistakesare not always mistakes because minorgrammatical mistakes do not frequently hinder communication and correct grammaris not the only requirement for successfulcommunication.

After reading and comparingthe principles ofcommunicative languagetest design put forth by Katsumasa and the OISE team, I prefer the framework of the OISE team because it involves the discussion of how to achieve beneficial wash back. Hughes (1989) indicated that “backwash is now seen as a part of the impacta test may have on learners and teachers,on educational systems in general, and on society at large” (p. 53). Therefore, theOISE emphasis on trying to achieve positive wash back is welcome.

How to Make a Good Test

Giving a test to the students is very prominent to measure whether the instruction has been delivered well by the teacher. Based on Applied, Skills, & Knowledge, LLC (2001) asserts that making tests which are fairly and accurately measure what you want to measure involves a series of simple steps and some quality control checks. Along with this, Hughes (1989) argues some following things that need to be considered in

creating tests to be more reliable; (1) take enough samples of behavior, (2) do not allow candidates too much freedom, (3) write unambiguous items, (4) provide clear and explicit instructions, (5) ensure that tests are well laid out and perfectly legible, (6) candidates should be familiar with format and testing techniques, (7) provide uniform and non-distracting conditions of administration, (8) use items that permit scoring which is as objective as possible, (9) make comparisons between candidates as direct as possible, (10) provide a detailed scoring key, (11) train scorers, (12) agree acceptable responses and appropriate scores at outset of scoring, (13) identify candidates by number, not name, and (14) employ multiple, independent scoring.

3. Conclusions

As teachers, we probably have to set and mark tests within our school as well as prepare candidates for external exams. There are many ways of evaluating learner performance and level. One of them is that give tests. A test is a sample of an individual’s behavior or performance on the basis of which inferences are made about the more general underlying competence of that individual. Language tests refer to any kind of measurement or examination technique which aims at describing the students’ foreign language proficiency, e.g. oral interview, listening comprehension task, free composition writing. Tests should be administered to student(s) in order to find out whether or not the teachers have been successful in giving instruction. And to show whether the instructional objectives have been achieved. Therefore, the teacher(s), as a facilitator are really demanded to give instruction effectively and should be able to create a good tests.

References

- [1] Applied Skills, & Knowledge, LLC. (2001). Making accurate & fair test. Retrieved from <http://www.appliedskills.com/White%20Papers/Making%20Accurate%20&%20Fair%20Tests.pdf>
- [2] Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.
- [3] Bailey, K. M. (1998). *Learning about language assessment: Dilemmas, decisions, and directions*. Pacific Grove, CA: Heinle & Heinle.
- [4] Brown, H. D. (2003). *Language assessment: Principles and classroom practices*. White Plains, NY: Longman.
- [5] Brown, H. D. (2004). *Principles of language learning and teaching*. New Jersey: Prentice-Hall Regents.
- [6] Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and learning. *Applied Linguistics*, 1(1), 1-47.
- [7] Carroll, B. J. (1983). Communicative language tests: Tasks, enabling skills, formats, and

- measurement criteria. *WorldLanguage English*, 2(1), 37-39.
- [8] Gonzalez, A. B. (1996). Testing English as foreign language: An overview and some methodological considerations. *RESLA*, 11, 17-49. Retrieved from <http://www.dialnet.unirioja.es/descarga/articulo/106212.pdf>
- [9] Harmer, J. (2001). *The practice of English language teaching*. Cambridge, UK: Longman. Retrieved from <http://www.scribd.com/doc/55944844/The-Practice-of-English-Language-Teaching-3rd-Ed-Jeremy-Harmer-2001>
- [10] Hughes, A. (1989). *Testing for language teachers*. Cambridge, UK: Cambridge University Press.
- [11] Hungerland, R. (2004). *Language proficiency testing: A critical survey*. Paper presented at TESL Canada National Conference, Canada, North America.
- [12] Katsumasa, S. (1997). Communicative language testing: Principles and problems. *English Review*, 12, 3-24. Retrieved December 1, 2007, from <http://ci.nii.ac.jp/naid/110004693443/en/>
- [13] Riley, R. W., & Cantu, N. V. (2000). The use of tests as part of high-stakes decision-making for students: A resource guide for educators and policy-makers. Retrieved from <http://www.ed.gov/offices/.../testing/TestingResource.doc>
- [14] Waugh, C. K., & Gronlund, N. E. (2013). *Assessment of student achievement*. New Jersey: Pearson Education, Inc.
- [15] Wesche, M. B. (1983). Communicative testing in a second language. *The Modern Language Journal*, 67 (1), 41-55.