

# Measurement and Assessment in Writing Skill

**Supriusman**  
Univeritas Riau

## Abstract

Measurement and assessment are two key words which have a very important role to play to help teachers indicate whether or not students have been successful in teaching and learning process. Both measurement and assessment have different characteristics. Assessment of students' language skill can be known when measurement has been conducted. According to the teachers and lecturers law No. 14, 2005, one of main duties of professional teachers is the ability of the teachers to assess their students. However, to assess students objectively needs adequate experience, understanding, knowledge and skill on assessment to avoid and reduce subjectivity in assessing receptive and productive skills, in particular writing skill. This paper describes how to assess writing skill, definitions of measurement and assessment, the characteristics of measurement and assessment, the purposes of assessment. Two types of assessment; traditional and alternative assessment are also described to increase especially teachers' insight in this topic.

**Keywords:** measurement, and assessment

## Introduction

It is generally accepted that the four language skills: listening, speaking, reading and writing can be assessed by language teachers. However, each skill has different assessment. This paper limits its discussion to assessing writing skill. The ability to assess writing in English language teaching is very important for English teachers because as one of language productive skills, writing cannot be assessed directly and objectively by unqualified and inexperienced English teachers. This is in line with the teachers and lecturers law No. 14, 2005 which states that one of main duties of professional teachers is to assess their students in teaching and learning process. In addition, it is generally accepted that the majority of teachers and learners also get confused to understand and make distinction between the terms of 'assessment' and 'evaluation'. Other important terms such as 'writing', 'measurement', and 'test' are also described in this paper. Therefore, this paper tries to describe them in details in order to get full understanding about them.

## The Purposes of Writing of This Paper

This paper is intended to answer the following questions:

1. What is measurement?
2. What is assessment?
3. What is evaluation?
4. What is the difference between measurement and assessment?
5. What are the purposes of assessment done by language teachers?
6. What is the difference between traditional assessment and alternative assessment?
7. How do language teachers assess writing skill?

From the above questions, the purposes of writing this paper are to be able to define measurement, assessment, and evaluation and to differentiate between measurement and assessment, to know the purposes of assessment done by language teachers, to know the

difference between traditional assessment and alternative assessment, and to know the ways and the indicators used to assess writing skill as one of the productive skills.

## Methodology

In describing and discussing this topic, library research is used to collect all relevant information needed especially to discuss and describe about measurement, assessment, evaluation, and test as mentioned in the following:

## Description And Discussion

The following are described, the definitions of measurement and assessment, evaluation, the purpose of conducting assessment, the difference between measurement and assessment, assessing writing as the productive skill and other important description about traditional and alternative assessment.

The success of the program can be known through measurement. Therefore, measurement has a very important role to play either in science and technology or in presenting information by policy maker. Basically, measurement is defined as the collection of data or information, usually in quantitative terms, of student performance. It can be also described that to measure means to convert information into numbers or figures. It can be seen from this definition that measurement needs figure or a certain scales such as nominal, ordinal, interval and ratio scales. Mardapi (2008) states that measurement compares between the result of observation and criteria. He also states that there are two measurement theories namely classical and modern theories. Classical theory means that the score gained by an individual can be described as a real score and an error score. Modern theory in the one hand means that an opportunity of someone to answer the question correctly is not determined by his opportunity to answer other item of the question. This theory tries to develop an analysis to determine someone's ability without an instrument (measurement) used. However, the instrument used should be valid and reliable. The validity of the measurement is determined by the correct or appropriate instrument used in measuring what is supposed to measure. For example, students are asked to write an essay to determine their ability in writing skill. The reliability of the measurement deals with the consistency or stability of the result gained even though it is measured again and again, and is measured by two or three raters especially in writing by using writing rubrics.

Assessment as described by Puhl (1997) is the process of collecting and integrating information about learners from various sources to help us understand these students and describe them. While Richards, Platt, and Platt (1992), in the one hand, define assessment as the measurement of the ability of a person or the quality or success of a teaching course, etc. Then, assessment according to Print (1993) involves the interpretation of measurement data. It makes sense of the data collected on student performance. It can be seen from the definitions that to know the information about what students can do or cannot do needs measurement by using the appropriate instrument. The general purpose of the assessment as described by Cohen (1994) can be categorized into administrative, instructional, and research functions as shown in Table 1.

Table 1: The Purpose of the Assessment

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

The next term is evaluation. In this paper evaluation refers to a process which consists of the sub-processes of measurement and assessment providing adequate and useful data upon which the final judgments are made. Puhl (1997) defines evaluation as 'the process of making a judgment of a product, a response, or performance based on criteria.' Evaluation in our schools is essentially concerned with two major approaches to making judgments: *Firstly*, product evaluation which means an evaluation of student performance in a specific learning context. Such an evaluation essentially seeks to determine how well the student has achieved the stated objectives of the learning situation. In this sense the student's performance is seen as a product of the educational experience. A school report is an example of product evaluation. The *second approach* is known as process evaluation which includes the experiences and activities involved in the learning process i.e. making judgments about the process by which students required learning or examining the learning experience before it has been concluded. In most cases, process evaluation is used when making judgments about school effectiveness, classroom interactions, the curriculum and about the effectiveness of specific programs. For example, process evaluation may be carried out upon the nature of student-teacher interaction, instruction methods, school curricula, and so forth.

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

The differences between assessment and measurement are as follows: Assessment is *interpretative* which means that assessment gives information about something through measurement, for example, people whose height are 165 centimeters are ideal heights for Indonesian people. Assessment is also *qualitative*. This means that the information of the assessment determines the quality of something such as excellent, sweet, tall, etc. The last characteristic of assessment is *subjective*. The information of the assessment is dependent on situation and condition of assessor, for example, a person with 165 centimeters tall is

categorized short for western people. Measurement, on the other hand, is *descriptive*, *quantitative*, and *objective*. Descriptive measurement refers to information shows the real fact of the object or person. For example, his brother is 150 centimeters tall. The second characteristic means that measurement provides information in figure or number. Measurement is objective because it presents information as it is without changing the information or data.

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

#### Assessing Writing as the Productive Skill

As has been mentioned earlier, as one of language productive skills, writing cannot be assessed directly and objectively by unqualified and inexperienced English teachers. writing in this paper is described as the ability of a person to express his feelings, ideas and opinions in written communication or form. Based on the writer's knowledge, experience, and observation, English teachers have difficulty in assessing writing. Hughes (1993) states that there are five components that must be taken into account namely grammar, vocabulary, mechanics, fluency and form (organization). Cohen (1994) scores writing skill in two different ways namely holistic scoring, and analytic scoring. Another way to assess writing as suggested by Brown (2004) is by looking at organization, logical development of ideas, mechanics, and style and quality of expression. To help teachers assess writing skill, it is used rubric. From this rubric, English teachers can see clear criteria to assess this skill as can be seen in TOEFL iBT Test Assessment. There are two writing rubrics namely independent and integrated writing rubrics. Each of them consists of score and task description as shown in Tables 2 and 3.

**Table 2: Independent Writing Rubrics**

Score	Task Description
5.	<p><b>An essay at this level largely accomplishes all of the following:</b></p> <ul style="list-style-type: none"> <li>• Effective addresses the topic and task</li> <li>• Is well organized and well developed, using clearly appropriate explanations, exemplifications and/or details</li> <li>• Displays unity, progression, and coherence</li> <li>• Displays consistent facility in the use of language, demonstrating syntactic, variety, appropriate word choice, and idiomaticity, though it may have minor lexical or grammatical errors</li> </ul>
4.	<p><b>An essay at this level largely accomplishes all of the following:</b></p> <ul style="list-style-type: none"> <li>• Addresses the topic and task well, though some points may not be fully elaborated</li> <li>• Is generally well organized and well developed, using appropriate and sufficient explanations, exemplifications, and/or details</li> </ul>

	<ul style="list-style-type: none"> <li>• Displays unity, progression, and coherence, though it may contain occasional redundancy, digression, or unclear connections</li> <li>• Displays facility in the use of language, demonstrating syntactic variety and range of vocabulary, though it will probably have occasional noticeable minor errors in structure, word form, or use of idiomatic language that do not interfere with meaning</li> </ul>
3.	<p><b>An essay at this level is marked by one or more of the following:</b></p> <ul style="list-style-type: none"> <li>• Addresses the topic and task using somewhat developed explanations, exemplifications, and/or details</li> <li>• Displays unity, progression, and coherence, though connection of ideas may be occasionally obscured</li> <li>• May demonstrate inconsistent facility in sentence formation and word choice that may result in lack of clarity and occasionally obscure meaning</li> <li>• May display accurate but limited range of syntactic structures and vocabulary</li> </ul>
2.	<p><b>An essay at this level may reveal one or more of the following weaknesses:</b></p> <ul style="list-style-type: none"> <li>• Limited development in response to the topic and task</li> <li>• Inadequate organization or connection of ideas</li> <li>• Inappropriate or insufficient exemplification, explanations or details to support or illustrate generalizations in response to the task</li> <li>• A noticeably inappropriate choice of words or word forms</li> <li>• An accumulation of errors in sentence structure and/or usage</li> </ul>
1.	<p><b>An essay at this level is seriously flawed by one or more of the following weaknesses:</b></p> <ul style="list-style-type: none"> <li>• Serious disorganization or underdevelopment</li> <li>• Little or no detail, or irrelevant specifics, or questionable responsiveness to the task</li> <li>• Serious and frequent errors in sentence structure</li> </ul>
0.	<p><b>An essay at this level</b> merely copies words from the topic, rejects the topic, or is otherwise not connected to the topic, is written in a foreign language, consists of keystroke characters, or is blanks.</p>

**Table 3: Integrated Writing Rubrics**

Score	Task Description
5.	<p><b>A response at this level</b> successfully selects the important information from the lecture and coherently and accurately presents this information in relation to the relevant information presented in the reading. The response is well organized, and occasional language errors that are present do not result in inaccurate or imprecise presentation of content or connections.</p>
4.	<p><b>A response at this level</b> is generally good in selecting the important information from the lecture and in coherently and accurately presenting this information in relation to the relevant information in the reading, but it may have minor omission, inaccuracy, vagueness, or imprecision of some content from the lecture or in connection to points made in the reading. A response is also scored at this level if it has more frequent or noticeable minor language errors, as long as such usage and grammatical structures do not result in anything more than an occasional lapse of clarity or in the connection of ideas.</p>

3.	<p><b>A response at this level contains some important information from the lecture and conveys some relevant connection to the reading, but it is marked by one or more of the following:</b></p> <ul style="list-style-type: none"> <li>• Although the overall response is definitely oriented to the task, it conveys only vague, global unclear, or somewhat imprecise connection of the points made in the lecture to points made in the reading.</li> <li>• The response may omit one major key point made in the lecture.</li> <li>• Some key points made in the lecture or the reading, or connections between the two, may be incomplete, inaccurate, or imprecise.</li> <li>• Errors of usage and/or grammar may be more frequent or may result in noticeably vague expressions or obscure meanings in conveying ideas and connections.</li> </ul>
2.	<p><b>A response at this level contain some relevant information from the lecture, but is marked by significant language difficulties or by significant omission or inaccuracy of important ideas from the lecture or in the connections between the lecture and the reading: a response at this level is marked by one or more of the following:</b></p> <ul style="list-style-type: none"> <li>• The response significantly misrepresents or completely omits the overall connection between the lecture and the reading.</li> <li>• The response significantly omit or significantly misrepresents important points made in the lecture.</li> <li>• The response contains language errors or expressions that largely obscure connections or meaning at key junctures or that would likely obscure understanding of key ideas for a reader not already familiar with the reading and the lecture.</li> </ul>
1.	<p><b>A response at this level is marked by one or more of the following:</b></p> <ul style="list-style-type: none"> <li>• The response provides little or no meaningful or relevant coherent content from the lecture.</li> <li>• The language level of the response is so low that it is difficult to derive meaning.</li> </ul>
0.	<p><b>An essay at this level merely copies sentences from the reading, reject the topic or is otherwise not connected to the topic, is written in a foreign language, consist of keystroke characters, or is blank.</b></p>

### Some Differences of Traditional and Alternative Assessment

**Brown** (2004) describes some differences of traditional and alternative assessment as shown in Table 4.

**Table 4: Traditional and Alternative Assessment according to Brown (2004:13)**

<b>Traditional Assessment</b>	<b>Alternative Assessment</b>
One-shot, standardized exams	Continuous long-term assessment
Timed, multiple-choice format	Untimed, free-response format
Decontextualized test items	Contextualized communicative tasks
Score suffice for feedback	Individualized feedback and wash-back
Norm-referenced scores	Criterion-reference scores

Focus on the 'right' answer	Open-ended, creative answers
Summative	Formative
Oriented to product	Oriented to process
Non-interactive performance	Interactive performance
Foster extrinsic motivation	Foster intrinsic motivation

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

**Table 5. Traditional and Continuous Assessment according Puhl**

<b>Dimension</b>	<b>Traditional Assessment</b>	<b>Continuous Assessment</b>
Purpose of test	Summative; it "sum up" what has been happening Judgmental; forces learner to study	Formative; it generates input to inform and guide teaching Developmental; diagnostic: directs instructional attention
Focus	Product of instruction Teacher-created activity Heavy on memorization De-contextualized	Process of instruction Learner-created activity Heavy on thinking, integration holistic
Feedback	A score or mark; final, no changing it	A range of comments from peers, teachers; happens the process while still time to change
Test task	Typically written work Medium: paper and pencil Narrow focus Exercises (for the future) Formal	Typically a range of tasks Multimedia Multidimensional Authentic (real life tasks for now) informal
Classroom management	Intrusive; interrupts class process "one-shot;" only one change to show competence Results need time to be determined Books closed	Integrated; part of class routine Overtime; change to revise, improve, add Feedback comes quickly References available
Frame of Reference	Norm-referenced Learner compared against norms based on other test takers	Criterion-referenced Learner compared against specified criteria of achievement
Stakeholders	Learners, parents, principal receive results	Learners, parents, principal invited to help assess

## Conclusions

Measurement and assessment have different characteristics. Three characteristics of measurement are descriptive, quantitative and objective. Three different characteristics of assessment, on the other hand, are interpretative, qualitative, and objective. Assessment can be done when measurement has been done. As one of the CTL components, assessment is known as authentic assessment which means a process of collecting information or data from various sources during and after teaching learning process. A test, in the one hand, is the instrument which is used to know the ability of student in language skills.

The teachers and learners can assess writing as one of the language productive skills by taking into account their grammar, vocabulary, mechanics, fluency, and form. In addition, they can use independent and integrated writing rubrics as shown in Table 2 and Table 3 respectively. In other words, the teachers can use the appropriate and acceptable rubric to assess their students' writing skill by taking into account their English proficiency level.

## References

- Anonim. (2012). *The Propell® handbook for the TOEFL iBT® Test Workshop*. New Jersey: Educational Testing Service.
- Brown, HD. (2004). *Language assessment: Principles and classroom practices*. San Francisco: Longman.
- Cohen, Andrew D. (1994). *Assessing language ability in the classroom*. Second Edition. Boston: Heinle & Heinle Publishers.
- Hughes, Arthur. (1993). *Testing for language teachers*. Great Britain: Oxford University Press.
- Mardapi, Djemari. (2008). *Teknik penyusunan instrumen tes dan non tes*. Jogjakarta: Mitra Cendikia Press.
- Puhl, Carol A. (1997). Develop, not judge continuous assessment in the ESL classroom in *English Teaching Forum*. April 1997.
- Print, Murray.(1993). *Curriculum development and design*. Australia: Allen & Unwin Pty Ltd.
- Richards, Jack C., Platt, John and Platt, Heidi (1992). *Longman dictionary of language teaching & applied linguistics*. Longman: Longman Group UK Limited.

