

## Large-Scale Educational Assessment against Classroom Assessment: Pedagogy and Measurement Paradigm in EFL Classroom

**Md. Jahangir Alam**

Faculty of Languages & Translation  
King Khalid University  
Abha, Saudi Arabia

**Tahmina Aktar**

Community College for Girls  
King Khalid University  
Abha, Saudi Arabia

### Abstract

*While the assessment principles focus on the large-scale educational assessment, the classroom assessment remains underexplored in teaching, learning and assessment paradigm. Classroom assessment with all its individual features is characterized by the unique context that it accompanies. However, because of the differences it is not easy to apply the assessment principles borrowed from the large-scale assessment on classroom assessment (Fulcher & Davidson, 2007). Taylor and Nelon (1996:17) described this as 'misfit of the measurement paradigm'. This essay will note these points of challenges before conducting a literature review that explores how classroom assessment can be utilized as both an assessment and learning medium against the backdrop of large-scale assessment principles. The discussion will then be focused on EFL classroom in Saudi Arabia to have a contextual understanding of the topic. In the later part of this essay, we will construct a list of possible activities and exercises based on our new knowledge and findings.*

**Keywords:** classroom assessment, large-scale assessment, EFL teaching

### 1. Introduction

The phenomenon of 'test' is in existence since when teaching and learning took the central stage of human endeavor towards development. Test is a vital and equally sensitive element in language education. A well-designed test can provide valuable information about a student's progress as well as his current state and condition as a learner. Test, in simple terms, is a method of 'measuring a person's ability or performance in a given domain' (H. Douglas, 2003:3). In EFL context, a language test will elicit performance that will correspond to the language use situation outside the testing context (Bachman & Palmer, 2010). A test is a method or technique that requires performance or skill on the part of its takers. To be a test, the method has to be structured and it should measure the performance of the test takers. A methodically sound test will imply the test-takers ability or linguistic competence in a given domain. Therefore, a test is a measurement-criterion of performance whereas the outcome of a test is the implication of test-takers competence.

Assessment on the other hand, is a continuous procedure that includes substantially a wider domain. Testing is only one component in the assessment or evaluation process. Assessment is crucial for the teachers because it can provide valuable information, which can be utilized for future guidelines of the classroom practice (Rea-Dickens & Germaine, 1992). In an ideal situation, a good teacher never stops assessing his students in the process of teaching. Whenever a student answers a question, offers an opinion or tries out a new structure, teacher makes an assessment and a point of reference about that student for his next step of development. Most of our classroom assessments are constructed with a view to 'forming' students' competencies. Michael Scriven (1967) notably termed this kind of assessments as 'formative evaluation'. According to him, all assessments can be summative and have the potential to serve a summative function, but only some have the additional capability of serving formative functions.

### 2. Large-scale assessment against classroom assessment

Classroom assessment is fundamentally different from large-scale educational assessment because of its contextual implications. For obvious reasons, it is not always feasible to take principles from large-scale assessment and apply them in classroom setting. Because of these complexities, teachers often question: whether testing is relevant to their day-to-day work (Fulcher & Davidson, 2007). For a teacher, the nature of the classroom allows him to have a great deal of understanding about the aptitude, competence and capability of his students without even conducting a formal assessment. Students on the other hand, can enhance their understanding and experience in the classroom with the process of interacting and dealing with real-life situation and tasks.

As Moss (2003) described, classroom offers a social situation, where students develop understanding by interacting with the environment including all the other elements that it accompanies. In classroom setting this environment or context plays a pivotal role in shaping students’ understanding by offering learning experiences that eventually lead to language acquisition. In large-scale educational assessment, tests are designed with certain skill or language properties. In classroom assessment on the contrary, lesson is developed with learning objectives where any task can be an occasion for assessment of the teaching and learning procedures.

In classroom, many activities are need-based, developed instantly out of the perception of learning demand. For the obvious reason, the outcomes of the classroom cannot be tested by the traditional standard, rather assessed in terms of how successfully the tasks were delivered and received. This approach to (classroom) assessment overlooks the traditional standard of grading or ranking rather, it sets a criterion of what constitutes a successful task achievement. Under this criterion-referenced assessment student undertake the next task provided that they attained the preceding learning objectives. This kind of assessment may not always be accepted by the students or institutions. They may prefer the system of grading that is apparently more tangible kind of assessment, which shows learners’ position among a larger group of students.

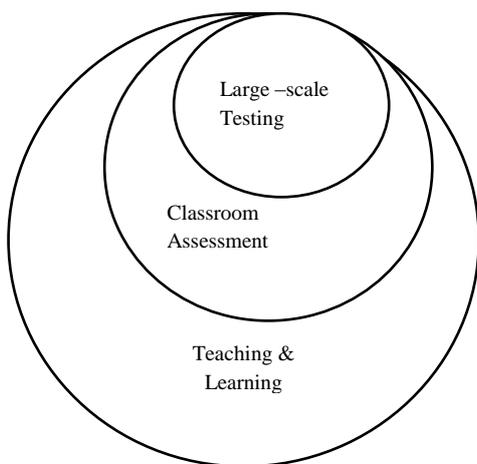


Figure: 1.1. Teaching, assessment and test. Adapted from D. Brown (2003), p.5

There is always this pressure of external large-scale testing criteria to be fitted in the classroom setting even though they do not represent the ideal form of assessment. In classroom the process of learning and personal experience of learners are two important components of assessment. Whereas, in large scale testing ‘the learner is a distant anonymous entity, about whom we need to collect as much evidence as possible in a very short period of time’ (Fulcher & Davidson, 2007:35). Classroom activities and assessment are mainly based on performance and the context of the classroom presents an environment similar to the outside world. Because of these characteristics, classroom assessment can play a vital role in language learning and testing paradigm. According to Bachman & Palmer (2010:40): one important feature of an ideal form of assessment should be... ‘the same characteristics of the language use situation that affect language use also need to be included in the kinds of assessment that we design or present to test takers’

**3. Tests and Assessments in EFL class**

Assessments are essential components and help shape language learning programs in Saudi Arabia. Saudi curriculum is rigid and prescriptive with no opportunity for teachers to develop, amend or improvise. However, classroom assessment constitutes a major component in the curriculum and syllabus design of EFL class. The main modes of assessment here being written assessments and structured oral assessments (Alhareth & Dighrir: 2014). In Saudi Arabia, assessments come in the form of traditional (standardized) tests which means, students usually do this type of test individually using paper and pencil to answer written questions which include gap fill, multiple choice, true or false and essays (Alsadaawi, 2010). Following table shows the assessment criterion for English language courses in one of the Saudi Arabian universities.

Semester-1				Total	Total Score
Continuous Assessment	Classroom Assessment	Class Tests Homework Quizzes	10%	50%	100%
	Midterm Exam		40%		
Final Exam	50%				

Table: 1.1 Assessment criterion (English language courses): KKU, Saudi Arabia

It is evident that, in an English language course, classroom assessment constitutes a significant part of the assessment standard and there is ample scope of improvement in this area. Based on the various studies conducted in this field, following observation can be pointed out which is universal to some extent:

Language teachers in many instances, lack adequate training in teaching field and they depend largely on the already created old assessment materials or the textbook based assessment. When no assessments are available teachers construct tests on their own, which are not in line with the standard assessment criterion and in some cases lack proper substance. Another problem with the semester-end-test is vagueness about the test pattern. Uncertainty about the topic on which students will be assessed is a universally common experience among students. A usual scenario is; they may have prepared themselves on a major assignment but teacher chooses a different one for assessment. As a result, students regard formal assessment as a matter of anticipation on which they do not have any control. Hence, there is a possibility of negative washback effect present in language classroom due to the large-scale assessment difficulties. As “*a poor test may be associated with positive effects and a good test with negative effects because of other things that are done or not done...*” in the classroom (Messick, 1996:242).

#### **4. Challenges with regard to classroom assessment**

There are a number of challenges that need to be addressed to make classroom assessment a success. Webb (2005) mentioned factors like organizational classroom set-ups, length of the periods, expectations of grade-based lesson content etc. Short duration of the classroom often poses challenges difficult to overcome and many teachers may succumb to this pressure and cut short the sustained learner engagement. In many occasions, a large class may prove to be a barrier on constructive feedback session at the end of the class. Additionally, in order to make classroom assessment effective, assessment principles (validity, reliability and fairness) should be upheld.

##### **I. Validity**

The most important and complex principle of an effective test is its validity. Validity in classroom assessment refers to the extent to which an assessment measures what it intends to measure (Maree, 2010). A valid reading test should measure reading comprehensibility not any other competence. In addition to this, three major types of validity should be maintained in classroom assessment. The first is *content validity*, which is an agreement in between curriculum objectives and assessment objectives. The second kind of validity is *criterion-related validity*, a classroom test designed to assess certain writing skill will have criterion validity if the test scores correspond to the subsequent observed behavior. The third form of validity is *consequential validity*, which refers to the way assessment is utilized in favor of teaching, learning procedure and learners. According to Fulcher and Davidson (2007), assessment has sequential validity to the extent that teachers can focus on activities, which will support learning and individual needs of students. McNamara (2000) pointed out, consequential validity is important because it emphasizes on the impact it can have on variables. Nevertheless, validity theory that has been originated mainly from large-scale test usually tend to overlook *context* that is essentially an inseparable and important part in classroom assessment. Context, under validity theory is therefore referred to as one part of *construct-irrelevant variance*. An alternative view to this is, context is not construct-irrelevant, but relevant to the assessment of the learners. Learners can better be assessed in relation to their involvement with the context (Fulcher & Davidson, 2007). Other important concerns in validity theory are, the *generalizability and comparability* of score meaning. That is, how far the test-score be meaningful beyond the test context. In other way, ‘lack of comparability across tasks is potentially serious threat to validity that could lead to invalid inferences concerning learners’ achievement (Brindley & Slatyer, 2002:370).

##### **II. Reliability**

Reliability is defined as the extent to which an assessment produces consistently similar results. An assessment can claim to be reliable when there is little or no contrast in students’ scores or ratings across different situations with different scorers (Brindley, 2003). As a result, reliability needs to be dependent on *performance* rather than scores that have no pre-set criteria (Lambert & Lines, 2000). In the large-scale assessment, it is expected that learners respond to as many different items as possible and the response to each item should be independent of the responses to others. Beside that, the scripts are expected to be graded anonymously as part of the standard procedure. Apparently, in classroom it is difficult to achieve because of the familiarity between learners and teachers. To face this challenge, teachers can gather much wider range of evidence for each student in a classroom.

##### **III. Fairness**

The issue of fairness is the most crucial among the challenges of classroom assessment. Lynch (2001) described ‘fairness’ as treating everybody in the same way and providing equal facility to everybody in contributing to an assessment process for demonstrating ability. In short, it is equality in designing and implementing an assessment procedure. While putting emphasis on fairness Brown (2005) pointed out, teachers would make sure that their personal belief should not interfere with the fairness of assessment. One example of fair assessment criterion is to maintain fairness in test; the content should be free of culturally unfamiliar ideas.

## 5. Suggested measures

- I. Assessment should be wide-range, which can provide a comprehensive view of students' competence. Assessment should not be an interruption to the teaching and learning process and be a goal in itself, as is currently the case in Saudi Arabia (Alhareth & Dighrir: 2014).
- II. Feedback should be embedded in classroom teaching and assessment. Teachers should facilitate learning by providing students with important feedback on their learning progress and by helping them identify learning problems (Bloom, et al, 1981).
- III. To generate a valid and reliable assessment procedure, teacher training needs to be brought in the forefront of teaching and assessment focus.
- IV. Teachers and administrators should work on the assessment procedures so that; it may have a *positive washback* effect on students. According to Messick (1996: 241) "For optimal positive washback there should be little, if any difference between activities involved in learning the language and activities involved in preparing for the test" Effective assessment procedure can produce positive washback: for example, an oral proficiency test can promote the teaching of speaking skills.
- V. Students, together with teachers and administrators should work to facilitate learning or competence building instead of simply craving for 'marks' or 'grades'. According to H. Douglas Brown: '*Grades and scores reduce a mountain of linguistic and cognitive performance data to an absurd molehill*' (2003:29)

## 6. Classroom assessment activities

Various types of traditional classroom assessments are in use in Saudi Arabia. For example: *Learner-Centered* classroom assessments are mainly to observe and improve learning, rather to improve teaching. *Teacher-Directed* assessments are dictated by teachers only, teachers decide what and how to assess and how to utilize the data received through assessment. *Mutually Beneficial* assessments are used when students give their opinion about the course content; teacher on the other hand, gains important insight about the content. *Context Specific* classroom assessments are utilized according to the requirement of a particular context. The *formative assessment feedback loop* involves perpetual and consistent interaction between teachers and students, with the ultimate goal being a metacognitive approach to teaching and learning. Both students and teachers should monitor learning progress and be involved in the process of data collection and analysis. Following are a few suggested classroom activities for EFL learners at intermediate level.

### Activity-1

#### Writing sample task:

Write a short paragraph explaining the life cycle of the *Camel*. Be sure to include a topic sentence. Use sensory details and domain-specific language to help your reader understand. Make sure your information is accurate and your sentences are complete. Use the task achievement guide to familiarize yourself with the checking process.

#### Task Achievement Guide

Sl.	Criteria	Status
1	Paragraph includes a topic sentence	
2	Sentences are complete	
3	Details address the topic	
4	Identifies main stages of the life cycle	
5	Sensory details provide clarity	
6	Includes accurate domain-specific language	
7	Link ideas using words, phrases and clauses	
*	Details are relevant	
*	Identifies sub-stages	

**Advanced/Exemplary:** Meets all the criteria including the star rated two criteria.

**Proficient:** Meets all the numbered (seven) criteria.

**Beginner:** Meets fewer than four criteria.

### Sharing Exemplar:

Teacher will distribute at least four sample paragraphs of similar kind that represent varying degree of quality (advanced, proficient, beginner) week before the actual test.

These can be students' assignment from previous years or teacher can produce sample from other sources. Students will be asked to match the assignment sample with the achievement criteria.

**Procedure:**

Teachers will explain the task and scripts-checking procedure in the previous class. They will explain further if any student has any query. On the day of classroom assessment, teachers will briefly discuss once again about the task. They will explain the sentence structure using adverb of sequence. Before handing the task over to the students, teacher will make it clear, what is expected of the students. During the whole procedure, teacher will concentrate on each student and try to judge the level of motivation among the students.

**Learning outcome**

Students are expected to master the technique of writing paragraph that is more than memorizing facts about something there familiar with. Students will know the criteria on which they will be tested and they will feel empowered in the classroom. Learners will be active and engaged participants in the classroom by being well informed about the assessment process.

**Rationale**

In this activity, students will write about a topic that they are familiar with, which is an important part of *content validity*. It can be easily assumed that the classroom condition (context) will be encouraging rather than intimidating for the students. Since, students know the topic of assessment and the evaluation procedure; they do not have to take extra pressure of the 'guessing game' of the large-scale assessment. Familiarity about the assessment process and motivation are expected to create *positive washback* among the students.

**Activity-2****Performance based speaking assessment:**

Provide students the statistic about road accident around the world. Tell the students to reflect on road accident in Saudi Arabia. Ask them to jot down their notes. After five minutes, teacher asks student: What are the possible reasons of higher rate of road accident in Saudi Arabia?

Road Death		
Country	Per 100000 people	Per 100000 vehicles
Saudi Arabia	49	146
USA	11.6	13.6
New Zealand	7.4	10.3
Australia	5.6	7.6

**Task assessment rubrics**

Criteria	Poor	Beginner	Proficient	Exemplary	Status
Task completion	inadequate	limited	adequate	extensive	
Fluency	limited	some	satisfactory	well developed	
Pronunciation	unclear	partially clear	mostly clear	very clear	
Vocabulary	limited	adequate	sound	extensive	
Accuracy	poor	fair	good	excellent	

**Procedure:**

Teachers have to establish a clear set of assessment criteria at the beginning of this kind of task. Teachers can work in a group with other teachers to set up the criteria. In this particular performance based task of speaking skill, teacher will try to elicit a satisfactory speaking attempt from the students. Teacher can do that by discussing on what would be satisfactory answer to this kind of task in the class. Teacher can explain each assessment standard and then distribute the task assessment rubrics. In this kind of assessment, teachers have to be cautious about the time and carefully plan the class so that they can have a feedback session at the end. At the same time, teachers will keep in mind that, this kind of tasks are prone to produce varied responses. Therefore, teachers have to be careful while they match assessment standards with performance.

**Learning outcome**

Learning outcome of this kind of formative assessment task is improvement in speaking competence by error analysis in the form of pronunciation, grammatical accuracy and fluency. By observing, other students' performance students will understand where they stand in the assessment standard. If motivated, students will set a goal of future improvement for themselves.

## Rationale

Students demonstrate knowledge and skill in performance-based classroom assessment. This is the application part of what they already know. In this kind of assessment, learning comes ahead of performance-testing. Teacher assess students to know the impact of assessment in the classroom or whether students are progressing or not. Since there is no right or wrong answer, students can come up with their own answer and teachers can have clearer picture of the class with all the diverse answers. The advantage of this kind of assessment is, teacher can assess students on various levels. One example of something that the teacher can assess in this procedure is how students interact with other students (Stiggins, 1994). Another aspect of assessing performance is the probability of making mistakes, which can be a platform for the teachers to work on improvement. Some assessment experts believe, students learn nothing from a performance, which is perfect. Rather, they learn best when their initial performance is less than perfect. For then, they can identify the required direction to which to improve (Wiggins, 1998).

## Activity-3

### One-sentence summary reading assessment:

Galileo Galilei was born in the year 1564 in the town of Pisa, Italy. When he was 20 years old, he was studying in Pisa. His father wanted him to be a doctor, but Galileo was bored with school except for math. Because math was the one subject where he was doing well, the court mathematician offered to tutor him privately so he could become a qualified mathematician. Galileo's father was disappointed, but he agreed. Because he needed to earn money, Galileo began experimenting with different things, trying to come up with some sort of invention that he could sell for money. He had a little bit of success with his invention that was like a compass that could be used to measure plots of land. He had already experimented with pendulums, thermometers, and magnets. When he heard that a Dutch inventor had invented something called a spyglass, but was keeping it a secret, Galileo decided to work on one of his own. Within 24 hours, he had invented a telescope that could magnify things to make them appear ten times larger than real life. One night, he pointed his telescope toward the sky, and made his first of many space observations: the moon was not smooth, like everyone thought. The moon was covered in bumps and craters. As technology has improved, first Galileo, and then many others, have made improvements on the telescope, the wonderful device that allows us to see from a distance.

Source: <http://www.k12reader.com/worksheet>

#### Question Matrix

Who?

When?

Does what?

Where?

Why?

#### Synthesize answers in one sentence

## Procedure

In this task, students will be given a reading text of a particular topic with few questions. Then they will be asked to find their answers in the reading passage and synthesize their answers in a single grammatically correct lengthy sentence. As the answers to the questions are varied so it is expected that the summary sentence will be diverse. Students can exhibit their creativity in answering the questions. Teachers can assess students on different categories. They can create an assessment matrix of their own to find out which question poses most difficulty to the students.

## Learning outcome

Students are expected to know how to connect ideas and concise a large passage in one sentence. They will know how to make grammatically correct complex sentences. They will have an idea about coherent and cohesive devices in sentence. Once they finish, their answers will be reviewed by their peers and later by the teacher.

## Rationale

This is an example of *formative* assessment. The purpose of this assessment is to impart knowledge rather than simply assessing. As Brookhart (2003:7) argues, assessment and learning are integrated within the classroom. On the other hand, students will receive *feedback* from their peers and from the teachers. Teachers will provide *descriptive feedback* in simple terms, so that students understand their weak points. Instead of simply labeling students' errors with red marks, feedback should guide students to better performance (Guskey, 2003)

## Conclusion:

Assessments are essential components of teaching and learning cycle and certainly, they play a crucial role in shaping individual learning. Assessments in Saudi Arabia are still perceived traditionally as a testing device and only criteria to measure students' achievement. Like all the other classrooms all over the world, classrooms here come with all its complex characteristics and present a good opportunity to assess for the purpose of learning.

## Appendix-1 Classroom assessment planning sheet

### References

- Alhareth, Y. A., & Dighrir, I. A. (2014). The Assessment Process of Pupils' Learning in Saudi education system: A Literature Review. *American Journal of Educational Research*, 2(10), 883-891.
- Alsadaawi, A., (2010), Saudi national assessment of educational progress (SNAEP), *International Journal of Education Policy and Leadership*, 5(11): pp. 1-14.
- Bloom, B. S., Madaus, G. F., & Hastings, J. T. (1981). *Evaluation to improve learning*. New York: McGraw-Hill.
- Brindley, G. & Slatyer, H. (2002). Exploring task in ESL listening assessment. *Language testing*, 19(4)
- Brookhart, S. M. (2003). Developing measurement theory for classroom assessment purposes and uses. *Educational measurement: Issues and practice*, 22(4), 5-12.
- Brown, H. D. (2003). *Language assessment: Principles and classroom practices*. NY: Pearson Education. p. 3
- Brown, J.D. (2005). *Testing in language programs. A comprehensive guide to English language assessment*. New York, McGraw-Hill ESL/ELT
- Fulcher, G & Davidson, F. (2007). *Language testing and assessment*. NY: Routledge
- Guskey, T.R. (2003). How classroom assessment improve. *Educational Leadership*, 60, 6-11.
- Lynch, B.K. (2001). *The ethical potential of alternative language assessment*. Cambridge University Press, pp. 228-239
- Maree, J.G. (2010). Critical appraisal of the system of education and prospects of meeting the manpower and development needs of South Africa. *Africa Insight*, 40, (2).
- McNamara, T.F. (2000). *Language testing*. Oxford: Oxford University Press
- Messick, S. 1996. 'Validity and washback in language testing'. *Language Testing*. 13/4:241-56.
- Moss, P. (2003) 'Reconceptualizing validity for classroom assessment.' *Educational Measurement: Issues and Practice* 22, 4, 13-25.
- Rea-dickins, P & Germaine, K. (1992). *Language teaching: Evaluation*. Oxford: Oxford University Press.
- Scriven, M. (1967). The methodology of evaluation. R.W. Tyler, R M. Gagne, M. Scriven (eds.), *Perspectives of curriculum evaluation*, pp.39-83. Chicago, IL: Rand McNally.
- Stiggins, R. J. (1994). *Student-centered classroom assessment*. New York: Macmillan Publishing Company.
- Taylor, C. S. and Nolen, S. B. (1996) 'What does the psychometrician's classroom look like? Reframing assessment concepts in the context of learning.' *Educational Policy Analysis Archives* 4, 17.
- Webb, D.C (2005). *Classroom assessment as a research context: variations on a theme of pedagogical decision making*. Madison. WI: University of Wisconsin Press.
- Wiggins, G. (1998). *Educative assessment*. San Francisco: Jossey-Bass.

**Appendix-1****Classroom assessment planning sheet****Planning**

A clear idea about what you are looking for will help you to determine which technique to use in the classroom, and how to interpret the results. Therefore, before using any technique in the class, complete a planning worksheet.

Sl.	Queries
1	What do I want to know from today's class? Answer:
2	Which assessment technique will I use to get this information? Why? Answer:
3	How will I implement this technique to my students? Answer:
4	How much class time will it take? Answer:
5	How will I know if the technique was successful? Answer:
6	What instructional changes will I make as a result of the information that I get? Answer: