

Classroom and Formative Assessment in Second/Foreign Language Teaching and Learning

Somaye Ketabi

Department of English, Faculty of Foreign Languages, University of Isfahan, Iran

Saeed Ketabi

Department of English, Faculty of Foreign Languages, University of Isfahan, Iran

Abstract—Formative assessment was first used by Scriven (1967), but studies focused on formative assessment in ESL and EFL classes started in 2000. Formative assessment is integrated in learning and teaching, so a great part of classroom assessment is formative. However, classroom assessment has been not defined clearly, as most of the texts about this type of assessment define it in terms of its formative or summative potentials (Fulcher & Davidson, 2007). This study is an attempt to define different types of assessment, compare formative and classroom assessment, and also identifies the areas which need more attention by researchers.

Index Terms—assessment, formative assessment, classroom assessment, English as a Foreign Language (EFL), English as a Second Language (ESL)

I. INTRODUCTION

Several researchers have shown that formative assessment is essential in ESL/EFL teaching and learning (Bachman, 1990; Gattullo, 2000; Rea-Dickins & Gardner, 2000). As Cheng, Rogers, and Hu (2004, p. 361) stated: “every model of the teaching-learning process requires that teachers base their decisions—instructional, grading, and reporting—on some knowledge of students’ attainment of and progress towards desired learning outcomes”. Therefore, teachers should continuously assess their students in order to help teaching and learning progress (Fulcher & Davidson, 2007). So, formative assessment is mainly conducted in the context of classroom. But what is done in the classroom (with the purpose of assessing the students) is not always in favor of forming students’ learning; in another word, assessment practices in the classroom are not formative all the time. For instance, as Brown (2004) categorized, classroom assessment can have diagnostic or achievement purposes, or even measure proficiency purposes. In order to discuss about different aspects of each kind of assessment, this article first presents definitions of assessment and then different types of assessment. At last, an overview of the effects of formative assessment on EFL/ESL teaching and learning will be presented.

II. DEFINITIONS OF ASSESSMENT

When students and some teachers think of assessment, they might remember the pressure of final exams. Unpleasant memories of several exams in a couple of days usually make students’ faces fall, and thinking about preparing and marking exams make teachers feel exhausted. But the fact is that, as Brown (2004) stated, testing and assessment are different.

Bachman (2004) defined assessment as “a process of collecting information about something that we are interested in, according to procedures that are systematic and substantially grounded” (pp. 6-7). The result of an assessment procedure can be a score or a verbal description. Ari Huhta (as cited in Spolsky & Hult, 2008) referred to assessment as “all kinds of procedures used to assess individuals (e.g., informal observations, self-assessments, quizzes, interviews, tests)” (p. 469). Teachers assess their students every session. However, testing is a way of conducting assessment which is technically associated with definite timing and settled procedures (Brown, 2004). As Huhta defined “tests denote a particular type of formal, often carefully designed instruments” (Spolsky & Hult, 2008, p. 469).

When teachers consider an assessment task, they usually have some questions in their mind such as: “*When and how often shall we assess the students?*”, or “*How should we conduct an assessment procedure?*” The question of “*What*” and “*Why*” rarely come to teachers’ mind (Bachman & Palmer, 2010). The reason of not asking *what*-question is quiet clear as teachers usually know what they want the learners to learn. However, it is somehow vague that why teachers do not ask *why*-questions; they either know the answer or seldom consider the reason of assessment (Ibid). *Why*-question is important since it will define the decision to be made about the outcomes of an assessment.

The very first use of language assessment is to make decisions for individuals (micro-evaluation), programs (macro-evaluation), and other stakeholders (Bachman & Palmer, 2010). It can be used to select individuals, place them into appropriate course of study, make changes in instruction, predict future performance of test-takers, make changes in educational programs (formative or summative decisions), to formulate new research questions, and modify the

understanding of a specific language phenomenon (Bachman, 2004). Therefore, the decision going to be made is so essential that it can define type of assessment.

In the following sections, classifications of assessment are presented based on the decisions which are going to be made according to their outcomes.

III. FORMAL VS. INFORMAL ASSESSMENT

Formal assessments are systematically planned and designed to get information about students' achievement in predetermined times. Brown (2004) described formal assessment as tournaments in tennis, and made a distinction between formal assessment and testing. He considered all kinds of tests as formal assessment, but this is not true vice versa; i.e. formal assessment is not necessarily performed as tests. He associated tests with time constraints which is not always the case in formal assessment. For example, systematic observation of students' oral performance in a kind of formal assessment, but it is hardly called a test which is limited to specific time and gathers limited pieces of information (Brown, 2004).

However, Harris and McCann (1994) used test as a synonym for formal assessment. Tests are taken in order to

- a) understand whether a student is ready to go to next level,
- b) know about problematic areas,
- c) figure out what the students have learnt, and
- d) compare the students. (Harris and McCann, 1994)

What is problematic about tests or formal assessment is that they are considered as enemies to students' competence for students are usually afraid that they will not perform successfully in tests (Cohen, 1996, as cited in Celce-Murcia, 2001). Students and teachers are suspicious of tests, and mostly they cannot see any relationship between formal tests and teaching and learning progress (Brown, 2004). But if it is justified to both teachers and learners that tests are employed as complements to other forms of assessment (e.g. informal assessment), they can be appropriate means to objectively assess students' knowledge (Harris and McCann, 1994). A fair and reliable decision of students' performance and learning cannot be made by using the results of testing or formal assessment, rather it can be obtained by the help of informal and self assessment (Ibid).

Informal assessment, based on Brown's definition (2004), includes occasional and unplanned comments and feedbacks. The teacher does not design informal assessment before the class. Results of this kind of assessment are not recorded and no judgment is made based on them. According to Brown, informal assessment consists of various types of feedback; from simply saying "*Nice job!*" to giving some detailed comments about students' performance (Brown, 2004). Informal assessment is merged with every second of teaching process as teachers are always giving feedback to students. So, informal assessment, according to Brown's definition, is more about giving feedback rather than deciding upon students' performance.

However, Harris and McCann (1994) presented a different definition for informal assessment:

Informal assessment is a way of collecting information about our students' performance in normal classroom conditions. This is done without establishing test conditions such as in the case of formal assessment. Informal assessment is sometimes referred to as continuous assessment. (p. 5)

Some phrases and words in this definition are of high importance. The first one is *collecting information*. The main aim of informal assessment is not collecting scores but rather gathering pieces of evidence about students' knowledge. Another important word is *performance*. Students actually perform the task and the teacher judge their performance by direct observation. This performance is not done under pressure; the students are not as anxious as they are at the time of formal tests. The third important phrase is *normal classroom condition*. Unlike formal assessment, informal assessment is done without establishing time limit and formal tests' rules. The last important word is *continuous*. Good teachers are always assessing their students. Formal tests are not the only reliable source for getting information, and additional data provided by day-to-day observation in a stress-free condition is needed to reach a reliable decision about students' performance.

It is essential to mention that, as Harris and McCann (1994) pointed out, unplanned and unsystematic informal assessment cannot be effective. Daily observation which is not systematic will result in a pile of useless information. Teachers should precisely define the abilities to be assessed, for it is impossible to assess students' performance every session. In addition, rough impressions are not enough to reach a trustworthy decision, and student should be assessed based on a criteria. Finally, the relationship between informal assessment and other kinds of assessment should be clearly defined (Harris and McCann, 1994).

IV. SUMMATIVE VS. FORMATIVE ASSESSMENT

Summative assessment, as its name suggests, summarizes what the students learnt during a course and it is usually done at the end of the semester (Brown, 2004). This kind of assessment shows what objectives have been accomplished, but it lacks feedback or any suggestion to improve performance. Final exams or proficiency tests are examples of summative assessment (Ibid). Alderson (2005) associated summative assessment with long traditional tests which were so stressful to students. Any kinds of test which lacks further feedback and the only possible use of it is gathering scores

in the eyes of students can be summative even if teachers have primarily designed the test to facilitate learning and teaching.

On the other hand, formative assessment, as Lewy (1990) confirmed, does not have a precise definition. Formative assessment takes place during learning and is aimed to help learning and teaching by giving appropriate feedback (Lewy, 1990). Nitko (1993) named two purposes of formative assessment: (a) selecting or modifying learning procedures, and (b) choose the best remedies for improving weak points in learning and teaching. Gattullo (2000) characterized formative assessment as “(a) it is an ongoing multi-phase process that is carried out on a daily basis through teacher–pupil interaction, (b) it provides feedback for immediate action, and (c) it aims at modifying teaching activities in order to improve learning processes and results.” (p. 279). Most of classroom assessment is formative and students *form* their knowledge by analyzing and internalizing teachers’ comments (Brown, 2004).

It seems that formative assessment has not always been the focus of attention in ESL/EFL studies. Before 2000 a number of studies were done about classroom assessment in regular school programs (Rogers, 1991; Wilson, 2000), but very few studies were conducted about this topic in EFL/ESL context (Cheng, Rogers, & Hu; 2004). Rea-Dickins and Gardner (2000) pointed to this neglect too and said that in compare to other topics in language testing, formative assessment had received less attention. However, it should be said that Bachman (1990) was one of the first scholars who discussed about the complexities and difficulties of formative assessment. He stated that types of feedback received by the students could affect the results of future formal tests. Bachman put more emphasis on was formal tests, and the construct of formative assessment were not discussed by him. In his later book with Palmer (1996), he focused on feedback and the relation between “formative evaluation” (p. 98) and formal tests. Shohamy (1995) was another scholar to take an initiatory step in discussion of formative assessment. He named some methods such as portfolios and projects which teachers used to put less reliance than on formal tests and to capture different aspects of language competence. However, the construct and practices of formative assessment in EFL/ESL context were roughly discussed before 2000 in spite the fact that its usefulness and help had been long been recognized by both teachers and researchers.

Students should use the language in order to learn it, and if they are graded all the time, they do not have the opportunity to do so. They should receive feedback, analyze it, and have the chance to test their hypotheses based on the feedback received. This is the very basic requirement to learn a language (Brown, 2004; Harris and McCann, 1994). Summative formal assessment makes use of traditional paper-and-pencil tests and is just followed by scores without any further feedback. They are usually done at the end of a course which is so stressful for students and teachers. Lack of feedback results in lack of diagnostic information, and students do not clearly know about their weak points. But, students’ performance highly depends on appropriate feedback from the teacher which is the defining feature of formative assessment. Teachers can make use of formative assessment to prevent negative washback effect of formal testing; which is the separation of teaching and learning in the eyes of students.

V. EXPLICIT VS. IMPLICIT ASSESSMENT

As Bachman and Palmer (2010) mentioned, sometimes the distinction between teaching and assessing is not well-defined. In a classroom condition, a teacher teaches, conducts assessment, and decides based on the outcomes, and then s/he starts teaching based on the decisions made and assesses the students. Students are not aware that teacher is assessing them continuously, and the teacher may not call this procedure an assessment. Meanwhile, this kind of assessment can help learning and teaching. When the learners (and sometimes even the teacher!) are not aware of assessment, implicit assessment is taking place (Bachman and Palmer, 2010).

Explicit assessment, nonetheless, occurs when teacher clearly announce the assessment (Ibid). The students and teacher both know that the process which is going on is an assessment. Explicit assessment can be either formative (in which the teacher is concerned about giving feedback) or summative (whether the objectives of the course have been satisfied or not).

TABLE 1:
IMPLICIT AND EXPLICIT ASSESSMENT

Mode	Characteristics	Purposes
<i>Implicit</i>	Continuous Instantaneous Cyclical Implicit: both teacher and students may be unaware that assessment is taking place	Formative decisions, e.g.: Correct or not correct student's response Change form of questioning Call on another student Produce a model utterance Request a group response
<i>Explicit</i>	Clearly distinct from teaching Explicit: both teacher and learners are aware that assessment is taking place	Summative decisions, e.g.: Decide who passes the course Certify level of ability Formative decisions, e.g.: Teacher: Move to the next level or review current lesson Teacher: focus more on a specific area of content Student: spend more time on particular area of language ability Student: use a different learning strategy

**Bachman & Palmer, 2010, p. 29*

VI. CLASSROOM ASSESSMENT

Defining classroom assessment is difficult since all the aforementioned types of assessment can be done in the classroom. Teachers can conduct proficiency, placement, or even aptitude tests in the classrooms (Brown, 2004). Observation of students' performance in the classroom can be graded or not; their interaction with teacher and students can be judged to rank them or not. Defining classroom assessment based on the means and decisions of assessment tasks. Using Bachman and Palmer's (2010) terminology, one can say that classroom assessment can be either explicit (concerned about summative decisions) or implicit (focused on formative decisions). Figure 1 roughly presents the relationship between summative, formative, and classroom assessment.

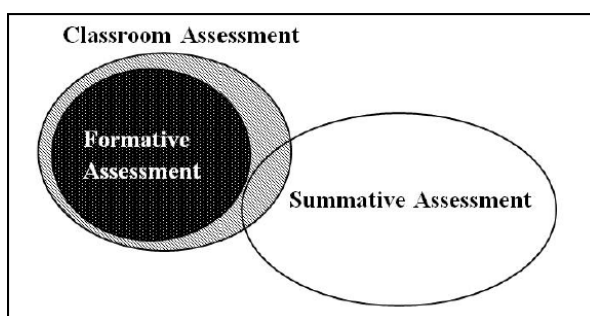


Figure 1: Summative, Formative, and Classroom Assessment

Summative assessment can be a part of classroom assessment if teachers consider gathering scores as the most important aim of assessment in the classroom and do not provide further feedback for the students. At times, when teachers are forced to collect scores for the administration center, or when the syllabus imposes final or midterm exams with the purpose of measuring attainment of achievement, summative assessment is conducted in the classroom. It might be assumed that even these tests are formative if providing feedback is focused, but stressful situation created by these exams hinders the students to perform as they want. This problem is mostly observed in crowded classes where teachers do not have enough time to observe all the students and provide feedback form them. In these classes the most feasible way to collect information on students' performance is supposed to be tests. Better understanding of classroom assessment is reached by the use of Black and William's (1998) terminology of classroom. They used *black box* to refer to classroom, since:

Certain *inputs* from the outside –pupils, teachers, other resources, management rules and requirements, parental anxieties, standards, tests with high stakes, and so on– are fed into the box. Some *outputs* are supposed to follow: pupils who are more knowledgeable and competent, better test results, teachers who are reasonably satisfied, and so on. (p. 1)

Therefore, different tasks are done in the classroom. The ideal condition is when all the assessment tasks in the classroom are formative and *for-learning*. As assessment is integrated with learning, most of classroom practices are formative. But there is some room for other kinds of assessment, too.

VII. EFFECTS OF FORMATIVE ASSESSMENT

Assessment is more apparent in the form of summative testing in educational world rather than informal formative tasks. Students look for ways to be the "top student" through assessment and they rarely try to see a connection between

assessment and learning. However, Black and William's extensive review of literature (1998) has changed the view towards assessment.

Black and William (1998) reviewed more than 250 articles and books to examine the effect of formative assessment on students' performance. They found that employing different methods to assess the student during the course would enhance students' success. The works on formative assessment suggest that new methods to increase effective feedback will change classroom practices and bring adjustment in learning and teaching. Motivation and self-assessment are also encouraged in formative assessment (Black & William, 1998).

Nevertheless, in spite of the clear benefits of formative assessment, techniques which enhance this type of assessment have usually been neglected because of practical problems in the classrooms. Marking problems (generous or unfocused), unrealistic practices, teacher-made tests which reflects standardized tests format, lack of negotiation among teachers about formative assessment, overemphasis on giving scores while useful feedback is neglected, norm-referenced assessment in which low-achieving students start to believe they cannot learn, and lack of care about students' performance in previous classes are just some problems mentioned in Black and William's review. Furthermore, if pressure of various external tests is added, there will be no room for actual formative assessment in classroom routine!

The problems mentioned in Black and William's study make scholars to examine formative for-learning assessment in all educational areas. Although scholars such as Bachman (1990) and Shohamy (1998) highlighted the role of feedback in teaching and feedback, studies which aimed at formative assessment in EFL/ESL classes have been started since 2000.

Rea-Dickins and Gardner (2000) are one of the first scholars who focused on features of formative assessment. They interviewed EAL (English as an Additional Language) teachers and observed their classes and found that decisions made based on students' performance in classroom were important. These decisions, if made carelessly, could not show the true level of students' language (Rea-Dickins & Gardner, 2000). They also questioned the reliability and validity of classroom assessment and mentioned that different procedures should be employed to assess these elements. In fact, "the validity of the inferences about the way in which the children use language seems to be equally important, whether the aim is *assessment for learning*, i.e., formative assessment or *assessment for establishing levels of achievement*, i.e., summative and evaluative" (Rea-Dickins and Gardner, p. 238, italics in original). Therefore, the construct of formative assessment and the skills teachers need to know in order to conduct this kind of assessment were first opened up for discussion.

Gattullo (2000) conducted a case study in which he observed four teachers' performance in applying formative assessment in their classes. He focused on children classes and reported the methods used. His study was a kind of survey in which the focus was just on children's classes and the techniques used in other age groups were not considered.

Assessment Reform Group (2001) and Qualifications and Curriculum Authority (2001) set some principles for classroom assessment practices which also influenced ESL/EFL world. They emphasized on continual nature of formative assessment and its integration with learning. Considering these principles, Leung and Mohan (2004) conducted a case study in EAL schools. Through observation of different classes, they concluded that teachers and students could decide on (rather than guess) further stages of learning and weak and strong points by the use of discussion, peer assessment, and formative techniques of assessment rather than standardized testing (Leung & Mohan, 2004). However, they mentioned that more investigation was needed about classroom assessment.

Cheng, Rogers, and Hu (2004) conducted a study on classroom assessment in three language settings (Canada, Beijing, and Hong Kong). These researchers suggested that less is known about classroom assessment of EFL/ESL instructors. This issue would become more essential if large number of students, influence of formal testing, and central role of this kind of assessment in learning and teaching processes is considered (Cheng, Rogers, and Hu, 2004). In their survey, purposes, procedures, and methods of classroom assessment were investigated. Through purposive sampling, university teachers from three different language settings (English dominant, bilingual, and Mandarin-dominant) were chosen. Altogether 461 questionnaires were gathered, but 267 questionnaires were returned to the researchers. The findings showed some complexities of classroom assessment which vary within different settings. Factors such as nature of the courses, teaching experience of instructors, levels of students, and the role of external testing can also influence classroom assessment (Cheng *et al.*, 2004). For instance, instructors in Hong Kong identified to have fewer assessment purposes and used less objective methods of scoring. Cheng *et al.* suggested that it could be explained by the fact that teachers in Hong Kong are more experienced than teachers in Canada and Beijing.

In another study, Ke (2006) first identified language skills of 222 Chinese adult learners and proposed a model for formative task-based language assessment. This model is criterion-referenced, skill-integrated, and driven from classroom activities (Ke, 2006). Moreover, the researcher took into account curriculum objectives and task-based instruction in order to design the model.

Wei (2010) conducted an action research and showed that formative assessment was significantly effective in promoting learners' motivation and performance. He suggested some hints to improve the quality of formative assessment such as performing needs analysis and formulating assessment plan. In this research too, there was no indication of better techniques in performing formative assessment.

VIII. CONCLUSIONS

Formative assessment has long been recognized as one of the most influential methods to improve learning and teaching (Bachman & Palmer, 1996; Shohamy, 1998; Rea-Dickins & Gardner, 2000; Davison, 2002; Cheng *et al.*, 2004; Ke, 2006; Wei, 2011). Many tools of formative assessment have been introduced since then; such as journals, portfolios, surveys, oral interviews, and presentations. Classroom assessment, on the other hand, has potentials to be formative and for-learning. However, Black and William (1998) showed that practical problems in classrooms can hinder effective formative assessment to be conducted. Therefore, by considering the importance of formative assessment in ESL/EFL teaching and learning, the difficulties and complexities teachers face in applying formative assessment in their classes, and lack of studies on this area, scholars should also focus their attention to the problems of making classroom assessment as formative as possible and suggesting ways to improve classroom assessment.

REFERENCES

- [1] Alderson, J. Charles. (2005). *Diagnosing foreign language proficiency: the interface between learning and assessment*. London: Continuum.
- [2] Bachman, L. (2004). *Statistical analysis for language assessment*. Cambridge: Cambridge University Press.
- [3] Bachman, L. and Palmer, A. (2010). *Language Assessment in Practice*. Oxford: Oxford University Press.
- [4] Bachman, L.F. & Palmer, A.S. (1996). *Language testing in practice*. Oxford: Oxford University Press.
- [5] Bachman, L.F. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.
- [6] Black, P. & Wiliam D. (1998). Inside the black box: raising standards through classroom assessment. *Phi Delta Kappan*.
- [7] Brown, H. D. (2004). *Language assessment: principles and classroom practices*. New York: Pearson Education.
- [8] Cheng, L., Rogers, T., and Hu, H. (2004). ESL/EFL instructors' classroom assessment practices: purposes, methods, and procedures. *Language Testing*, 21 (3), 360-389.
- [9] Cohen, A. D. (1996). *Second Language Assessment*. In Celce-Murcia, M. (Ed.). (2001). *Teaching English as a second or foreign language* (3rd ed.). New York: Heinle & Heinle.
- [10] Fulcher, G. & Davidson, F. (2007). *Language testing and assessment: an advanced resource book*. New York: Routledge.
- [11] Gattullo, F. (2000). Formative assessment in ELT primary (elementary) classrooms: an Italian case study. *Language Testing*, 17 (2), 278–288
- [12] Harris, M. & McCann, P. (1994). *Assessment (Handbook for the English classroom)*. Oxford: Heinemann Publishers.
- [13] Huhta, A. (2008). Diagnostic and Formative Assessment. In Spolsky, B. and Hult, F. M. (Ed.) (2008). *The Handbook of Educational Linguistics*. Oxford: Blackwell.
- [14] Lewy, A. (1990). Formative and summative evaluation. In Walberg, H. & Haertel, G. (Eds.), *The International Encyclopedia of Educational Evaluation*, 26-28.
- [15] Nitko, A. J. (1993). Designing tests that are integrated with instruction. In Linn, R. L. (Ed.), *Educational Measurement*, 447-474. Phoenix, Arizona: Oryx Press.
- [16] Rea-Dickins, P. & Gardner, Sh. (2000). Snares and silver bullets: disentangling the construct of formative assessment. *Language Testing*, 17(2), 215-243.
- [17] Rogers, T. (1991). Educational assessment in Canada: Evolution or extinction? *The Alberta Journal of Educational Research*, 37(2), 179-192.
- [18] Scriven, M. (1967). The methodology of evaluation. In Tylor W. R., Gagne, R. M. and Scriven, M. (Eds.). *Perspectives of Curriculum Evaluation*. Chicago: Rand McNally.
- [19] Shohamy, E. (1995). Performance assessment in language testing. *Annual Review of Applied Linguistics*, 15, 188-211.
- [20] Wilson, R. J. (2000). A model of assessment-in-practice. In Cheng, L., Rogers, T., and Hu, H. (2004). ESL/EFL instructors' classroom assessment practices: purposes, methods, and procedures. *Language Testing*, 21 (3), 360-389.

Somaye Ketabi received her BA degree in English Translation from the University of Isfahan in 2010. She is currently an MA student in Teaching English as a Foreign Language (TEFL) at the University of Isfahan. Her areas of interest include language testing and classroom discourse.

Saeed Ketabi is an associate professor of applied linguistics at the University of Isfahan. He has published numerous articles in the area of English language teaching and learning. His main areas of interest are English Teaching Methodology and Materials development.