A Synoptic Review on Alternative Assessment in Foreign Language Teaching and Learning

Layla Aknouch

ORCID: 0000-0002-3534-6867

*Email: Layla.aknouch@usmba.ac.ma

Abstract

The current paper aims to review the literature and establish a theoretical framework for alternative assessment. Specifically, the paper seeks to clarify what is meant by alternative assessment, define its characteristics, as well as explain its advantages in teaching and learning English as a second or foreign language. The paper also discusses some of the challenges that face the implementation of different alternative assessment forms, and suggests ways of dealing with such challenges. Generally speaking, the present paper seeks to expand awareness of alternative assessment among novice English as a foreign language (EFL) practitioners arguing that the learning of English as a second or foreign language is better enhanced when both alternative and traditional assessment methods are used congruently. In this way, data that are gathered by means of traditional assessment tasks should be used to complement and consolidate data that are collected using alternative assessments.

1.   Introduction

The shift in teaching methodologies towards learner centered learning has created new challenges for foreign language teaching. The main challenge is to engage learners in all aspects of the teaching and learning process, making them active participants rather than passive recipients. One aspect of language teaching and learning that has received considerable attention as a result of this shift is assessment.

Conventional methods of assessment like fill-in-the gap tasks, multiple-choice exercises, and comprehension questions were considered to be the most effective methods of assessment in the EFL context, in the past. This is because these techniques paralleled the classroom practices adopted. Nevertheless, traditional assessment techniques have become incongruent with today's EFL teaching methodologies which recommend more students' participation, increase students' motivation and provide a clear understanding of students' achievement. A number of researchers in the field of education (Bailey 1998, Baker 2010, Hamayan 1995, Brown & Hudson 1998, Stiggins 2002, etc.) have argued that traditional assessment techniques are inadequate in assessing the numerous dimensions of language learning. This paved the way to the introduction of assessment techniques that are consistent with learner/learning centered approach. As a result, new types of educational assessment rose to prominence during the 1990's in an attempt to compensate for the limitations of traditional assessment. Hence, the term "alternative assessment" was suggested as a reforming movement and a necessary move from conventional assessment. Accordingly, this paper aims to answer the following questions:
1.   How is alternative assessment defined in the literature?
2.   What are the main features of alternative assessment?
3.   What are the benefits of alternative assessment?
4.   What challenges hinder the use of alternative assessment?

To answer these questions, the paper starts by explaining what is meant by alternative assessment and listing its distinguishing features. Moreover, the paper stresses the importance of alternative assessment by discussing some of its benefits for the language learner and for the teaching and learning processes in general. The paper also summarizes the challenges that hinder the effective incorporation of alternative assessment forms in the language classroom and suggests ways of handling those challenges.

2.   Definition of Alternative Assessment

O'Malley & Pierce (1992) define alternative assessment as "any method of finding out what a student knows or can do that is intended to show growth and inform instruction and is not a standardized or traditional test" (p. 2). For Hancock (1994, p. 3), alternative assessment is "an ongoing process involving the student and teacher in making judgments about the student's progress in language using non-conventional strategies". Similarly, Berry (2008, p. 93) describes alternative assessment as "any method that differs from conventional paper-and-pencil tests, most particularly objective tests".

As such, alternative assessments are the procedures and techniques which can be used within the context of teaching and learning to assess students' performance and achievement, and that can be easily incorporated into the daily classroom activities. It is an alternative to standardized testing and all the weaknesses that such tests proved to have. A variety of assessment information is continually offered to teachers during the teaching and learning process. Hence, a teacher's record of any formal or informal assessment data that can be generated from classroom and school activities is alternative assessment.

There has been a variety of labels which were developed as a way of distinguishing between alternative assessment and conventional testing. The terms that are frequently used in the literature are 'alternative assessment', 'authentic assessment', and 'performance assessment'. When this type of assessment was first introduced as an alternative to traditional standardized assessments, many researchers, including Wiggins (1989) who can be regarded as a pioneer in this movement, used the label 'authentic assessment'. While, sometimes, the terms authentic assessment and alternative assessment are used interchangeably, O'Malley & Pierce (1996) argue that there is a slight difference between the two labels. Alternative assessment is a broader term describing different types of assessment that provide data about learning, motivation, attitude of students towards classroom activities, and achievement. Authentic assessment, on the other hand, is a term that is used to "connote the relationship of an assessment task to application in everyday situations" (Garcia & Pearson, 1994, p. 21).

Therefore, authentic assessment is specifically associated with the authenticity of a given task or to the degree to which the assessment tasks and activities mirror real-life situations. It "simulates, as far as possible, the authentic behavior which learners will need to enact in real situations" as Lynch (2000, p. 7) puts it. That is, with authentic assessment, those skills that are necessary outside of the school setting and that are needed in the real world are stressed within the classroom. The major aim behind conducting authentic assessment is to show not only what a student knows, but also whether the student has the skill to make use of that knowledge in a real world situation.

As for performance assessment, it is defined as "assessments in which the teacher observes and makes a judgment about a pupil's skill in carrying out an activity or producing a product" (Cizek, 1993, p. 3). It requires students to perform or demonstrate specific skills while actively accomplishing complex and significant tasks that are observed and evaluated simultaneously. In other words, performance assessment is a task whereby a student shows certain competencies and skills in "creating a product, constructing a response, or making a presentation" (Mcmillan, 2017, p. 269). Therefore, performance assessment reflects students' performance on classroom tasks. Whether it is authentic assessment or not depends on the nature of those tasks.

To further clarify the distinction between traditional assessment, authentic assessment, and performance assessment, Cunningham (1998) provides the example of a fourth grade teacher who tries to determine whether his students know how to use money and make change. A traditional approach to assessment would provide pupils with a worksheet which includes the following situation:

> Mark the right answer. If a candy bar costs 35 ¢ and you give a store clerk 1 $,
>   how much change should you get?
>   1.   Two quarters
>   2.   Two quarters and two dimes

3.   Two quarters a dime and a nickel
4.   Three quarters (Cunningham, 1998, p. 122)

In order to make the above mentioned task a performance one, the teacher would use play money and ask pupils to come to the desk and buy an object for which they are going to get change. The same task would be authentic when the pupils were given real money, brought to a shop where they would actually buy the object, counted the change they were given and decided if it was the right change.

Based on these distinctions, one can come to the conclusion that alternative assessment encompasses the two other assessment terms; authentic assessment and performance assessment. Garc´ıa & Pearson (1994) include other labels such as informal assessment, portfolio assessment , assessment by exhibition, and situated assessment. While the strategies used in each assessment may be diverse, they share a common vision which is requiring learners to generate a response rather than simply choose one.

It is worth mentioning, however, that alternative assessment does not present a substitution to standardized testing as the term suggests. Rather, such new forms of assessment are meant to be used in addition to the traditional one to compensate for any limitation that the latter might have. Following the same line of reasoning, Brown & Hudson (1998) explain that the term 'alternative assessments' implies three things. The first implication is that these techniques of assessment are a completely new method of assessing learning. The second implication is that such procedures are somehow totally independent and different from traditional assessment. The third implication the authors discuss is that alternative assessments, in a way, do not require responsible construction of tests or decision making.

These implications are viewed by the authors as being "somewhat destructive" (p. 6). This means that no matter how exciting and interesting these new approaches to assessment are, one should not view them as being entirely different and disconnected from traditional assessment. Brown & Hudson(1998), thus, refer to procedures like presentations, portfolios, journals, peer assessments, and self-assessments as "alternatives in assessment" (p. 6) rather than 'alternative assessments'. Such assessments should, accordingly, complement rather than substitute the traditional ones. This is also supported by Berry (2008) who contends that "alternative assessment does not have to be a replacement of tests. The information collected can be used to support evidence collected from tests, notwithstanding that it can be used independently for understanding student learning" (p. 93).

3.   Characteristics of Alternative Assessment

The ultimate objective of alternative assessment is to "gather evidence about how students are approaching, processing, and completing 'real-life' tasks in a particular domain" (Thomas & Tippins, 1993, p. 2). Therefore, alternative assessment differs from conventional testing in the sense that it actually requires students to demonstrate what they are able to do. Therefore, assessing students is primarily based on what they can produce instead of what they simply can remember or reproduce. Jacobs & Farrell (2001) mention three principles of alternative assessment. The first principle is that with alternative assessment techniques the focus is on meaning and not on form. The second principle is that such tools of assessment aim to investigate the process of learning as a whole rather than the end product. Third, alternative assessments emphasize the 'social nature of learning'. This is achieved through techniques such as peer assessment and group work activities.

Brown & Hudson (1998, p.3) provide a thorough explanation of the main distinguishing features of alternative assessment. To begin with, alternative assessment is being viewed as a way of encouraging students to be actively involved in problems and challenges that are similar to the ones people face in real-world situations. This allows for the provision of rich information concerning the complex thinking skills of the learner. Furthermore, alternative assessment procedures do not necessarily need a separate block of time to be implemented, as is the case with traditional tests. That is why they are seen as non-intrusive to classroom practices. Besides, alternative assessment is carried out on the basis of the same activities which the learner does in the classroom. This means that the implementation of alternative assessment requires little change, if any, in classroom activities and routines. Due to the fact that alternative assessment is based on the day-to-day classroom activities, it is consistent with the curriculum that is being taught.

Moreover, alternative assessment reveals the strengths and the weaknesses of each student. As the data collected from alternative assessment are based on tasks that mirror real life, such data may provide the teacher with insights into what a student's strong points are and what areas need additional work. Additionally, alternative assessment does not rely on one single method of assessment; it rather offers a rich repertoire of possibilities.

Accordingly, a student's progress can be more reliably assessed since the process of assessment includes information from varied sources. Being multi-culturally sensitive is another feature characterizing alternative assessment procedures. In other words, alternative assessment procedures lend themselves to be used with the diverse student populations within a given classroom. Finally, alternative assessment redefines the roles of teachers and students within the classroom. Differently put, through the use of such assessment, teachers, instead of being in full control, they give students the opportunity to be active participants in their learning.

It is worth highlighting, however, that no matter what differences might exist between these two assessment paradigms, one should not be biased towards any paradigm and view it as the perfect choice for all language assessors in all situations. Rather, whatever assessment types are available to us ought to be "valued and utilized for the functions that they provide" as warned by Brown (2004, p. 24). That is to say, the choice of the most appropriate form of assessment to conduct depends mainly on the objectives sought as well as on the role that this assessment form is supposed to fulfill.

4.   Benefits of Alternative Assessment

The benefits of incorporating alternative assessment techniques in the language classroom are numerous. Several researchers in the field of education have tackled this issue and argued for its importance. Accordingly, this section discusses the main advantages of alternative assessments and how such assessments could benefit the learning of languages.

In an article entitled "The Effect of Using Alternative Assessment Activities on Students' Success and Attitudes", Kirikkaya & Vurkaya (2011) listed a number of studies that were carried out to investigate the importance of alternative assessment tools. For instance, in a study conducted by Gamor (2001), the findings revealed that alternative assessment practices have facilitated students' construction of knowledge and phenomena compared to the conventional assessment methods. Kartal & Buldur (2007) carried out a study in an attempt to determine in which alternative assessment technique(s) students were more successful. They also compared students' scores in alternative assessment techniques with those in traditional assessment methods. Their study concluded that students were more successful in assessment that is devised through alternative techniques than that devised through traditional ones. Similarly, Karahan (2007) investigated the effect of alternative assessment techniques on achievement. He concluded that these techniques of assessment are closely related to learning and have a positive influence on students' achievement level.

Advocates of alternative assessment cite several advantages of such assessment. Opp-Beckman & klinghammer (2006) argue that good alternative assessment tasks are important in a number of ways. To begin with, helping learners develop communicative competence is the ultimate goal of any language class. Communicative competence refers to one's ability to express himself or herself and to understand others effectively (Hymes, 1972). That is to say, being communicatively competent means having the ability to adapt one's language to specific contexts and with specific audiences.

Implementing alternative assessments in the language classroom is thought of as a means to enhance learners' ability to engage in meaningful conversations and communicate effectively. Wiggins (1989) asserts that "a true test of intellectual ability requires the performance of exemplary tasks …that replicate the challenges and standards of performance typically faced in the real world" (p. 2). Since communicative competence is context bound, learners should be provided with different communicative opportunities that allow them to become aware of the varied uses of language in their speech community. This means that assessment should be consistent with the needs of learners and their environment. Yet, it is not always an easy job for teachers to decide the extent to which the spoken language of their students is developing. Hence, classroom performance assessments that enable students to show directly how well they are able to communicate are needed.

In this respect, alternative assessment allows teachers to assess learners' ability to apply their knowledge to different situations and also demonstrate and communicate their understanding in different ways. Through alternative forms of assessment, learners are exposed to a wide range of communicative opportunities within the classroom. They get involved in tasks that reflect day-to-day situations in meaningful and realistic contexts. In this way, alternative assessment helps students identify areas of weakness and become successful communicators. Besides, when learners are given the chance to communicate, ask for elaborations and explain their answers and ideas, teachers get a clear picture of how their students learn. Teachers, in other words, get clear insights into how learners understand and learn and eventually, adapt their teaching practices to make any necessary amendments and meet their learners' needs (Black & Wiliam, 1998).

Additionally, alternative assessments "eradicate the lines that separate assessment, instruction, and curriculum" (Thomas & Tippins, 1993, p. 4). This allows both products and processes to be valued equally. Alternative assessments can be regarded as instruments capable of measuring students' progress towards meeting the broad aims and purposes of learning a language. Hamayan (1995) adds that the move towards alternative assessment is related to an increased importance of the relationship between assessment and the process of learning on the one hand, and more sophisticated and higher educational goals on the other. In other words, assessment works best when it is regarded as a daily and continuous process and an important part of the learning process. The data gathered from this process can then be helpful for teachers to evaluate the effectiveness of their practices and make informed decisions about subsequent courses. When teachers gain more awareness of the learners' needs, interests, weaknesses, and strengths, they become in a better position to understand what their students are able to do, and what they are attempting to achieve. This allows the teacher to comment constructively and to provide formative feedback. Students in turn would benefit more from feedback that is provided to them constantly throughout their learning process than from grades obtained on a test which is one-shot and mainly measures what a student knows or can do at a certain time.

More importantly, alternative assessments teach students how to "identify successful behaviors, and be in charge of their own outcome on class assignments" (Butler & Lee, 2010, p. 5). Unlike tests which are usually designed, administered, and graded by the teacher only, with some forms of alternative assessment learners are given a role and are encouraged to become actively involved in monitoring and assessing their own progress. They become able to recognize their main sources of difficulty and identify ways to overcome those difficulties. Learners become aware of the actions necessary to improve both their scores and their performance. This type of involvement encourages learner autonomy and places students at the center of the teaching and learning process. Along with enhancing higher levels of autonomy and active learning, the use of alternative assessment in the foreign language classroom encourages interaction between teachers and students and between students themselves. This interaction in turn increases learners' self-esteem and motivates them to learn (Berry, 2008). When learners are given a role in assessing themselves and their peers, for instance, they become more encouraged to strive for improvement. Moreover, traditional standardized tests can cause learners much anxiety affecting their motivation to learn (Shohamy, 1982). Therefore, learners need to learn and be assessed in an environment where anxiety is minimized to the utmost degree. This can be achieved if learners see assessment as a part of the teaching and learning process and not an independent component whose objective is to make judgments concerning their abilities in comparison with their peers. Using alternative forms of assessment can help alleviate the level of anxiety that is caused by focusing on linguistic accuracy that characterizes most tests, and enhance students' motivation by stressing more communicative skills. Furthermore, it is a given fact that learners do not learn in the same way. In this regard, alternative assessments take into consideration learners' individual differences and variations by offering a broad range of assessment opportunities to cater for individual differences, learning styles, and multiple intelligences of learners.

Another major advantage of implementing alternative assessments is their focus on the type of knowledge and competencies that the student is supposed to demonstrate instead of simply expecting students to provide an answer to a given question or set of questions. Traditional methods of assessment emphasize memorization of facts and development of skills that require a "low level of cognitive effort" only (Dikli, 2003, p. 1). Alternative assessments, on the other hand, "measure students' cognitive thinking and reasoning skills and their ability to apply knowledge to solve realistic, meaningful problems" (Hammond & Adamson, 2010, p. 7). In other words, since alternative assessment methods encourage students to generate responses, such methods are believed to entice students to resort to the use of "higher-order skills" as opposed to "lower-level skills". Instead of tests that often measure content knowledge and assign a grade, alternative assessment methods allow teachers to assess multiple dimensions of students' learning. It is worth mentioning, however, that alternative assessments can also be used to evaluate student understanding and assign grades. Yet, their uniqueness lies in that they give learners more opportunities to learn and improve their learning. They also guide future teaching practices through the provision of feedback. In this respect, Berry (2008) claims,

> By integrating different forms of alternative assessment strategies in their instruction, such as portfolios, observations, experiments, projects, simulations, interviews, performances, presentations, peer assessment, and self-assessment, teachers can achieve a deeper understanding of students' learning and promote deeper learning on the part of the learners (p. 92).

All in all, alternative assessment proved to have overwhelming benefits. It seeks to promote higher levels of thinking among students, stresses the importance of both process and product, and considers individual differences and preferences. This, in turn, boosts learners' level of motivation. However, the implementation of such assessment tools faces a number of challenges and difficulties.

5.   Challenges of Using Alternative Assessment

Despite the many advantages alternative assessment proved to have. There are a number of challenges facing its implementation. Maeroff (1991) confirms this by claiming that "for all its attractiveness, alternative assessment is fraught with complications and difficulties" (p. 4). Generally speaking, the main criticisms voiced against alternative assessment are issues related to validity and reliability.

In her article entitled "Assessing Alternative Assessment", Maeroff (1991) cites the example of Vermont school in the United States. This school implemented portfolios as an alternative assessment technique in mathematics and writing, but it was difficult for teachers to consistently score students' portfolios and compare between them in an accurate manner. The major problem, Maeroff says, was that the portfolios lacked the element of standardization. The article also mentions the example of an elementary school in Colorado which started using an assessment system based on students' actual performances, but which found problems developing criteria for scoring students' performances. Therefore, the shortcomings of alternative assessments are mainly subjectivity in scoring, and the difficulty in ensuring the reliability and validity of the scores. Moreover, alternative assessments require students to construct their own answers. Consequently, the results of such assessment vary from one student to another which makes it difficult for teachers to objectively assess students' performances or achievement.

Another weakness of alternative assessments is that they generally take more time than traditional assessments both in their design and in their evaluation especially in large classes. Although there is the possibility of being systematic when using alternative assessment, it is obviously not easy to score a big number of alternative assessments such as portfolios, demonstrations, interviews, or even compositions. In this respect, Maeroff argues that

> Speed and low cost were the silver bullets that enabled the norm-referenced test - with its multiple-choice responses - to conquer the world of education and hold it in thrall. Meanwhile, alternative assessment, which is not so new an idea as some people think, tends to be a time-consuming, labor intensive, imprecise exercise (Maeroff, 1991, p. 5).

So, given that alternative assessment is time-consuming, a central problem with incorporating it is finding ways to implement the assessment in a controlled time period. In addition to issues of reliability and time, Hammond & Anderson (2010) add another challenge facing alternative assessment in the United States. This is related to some "energized stakeholder groups who objected to aspects of the assessments or the manner in which they were implemented" (p. 22). In certain states, the authors explain, people objected the use of alternative assessment because this type of assessment was unfamiliar to them.

Similarly, Hughes (1993) argues that one should not take it for granted that teachers would move to alternative assessment easily. This is due to the fact that many teachers who use alternative assessment procedures are not well familiarized with these kinds of assessments. Most states, Hughes adds, do not require training in assessment for teachers' certification, and even when teachers are trained in assessment, the trainings offered are not sufficient in providing teachers with the necessary skills and knowledge to become "assessment literates" (Maeroff, 1991, p. 10).

The challenges facing alternative assessments are explained by Hughes (1993) through another example. In March 1991, several experts in measurement from all over the United States gathered to discuss the difficulty of meeting the rapid demands of policy makers to shift to alternative assessment methods. The experts argued that the main difficulties are the time needed both to administer the assessments and to ensure validity and reliability of the results obtained as well as the need to provide teachers with adequate training in using alternative assessment methods.

In discussing some of the challenges facing the implementation of alternative assessment, Thomas, *et al.* (1996) state that "the advocated shift away from treating assessment as independent of curriculum or instruction and

toward aligning assessment with curriculum and instruction comes with a price" (p. 3). The price would be the additional workload that these techniques of assessment pose on teachers when planning their lessons and assessing the performances of their students. The authors mention four issues that they believe to be central to most teachers when implementing alternative assessments. These issues are loss of predictability in the classroom, content coverage, dealing with increased demands on time, and communicating with parents.

It is difficult, if not possible, to ensure what a student's performance or response to an alternative assessment tool is going to be. This is likely to "create a climate of unpredictability" (Thomas, *et al.* 1996, p. 3) which can be a challenge for some teachers. Besides, some teachers were found to develop some sort of reservation with respect to implementing alternative assessment methods because they are more concerned with covering the content of their courses. This concern is particularly intense for teachers of upper-level courses, which highly emphasize preparing students to score well on standardized tests. Moreover, the concern over time makes the use of alternative assessment appear overwhelming for some teachers. Communication with parents is another major issue preventing teachers from conducting alternative assessments. When alternative assessment is used, some parents, who are not familiar with such type of assessment, find it difficult to interpret their children's grades. They, thus, criticize the grades for being too subjective (Thomas, *et al.* 1996).

Huerta-Mac´ıas (1995) mentions two major challenges that teachers encounter when implementing alternative assessments. To begin with, only a small number of teachers have experience in using such assessment forms. Besides, alternative assessment procedures will require more time to prepare, administer, and rate than standardized tests would. The author also found that portfolio assessment failed to provide rich and accurate information about learners' achievement in primary levels even after three years of development and improvement. Reasons for these results are more or less the same found by other researchers, viz, lack of standardization, issues with the scoring rubrics, and insufficient training.

Accordingly, it is obvious that there is much about alternative assessment that remains problematic. Nevertheless, research suggests that such problems could be handled, and that alternative assessment does play a crucial role in helping students learn the skills that they need.

6.   Suggestions for Dealing with the Challenges

The literature addresses issues related to reliability and validity of alternative assessment using what is referred to as "trustworthiness". According to Huerta-Macias (1998), "an instrument is deemed to be trustworthy if it has credibility and auditability" (p. 3). If, in other words, the instrument measures what it's expected to measure and whether the instrument is consistent giving similar results when it is repeated. The author further argues that since alternative assessment looks at students' performance on tasks that are real-life like, such as playing roles, participation in pair or group work, and presenting before a class, "these procedures, in and of themselves are, therefore, valid" (p. 3). The author gives the example of samples of students' writing to explain how alternative assessment is valid. Students' written work samples will reveal the ability of the student to express and organize their ideas in writing, to use appropriate mechanics, transitions, and vocabulary. They will, at the same time, show to what extent the student is able to write in the language. Hence, alternative assessments lend themselves to be valid since they employ a wide range of strategies and rely on varied sources of information.

Baker (2010) suggests a set of criteria for evaluating the validity of alternative assessments. He classifies these criteria into the following two main categories:

Internal Validity Criteria: The first internal criterion for judging the validity of alternative assessments is 'Cognitive Complexity'. This means that assessment should not be restricted to the surface features only, but should also consider "the intellectual demands of the task" (Baker, 2010, p. 15). This also requires devising assessment tasks that are new for the student and that the student had not memorized and practiced before. The second internal criterion proposed by the author is 'Meaningfulness'. This means that the instructions for the task should be expressed clearly in a way that is comprehensible to all students. This criterion is intended to motivate all students to perform well and also to ensure fairness and avoid bias. The third internal validity criterion is 'Quality'. This is related to how accurate the content of the assessment is and to what extent it covers what has been studied. Differently put, the subject matter of the assessment should be aligned with the learning objectives. The last criterion in this category is 'Transfer' and 'Generalizability' which means that the assessment used should allow "the generalizability and transfer for the particular performance to other related topics or domains" (Baker, 2010, p. 16). This is achieved by using tasks that are relevant to a wide range of applications and by establishing assessment criteria that can be applied across different tasks.

External validity Criteria: 'Consequences' is the first criterion of external validity, and it is a term used by Messick (1989) who argues that the concept of validity, rather than focusing on content and construct related validity, should include recognizing the consequences of assessment as well. Enhanced Students' interest and involvement in the learning process could be considered a positive consequence of assessment. Consequences could also describe the positive effect that the assessment has on teaching strategies. When teaching strategies are directed towards developing problem solving skills, assessment could be said to have a positive consequence (Cunningham, 1998). The second criterion, which is closely related to the first, is 'Fairness'. Assessment should be independent of external factors such as the student's social class, gender, ethnicity, or race. Also, fairness requires that students are given the chance to learn and prepare well for the assessment. More importantly, fairness is about ensuring that scorers and the scoring criteria used are bias free. The last criterion for judging alternative assessments is 'cost and efficiency'. As mentioned earlier, alternative assessments require much time and resources to design, administer, and grade. This requires ensuring that the performance measured by these assessments is more credible, and that the adequate data about the student's knowledge and ability is provided by the assessment used in order to justify the cost.

As for reliability, it is believed to be of less concern in classroom assessment because in the classroom context students are assessed on a continuous basis. This helps the teacher get new insights, and hence make any necessary adjustments. That is, if assessment decisions in the classroom seem to be wrong, they can easily be reconsidered (Shillingburg, 2016). In the same vein, Olfos & Zulantay (2007) state that authentic assessment happens in contexts that are similar to the real world, so generalizations are not of much importance, and thus replicability shouldn't be much of a concern. Similarly, Brown & Hudson (1998) add that "the issues of reliability and validity must be dealt with for alternative assessments just as they are for any other type of assessment-in an open, honest, clear, demonstrable, and convincing way" (p. 5). Thus, to ensure that the assessments are valid, alternative assessment users and designers should be careful in structuring the ways they use the alternative assessment forms as they structure any other forms of assessment. However, Cinningham (1998) argues that it is not really important to adhere to the same rules of conventional tests when dealing with performance assessments especially when such assessments are used for purposes related to the classroom and decisions are not of high stakes. According to the author, the use of performance tasks is easily justified instructionally so long as students benefit educationally and teaching is enhanced. Huerta-Macias (1995, p. 3) suggests that reliability in alternative assessments could be ensured by designing tasks that are different but that yield similar results, using criteria that are clearly established and using judges who are well trained, and finally monitoring on a regular basis to make sure that the criteria used are consistent among raters.

Triangulation is another means to ensure the reliability or consistency of assessment. When used in qualitative research, triangulation means "the combination of methodologies to strengthen a study design" (Huerta-Macias, 1995, p. 3). When applied to alternative assessment, triangulation means collecting information about students' learning from three sources. Triangulation in alternative assessment in an EFL/ESL context can be applied in different ways. For example, if a teacher likes to assess, say, students' level of literacy in the language, the sources of information could be, the students themselves, the teacher, and the parents. The purpose is to have enough information for getting a clear and holistic picture of students' development. In such case, interviews or surveys for example could be used to obtain information regarding the student's background. As for data that is gathered from students, the teacher could use portfolios that contain samples of students' written work, audiotapes of exchanges that they engaged in, videos of the students while playing roles, and self or peer assessment sheets. Then, the teacher would add information from class observations, performance checklists, or journals.

Using rubrics is another solution suggested with respect to the issues of objectivity, validity, and reliability (Baker, 2010). The use of rubrics is encouraged as they contain essential qualities necessary for acceptable performance and for more effective instruction by teachers. Raters should establish carefully designed scoring criteria in advance and visit them throughout the scoring process. This is likely to address the issue of intra-rater reliability and maintain consistency of the results. Scoring rubrics are also believed to reduce variations between scorers. This is achieved when the criteria are formalized at each score level and are used to guide the assessment process (Moskal & Leydens, 2000).

As regards the issue of time constraints, Hughes (1993) states that "in classrooms that stress hands-on activities and student involvement, instruction and evaluation can result from the same tasks" (p. 6). In other words, the time constraints that alternative assessments are claimed to pose diminish if teachers make data collection for alternative assessment a part of their day-to-day teaching routines. The author adds that "authentic assessment puts more emphasis on whether the student's progress is reasonable for his or her age and experience and less

emphasis on comparison with other students" (p. 7). Therefore, when teachers establish a classroom environment with more students' involvement in learning, time constraints would not be an issue.

Concerning the issue of lack of objectivity in alternative assessment, Huerta-Macias (1995) ascertains that alternative assessment should not in the least be considered as less objective than conventional testing. That is, the notion of objectivity has been challenged even with standardized tests. She explains this by the fact that all humans have an inclination to be biased whether we are aware of that or not. The fact that standardized tests require that a number of people agree on scoring criteria, content, and format does not in the least mean that such tests are objective. Rather, they just "collectively share the same biases" (p. 4). Accordingly, a standardized test should not be regarded to be more objective than a method of alternative assessment.

7. Conclusion

This paper provided a synoptic review of the literature related to alternative assessment, its characteristics, its benefits, as well as the challenges faced when using it. The paper also discussed what the literature suggests as strategies for dealing with the challenges in order to reap the benefits of alternative assessment. In conclusion, one would say that it is a prerequisite that language teachers take the above discussed suggestions into account when trying to implement alternative assessment methods with their students. This is for the many advantages and positive effects that these methods of assessment proved to have. In light of these advantages, it is recommended that EFL teachers employ different alternative assessments in addition to the conventional assessment methods to ensure effective learning.

References

Baker, L. (2010). What Probably Works in Alternative Assessment. *National Center for Research on Evaluation, Standards, and Student Testing Graduate School of Education & Information Studies* (CRESST).

Berry, R. (2008). *Assessment for Learning*. Hong Kong University Press.

Black, P., Wiliam, D. (1998). Assessment and Classroom Learning. *Assessment in Education: Principles, Policy and Practice. 5(1), 7–74*.

Brown, J. D., Hudson, T. (1998). The Alternatives in Language Assessment. *TESOL Quarterly, 32(4), 653-675*.

Butler, W.G., Lee, J. (2010). The Effects of Self-Assessment among Young Learners of English. *Language Testing*, 27(1), 5-31.

Cizek, J. (1993). Alternative Assessments: Yes, But Why? *Educational Horizons, 72(1)*, 36- 40.

Cunningham, G. K. (1998). *Assessment in the Classroom: Constructing and Interpreting Texts*. Psychology Press.

Dikli, S. (2003). Assessment at a Distance: Traditional vs. Alternative Assessments. *The Turkish Online Journal of Educational Technology, 2(3), 13-19.*

García, G. E., Pearson, P. D. (1994). "Chapter 8: Assessment and Diversity". *Review of Research in Education, 20(1), 337-392)*.

Hancock, C. R. (1994). Alternative Assessment and Second Language Study: What and Why? *ERIC Digest.*

Hammond, L., Adamson, F. (2010). *Beyond Basic Skills: The Role of Performance Assessment in Achieving 21st Century Standards of Learning*. Stanford Center for Opportunity Policy in Education.

Hamayan, E. V. (1995). Approaches to Alternative assessment, *Annual Review of Applied Linguistics, 15, 212-226*.

Hughes, S. (1993). What Is Alternative / Authentic Assessment and How Does It Impact Special Education? *Educational Horizons, 72 (1)*, 28-35.

Huerta-Macias, A. (1995). Alternative Assessment: Responses to Commonly Asked Questions. *TESOL Journal, 5, 8-27.*

Jacobs, G. M., Farrell, T. S. (2001). Paradigm Shift: Understanding and Implementing Change In Second Language Education. http://www.zait.uni- bremen.de/wwwgast/tesl_ej/ej17/a1.html

Kirikkaya, E.B. & Vurkaya, G. (2011). The Effect of Using Alternative Assessment Activities on Students' Success And Attitudes in Science and Technology Course. *Educational Sciences: Theory & Practice - 11(2), 997-1004*.

Lynch, R. (2000). Authentic, Performance-Based Assessment in ESL/EFL Reading Instruction. *Asian EFL Journal*. 5 (4), 1-28.

Maerhoff, G.I. (1991). Assessing Alternative Assessment. *The Phi Delta Kappan, 73(4), 272- 281*.

McMillan, J. H. (2017). *Classroom Assessment: Principles and Practice That Enhance Student Learning and Motivation* (7th edition). Pearson.

Messick, S. (1989). Meaning and Values in Test Validation: The Science and Ethics of Assessment. *Educational researcher,18(2), 5-11.*

Moskal, B. M., & Leydens, J. A. (2000). Scoring Rubric Development: Validity and Reliability. *Practical Assessment, Research, and Evaluation*, 7(1), 10.

Olfos, R., Zulantay, H. (2007). Reliability and Validity of Authentic Assessment in a Web Based Course. *Educational Technology & Society, 10 (4), 156-173.*

Opp-Beckman, L., Klinghammer, S. J. (2006). *Shaping the Way we Teach English: Successful Practices around the World* (Instructor's manual). Office of English Language Programs, Washington DC, USA.

Shillingburg,W. (2016). Understanding Validity and Reliability in Classroom, School-Wide, or District- Wide Assessments to Be Used in Teacher/Principal Evaluations. *Journal of Behavioral Education*, 10 (4), 205-212.

Shohamy,L. (1982). Affective Considerations in Language Testing. *The Modern Language Journal, 66, (1), 13-17.*

Thomas M. & Tippins, J. (1993). Considering Alternative Assessments for Middle Level Learners. *Middle School Journal*, 25 (2), 3-5.

Thomas, J. C., Bell, K., Fisher-Cauble, D., & Sanchez, W. B. (1996). The Demands of Alternative Assessment: What Teachers Say. *The Mathematics Teacher, 89 (6)*, 484- 488.