

Assessment Methods.

Mardanova Firuza Subhanovna¹,Kultaeva Feruza Ergashevna²,
Sultanova Muhabat Shamsievna³

¹Samarkand State Architectural and Civil Engineering Institute

²Samarkand State Architectural and Civil Engineering Institute

³Samarkand State Architectural and Civil Engineering Institute

Article History: Received: 10 January 2021; Revised: 12 February 2021; Accepted: 27 March 2021; Published online: 28 April 2021

Abstract

Alternative assessment is type of assessment practice that seeks to assess learners' language abilities, in ways other than how traditional standardized tests did. The present article reviews the issues related to alternative assessment and adopts a partly critical position regarding some of the related arguments. The advantages as well as quality concerns are discussed and a number of challenges that need to be dealt with are considered. Finally suggestions are made to ameliorate practice.

Keywords: Alternative Assessment, Traditional Tests ,Congruent assessment

1. Introduction

It is the common practice among students (language students or other majors) to raise their objections to their scores when they fail to get a criterion score or when they feel they have not gotten the score they deserve. Even somehow similar attitudes can be found among those who have gotten the required scores and so do not have any problem of moving on to the next level, grade, class etc. Where is the problem? Does the sheer act of taking a test and getting a score produce bitter feelings in students? Most likely no. Do they feel their abilities are underestimated? Or simply, they just want to have their scores changed for higher ones? Although the answer as to why of this customary practice is not clear, what is clear is that testing has kept open the channels that produce such judgments of inefficiency of testing from the part of test takers. If this is the case, what is the problem with tests and/or testing that results in attitudes like these? One major answer is that tests – or at least tests that are used in such situations – do not allow the students to display their full range of abilities, especially if they are language students. Individuals who sit in one corner of the test-taking hall, and provide answers for a number of questions are well excused for voicing their discontent with such tests, if the phenomenon to be tested is an ability rather than knowledge.

One way to overcome the problem referred to above, is to use methods of testing other than what has been practiced so far. In other words, practitioners are advised to try alternative forms or methods of testing; those that reach out to probe more into language learners' abilities. Alternative assessment is one solution. It is a form of assessment that claims to break away with the traditional tests, and provide a better picture of what learners can do with language. These claims are appealing. Provided that alternative assessment stands up to the claims it makes, it can be of much help to remedy the limitations of traditional tests. But, does it? It is to this that I now turn to in the following section.

2. Alternative Assessment: What is it?

From the very start the difficulties associated with alternative assessment manifest themselves. The term "alternative assessment" is suffering from a number of shortcomings. I see two problems with this label; one is the fact that this label is not used unanimously by scholars for example, names some of the terms that have been used by different writers, all of whom referred to alternative assessment. These terms are "alternative assessment", "informal assessment", "authentic assessment", "performance assessment", "descriptive assessment" and "direct assessment". The second problem with the label is that it does not refer to a coherent group of activities that can be put together under one term as a representative . Alternative assessment includes a variety of activities, very different in nature, aimed at generating data about individual learners. Although all of the activities share the common goal of producing information about an individual's language competence, considering them as homogeneous with shared features is misleading. More discussion about what these activities are is

provided below. It also expresses the vagueness she perceives is present in the term. Further, she proposes the use of the term congruent assessment instead, which captures the "...process by which the assessment fits the goals of the adult education program".

Whatever label is used, alternative assessment does appear to have the consensus of scholars about what the *notion* refers to. The *Longman Dictionary of Language Teaching and Applied Linguistics* defines alternative assessment as: "various types of assessment procedures that are seen as alternatives or complements to traditional standardized testing, pp. 23". As such, alternative assessment can be taken to refer to a group of procedures that aim to evaluate a learner's overall performance. It gives a role to "the context of instruction" and relates alternative assessment to the daily activities of the school or classroom. A variety of procedures have been suggested in the literature, includes in the list such practices as checklists of student behaviors or products, journals, reading logs, videos of role-plays, audiotapes of discussions, self-evaluation questionnaires, work samples, and teacher observations or anecdotal records. The list and its inclusion of very disparate activities implies that whatever source that yields data about a learner's competence is pertinent. These data are used to evaluate the learner's language ability.

3. Advantages

Proponents of alternative assessment postulate a number of advantages for it. The most noticeable and promising of those advantages is that alternative assessment provides a more thorough picture of a learner's abilities, achievements and weaknesses, than did multiple-choice tests. Unlike traditional paper-and-pencil tests that give limited information about individuals' theoretical knowledge of grammar and usage only, alternative assessment allegedly taps into what the learner is able to do as a result of instruction. The use of multiple sources of information is appealing, as it takes into account both knowledge and ability, and there does not appear to be any objections to the very idea, since it has been established that communicative competence is composed of many delicate knowledge components with intricate interactions within, and a simple sheet of questions is unable to reflect such a complicated web of knowledge.

One other merit of alternative assessment which is related to the one just mentioned, is its awareness of the context in which a learner is to demonstrate his/her competence. Traditional standardized tests, assess learners' discrete knowledge of components of language, a statement which has become trite over the years. Such testing does not give a realistic picture of a learner's true abilities to perform real-world tasks. Alternative assessment, on the other hand, contextualizes learners' performance, and requires them to demonstrate their abilities by doing meaningful and purposeful tasks, embedded in a social context.

Another improvement of alternative assessment is its inclusion of learners in the process of assessment. Alternative assessment gives a role to learners and requires their contribution to the estimations and decisions made of and about them, thereby introducing democraticity to the enterprise of language teaching.

Alternative assessment allows for the adoption of a holistic or integrative view of language. In contrast to traditional standardized tests, that examine the learner from a single dimension, alternative assessment adopts a multi-dimensional perspective in that it employs other sources of data, in addition to test answers, each of which brings its own piece to the overall puzzle of a learner's language knowledge. A holistic view of language also means to look at language in its entirety, rather than as a sum of separate aspects like phonology, grammar and vocabulary.

As with any other language assessment measure, the most important quality concern with alternative assessment is its validity and reliability. Much has been written about the validity of traditional language tests, and a great deal of improvement has been made regarding our understanding of this concept. However, most claims about the quality of alternative assessment are intuitive in nature, and little, if any, attempts has been made to verify such claims. Huerta-Macías (2002), for example, is very optimistic in terms of the quality of alternative assessment, and assumes an automatic validity for it due to its nature, and, therefore, it is easy to criticize her for her stance towards validity of alternative assessment. Although she admits the cruciality of validity in alternative assessment, she believes that "The procedures [of alternative assessment] in and of themselves are, ..., valid". It is true that in alternative assessment the measuring instrument is closely linked to the phenomenon

under observation, as Huerta-Macías reasons, but this alone does not guarantee validity. What remains to be ensured is the subjectivity of the evaluator's judgments. She further argues that all human judgments, including those made in standardized tests' validity, is subjective, and so she appears to circumvent the problem of subjectivity. The problem is not that standardized tests are also subjectively validated; it is how we can ensure the validity of alternative assessment in the face of the assessor's subjective attitudes. Accountability requires that those who are in the position of decision makers for others' lives, assume responsibility and make the best of their efforts to gather validity evidence in order to demonstrate the validity of their assessment. Practitioners of alternative assessment are not at all an exception in this case. They also are required to ascertain their addressees in terms of the quality requirements that are needed of their assessment endeavors.

There are a number of measures to be taken to ensure the validity of alternative assessment. Huerta-Macías suggests these: 1. designing multiple tasks that produce similar outcomes, 2. using trained judges, 3. setting up clear criteria, 4. monitoring raters for consistency, and 5. triangulation. As it is apparent, these strategies are not new, as claimed by Huerta-Macías, and what is new is the idea that carrying them out is enough, a view held by Brown & Hudson (1998). They argue that it is to be irresponsible to take the validity of assessment for granted, once these procedures are practiced. Since assessment affects the lives of individuals, they further hold, it is the duty of the assessor to make every effort to demonstrate and enhance the validity of assessment (pp. 656).

The mainstream test types used in language classes is mainly focused on the end product of instruction or on an individual's proficiency as it is related to what s/he has achieved. This product-orientedness of standardized tests is what is contrasted with alternative assessment's process-orientedness (Alderson & Banerjee, 2001; Hancock, 1994; Hamayan, 1995). Alternative assessment through bringing to the forefront those aspects of the teaching/learning process that mainly go unnoticed by multiple-choice tests, because of their concern with the end-of-the-term product of instruction, gives information about the processes that learners use or undergo to learn language. A think-aloud protocol, for example, provides valuable information about what strategy a learner uses to comprehend a text, or it guides us in recognizing what a particular learner's attitudes are towards the target community; information that a multiple choice test of linguistic knowledge hardly, if ever, can provide. However, some may raise their objection to this argument by pointing to formative assessment and say that it is aimed at detecting learners' needs and evaluating instruction during the term. I would argue that although formative assessment's goal is to provide information for teachers about how the course is progressing, it employs largely product-oriented means to evaluate the process. The kind of data that formative assessment provides is of the type summative assessment produces. What is different between these two types of tests, formative and summative assessments, is not very much the kind of data that they produce; rather it is the interpretations that are made of the data that distinguishes the two. However, alternative assessment, be it used formatively or summatively, produces information that is different in nature. However, it is erroneous to think that alternative assessment should replace objective tests. There is a danger to this assumption. We should not think radically and fall from the other side of the line. Objective tests do produce invaluable data, but that the data is not enough. We can boost the advantages by integrating the two: to consider multiple choice kind of tests as an alternative and accommodate it *within* the framework of alternative assessment. In other words, it is possible to utilize this kind of tests along with other measures, i.e. teacher observations, diaries, written samples, etc., and let each of them play their own roles to contribute to our evaluations of learners.

Another characteristic of alternative assessment is the claim that it is mainly used for low-stakes situations, and when evaluation is needed for a high stakes situation, traditional tests are used (Alderson & Banerjee, 2001). Alderson & Banerjee (2001), for example make the point that alternative assessment is used for classroom-based assessment, focusing on the progress of the instructional program. One reason seems to be for teachers, administrators and authorities to be unwilling to take risks. They are reluctant to stop doing what they have practiced so far, and opt for what they do not know exactly. If this is the case, we are not allowed to blame them for such an attitude, since, if we be optimistic and look at the story by seeing the good, they want to take responsibility for their learners and want to do what has been established and accepted, not something that would jeopardize the benefits of their learners for the sake of trying a new thing, only. It is, then, the responsibility of those working in the theoretical camp, i.e. researchers and theory makers of the

field, to unravel the benefits of alternative assessment for the practitioners through holding seminars, publishing specialized material for teachers and similar measures, certainly with the help of other authorities and agencies.

The failure to utilize alternative assessment by teachers brings up the question of the gap between theory and practice. This problem is not unique to alternative assessment and is the case for most topics in the field of applied linguistics. However, the problem is not unidimensional. It is not only the teachers that do not use alternative assessment; it is both under-practiced and under-researched. To the best of my knowledge no study has been done to explore alternative assessment to add to the current state of knowledge. Although studies have been done to investigate the different activities launched inside the framework of alternative assessment, such as diary studies, such attempts do not reveal much about alternative assessment as a whole. Little is known about alternative assessment, and there are many facts still waiting to be explored. To be humorous, the only thing about which we are sure regarding alternative assessment is that it exists! Certainly, it is not enough. More is needed, before we can make strong claims as to the benefits it provides for teachers and learners. Attempts need to be made and researches needs to be done to investigate the different elements of alternative assessment.

A less recognized concern is a meta one. As pointed above, different labels are used to refer to alternative assessment. This is not a satisfactory situation. If our understanding of alternative assessment is to be enhanced, we need a label that is unanimously used by scholars, so that discourse is facilitated. Some of the labels in some contexts point to a different phenomenon, altogether. For example, performance assessment is used as a label to refer to alternative assessment by some writers, while it is used to refer to an entirely distinct concept by some others. If there is no consensus about what term to use to refer to alternative assessment as a concept in language testing, then how are we to reach an agreement on other more grave matters like how to validate assessment. Also, different terms mean that there will be ambiguity in discussions related to this concept. Therefore, the first step of all is for me to discard all the other *alternative* terms and use one more familiar; my suggestion is alternative assessment.

Currently we know nothing about teachers' perceptions about alternative assessment. Teachers' ideas may shape their practice and hence influence the lives of their learners. Therefore, if we are to exploit the benefits alternative assessment provides, we have to know more about what teachers think about alternative assessment, how it can contribute to learners' success, how it is carried out, how it is different from traditional paper-and-pencil tests, how are its effects differentiated from the effects of standardized tests, what types of evidence are required to validate it and so on. However, before probing into teachers' attitudes, we first need to introduce it to teachers, as stated above. Until then the task is even more cumbersome.

Another responsibility for researchers is to reveal what procedures are best suited for specific situations. The procedures that are employed in alternative assessment such as teacher observation, examination of sample writing, etc., each sheds light on a different aspect of a learner's competence. On the other hand, each assessment situation has its own agenda and goals, and it is logical to think that not every combination of the procedures is appropriate for every situation. In other words, determining what procedures to employ for a particular situation seems to be necessary, depending on the goal of assessment and the type of audience we have. Before such studies, it remains for teachers' intuition to determine what course of action to follow. If the appropriate combination is not employed, the whole assessment may lack validity and produce results that are far from representative of learners' true abilities; it may result in the same pitfalls that are being discussed is the case with traditional type of tests. The decisions made on the basis of the information produced by such assessment may have unfavorable repercussions for the society.

Another challenge for research is the fact that the exact methodology to employ is not clear and standardized (Balliro, 1993). The question is a very general one: what should teachers do? Where should they start from? What stages should they follow and in what sequence? What should they do with their findings or how should they interpret the results? What resources are available for them to consult? What types of decisions can be made on the basis of assessment? A related question pertains to the data that assessment procedures produce. Each of the procedures yields data that are inherently different from others. A checklist that is used to rate the frequency of learners' behavior, for example, gives data that cannot be compared to information obtained from examining learners' writings. Both are different from learner diaries with regard to the data produced; none is comparable to what is

obtained from think-aloud protocols. Although for Hamayan interpreting the information obtained is easy, the disparity of the data and the consequent interpretations required will most likely be unmanageable for teachers who are not prepared for the task. Researchers should step in and provide guidance by outlining teachers' roles and the precise path to go.

Conclusion

In writing a concluding paragraph for the paper, a term which was reverberating in my mind was a frightening one: change! Changes always entail leaving your safe grounds in search of unexplored territories. Not everyone has the courage to do it. But in the case of our shared discipline we have to embark on such a demanding task, at least regarding the fact that we are dealing with knowledge and its consequences, the consequences in this case being the effects that our practices leave on the lives of our learners. Change does not always require discarding previous objects, beliefs, practices and behaviors altogether. Rather it means, if we mean to improve, to modify those aspects of the target phenomenon whose inefficiency is apparent. Alternative assessment resonates with this notion of change absolutely.

REFERENCES

1. Alderson J. C., & Banerjee, J. (2001). Language Testing and Assessment. *Language Teaching*, 34, 213-236.
2. Aschbacher, P. R. (1991). Alternative Assessment: State Activity, Interest, and Concerns. UCLA Center for Research on Evaluation, Standards, and Student Testing.
3. Balliro, L. (1993). What Kind of Alternative? Examining Alternative Assessment. *TESOL Quarterly*, 27 (3), 558-560.
4. Brown, H. D., & Hudson, T. (1998). The Alternatives in Language Assessment..
5. Hancock, C. R. (1994). *Alternative Assessment and Second Language Study: What and Why?* Eric Digest. Retrieved June 25, 2013 from
6. <http://www.cal.org/resources/digest/hancoc01.html>Huerta-Macías, A. Lynch, B. K. (2001). Rethinking Assessment from a Critical Perspective. *Language Testing*, 18 (4) 351–372.
7. Methods and strategies in teaching English. D .Sh. Ibragimova
European Journal of Molecular & Clinical Medicine ISSN 2515-8260 Volume 07,
8. Richards, J. C., & Schmidt, R. (2010). *Longman Dictionary of Language Teaching & Applied Linguistics*. Pearson Education Limited.