

Chapter 6

Beyond Tests: Alternatives in assessment

Book name: Language Assessment Principles and Classroom Practices

Written by: Brown and Abeywickrama

Summarized by: Saeed Mojarradi Ph.D. Candidate Professor: Dr.Zoghi

In the public eye, tests have acquired an aura of infallibility. (Without errors) Everyone wants a test for everything, especially if the test is cheap, quickly administered, and scored as soon as possible.

Some believe that all testing is invidious. But tests are simply measurement tools. It is clear by now that tests are one of a number of possible types of assessment.

The distinction:

An important distinction was made between testing and assessing. Tests are formal procedures, usually administered within strict time limitations, to sample the performance of a test – takers in a specific domain. Assessment has a much broader concept in that most of the time when teachers are teaching, they are also assessing. Assessment includes all occasions from informal impromptu observations and comments up to and including tests.

In the decade of the 1990s, when a number of educators questioned the notion that all people and all skills could be measured by traditional tests, a novel concept emerged that began to be labeled

“alternative assessment.”

As teachers and students were becoming aware of shortcomings of standardized testing and the problems found with such testing was proposed.

That proposal was to assemble additional measures of students-portfolios, journals, observations, self-assessments, peer-assessments, and the like.

Why should we even refer to the notion of alternatives when assessment already encompasses such a range of possibilities?

Brown and Hudson responded: they noted that to speak of alternative assessment is counterproductive because the term implies something new and different that may be exempt from the requirements of responsible test construction.

We remembered that all tests are assessments but that, more importantly, not all assessments are tests.

The characteristics of various alternatives in assessment were summed up by Brown and Hudson:

- Require students to perform , create , produce , or do something
- Use real –world contexts or simulations
- Are not invading in that they extend the day-to-day classroom activities
- Allow students to be assessed on what they normally do in class every day.
- Use tasks that represent meaningful instructional activities.
- Focus on processes as well as products
- Higher – level thinking and problem – solving skills
- Provide information about both the strengths and weaknesses of students
- Are multiculturally sensitive

- Ensure that people, not machines, do the scoring, using human judgment.
- Encourage open disclosure of standards and rating criteria
- Call on teachers to perform new instructional and assessment roles

The dilemma of maximizing both practicality and wash back

Tests, especially the large – scale standardized tests tend to be one-shot performances that are timed, multiple-choice, decontextualized, norm – referenced, and that foster extrinsic motivation. On the other hand, tasks like portfolios, journals, and self-assessments are:

- Open-ended in their time orientation and format
- Contextualized to a curriculum
- Referenced to the criteria of that curriculum
- Likely to build intrinsic motivation

Formal standardized tests are almost by definition highly practical, reliable instruments. They are designed to minimize time and money on the part of test designer and test-taker and to be accurate in their scoring.

Alternatives such as portfolios, conferencing with students on drafts of written work, or observations of learners over time all require considerable time and effort on the part of the teacher and the student.

Looking at a figure on page 124 we see that as a technique increases in its wash back and authenticity, its practicality and reliability tend to be lower. Conversely, the greater the practicality and reliability, the less likely you are to achieve beneficial wash back and authenticity.

In that case there has been placed three types of assessment on the regression line to illustrate.

- Large-scale, standardized multiple-choice tests
- In-case, short-answer essay tests
- Portfolios, journals, and conferences
- Large – scale multiple-choice tests cannot offer much wash back or authenticity, and portfolios and such alternatives cannot achieve much practicality or reliability.

A number of approaches to accomplishing the end in which we can transform otherwise inauthentic and negative-wash back producing tests into more pedagogically learning experiences. **They include:**

- Building as much authenticity as possible into multiple-choice
- Designing classroom tests having both objective-scoring and open ended response sections.
- Turning multiple – choice test results into diagnostic feedback
- Maximizing the preparation period before a test
- Teaching test-taking strategies
- Helping students to see beyond the test not teaching to the test
- Information on a student before making a final assessment of competence

Performance-based assessment

The other name for this performance based assessment, sometimes merely called performance assessment.

Is this different from what is being called alternative assessment?

The push toward more performance based assessment is part of the same general educational reform movement that raised strong objections to using standardized test scores as the only measures of student competence.

Performance based assessment would require the performance of the actions or samples which would be systematically evaluated through direct observation by a teacher.

Performance based assessment, according to Norris (1998) involves test takers in the performance

of tasks that are as authentic as possible and that are rated by qualified judges.

J.D. Brown (2005) noted that a related concept , **task – based assessment** , is not so much a synonym for performance based assessment as it is a subset in which the focus of assessment is explicitly on particular tasks or types.

According to Pierce (1996) performance based assessment is a subset of authentic assessment. In other words, not all authentic assessment is performance based.

One could infer that reading,listening and thinking have many authentic manifestations, but because they are not directly observable in, they are not performance based.

Again according Pierce and Malley the characteristic of performance based assessment are:

- Students take a constructed response(selecting answer from options)
- They engage in higher-order thinking with open-ended tasks
- Tasks are meaningful, engaging, and authentic.
- Tasks call for the integration of language skills
- Both process and product are assessed.
- Depth of a student’s mastery is emphasized over breadth.

Performance based as assessment procedures need to be treated as traditional tests:

- State the overall goal of the performance
- Specify the criteria of performance in detail
- Prepare students for performance in step-wise progressions
- Use a reliable evaluation form, checklist and rating sheet.
- Treat performances as opportunities for giving feedback

Rubrics

Rubrics that teachers engage in their day-to-day based assessment procedures are not a separate alternative in assessment but rather a virtually indispensable tool in effective, responsible, performance based assessment.

A rubric is a device used to evaluate open-ended oral and written responses of learners. Some rubrics involve scaling, that is, the assignment of numbers (a numerical scale) to the described levels of performance.

In recent years, with a marked increase in the use of alternatives in classroom based assessment, rubrics have taken a front seat in teacher’s evaluative tools.

Creating effective rubrics requires effort, care, and precision on your part.

Consider the following steps (adapted from Andrade and others) :

- 1- Clearly list the objectives of assessment instrument
- 2- Describe levels of performance that you expect.
- 3- Use task – specific performance
- 4- Make your statements concise and avoid dysfunctional performance
- 5- Consult other models of rubrics
- 6- Consider revisions that will improve evaluation.

Portfolios

One of the most popular alternatives in assessment, especially within a framework of communicative language teaching is portfolio development.

According to Genesee and Upshur (1996) a portfolio is a purposeful collection of students' work that demonstrates their efforts, progress, and achievement in given areas.

Portfolios include materials such as:

- Essays and compositions in draft and final form
- Reports , projects , and presentations outlines
- Poetry and creative prose
- Artwork, photos, newspaper or magazine clippings
- Audio and or video recordings of presentations, demonstrations, etc.
- Journals , diaries, and other personal reflections
- Tests, test scores , and written homework exercises
- Notes on lectures
- Self-and peer-assessments – comments, evaluations , and checklists

Portfolios were thought to be applicable only to younger children who assemble a portfolio of artwork and written work for presentation to a teacher and or a parent.

Gottlieb (2000) suggested a development scheme for considering the nature and purpose of portfolios using the acronym CRADLE to designate six possible attributes of portfolio:

Collecting
Reflecting
Assessing
Documenting
Linking
Evaluating

- As **collection**, portfolios are an expression of students' lives and identities the appropriate freedom of students to choose what to include should be respected.
- **Reflective** practice through journals and self- assessment checklists is an important ingredient of a successful portfolio.
- Teacher and student both need to take the role of **Assessment** seriously as they evaluate quality and development over time.
- We need to recognize that a portfolio is an important **Document** in demonstrating student's achievement.
- A portfolio can serve as an important **Link** between student and teacher, parent, community, and peers.
- **Evaluation** of portfolios requires a time –consuming but fulfilling process of generating accountability.

The CRADLE acronym offers the possibility of constructing a rubric for evaluating student's performance on their portfolios.

Portfolios can fail:

- If objectives are not clear
 - If guidelines are not given to students
 - If systematic periodic review and feedback are not present
- Sometimes asking students to develop a portfolio seems daunting (with fear), especially for

new teachers and for those who have never created a portfolio on their own.

Successful portfolio development depends on following a number of steps and guidelines:

- 1- State objectives clearly
- 2- Give guidelines on what materials to include
- 3- Communicate assessment criteria to students
- 4- Designing time within the curriculum for portfolio development
- 5- Establish periodic schedules for review and conferencing
- 6- Designate an accessible place to keep portfolios
- 7- Provide positive wash back-giving final assessment

For more details about these seven items, refer to pages131-133

Until page 133
Mojarradi