Alternative Assessment: Growth, Development and Future Directions

Vino Sarah Reardon

Abstract This chapter will attempt to focus on *alternative assessment* from a pedagogical perspective (i.e. in the context of learning environments that address students with and without *special needs*). Thus, the purpose of this chapter is threefold: (1) to investigate how alternative assessment has been used in the United States and elsewhere, (2) to define the general characteristics of alternative assessment (in particular, portfolios), and (3) to examine the direction that alternative assessment has taken in recent years within *SPED* (Special Education) programs, *ESL* (English as a Second Language) and/or *EFL* (English as a Foreign Language) classrooms. This chapter offers a significant contribution to the field of ESL/EFL by demonstrating that alternative assessment best serves the needs of English Language Learners (ELLs) with and without speech impairments.

Keywords Assessment · Alternative · Traditional · SPED · ESL · EFL

1 Introduction

Historically, students who have special needs (i.e. students who are physically or mentally challenged because of physical or cognitive impairments) have been placed in what is commonly referred to as SPED programs in the United States. No doubt, the goal has generally been to offer this *learning disabled* (LD) student population adequate academic support and preparation to *mainstream* these students into general education classrooms, if and when possible.

V.S. Reardon (✉)
Acaciawood College Preparatory Academy, CA, USA
e-mail: vinoreardon@gmail.com

2 Background

In this context, alternative assessment (or *alternate assessment*) which is characterized by an untimed, free-response or open-ended format (Brown, 2004) has been found to be particularly useful in special education programs because it is better suited to individualized feedback. Since it is *formative* in nature (rather than *summative* as in the case of *traditional* forms of assessment), it is more oriented to *process* (as opposed to being *product-focused*), thus making it easier to assess and help students who may have special needs. Because of the extent to which these special needs students are at risk for academic failure, teachers are expected to assume "new" assessment roles, helping their LD students to *self-assess* themselves (for example, within the context of "multiple intelligences") and hence engage in tasks that are more meaningful to them (Brown, 1999, 2002; Gardner, 1983).

It would be helpful to address the fact that an alternative or alternate assessment should not be perceived as either a traditional large-scale assessment or a diagnostic assessment that has been individualized. More importantly, it must be understood that alternate assessments can be administered to students who "differ" greatly in their ability (when responding to stimuli and/or providing responses) in the same way that the general student population does.

In the United States, the federal *Individuals with Disabilities Education Act* (IDEA, 2004) (Wright & Wright, 2007) has stipulated that students with disabilities participate in statewide assessments which include an appropriate form of alternative assessment (this type of alternate assessment subscribed to "modifications" earlier but is currently seen to favor "accommodations"), especially in the areas of reading and mathematics but also in English, as in the case of ELLs. Prior to the 2004 reauthorization of *IDEA 2004*, the term *modifications* referred to *changes* in terms of instructional level or delivery of content in relation to district-wide or statewide tests for students receiving special education services. As a result, modifications resulted in scaling down the standards by which these students were evaluated. Modifications indicated that curriculum and/or delivery of content had been altered to a large extent. It should be understood that when modifications were made, LD students were not expected to master the same academic content as others in a regular classroom setting. In such a situation, grades did not seem to reflect the true abilities of these LD students. For example, a student who was incapable of performing well in the content area of "complex fractions" in a math class might only have been capable of working in the area of "additions" in a very simplified format.

What then was the underlying implication here? It simply indicated that the student's *instructional level* had changed considerably. Therefore, when one looked at the *grade* of the student, it became important to determine whether the student had received this grade in the standard curriculum for his grade level or in a modified curriculum. Consequently, the term "modifications" emphasized the *simplified* versions of these tests. This led to a rather complex situation.

Thus, beginning with *IDEA 2004*, the term "modification" has no longer been used in relation to district-wide and statewide testing because the federal *No Child Left Behind Act* (NCLB) mandates that students with learning disabilities be tested using the same standards as those without the same learning disabilities. According to some of the final regulations in the NCLB Act of 2007, students with disabilities are expected to have access to tests that allow them to demonstrate their competence on those tests (Elliot, 2010). Thus, to be included in an educational "accountability system" (i.e. by generating information about how a school or district or state is doing in terms of overall student performance), a student who is physically or mentally challenged would need to have access to a mechanism which provides *alternate* ways of assessing him/her. This can usually be achieved by administering the test in a different setting and changing the presentation of test items or response format.

Thus, in a research study conducted by Elliot (2010), the findings indicated that students with disabilities performed far better when fewer distractors were used or when more white space was added to the page. Based on these findings, the researchers were able to conclude that modified tests were helpful to students with disabilities as long as those who designed the modified tests were trained in creating tests that could better reflect learning outcomes. This study contributed to the understanding that students who qualified for *AA-AAS* (Alternate Assessment based on Alternate Achievement Standards) did *not* perform as well on multiple-choice tests while students who qualified for the *AA-MAS* (Alternative Assessment based on Modified Achievement Standards) demonstrated that they would need more time to make the required progress.

A study by Kearns, Towles-Reeves, Kleinert, Kleinert, and Thomas (2011) indicated that students who required augmentative and alternative communication devices needed additional time to acquire and maintain the skills; that is, when the subjects were severely impaired, only 50 % of that population could access an alternate assessment. Of equal and related significance, the objective of the study conducted by Kettler (2011) attempted to find a reliable screening method that would identify the small percentage of students that would qualify for AA-MAS. The Computer-Based Alternate Assessment Screening (C-BAAS) test was deemed appropriate to screen students who needed a modified assessment. It became increasingly clear that the results of this test could be used to help teachers and *Individualized Education Program* (IEP) teams offer the instructional support that some students (i.e. those who could *not* meet the required proficiency level within one year in the classroom) needed desperately (Kettler, 2011).

In the United States, an IEP team is expected to determine whether a student is eligible to participate in an alternate assessment (Musson, Thomas, Towles-Reeves, & Kearns, 2010). It is true that if identified inappropriately, students who qualify for the program may fall through the cracks; conversely, a student who qualifies may not become eligible to use the service when misdiagnosed.

However, when a student is identified appropriately, an IEP team becomes responsible for considering the SPED student's unique needs in designing assessment *accommodations*. What are accommodations? Accommodations are best

envisioned as adjustments that are made to ensure that the learning disabled students have equal access to the curriculum so that they can have successful learning outcomes. Therefore, a student with LD can learn the same material but in a different way; in other words, a student who is a slow reader may have the ability to listen to the audio version of the same book (and be exposed to the same content) to allow them to better demonstrate the learning outcomes.

In a related situation, what is of importance is also how a disability intersects, for example, with second language acquisition as in the case of ELLs. It should not be assumed that accommodations designed for typical ELLs (without disabilities) would necessarily give adequate support to ELLs with disabilities (cf., Reardon & Nagaswami, 1997). The issue of disability and how it is affected by language proficiency (or the lack thereof) will be revisited later in this chapter.

In such a situation, the IEP has to be structured in a more specific way so that it can indicate accurately the knowledge and skills of the special needs ELL population. It must be clearly understood that the primary purpose of using alternative assessments is to evaluate students who are not able to participate in general state assessments even when they are provided *accommodations*, for example, by changing the seating arrangement of a student. It is generally the case that through the IEP, accommodations can be developed both formally and informally. In the United States, the 50 states vary a great deal in the specific accommodations that they allow; all 50 states have written guidelines that indicate allowable assessment accommodations for all students with disabilities, including ELLs. What should be remembered is that for any assessment to successfully measure a student's learning outcomes, the assessment needs to be both *valid* and *reliable*. The issues of *validity* and *reliability* will be revisited in a later section of this chapter.

No doubt, accommodations can offer the possibility for the LD student to use "assistive technology" (e.g. a tape recorder or a computer) to complete his/her project on "air pollution" successfully, for example, by not having to struggle with pencil and paper. Moreover, the student's IEP can outline these accommodations clearly, if necessary. For instance, what about those students who are identified as "deaf"? In a study conducted by Cawthon (2011), the researchers concluded that the communication *mode* seemed to affect assessment outcomes, so it was important to use the same mode that the students had been familiarized with in classroom instruction. In such a context, especially when attempting to establish which students needed to take the AA-MAS, a posttest questionnaire seemed to shed light on how items should be presented in a test (Roach, Beddow, Kurz, Kettler, & Elliott, 2010) to make alternative assessment more viable in terms of producing desirable learning outcomes. Thus, investigating if variables such as "visuals" or "bolded items" or "fewer distractors" made a difference in test outcomes became the focus of the study. The results of the test showed that some of these differences proved to be useful to those students who *perceived* them to be helpful.

Thus, this led to the premise that a student's individual learning style was critical in determining how instruction would need to be tailored. It should be underscored that students with disabilities vary greatly in terms of *how* they learn even when they are exposed to the same content. The inherent assumption here is that content

should be delivered differently to suit the needs of the individual learner in order to have appropriate learning outcomes. This leads us to the following question: what pedagogical implications does this have for classroom instruction?

3   Implications for Classroom Instruction

Evidence indicates that instructional planning would influence how delivery is sustained within the broader scope of implementing and evaluating the IEPs of students who have disabilities (Johnson, 2012). Since school districts are accountable for the learning outcomes of *all* students, those that have disabilities are not exempted from this group of students that get monitored for growth and progress on a yearly basis. What this implies is that test scores would need to be made more *valid* within the overall context of using modifications and/or accommodations. In a study conducted by Hager and Slocum (2011), the researchers examined various ways (by conforming to federal guidelines) through which alternative assessment could be used so that students could have successful learning outcomes.

Thus, at this juncture, it would also be helpful to emphasize that alternate assessment has been used with students whose learning disabilities are not too severe and who have responded more favorably to *general assessment tools* (i.e. the general assessment instruments that are used in regular classroom situations when these students are placed in "pull-out" K-12 programs). It should be mentioned that these students are *mainstreamed* as and when it is deemed appropriate.

In addition, it is important to focus on the fact that alternate assessment has been used in the United States even in "general education" classrooms (i.e. in regular classrooms as opposed to "special education" classrooms) as a way to extend day-to-day classroom activities and to promote problem-solving skills. It should be noted that the notion of *alternative assessment* has become more popular, especially in ESL and/or EFL classrooms, since the early 1990s as a result of teachers realizing that not all students and not all skills can be measured by traditional tests. (Brown, 2004). Alternate assessment is often seen as a continuous long-term assessment which fosters intrinsic motivation and uses criterion-referenced scores (the criterion is often established by the teacher, who uses portfolios, for example, as a form of alternative assessment). In contrast, traditional assessment is often perceived as one favoring a one-shot exam (e.g. a standardized exam) which is timed and promotes extrinsic motivation, often using norm-referenced scores (Brown, 2004). Thus, the term *alternative* was proposed (Huerta-Macias, 1995) based on what teachers felt were conceivably shortcomings displayed by traditional tests.

Proponents of alternative assessment believe that alternative assessment is a more equitable form of assessment than traditional assessment (Lynch, 2001) because it includes a variety of measures or instruments to assess students. For example, by using tools such as journals, performance-based assessment, student-teacher conferences, self and peer assessments, observations, and portfolios,

students can be measured in various ways. Therefore, Brown and Hudson (1998) felt that it would be more appropriate to refer to these measures as "alternatives" in assessment, which could be considered a subset of assessment (and not as something that would be excluded from test design or construction).

More importantly, in recent times, *alternatives* in assessment have been embraced by those educators in the mainstream who feel that traditional assessment alone cannot offer the panacea to account for students' learning outcomes at the post-secondary/tertiary level. It appears that alternative assessment is more sensitive to heterogeneous populations from different backgrounds; that is, ESL and EFL students with and without disabilities (as well as those in SPED programs). As Hamayan (1995, p. 213) stated, "Alternative assessment refers to procedures and techniques which can be used within the context of instruction and can be easily incorporated into the daily activities of the school or classroom." More specifically, it would be useful to point out that it is particularly suited for those with limited English skills because as Huerta-Macias (1995, p. 9) claimed, "students are evaluated on what they integrate and produce rather than on what they are able to recall and reproduce."

4   The Defining Characteristics of Alternatives in Assessment

It is important to underscore that alternatives in assessment need to use authentic or real-world contexts (Brown & Hudson, 1998). It would also be useful to point out that students should be assessed on what would be considered normal everyday classroom activities where both *process* and *product* are equally emphasized. Ideally, these alternatives should be able to provide pertinent information about the students (i.e., their strengths and weaknesses) within the overall context of developing their critical thinking skills. At this point, it may be helpful to examine why the following kinds of student performance should be observed, particularly in a language classroom: discourse-level skills, sentence-level oral production skills (e.g. grammatical features), responses to tasks, interaction with other students, and evidence of listening comprehension or nonverbal behavior, amongst other forms of behavior.

It is important to emphasize that in these types of learning environments, trained instructors (and not machines) should do the scoring, which implies that scoring criteria will need to be determined and raters will need to be trained so that the issue of *reliability* (i.e., consistency and accuracy) would not be compromised (Bachman, 1990). In this context, it seems more appropriate to use "alternatives in assessment" with our special needs students and EFL/ESL students (with or without disabilities) to better reflect the learning outcomes with the following forms of alternative assessment (Brown, 2004):

• Journals
• Conferences and interviews

- Observations
- Self- and peer-assessments
- Performance-based assessment
- Portfolios

Thus, at this juncture, it would be useful to examine each of these *alternatives in assessment* to gain a deeper understanding of how each type of alternative in assessment can be employed to better suit the needs of atypical student populations.

For example, *journals* (Brown, 2004; Staton, Shuy, Peyton, & Reed, 1987) could be employed as assessment tools in the following manner: as classroom-oriented dialogue journals (emphasizing interaction between teacher and student), language-learning logs (self-monitoring one's own achievement in one or more skills), diaries of feelings or attitudes, reflections, grammar journals (monitoring one's own errors), reading response journals, and so on.

As far as *conferences and interviews* (Brown, 2004) are concerned, they could be seen to have the following functions in the classroom: commenting on essays, responding to portfolios, focusing on aspects of oral presentations, monitoring research or project proposals, assessing general progress in a course, advising a student (in relation to a student's academic plan), setting personal goals for the course, and many such related functions. It should be noted that *systematic observations* (not just informal observations) can also be useful in terms of the information that can be gleaned about students (Spada & Frolich, 1995).

Additionally, it is helpful to see how Brown (2004) categorizes *self and peer assessments* into: direct assessment of performance (e.g. student monitors himself after an oral presentation); indirect assessment of performance (for instance, self- or peer-assessment targets larger amounts of time in the form of questionnaires); and metacognitive assessment (more strategic in nature) by setting goals/strategic planning. Students are expected to perform, produce or create something in real-world contexts/situations where classroom activities are an extension of day-to-day life. Students are also trained to tap into critical thinking skills. Thus, as mentioned, this kind of assessment is likely to provide information about students' weaknesses and strengths. Of course, such an assessment would require an open disclosure of standards and rating criteria so that teachers/students are in a better position to make informed decisions.

Clearly, *performance-based assessment* can include interviews, role plays, story retelling, oral speeches or presentations, oral summarizing of articles, oral reports, and so on. Pierce and O'Malley (1992) suggest that teachers could elicit stories from students by offering them a set of pictures (it would be best if teachers could give them visual cues). Role plays are also seen as a way to get students to produce language by transforming themselves into "characters" (Kelner, 1993).

Now, finally, this brings us to the most popular alternative in assessment, namely *portfolios*, which are used to maximize the positive *washback effect* (i.e., the positive effect of testing on instruction and learning).

5 The Use of Portfolios in the United States and Other Countries

For a number of years, portfolios have been used successfully in the United States with mentally challenged students in the field of special education. However, more recently, portfolios are being increasingly used with English language learners in ESL/EFL classrooms in different parts of the world. Portfolios can be envisioned as both *formative portfolios* (this emphasizes process) and *summative* (this emphasizes the learning outcomes) *portfolios* (Cooper & Love, 2001, cited in Ali, 2005). Genesee and Upshur (1996) believe that a portfolio is reflective of a student's growth and achievement in a certain skill or area. What needs to be remembered is that successful portfolio development is linked to well-stated objectives, clearly developed guidelines and criteria for assessment where a variety of entries included in the portfolio assume significance (Yawkey, Gonzalez, & Juan, 1994). The reflective nature of portfolios needs to be underscored in the context of *formative* (i.e., ongoing assessment during instruction) assessment, where the issues of positive *washback effect* and *validity* (i.e., whether a test measures what it is supposed to measure) assume significance.

The acronym CRADLE (i.e., Collecting, Reflecting, Assessing, Documenting, Linking, Evaluation) was suggested by Gottlieb (1995) to highlight the dynamic nature of portfolios. Portfolios are seen to have numerous advantages (O'Malley & Pierce, 1996; Weigle, 2002). Benefits include, for example, a student taking control of the learning process (the teacher is envisioned as a facilitator) because the portfolio becomes the evidence of this student's progress in a course. Furthermore, portfolios can accommodate assessment on a multidimensional basis because there are many different kinds of portfolios that are used for various purposes. From this perspective, portfolios are seen to be employed in much the same way in both the United States and Russia (Reardon & Kozhevnikova, 2008). As Brown (2004) underscores, a successful portfolio would need to subscribe to some of the following guidelines: stating objectives clearly, giving clear guidelines, communicating assessment criteria openly to students, designating time within the curriculum for portfolio development, establishing schedules for conferencing, and providing positive *washback* when generating final assessments.

It might be best if teachers could maintain *anecdotal records* (Tierney, Carter, & Desai, 1991) to keep track of their regularly scheduled conferences with their students so that students would be able to critique their own portfolios. As is commonly understood, portfolios could include numerous kinds of materials (Brown, 2004; Tierney, Carter, & Desai, 1991) including: tests and quizzes, essays, different kinds of writing assignments, reports, projects, poetry, creative prose, artwork, photographs, audio/video recordings of presentations, journals, book reports, grammar logs, checklists created by teachers and students, self- and peer-assessments, reflections, and so on.

Currently, there are five major types of portfolios, which are used in schools in different parts of the world: the *European Language Portfolio*, the *Collections*

Portfolio, the *Showcase Portfolio*, the *Assessment Portfolio*, and the *Electronic Portfolio*.

The *European Language Portfolio* (ELP) is perhaps the first one to be introduced to different systems of education. It appears to be the most popular one in Russia (Reardon & Kozhevnikova, 2008). Based on its widespread use, the ELP is seen to foster student autonomy, with self-assessment being a major aspect of it. Based on common perception, the ELP can be seen to be characterized by the following components:

- It expresses the student's "linguistic identity" (or "passport") by accounting for the second language that has been learned and the student's assessment of his or her current competence (the *passport* could contain certificates that the student has obtained to indicate competence) in the target language;
- It can also be viewed as a language "biography" that can account for one's goals and objectives (the *biography* could contain "self-assessment" materials) in the target language;
- It can be considered a "dossier" which contains a selection of work that monitors the student's growth and progress in the target language (Little, 2005).

The *self-assessment checklists*, which constitute an inherent part of the ELP, are derived from the "Common European Framework of Reference" (CEFR). Students can reflect on their experiences and assess their accomplishments against the checklist. They can also set themselves "goals" and assess their progress by taking "ownership" of the learning process. No matter how beneficial the use of the ELP is, some teachers and students find the ELP cumbersome (the storage of portfolios could prove to be problematic) even though the language portfolio provides clear-cut learning objectives and a way to record growth and progress.

The *Collections Portfolio*, an "aggregate of everything a student produces" (Moya & O'Malley, 1994), consists of a student's entire work, so it is popular at the elementary/primary school level because it represents a variety of daily assignments that a young learner deals with in the classroom and at home. At this age, students do not produce many entries in terms of daily assignments that can go into a portfolio, so the portfolio is quite manageable. For example, portfolios could include a young student's art work, tests, photographs, and even handicrafts. They are almost always kept in the classroom, and are easily accessible to anyone who would like to see these portfolios (e.g. parents or school officials). Maintaining collection portfolios can be very motivating both for young students and teachers because the students can view their own achievements and also the accomplishments of their peers (there can be a big difference in the way these portfolios are personalized and designed). Teachers, on the other hand, can use portfolios to measure each student's progress on a daily or weekly basis.

The *Showcase Portfolio* is often used to exhibit a student's best work (O'Malley & Pierce, 1996). Entries could include (but are not limited to) short stories, essays, poems, audio and video samples. Samples in the portfolio are very carefully selected to illustrate a student's achievement in the classroom and outside the

classroom (i.e. samples that a student considers most representative of his best work). Sometimes, university professors also favor the idea of maintaining such portfolios.

The *Assessment Portfolio* is becoming more popular with teachers all over the world because it is so manageable in that it can use scoring guides and rubrics (i.e. predetermined criteria) to measure a student's work (O'Malley & Pierce, 1996). More importantly, it helps a student to become a responsible learner by taking control of his/her learning process. Using assessment portfolios also gives the teacher opportunities to monitor a student's growth over a period of time, which is an important factor in measuring educational success. Portfolios provide good evidence of a student's achievements for school officials at the secondary level.

The assessment portfolio can be particularly useful in EFL classrooms where a teacher can account for both students with and without *special needs*. Not too long ago, in my own EFL classes (Reardon, 2014) in the Sultanate of Oman, the majority of students claimed that they found *assessment portfolios* to be useful as reflective tools. By looking at their portfolios, they were surprised at how far they had come in acquiring writing skills, for example. Thus, when they looked at the third drafts of their essays, they were actually able to see the progress that they had made. No doubt, this can be especially significant in the case of LD students. For the teacher, it is also a time to reflect on how things could be improved in the future.

The *Electronic Portfolio* is also gaining in popularity these days. As Barrett (2000, cited in Ali, 2005) stated, it consists of the "use of electronic technologies that allow the portfolio developer to collect and organize artifacts in many formats (audio, video, graphics, and text). A standards-based electronic portfolio uses hypertext links to organize the material to connect artifacts to appropriate goals… not a haphazard collection of artifacts (i.e. a digital scrapbook) but a reflective tool". Thus, the electronic portfolio is something that most teachers find useful these days.

6   Portfolio Model and Implementation

What about a portfolio model that teachers can readily access? Moya and O'Malley (1994) offer a portfolio model that has the following characteristics: identifying the purpose (e.g. plan/focus), planning the contents (including how often to assess), designing an analysis of the portfolio (e.g. standards and criteria), preparing for instruction (e.g. giving feedback to students), and verification of procedures (establishing a system for validating decisions). This seems like a good starting point for a teacher who is looking for guidelines to make the portfolio more manageable.

What are some steps that can be followed in implementing portfolios that are manageable? Huang (2012) recommends the following seven steps: planning the assessment purpose, determining portfolio tasks, establishing criteria for assessment, outlining organizational steps, preparing the students, monitoring the portfolio, and assessing the portfolio.

Thus, by using a good model and following these steps, it is possible to make the portfolio a successful assessment instrument and allow the implementation of the portfolio with greater success in all kinds of learning environments.

7   Portfolio Objectives, Criteria, and Assessment

As teachers get students' consent (Ali, 2005) in developing objectives and criteria for portfolio assessment, they are in a better position to specify the time frame and portfolio organization because the latter become actively involved in their own learning and start *reflecting* (Ali, 2002) on their learning styles or strategies, their successful practices or their failures. Learners understand that they can take control of their learning. Self-assessment and peer assessment constitute an important part of portfolio assessment: they increase students' ability to monitor their own progress and set improvement goals. By getting students involved in peer review and including a feedback session midway through the course (Ali, 2005), the teacher can receive a more in-depth knowledge of the student as a learner in terms of his or her learning styles and strategies, and strong and weak points, which means that the teacher can individualize instruction for the student and become more involved in adjusting the curriculum objectives, if necessary.

As mentioned earlier in this chapter, the issues of *reliability* and *validity* are crucial for any assessment tool. Based on our common understanding of testing constructs, *reliability* refers to the consistency and accuracy of the assessment tool (Coombe, Folse, & Hubley, 2007). For a portfolio to be considered reliable, the teacher should develop clear-cut goals and objectives, detailed criteria, explicit and consistent ways of scoring a student's portfolio. The students should fully understand what they are supposed to do and how their work is going to be assessed. Furthermore, the student and the teacher can negotiate the goals and objectives of the portfolio, its content, organization and the expected learning outcomes. Thus, students can develop the feeling of ownership and responsibility when designing their portfolios.

The *validity* of an assessment tool is related to how well this tool does what it is intended to do; that is, whether the tool succeeds in informing the teacher about the student's progress toward some goal in a curriculum or course of study (Bachman, 1990). *Content or curriculum validity* should ensure correspondence between curriculum objectives and the contents of a portfolio. Therefore, the teacher should make sure that the goals of the portfolio are in line with curriculum objectives, the contents of the portfolio match its goal and organization, and the criteria developed are related to the objectives (Delett, Barnhardt, & Kevorkian, 2001). This way, the teacher will integrate instruction and assessment, with portfolios having high *content validity*. One other type of validity relevant to portfolio assessment is *face validity*. This refers to the way the students perceive the portfolio as an assessment tool. If learners do not have a clear understanding of the benefits of the portfolio for their learning outcomes, they will not regard it as a credible assessment tool, and

portfolios will have low *face validity* (Bernhardt, Kevorkian, & Delett, 1999). Thus, before introducing the portfolio as an assessment tool, teachers should carefully explain to the students its benefits and challenges. Then, as long as teachers can establish clear guidelines for how portfolios should be developed and evaluated, portfolios can become extremely *reliable*. No doubt, from the perspective of *washback effect*, portfolios can be considered *valid* tools of assessment.

When portfolios are appropriately designed, they are seen to have the following benefits: they can offer a multi-dimensional perspective of a student's progress over time, promote self-reflection or learner autonomy, and integrate learning, teaching and assessment. Thus, a portfolio can help the student to extend learning and the teacher to gain insight about the student's learning process (Banfi, 2003; Delett et al., 2001).

8   Rationale for Using Portfolios in EFL/ESL Classrooms

The rationale derives from the necessity for developing adaptable alternate assessment techniques that will best suit the needs of diverse bilingual student populations with or without learning disabilities. Typically, with limited English proficiency, EFL/ESL students score poorly on traditional tests (as in standardized tests). Thus, "single-measure approaches" should be avoided by using portfolios which could be used for ongoing "continuous assessment" (Haney & Madaus, 1989).

In this context, it is also crucial to point out that the problem of limited English proficiency is exacerbated when the ESL/EFL bilingual student has learning disabilities which affect learning outcomes adversely. Thus, in considering the following scenario where disability intersects with language proficiency, a *differential diagnosis* is often necessitated, for example, to distinguish between "normal disfluency" (associated with a bilingual student who is considered a "beginner" learning English) and "stuttering" (associated with a bilingual student who has a speech-language disability). Therefore, if an EFL/ESL student has such a speech-language disability (e.g. stuttering) it can make the issue of measuring learning outcomes rather complex. In this context, how can a teacher appropriately evaluate the problem and offer the proper intervention? Clearly, it is important to understand that in diagnosing nonnative speakers of English who stutter, clinicians use an established assessment instrument to measure speech-language proficiency (for example, the SPEAK test) that includes an assessment of *normal disfluency* while conducting a parallel assessment of *stuttering* (for example, the SSI test).

9   Normal Disfluency and Stuttering

What makes the issue of distinguishing "true" stuttering from normal disfluency rather complex is, as Van Riper (1982) notes, the fact that nonstutterers occasionally exhibit similar behaviors as stutterers although stuttering is almost always characterized by "struggle". According to Peters and Guitar (1991, p. 73), some of the features that distinguish normal disfluency from stuttering are "…the *amount* of disfluency, the *number* of units in repetitions and interjections, and the *type* of disfluency…" (italics added).

Commonly, the speech of nonstuttering bilinguals is also characterized by disfluency in the form of false starts, interjections, and repetitions, often considered the hallmark of second language acquisition. Thus, this type of normal disfluency that is triggered by "interference" or "transfer" gives rise to a variety of possibilities. Conceivably, normal disfluency or stuttering might occur in a similar way in the two languages (Yudaken, 1975, cited in Jankelowitz & Bortz, 1996; Woods & Wright, 1998); or, stuttering may occur in only one of the two languages. Then again, it might occur in both languages but not to the same degree (i.e., depending on the "level" of proficiency), thus underscoring the need to investigate the relationship between level of proficiency and degree of stuttering.

10   Level of Linguistic Proficiency and Degree of Stuttering

When Jankelowitz and Bortz (1996) assessed the language proficiency of an Afrikaans-English bilingual stutterer who used the two languages interchangeably, they reported that linguistic competence influenced frequency, distribution, and nature of stuttering. Findings demonstrated that the subject, a 63 year-old English-Afrikaans speaker, who was a *compound bilingual* (i.e., a bilingual who learns two languages at the same time), stuttered less in Afrikaans, which was his dominant language. Apparently, although both languages were spoken at home, Afrikaans predominated. Clearly, as the authors indicated, the findings demonstrated that linguistic competence influenced stuttering.

In addition, a pilot study conducted by Reardon (1998), involving a Russian-English *coordinate bilingual* (i.e. a bilingual who learns the second language later) stutterer, served to clarify these previous findings by demonstrating that the subject, a Russian-English bilingual, stuttered more in his second or less dominant (less proficient) language, English. Thus, both case studies contributed to the hypothesis that language proficiency seemed to play a role in promoting stuttering; that is, the less proficient a stutterer is in a particular language, the more likely it is that he would stutter.

Thus, Reardon's (2000) investigation sought to clarify and extend previous research that used very small samples ($N = 1$) in investigating the interaction

between stuttering and language competence in the context of bilingualism. As with past studies (Reardon, 1998), the findings of this investigation (Reardon, 2000) served to support an inverse relationship between language proficiency and stuttering ($r = -0.56$). That is, the greater the proficiency, the less the stuttering. Conversely, the less the language proficiency, the greater the stuttering.

Because the findings of this study indicated that the greater the stuttering the less the proficiency, it led to the following conclusion: as a stutterer gains proficiency in a language, it is less likely that he/she will stutter to the same degree either qualitatively or quantitatively in that language. Interestingly, this finding seems logical in the context of stutterers who engage in substitutions to avoid stuttering. It makes sense that stutterers who are more proficient linguistically would be more able to substitute words that they anticipate as problematic.

Clearly, this finding has pedagogical implications for intervention in EFL/ESL classrooms. It is strongly recommended that a student's speech therapist collaborate with a language specialist (i.e., an EFL/ESL specialist in the case where English is spoken as a foreign or second language) in order to bring about effective intervention. In such a scenario, it is quite possible that as the stutterer becomes more proficient in the language that he stutters in, his stuttering would decrease in that language.

11   Conclusion

Thus, from a pedagogical perspective, it seems crucial that in designing an IEP for a student who stutters and is being offered alternative assessment, that speech-language pathologists understand the differences among EFL/ESL students who are culturally and linguistically different so that they can "accurately discriminate an actual language disorder from the natural progression of second language acquisition and acculturation," (Morsink, Thomas, & Correa, 1991, p. 282). Therefore, it is critically important that the EFL/ESL instructor and the speech-language pathologist work together as an interactive IEP team to enhance learning outcomes in the classroom by using the appropriate alternative assessment tools, for example, portfolios.

In summation, does alternative assessment work in EFL/ESL classrooms? In general, the benefits of alternative assessment are well supported in the literature, especially in relation to portfolios (Allen, 2004; Balhous, 2008; Caner, 2010; Lo, 2010). Alternative assessment provides the teacher with extensive information about students as autonomous learners. As research indicates, most students have positive attitudes toward alternative assessment, in particular toward portfolios.

Clearly, it can be demonstrated that alternative assessment can be successfully used with all students (with and without special needs) in EFL/ESL classrooms at all levels of education. Thus, I do believe that this chapter has contributed to the understanding that alternative assessment can best serve the needs of ESL/EFL populations with and without speech-language disabilities.

References

Ali, S. Y. (2002). Portfolio assessment: Motivating ESP students to write better through continuous assessment. In S. Troudi, S. Riley & C. Coombe, C. (Eds.), *EFL challenges in the new millennium*. TESOL Arabia 2001 Conference Proceedings (Vol. 6).

Ali, S. Y. (2005). An introduction to electronic portfolios in the language classroom. *The Internet TESL Journal, 11*(8). Retrieved from http://iteslj.org.

Allen, L. Q. (2004). Implementing a culture portfolio project within a constructivist paradigm. *Foreign Language Annals, 37*(2), 232–39.

Bachman, L. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.

Balhous, R. (2008). The self-assessed portfolio: A case study. *Assessment and Evaluation in Higher Education, 33*(4), 381–393.

Banfi, C. S. (2003). Portfolios: Integrating advanced language, academic, and professional skills. *ELT Journal, 57*(1), 34–42.

Bernhardt, S., Kevorkian, J., & Delett, J. (1999). *Portfolio assessment in the foreign language classroom*. Washington, DC: National Capital Language Resource Center.

Brown, H. D. (1999). *New vistas: An interactive course in English*. White Plains, NY: Pearson Education.

Brown, H. D. (2002). *Strategies for success: A practical guide to learning English*. White Plains, NY: Pearson Education.

Brown, H. D. (2004). *Language assessment: Principles and classroom practices*. New York: Longman.

Brown, J. D., & Hudson, T. (1998). The alternatives in language assessment. *TESOL Quarterly, 32*, 653–675.

Caner, A. (2010). Students' views on using portfolio assessment in EFL writing courses. *Anadolu University Journal of Social Sciences, 10*(1), 223–236.

Cawthon, S. (2011). Making decisions about assessment practices for students who are deaf or hard of hearing. *Remedial and Special Education, 32*(1), 4–21.

Coombe, C., Folse, K., & Hubley, N. (2007). *A practical guide to assessing English language learners*. Ann Arbor: The University of Michigan Press.

Delett, J. S., Barnhardt, S., & Kevorkian, J. A. (2001). A framework for portfolio assessment in the foreign language classroom. *Foreign Language Annals, 34*(6), 559–68.

Elliot, S. (2010). Effects of using modified items to test students with persistent academic difficulties. *Exceptional Children, 76*(4), 475–495.

Gardner, H. (1983). *Frames of mind: The theory of multiple intelligences*. New York: Basic Books.

Genesee, F., & Upshur, J. A. (1996). *Classroom-based evaluation in second language education*. Cambridge: Cambridge University Press.

Gottlieb, M. (1995). Nurturing student learning through portfolios. *TESOL Journal, 5*, 12–14.

Hager, K., & Slocum, T. (2011). Using alternate assessment to improve educational outcomes. *Rural Special Education Quarterly, 30*(1), 24–29.

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

Haney, W., & Madaus, G. (1989). Searching for alternatives to standardized tests: Whys, whats, and whithers. *Phi Delta Kappan, 70*, 683–687.

Huang, J. (2012). The implementation of portfolio assessment in integrated English course. *English Language and Literature Studies, 2*(4), 15–21.

Huerta-Marcia, (1995). Alternative assessment: Responses to commonly asked questions. *TESOL Journal, 5*, 8–11.

Jankelowitz, D. L., & Bortz, M. A. (1996). The interaction of bilingualism and stuttering in an adult. *Journal of Communication Disorders, 29*, 223–234.

Johnson, P. (2012). The benefits and administration of alternate assessments. *New Horizons of Learning, 10*(1). Retrieved from http://education.jhu.edu/PD/newhorizons/Journals/Winter2012/Johnson.

Kearns, J., Towles-Reeves, E., Kleinert, H., Kleinert, J., & Thomas, M. (2011). Characteristics of and implications for students participating in alternate assessments based on alternate academic achievement standards. *Journal of Special Education, 45*(1), 3–14.

Kelner, L. B. (1993). *The creative classroom: A guide for using creative drama in the classroom, preK-6*. Portsmouth: Heinemann.

Kettler, R. (2011). Computer-based screening for the new modified alternate assessment. *Journal of Psychoeducational Assessment, 29*(1), 3–13.

Little, D. (2005). The Common European Framework and the European Language Portfolio: Involving learners and their judgments in the assessment process. *Language Testing, 22*, 321–336.

Lo, Y. (2010). Implementing reflective portfolios for promoting autonomous learning among EFL college students in Taiwan. *Language Teaching Research, 14*(1), 77–95.

Lynch, B. K. (2001). The ethical potential of alternative language assessment. In C. Elder (Ed.), *Experimenting with uncertainty: Essays in honour of Alan Davies* (pp. 228–239). Cambridge: Cambridge University Press.

Morsink, C. V., Thomas, C. C., & Correa, V. I. (1991). *Interactive teaming: Consultation and collaboration in special programs*. New York, NY: Macmillan.

Moya, S. S., & O'Malley, J. M. (1994). A portfolio assessment model for ESL. *The Journal of Educational Issues of Language Minority Students, 13*, 13–36.

Musson, J., Thomas, M., Towles-Reeves, E., & Kearns, J. (2010). An analysis of state alternate assessment participation guidelines. *Journal of Special Education, 44*(2), 67–78.

O'Malley, J. M., & Pierce, L. V. (1996). *Authentic assessment for English language learners: Practical approaches for teachers*. White Plains, NY: Addison-Wesley.

Peters, T. J., & Guitar, B. (1991). *Stuttering: An integrated approach to its nature and treatment*. Baltimore, MD: Williams and Wilkins.

Pierce, L. V., & O'Malley, J. M. (1992). *Performance and portfolio assessment for language minority students*. Washington, DC: National Clearinghouse for Bilingual Education.

Reardon, V. S. (1998). *Disfluency: A cross-linguistic comparison*. Unpublished manuscript, University of Kansas.

Reardon, V. S. (2000). *Stuttering in the context of bilingualism: The relationship between language proficiency and degree of stuttering*. Unpublished doctoral dissertation, University of Kansas, USA.

Reardon, V. S. (2014). *Special needs learners, diagnosis and intervention: Alternative assessment*. Paper Presented at the Oman 14th International ELT Conference, Muscat, Oman.

Reardon, V. S., & Kozhevnikova, L. (2008). Alternative assessment and the use of portfolios in the United States and Russia. In *International Academic-Applied Conference, Moscow, Russia*. Moscow: Cambridge University Press.

Reardon, V. S., & Nagaswami, G. (1997). *Attenuating speech disability with e-mail*. Paper Presented at the Thirty-first Annual Convention & Exposition of the Teachers of English to Speakers of Other Languages, Inc., Orlando, Florida.

Roach, A., Beddow, P., Kurz, A., Kettler, R., & Elliott, S. (2010). Incorporating student input in developing alternate assessments based on modified academic achievement standards. *Exceptional Children, 77*(1), 61.

Spada, N., & Frolich, M. (1995). *Communicative orientation of language teaching observation scheme*. Sydney: National Centre for English Language Teaching and Research, Macquarie University.

Staton, J., Shuy, R. W., Peyton, J. K., & Reed, L. (1987). *Dialogue journal communication: Classroom, linguistic, social and cognitive views*. Norwood, NJ: Ablex.

Tierney, R. J., Carter, M. A., & Desai, L. E. (1991). *Portfolio assessment in the reading-writing classroom*. Norwood: Christopher Gordon.

Van Riper, C. (1982). *The nature of stuttering*. Englewood Cliffs, NJ: Prentice-Hall.

Weigle, S. C. (2002). *Assessing writing*. Cambridge: Cambridge University Press.

Woods, D. W., & Wright, L. W., Jr. (1998). Dismantling simplified regulated and experimental breathing: A case of a bilingual stutterer. *Journal of Behavior Therapy Psychiatry, 29*, 179–186.

Wright, P., & Wright, P. (2007). *Wrightslaw: Special education law* (2nd ed.). Virginia: Harbor House Law Press Inc.

Yawkey, T. D., Gonzalez, V., & Juan, Y. (1994). Literacy and biliteracy strategies and approaches for young culturally and linguistically diverse children: Academic excellence P.I.A.G.E.T. comes alive. *Journal of Reading Improvement, 3*(3), 130–141.