

ASSESSING AND EVALUATING STUDENTS IN A FOREIGN LANGUAGE LEARNING

Adrianto Nasution¹ & Ayu Bandu Retnomurti²

Universitas Negeri Padang¹ & Universitas Indraprasta PGRI²
adrianto_nasution@yahoo.com
081374842103

Abstrak

Penilaian dan evaluasi peserta didik dalam pembelajaran bahasa kedua ditujukan bagi para pendidik profesional yang sering bertindak sebagai (atau bekerja sama dengan) evaluator dalam program pembelajaran bahasa kedua. Menurut undang-undang RI No. 20 tahun 2003 tentang sistem pendidikan nasional pasal 57 ayat 1, evaluasi dilakukan dalam rangka pengendalian mutu pendidikan nasional sebagai bentuk akuntabilitas penyelenggara pendidikan kepada pihak-pihak yang berkepentingan, di antaranya terhadap peserta didik, lembaga, dan program pendidikan. Di sisi lain, penilaian dan evaluasi juga merupakan tolok ukur bagi pendidik yang berfungsi sebagai instrumen untuk menolong peserta didik dalam memahami ilmu pengetahuan yang diidentifikasi sebagai "bahasa Inggris". Penilaian hasil belajar pada dasarnya adalah mempermasalahkan bagaimana pendidik dapat mengetahui hasil pembelajaran yang telah dilakukan. Artikel ini menganalisis sejumlah prinsip penilaian dan evaluasi beserta prosesnya yang merupakan bagian tugas dari seorang pendidik yang mengarah pada tujuan tercapainya kesuksesan pengajaran dan pembelajaran.

Kata Kunci: Penilaian, evaluasi, pembelajaran bahasa kedua, kurikulum, pendidikan.

Abstract

Assessment and evaluation of learners in learning a second language is intended for professional educators who often act as (or work with) evaluators in second language learning program. According to the law of Republic of. 20 of 2003 on the national education system of article 57, paragraph 1, the evaluation done in order to control the quality of national education as a form of accountability organizers of education to the parties concerned, including the learners, institutions, and educational programs. On the other hand, assessment and evaluation is also a benchmark for educators who serve as an instrument to help the students in understanding science identified as "English". Assessment of learning outcomes is basically concerned about how educators can find out the learning outcomes that have been done. An educator must know the extent to which students have understood the material that has been taught or the extent to which objectives / competencies of learning activities that are managed can be achieved. This article analyzes a number of principles of assessment and evaluation processes and tasks that are part of an educator that lead to successful achievement of the goal of teaching and learning

Keywords: Assessment, evaluation, second language learning, curriculum, education.

INTRODUCTION

Assessment and evaluation are essential components of teaching and learning in English language arts. Without an effective evaluation program, it is impossible to know whether students have learned, whether teaching has been effective, or how best to address student learning needs. The quality of the assessment and evaluation in the educational process has a profound and well-established link to student performance. Research consistently shows that regular monitoring and feedback are essential to improving student learning.

What is assessed and evaluated, how it is assessed and evaluated, and how results are communicated send clear messages to students and others about what is really valued—what is worth learning, how it should be learned, what elements of quality are most important, and how well students are expected to perform. Although the terms assessment and evaluation are often used interchangeably, in actuality they are two parts of the same process. Assessment is the process of gathering evidence of what the child can do. Evaluation is the process that follows this collection of data, including analysis and reflection, as well as decisions based on the data.

The assessment and evaluation of literacy needs from the learner's perspective is an important part of an instructional program. Learners come to adult English as a Second Language programs for diverse reasons. Although they may say they just want to "learn English," they frequently have very specific learning goals and needs; for example, to be able to read to their children, speak with their children's teachers, or to get a job. If their needs are not met, they are more likely to drop out than to voice their dissatisfaction. Therefore, using informal, self-assessment tools to gauge learner needs and goals is important. Also important, of course, is using formal assessment tools to gauge learner progress.

The needs assessment and evaluation process can be used as the basis for developing curricula and classroom practice

that are responsive to learners' needs. It encompasses both what learners know and can do and what they want to learn and be able to do. Learners also need opportunities to evaluate what they have learned—to track their progress toward meeting goals they have set for themselves in learning English.

RESEARCH METHODOLOGY

This is a qualitative research and the research is all about exploring issues, understanding phenomena, and answering questions which aims to gather an in-depth understanding of human behavior and the reasons that govern such behavior.

Procedure:

Planning the Assessment

In planning assessments to be taken by the general student population, including ELLs (English Language learners), the general principles of good assessment practices apply.

Test Purpose

The purpose of a test must be clear in order for valid interpretations to be made on the basis of the test scores. Tests have different purposes. For example, one test may be used to evaluate students' readiness to advance to the next grade, while another evaluates students' need for remediation. It is also important to outline the specific interpretations that will be made based on the scores. For example, tests used as a criterion for high school graduation will affect students differently than tests designed to inform instructional decisions.

Developing Test Items and Scoring Criteria

Matching the Task to the Purpose

The first step in developing a test item should be to link, directly to the test specifications and content standards, the content and skill that the item is supposed to measure. If the items require a high level of English proficiency, unrelated to the construct as defined, this will likely affect

the scores for ELLs as well as students in the general population. For content area assessments, only include items that require high degrees of English proficiency if they are consistent with the assessment specifications. Examples of items that require a high degree of English proficiency are those that ask examinees to identify or provide specific definitions or terminology in English that are unrelated to the construct, or items that are evaluated based on the quality of the language in a constructed response.

Item writers and reviewers should work to ensure that all test items maintain specificity in their match to content guidelines. As part of the process of creating and reviewing test material to ensure that it is appropriate and accessible to examinees, it is important that item developers, state content review staff, and state review committees analyze each item critically to ensure that it only measures the intended construct.

Defining Expectations

Because ELLs—just like students in the general population—come from a wide variety of cultural and educational backgrounds, item writers should not assume that students have had any previous experience with given tasks. For example, students should be told explicitly what type of response is acceptable for a constructed-response question, whether it is a paragraph, complete sentence, list, diagram, mathematical equation, and so on. Likewise, the criteria for the evaluation of the response should be made clear to the student. As this may add a significant reading load to the directions, information about how responses will be scored may be especially helpful if students receive it prior to the test.

Using Accessible Language

Using clear and accessible language is a key component of minimizing construct-irrelevant variance. However, do not simplify language that is part of the construct being assessed (e.g., the passages on a reading comprehension test or challenging

vocabulary that is part of the construct of a subject area test). In other cases, though, the language of presentation should be as simple and clear as possible. Some general guidelines for using accessible language are provided below:

- Use vocabulary that will be widely accessible to students. Avoid colloquial and idiomatic expressions, words with multiple meanings, and unduly challenging words that are not part of the construct.
- Keep sentence structures as simple as possible to express the intended meaning. For ELLs, a number of simple sentences are often more accessible than a single more complex sentence.
- Avoid use of negatives and constructions utilizing not in the questions' stems and options as they can cause confusion, especially for ELLs.
- When a fictional context is necessary (e.g., for a mathematics word problem), use a simple context that will be familiar to as wide a range of students as possible.
- A school based context will often be more accessible to ELLs than a home-based context.

Technique

Divide the course into three phases:

- The first phase consists of sessions on effective teaching and introduction to Communicative Language Teaching (CLT); skills and sub skills of listening, speaking, reading and writing; grammar; vocabulary; Classroom management, educational technology and assessment. The sessions are conducted daily from 10 am to 5 pm. After the sessions, participants work in groups (each group consisting of four or five participants) and plan and prepare lessons for peer/team teaching with the help of a supervisor.
- The second phase is peer teaching, where trainees teach sections of their prepared lessons to their fellow students

who act as students. The lessons are observed by a trainer or guest observer; and are followed by a feedback session.

- The third phase is ‘real teaching’ where participants teach a lesson with a group of proficiency course learners, and the lessons are observed by a trainer and followed by a feedback session.

Strategies for Collecting Data

The assessment/evaluation process involves the use of multiple sources of information collected in a variety of contexts. At the primary level, this research uses observation, work samples, and self evaluation as tools in the process of assessment and evaluation.

Observation

Observation is the careful consideration and analysis of students’ behavior and performance based on a broad range of contexts. In order to use observation effectively, teachers need to know a lot about students, language, and how students learn language, and they need to be able to interpret what they are observing.

Students demonstrate what they think, know, and can do as they engage in various classroom activities that require the application of language processes and learning strategies. Teachers can learn a great deal about students by observing them engaged in such processes as reading, writing, and interacting with others.

Teachers who have not been accustomed to using observation as an assessment tool are sometimes uncertain about what they should be looking for. The key-stage and specific curriculum outcomes provide a framework for teachers to use in their observations.

Gathering Observational Data

To make classroom observation manageable and effective, teachers need to focus their observations. Many teachers develop a systematic, rotational schedule. As well, they might choose a particular focus

for their observations in each of these contexts as they work through their class.

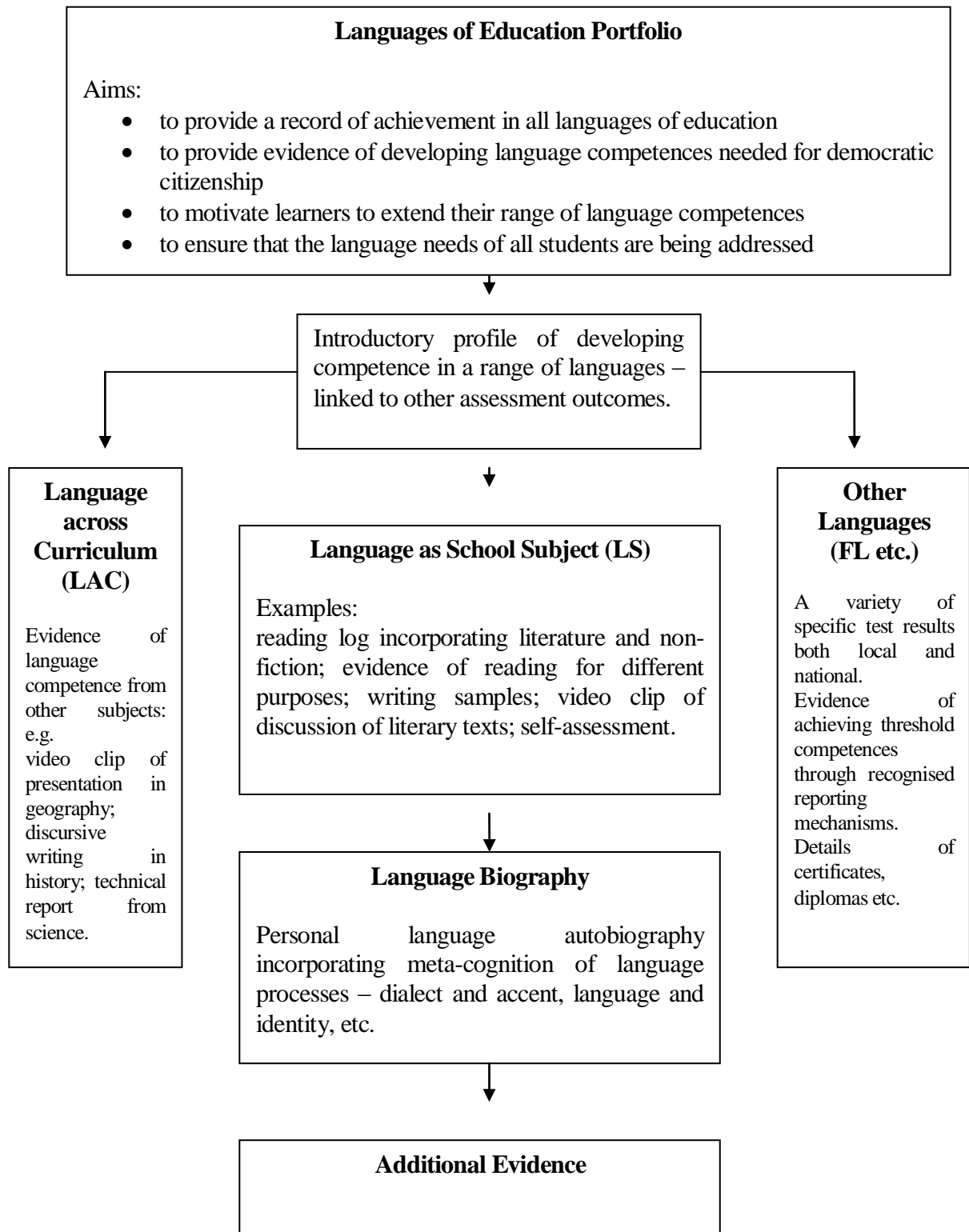
Important and relevant information can also be gathered more incidentally. Teachers can also gather important data about students’ attitudes, understandings, and knowledge through questioning students and talking with them. Although students may know or be able to do more than they put into words, conferences and other forms of dialogue can provide a wealth of information about students and their learning.

DISCUSSION

Teachers of language as school subject are sometimes hostile to the idea of large-scale or formal testing on the grounds that it diminishes the subject and ignores the significance of context. This view needs to be considered.

Basic Principles of Assessment/Evaluation

- The primary purpose of assessment and evaluation is to inform teaching and to promote and encourage learning—to promote optimal individual growth. In order to provide information vital to the teachers, assessment and evaluation must be an ongoing and integral part of the teaching/learning process. It is one continuous cycle consisting of collecting data, interpreting data, reporting information, and making application to teaching.
- Assessment and evaluation must be consistent with beliefs about curriculum and classroom practices, and clearly reflect the various outcomes of the Kindergarten–3 English language arts curriculum, including those areas that cannot easily be assessed with pencil and paper (e.g., processes, attitudes, and values).
- The assessment/evaluation process involves the use of multiple sources of information collected in a variety of contexts.



What is Needs Evaluation and Assessment?

Needs assessments and evaluation with adult English language learners examine the following aspects from the perspective of the learner:

- English language proficiency
- Native language literacy
- Literacy contexts in which the learner lives and works
- Learner need for native language translation or aid of an interpreter
- Learner wants and needs for functioning in specified contexts
- Learner expectations from the instructional program

The needs assessment and evaluation process focuses and builds on learners' accomplishments and abilities rather than deficits, allowing learners to articulate and display what they already know (Holt & Van Duzer, 2000). It is a continual process and takes place throughout the instructional program. The process can influence student placement, materials selection, curriculum design, and instructional practice (TESOL, 2003).

At the beginning of the program, needs assessment and evaluation might be used to determine course content, while during the program, it assures that learner goals and program goals are being met and allows for necessary program changes. At the end of the program, needs assessment and evaluation can be used for planning future directions for the learners and the program (Marshall, 2002). These same tools also may be used as a way to measure progress at the end of the year. However, for reporting outcomes to funders and external stakeholders, standardized assessments must be used.

Purposes of Assessment

In its simplest formulation, assessment provides information on whether teaching/learning has been successful. However the information it provides has a number of potential different audiences whose precise requirements may vary.

Classroom teachers need regular information on how students' knowledge, skills and understanding are develop, both to inform how they should adjust their teaching and to determine what kind of feedback is needed to improve students' learning. On the other hand, school principals and policy makers need additional, broader information on the quality of education in a school or country.

The sort of comparative data required for this purpose needs a high level of reliability and uniformity. In the case of language as school subject this requirement is challenging because it is difficult to create tests which are manageable but at the same time faithful to the aims of the subject. Employers and society at large also need reliable information which can help certify achievement and provide a basis for selection. Parents too require information which can help them understand their children's achievements and limitations. Learners themselves need to know how they are progressing and how to improve their performance but they may need to be protected from the potentially demotivating effects of negative assessment.

The concept of 'accountability' when used in relation to assessment usually refers to the imposition of systems of assessment external to the learning process as a form of 'policing' of standards to ensure that the education system is functioning effectively. But the term may be employed more broadly and more positively than this, referring to the different obligations that are relevant to all.

Teachers have a responsibility to the learner but also to the needs of the wider society. Policy makers clearly have a duty to the public and need to ensure that the education system is delivering results but they also have responsibilities to the individual learners and need to consider consequences of policies in those terms. The concept of accountability interpreted in this way will take people outside of vested interests in order to see the larger context.

Accountability needs to be linked with a process of sharing perceptions and fostering understanding. It is important

therefore not to exaggerate differences between different potential 'stakeholders'; what all parties have in common is a fundamental concern that assessment should help raise achievement and improve learning. A starting point for resolving tensions related to matters of assessment is to develop understanding of other points of view. A key challenge is to develop a system of assessment that acknowledges the different functions of assessment and it helps to see these as complementary rather than being in opposition to each other.

Approaches to Assessment

The different purposes of assessment lead to different approaches to assessment. Traditionally assessment of language as school subject took a very simple form: students were given a narrow written task which was then awarded a grade or mark. This allowed them (and the teacher) to make a judgement of how they ranked in relation to the rest of the group (normative assessment).

However the absence of clear criteria meant that the information rarely gave an indication of how they could make progress in their learning. Also the test itself often embodied a very narrow conception of what competence in language entailed; it often centred on knowledge of language form and structures (syntax and grammar) and a narrow range of language uses (often only a written form of essay).

The implicit understanding of what reading literacy involved was also very narrow, often involving just decoding and literal understanding. In traditional approaches of that kind there was unlikely to be any attention to oral work, to a range of writing purposes, to a wide range of reading and response to reading.

On the other hand, the advantage of a fairly narrow approach to assessment was that it was easier to provide reliable outcomes; the more complex the system of assessment becomes, the more difficult it is to ensure that the award of grades or marks for particular outcomes are consistent.

In approaches to assessment, two central tendencies emerge which are relevant to language as subject. One places emphasis on the assessment *of* learning where reliable, objective measures are a high priority. The focus here is on making summative judgements which in practice is likely to involve more formal examinations and tests with marks schemes to ensure that the process is sound.

An alternative approach is to change the emphasis from assessment *of* learning to assessment *for* learning, implying a more formative approach where there is much more emphasis on feedback to improve performance. The approach here might be through course work and portfolio assessment in which diverse information can be gathered which reflects the true broad nature of the subject.

Portfolio assessment has a number of advantages for language as school subject; for example, it can motivate and empower the learner, it can provide samples of performance collected over time, evidence of use and awareness of process. Portfolio assessment incorporates evidence derived from more realistic tasks in meaningful contexts, rather than relying on artificial, decontextualised tasks undertaken in timed conditions.

A further advantage of this approach is that it can embody different forms of self-assessment which can also be helpful ways of motivating learners and having them reflect on their progress. Self-assessment encourages students to take responsibility in the learning process although it is advisable for them to be trained in self-assessment techniques for this to work effectively. The difficulty with portfolio assessment if it is conceived *only* as the accumulation of evidence produced in informal settings is that it does not easily satisfy demands for reliability. Work which has been produced over an extended period of time, with formative guidance from the teacher and collaboration with classroom peers is not always convincing evidence of competence.

The broader the approach to assessment (incorporating the judgement of a range of different performances in different contexts), the more it can be said to constitute a meaningful assessment of performance in the subject. However, as suggested, tension emerges because it is sometimes difficult to compare, with any degree of accuracy, the results drawn from broad approaches to assessment. The quest for 'objective' and reliable methods of assessment driven by narrow ideas of accountability brings with it a number of dangers. So called 'teaching to the test' may not be a problem if the tests are sophisticated and wide-ranging but there may be practical difficulties in administering those that are too complex. If the tests are too narrow and simplistic then this may have an adverse effect on the teaching. Here then is one source of polarised opinions, one stressing the importance of objectivity, reliability and summative judgement, the other more tolerant of subjectivity in order to ensure that the assessment approach is faithful to the complexity of the aims.

A key concept is embodied in the notion of 'transparency', the view that those being assessed are aware of the criteria which are being used to make judgements about them and how those judgements are made. Knowledge of criteria can help performance and improve motivation but once again, in the context of language as subject, the issues are more complex than they first seem. A common assumption is that students learn best when they know what they are trying to achieve and why. While this view is largely true, there are exceptions. Because the development of language can in some ways be described as a 'natural' process learners do not always need to be fully focused on specific aspects of their performance in order to improve. In fact too much focal awareness on performance can make them too self-conscious: speakers can appear too groomed and artificial; the writer who has been told to strive for effect by using more adjectives may develop a highly artificial and awkward

style. These insights do not negate the importance of transparency as a principle but highlight the fact that in pedagogical practice the principle needs to be interpreted and implemented with care.

The use of competences to describe outcomes which can be assessed has developed considerably in recent years. It is worth noting that the term 'competence' is used by writers in different ways which can be a source of confusion. Sometimes it is used in a very general way as a synonym for 'ability' or capability as in 'language competence'. Other writers use the term to describe broad language modes or domains such as reading, writing, speaking and listening. More commonly however 'competences' refers to the specific actions which a learner must perform and which in turn can be assessed to demonstrate achievement in a subject.

Advocates of using competence statements for assessment purposes and syllabus design see their value largely in bringing clarity and transparency to the specification of learning outcomes. Critics of a competence approach take the view that performance statements are too narrow and specific, and do not reflect the range and subtlety of what is involved in language development. There are parallels here with differences of opinion over the relative merits of portfolio and formal testing. The balance of advantages and disadvantages needs to be considered.

One of the challenges posed by assessing language as school subject (LS) is that the content is so varied and complex which is in turn a reflection of the complexity of the aims. Each mode of 'writing', 'reading', 'speaking' and 'listening' can be broken down into further areas.

A subject that is so multidimensional raises the question as to whether an assessment task in one area is representative of achievement in the subject as a whole. For example, it is fairly safe to assume that performance in speaking is not necessarily indicative of reading competence.

On the other hand it is less clear whether it is necessary to assume that reading ability varies in relation to texts of different types (fiction, non-fiction, media). The assessment of reading can easily remain at a surface level only addressing recall or literal comprehension rather than deeper understanding. Multiple choice questions on a text are easy to mark and may yield high reliability (in the technical sense) but are less equipped to assess the learner's deep and individual response to a text. Even with very young children the reading process is more than simply decoding text and any system of assessment needs to reflect that fact. It is the complexity of the subject that accounts for the hostility teachers of languages as school subject sometimes express towards large-scale and formal testing because only a fairly narrow range of competences can be assessed in a single test.

A similar question arises in relation to writing. Do students need to be assessed on a range of different tasks reflecting the fact that writing exists for different purposes and for different audiences? The assessment of writing appears to be more straightforward than that of reading and speaking because at least there is always a product which can be referred to after the event. But there is a major challenge in determining what criteria should dominate in making a judgement (for example the accuracy of the writing as opposed to the impact of the content) and whether the criteria should change in relation to different types of writing. It is also wrong to assume that the only response to students' writing comes when it is completed. Formative assessment in the form of a dialogue about the work in process is an important way of improving standards. Assessment is always a selection and therefore can unwittingly value some aspects of the subject more than others.

Speaking and listening is extremely difficult to assess because, even more than other aspects of LS, performance varies with the theme, context and level of motivation. Poor performance in oral work is often to do with the nature of the task which has been

devised, the atmosphere of the classroom and the dynamics of the group rather than the competence of the students; to provide a valid assessment of speaking and listening, evidence needs to be drawn from a variety of situations. Some might argue that the assessment of speaking and listening is so complex and context specific that it should not be assessed formally. However, because assessment so often determines the curriculum and the way it is taught there are arguments to suggest that speaking and listening should be assessed despite the difficulties; the ability to articulate a point of view orally and to argue a case are essential skills for meaningful participation in a democracy.

An ideal assessment system would reflect the full complexity of language as school subject (LS), and would motivate learners by giving useful feedback, while also providing other stake-holders (e.g. policy-makers and employers) with the information they need. An integrated approach to assessment would ensure that the different purposes and approaches are balanced so that no one priority has adverse and undue influence on the system as a whole.

Language(s) of Education (LE)

So far consideration has been given to the challenges presented by assessment of 'language as school subject' (LS). Is there any advantage in considering the assessment of LS within a broader framework of language(s) of education (LE) which incorporates 'language across the curriculum' (LAC) and foreign language learning (FL)? 'Languages(s) of education conceived in that way is not a subject but an umbrella construct; it is fairly clear therefore that the notion of examining or testing language(s) of education is entirely inappropriate. However it is conceivable to profile a student's competence in language(s) of education by assembling their competences and achievements in a range of domains.

The present Council of Europe's European Language Portfolio (ELP) is a useful model to demonstrate the type of instrument which could be considered. This approach has a number of potential advantages. It takes seriously the concept of plurilingualism in the school context: a student may be a low achiever in language of schooling taught as school subject but be fluent in two or more other languages.

Thus the deficit model which is often applied to assessment of LS is replaced by a more positive recognition of achievement. It would still be important to be able to identify an individual's competence in aspects of LS as a component within a broader profile, but being able to situate that description of competence in a larger context could have a positive impact on motivation and self-esteem.

Could portfolio assessment extend to embrace language competence in other subjects, to include the dimension of language across the curriculum? This is a practical challenge but perhaps not insurmountable. Some uses of language (e.g. giving presentations, writing formal reports, reading for information) are clearly required and demonstrated within different subjects and it is not inconceivable that those subjects should make a contribution to a student's overall language profile.

There is an argument to suggest that if developing language competences across the curriculum is to be taken seriously it must have some impact on how language use is assessed. Underlying questions have to do with subject boundaries and the degree to which competence in language use can be easily separated from the context in which it is used; learners of history could also in some sense be said to be learning the language of history.

Integrated Assessment

As described earlier, a key challenge and source of tension in relation to assessment is to satisfy the different purposes assessment is expected to fulfil and

to provide appropriate information for the different interest groups.

As suggested, teachers of language as school subject are sometimes hostile to large-scale, formal testing because of the perceived narrowing of the subject. There are a number of points to consider in relation to this view. The testing 'industry' has become increasingly sophisticated in recent years and designers of assessment tasks and items are more adept at addressing issues of validity and reliability. That does not mean that formal tests can address all language competences but once that fact is recognised then the results of such tests may make a useful *contribution* to the overall profile of a students' competence in language use.

Large-scale testing both at national and international level is a fact of modern life. Policy makers need information on levels of proficiency achieved by groups in schools, education authorities or countries. It is possible to discharge this 'evaluative' function of assessment by a process of sampling rather than by using summative data derived from an entire population, but it is unlikely that the quest for information of this kind will diminish.

When assessment data is used to compare the progress of different cohorts of students, the use of a 'value-added' approach which takes account of the different base-lines from which the students are progressing is becoming more common. All of this information can be informative for teachers as long as the limitations of the data are recognised.

The concept of an integrated approach to assessment is intended to counter the tendency towards polarised views. As suggested above, portfolio assessment has many advantages but without information based on some form of 'controlled' performance it will struggle to make convincing claims to reliability. The incorporation of specific results derived from appropriate tests designed to assess specific (not all) language competences into a portfolio can help strengthen both the validity and reliability of the assessment

approach. The key may be to think broadly in terms of an assessment strategy which makes use of a variety of assessment tools, rather than assuming that one assessment tool will fulfil all purposes.

Internal Evaluation

Students are assessed separately in each subject by their teachers. The results of the formative assessment carried out during a school year are taken into account in the end-of-year summative assessment.

In grades 1-3 of primary school the evaluation is only descriptive and applies to educational attainments and behaviour (conduct) separately. Starting with grade 4, each subject teacher has the following scale of marks at his/her disposal: 6 - excellent, 5 - very good, 4 - good, 3 - satisfactory, 2 - acceptable, 1 - unsatisfactory. School marks as well as the criteria on which they are based should be open to the student and his/her parents.

Students also receive marks for their behaviour (conduct) according to the following scale: excellent, very good, good, acceptable, unacceptable, and inadmissible. The mark for behaviour (conduct) should not influence subject marks or promotion to a higher grade.

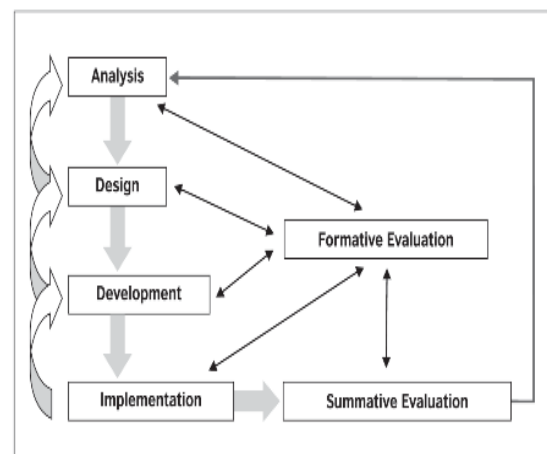
Students may take a verifying examination if the mid-year or end-of-year summative mark, given by the teacher, is too low in their or their parents' opinion. This kind of exam is also commissioned for students whose results cannot be assessed for other reasons, such as an individual study programme, out-of-school study or a long period of illness.

External Evaluation

The new system of external assessment, implemented gradually from 2002 on, aims at measuring the achievements and identifying shortcomings of students' education, assessing the effectiveness of teaching, and comparing objectively current certificates and diplomas irrespective of the place where they have been issued.

Ways to Bridge the Gaps between Planning, Implementation and Evaluation

In cases when a course is run in a third country such Indonesia, all the stakeholders responsible for selection of participants from different countries need to cooperate before, during, and after the course is over. This will help to assess whether the course has been effective or not. Such collaboration can be achieved by using the ADDIE model as the basis for shared planning and understanding going forward. It means that needs analysts, designers, materials developers, trainers and evaluators (trainers and country representatives) work together in order to run an effective trainer training program.



ADDIE Model (diagram adapted from Steven J. McGriff, *Instructional Systems*, College of Education, Penn State University)

The ADDIE model's focus on formative evaluation of the course will help to build in immediate remedies during the course. To understand the longer term actual impact of the training (which consists of 4 levels: 1. React, 2. Learn, 3. Change, and 4. Impact) this model can be adapted. It implies that the concerned stakeholders and the course participants should continue to work together after the course is over.

Evaluations can be macro or micro in scale and can be carried out for either accountability or developmental purposes or both of these. In macro evaluation, various administrative and curricular aspects are examined (e.g. materials evaluation, teacher evaluation, learner evaluation), while micro

evaluation focuses on the specific aspect of the curriculum or the administration of the program such as evaluation of learning tasks, questioning practices, learners' participation etc. (Ellis, 1998). The evaluation in language teaching has been primarily concerned with the macro evaluation of programs and projects (Ellis, 1998), and most evaluation studies have been conducted in order to measure the extent to which the objectives of a program have been met, and to identify those aspects that can be improved. As Ellis (1998) observes, this kind of analysis is obviously of interest to teachers as they learn whether or not the goals have been accomplished and whether any changes should be made to the program. However, most teachers are less likely to be concerned with the evaluation of the program as a whole, and more concerned with the extent to which a particular textbook, or a teaching activity is effective in their teaching context.

The evaluation of teaching materials may be done before they are used in the classroom in order to determine whether they suit the needs of the particular group of learners (predictive evaluation), or after the materials have been used in the classroom in order to evaluate their effectiveness and efficiency, and teachers' and learners' attitudes towards them (retrospective evaluation). This paper will introduce a systematic procedure for conducting the predictive evaluation of language teaching tasks.

Peer Observation

The main objective of peer observation is to give feedback to the trainee on how well she/he has performed in teaching a particular lesson in a given class. Very often this main objective is not achieved as the peer group is not guided well about how and what they have to observe. A sample observation schedule is presented below:

Sr. No.	Parameters	Scope for improvement	Average	Effective	Very effective
	Presentation	1	2	3	4
a.	Details on class board, lesson title, learning outcomes, key words are prominently displayed				
b.	Uses relevant vocabulary				
c.	Exhibits fluency of language				
d.	Uses suitable tone of speech				
e.	Oral and written instructions are clear				
f.	Teaching aids and learning materials are appropriate, well-organised, resourceful and stimulating				

Sample Tasks:

Task 1.

What are the specific objectives of the Presentation stage? a. b. c.

Task 2

a. How does the teacher arrange the class before the session?
b. How does the teacher motivate the class?

Task 3

What are the different classroom interaction patterns that the teacher has achieved?
a.
b.
c.
d.

Task 4

Grammar	
What new language item does the teacher present?	What techniques does the teacher use to present the new language item? What types of tasks does the teacher give the class to use the new language item?

Task 5

What types of errors does the class make?	How does the teacher deal with the errors?
---	--

Task 6 Teaching and learning roles

Lesson Stage	Teacher's role	Learner's role	Comments
a.			
b.			
c.			

Task 7. The teacher's meta-language

What does the teacher say?	What is the communicative purpose?	What is the immediate context?	How might this be said to a native speaker?
"Look at the picture. Can you see the boys sitting under the tree?"	Giving instructions Directing	The teacher is setting up a task with a visual aid (pictures/charts)	Can you see where the boys are sitting (in the picture)?

Adapted from: Ruth Wajnryb (1992)

- 20% for a book report which was marked for content, conventions, grammar, process, understanding and spelling
- 20% which reflects the mean average of 3 quizzes given
- 20% for an oral presentation for which I was evaluating reasoning, oral communication and organization
- 10% Weekly spelling and grammar quizzes

The most important element of assessment and evaluation is to provide information for improved student performance. When one looks at a mark of 65% or a C in Language/English, it is important to know how the student can improve to receive a 75% or a B next time. Look at the evidence and work with the teacher to determine the areas of strength and the areas of weakness to improve overall learning.

Alternative assessment practices (assignments, term papers, projects) in teacher education courses do focus on these aspects and are excellent evaluation tools, but there is no guarantee that they will serve as true indicators of individual performance and by inference, capability. One cannot prevent students from discussing assignments and term papers with more experienced colleagues or more enabled peers.

Scaffolding is not a necessary evil to be lived with and should be encouraged to enable learning to happen. However scaffolded performance is not desirable where individuated certification is involved. In this context, open book examinations need to fulfill the requirements of summative evaluation and also ensure that knowledge is not merely reproduced but applied, synthesised and evaluated.

An open book examination (OBE) is a valid method of testing how much a student has learnt and internalised from a course; it is an examination that does not test knowledge of theory but instead focuses on practical applications of theory, in new/authentic

Assessment, Evaluation, and Report Card Marks or Final Grades

Ongoing assessment and certain evaluations will make up final marks and/or report card grades. For instance, let's say a teacher is ready to give a final/report card mark for language/English. Here is one example on how she/he would come up with that mark:

- 15% for notes done in class
- 10% for participation in group work
- 5% for homework completion

contexts. The questions/tasks are usually context specific and may also carry the possibility of open ended responses. It is an examination where students can have access to their prescribed textbooks and even other resource material. Since such referring is possible, meaningful and deep learning can be fostered.

There are many kinds of OBEs, ranging from allowing access to restricted resources, access to unrestricted resources, as well as open resource examinations where access to an entire library is permitted. In a restricted OBE printed documents such as logarithmic tables or dictionaries are permitted.

A few primary sources may also be accessed, but only with the prior approval of the course instructor/examiner. An unrestricted OBE is one where lecture notes, worksheets, handouts and even books can be brought in. The choice of such material is left completely to the student. The open resource/open library OBE is like a take home examination; but in such contexts the questions/tasks set are likely to be individualised and the evaluation stringent to avoid mass copying or plagiarism.

Teacher's Self-Evaluation

Very few teachers can put hand on heart and say that they enjoy being observed. But why is this? There may be several reasons, but one of the most likely ones is that it is perceived by an observee as an extremely high stakes event.

The decision to become a teacher is an existentially significant one, and the thought that one's professional judgement and skills are under scrutiny by an 'outsider' to the class community one has striven to mould and develop is usually seen as scary, perhaps even more so because a single observation can never be more than just a 'snapshot' of the longer term endeavours which are crucial in teaching and learning.

The potential for misunderstandings seems, in the mind of the observee, to be immense. Yet it is the observer who is the 'senior partner' in this process and I will

now look at ways in which observers can be more self-aware and better prepared for the task they are charged with, whether they are observing for assessment, for training, for research or for professional development.

A teacher trainer has to be able to motivate, guide and facilitate participants' learning; observe classes and give constructive feedback to enable participants to grow by reflecting upon themselves and their classrooms – in other words, the trainer needs to understand the human dimension of their work and deal with trainees' affective factors and belief systems to enable them to find their own unique ways of teaching.

Teacher educators should possess theoretical knowledge of different teaching and learning principles, concepts and processes; as well as training models; and be familiar with the academic literature that underpins their work. This linking of theory and practice is essential to producing well-prepared trainee teachers and informing their thought and practice about teaching.

In order to cater to the demands of education assessment and evaluation, the nature of teacher education has to change; it has to help teachers acknowledge diversity of learning spaces and curriculum styles and also help them build the capacity in the child to construct knowledge. Teacher education programs must, therefore, include reflective practice for unless the teacher is able to reflect and become critical, he/she cannot enable such thinking in the mind of the child.

Teacher's self-evaluation should be based on the teacher's action-research into the process of teaching-learning. This approach involves planning aims and methods before teaching, acting on the plan in class, observing what students do in class, and analyzing why students behaved in this way.

You can write up your lesson plans, observation and analysis in a Lesson Record. It is useful to ask questions of yourself under each heading in order to structure your record:

1) Planning:

- What are the aims of this lesson?
- What will I expect children to do by the end of the lesson which will show that the aims have been achieved?
- Are there opportunities to develop interactive social skills?
- What problems could be anticipated? How could these be overcome?
- How will I be able to support the less able children/challenge the more able children?

2) Observations

- Where children absorbed in the activities? (Give examples of what students did.)
- Where any children not involved? Why?
- Did any children show particular abilities (e.g. producing spontaneous language)?
- Was pair/group work carried out effectively? Were there any practical problems (e.g. problems with materials, space, noise)?

3) Analysis

- To what extent were my aims achieved? - Which aspects of the lesson were most successful? Why?
- Which aspects were less successful? Why?
- How can I reinforce students' learning?

Students' self-evaluation

It is important to give students a sense of the progress they are making in class. One way of doing this is by means of a *Progress Diary*. It is normally used at the end of each unit and provides an opportunity for the student to reflect upon his or her learning process. Although the questions are usually in English, the answers may be given in L1. This diary concentrates on the particular objectives of that unit, and encourages students to comment on the extent to which they feel they have achieved

those objectives. The diary is used for self-evaluation, but it is also a useful tool for the teacher, who can use it to assess the student's progress. The diaries should be collected regularly, read, and account taken of the information in them.

Teachers can use self-evaluation to analyze the process of teaching and learning. It will also act as motivation for the students. They will develop a positive attitude towards the foreign language if they have an opportunity to show what they know and what they can do. We can encourage our students to evaluate themselves by means of simple conversations in their mother tongue and activities that can help them to reflect upon their learning of English and the strategies they use to learn it.

Explanation of the Attainment Targets of Oral Communication

Attainment target	The pupil is able to ...	In activities such as...
Can understand specific information and simple oral instructions.	understand the full meaning of oral texts and instructions needed to carry out a task he or she is required to do.	Listen and complete the picture
Can grasp the overall sense of oral texts.	grasp the topic or type of oral document presented to him. In particular, the pupil is able to follow the story line of a tale told by the teacher.	Listen and order the pictures
Can grasp specific details within an extended oral text.	choose the information that he or she needs, and disregard that which is not understood or needed.	Listen and choose the right picture
Can understand and use social expressions.	use social expressions learnt in previous years or which appear in this year for the first time.	Ask and answer

Can recognize and reproduce the sounds, intonation, and rhythm of English.	imitate the pronunciation and rhythm of the English language in oral activities and particularly in songs and chants.	<i>Sing the song</i>
Can reproduce short messages based on given models.	carry out a series of oral activities (e.g. comic stories, interviews, etc) which require the use of particular phrases.	<i>Listen to the story. Then, act it out.</i>
Can use English to communicate with the teacher and class-mates in everyday situations.	use the English language in normal communication situations, although this use may not be free of errors.	
Can engage in oral interchanges using and experimenting with language already known.	use words and phrases already studied in a creative manner.	<i>Play the game</i>

in fact they still have difficulty with. The tests should also bring to the teacher's attention any students who are not keeping up with the rest of the class, and who may need extra help.

It is important that these tests reflect the language subject of study and the type of activities they do normally in class. The tests have to reflect the difference between the courses and cycles we divide Primary Education. In the first years they include a word or sentence approach. In the last years of Primary they must include a more discursive approach. The tests are the quantitative element of the evaluation process. In Primary, it is important that it is combined with qualitative observation.

Explanation of the Attainment Targets of Written Communication

Attainment target	The pupil is able to ...	In activities such as...
Can understand specific information and simple written instructions.	understand the written instructions used in the book on how to carry out an activity.	
Can grasp the overall sense of written texts.	skim-read a text and identify its main idea or topic.	<i>Read and match the description with the pictures</i>
Can grasp specific details within an extended written text.	read a text and extract specific information to solve a problem or answer some questions.	<i>Read the text and answer the questions.</i>
Can reproduce sentences and short texts following given models.	write a series of phrases following a fixed model that he or she will have to adjust.	<i>Look and write</i>
Can produce short texts, using and experimenting with language already known.	use phrases and words that are already known to make new combinations that help the pupil learn to express him/herself in writing in English.	<i>Write about the capital of your region</i>

Qualitative Observation

The qualitative observation can be done with observation sheets that can help us in the process. The sheets we propose can be organised by the teachers as they like and include aspects they can observe every day in the classroom. We think that those sheets can be a valuable register of the linguistic and socio-cultural progress of our students. These sheets can be used as internal reference and also with meetings with parents. The observation sheets could be divided in categories that each teacher can complete in the ways he or she likes. It could include the following categories.

- Communicative abilities and uses of English Attitude towards learning
 - Behaviour in the classroom
 - Socio cultural and intercurricular themes
- Each category can be covered with items such as:

Attainment tests

The main aim of these tests should not be only to label the student in terms of pass/fail, but to collect information in order to check aspects of learning that are difficult to assess by classroom observation alone, and to gain a more rounded view of the student's achievements.

The information collected in this way should have a direct effect on the planning of future classes, for example by highlighting aspects of the language the students were thought to have mastered, but

1.1.	Recognises the vocabulary of the unit
1.2.	Tries to use English to communicate in class
1.3.	Reads and understands short stories
2.1.	shows a positive attitude towards the foreign language
2.2.	shows interest in communicating in English
2.3.	begins to use language learning strategies consciously
3.1.	answers with good behaviour
3.2.	brings his/her school books to the class
3.3.	works well in pairs or groups
4.1.	non-sexist education
4.2.	environmental education
4.3.	peace education

Needs assessment and evaluation may take many forms and may be carried out at different times during the instructional process. Whatever the focus and format, the basic purpose is to determine what learners want and need to learn. When learners know that educators understand and want to address their needs and interests, they are motivated to continue in a program and to learn.

To assert that second language programme evaluation is a practical activity is not, however, to consign it wholly to the realm of the arbitrary, the make shift and the convenient. An evolving "discipline" of second language evaluation seeks to establish standards of practice and to afford its own socially constructed perspective, wider in space and time yet narrower in "professional" focus, upon this area of activity.

CONCLUSION

The overall goal of assessment and evaluation is to improve student learning. Assessment and evaluation provides students, parents/guardians, and teachers with valid information concerning student progress and their attainment of the expected curriculum. Both of them should always be viewed as information to improve student achievement.

Assessment and evaluation measure whether or not learning and/or learning objectives are being met. One could look at assessment and evaluation as the journey

(assessment) versus the snapshot (evaluation). Assessment requires the gathering of evidence of student performance over a period of time to measure learning and understanding. Evidence of learning could take the form of dialogue, journals, written work, portfolios, tests along with many other learning tasks.

Evaluation on the other hand occurs when a mark is assigned after the completion of a task, test, quiz, lesson or learning activity. A mark on a spelling test will determine if the child can spell the given words and would be seen as an evaluation. Assessment would be a review of journal entries, written work, presentation, research papers, essays, story writing, tests, exams etc. and will demonstrate a sense of more permanent learning and clearer picture of a student's ability. Although a child may receive high marks in spelling test, if he/she can't apply correct spelling in every day work, the high spelling test marks (evaluations) matter little.

Effective teachers will use both assessment and evaluation techniques regularly and on a daily basis to improve student learning and to guide instruction. In a teacher training programme it is necessary to introduce a task-based peer observation scheme as the teacher-trainees are inexperienced in observing and making inferences. It also decreases the psychological pressure which they might have when they are asked to observe a lesson without any guidance. A task-based observation and feedback scheme will help provide a conducive climate in which both the givers and receivers of feedback can deepen their understanding of effective language teaching practices that promote quality learning

REFERENCES

- Holt, D. and Van Duzer, C. (2000). Assessing success in family literacy and adult ESL. Retrieved from <http://calstore.cal.org./store>
- Ellis, R. (1998). First steps in Reading: A Teacher's Handbook for Using Starter Readers in the Primary School. Oxford: Heinemann.
- Marshall , B. (2002). Preparing for success: A guide for teaching ault ESL learners. Retrieved from <http://calstore.cal.org./store>
- McGrift, Steven J. (2000). Retrieved form http://en.wikipedia.org/wiki/ADDIE_Model
- Teaching of English to Speakers of Other Languange, Inc. (2003).Standards for adult education ESL programs. Retrieved from <http://www.tesol.org>