

INTERNATIONAL BUREAU OF EDUCATION

Training Tools for Curriculum Development

Inclusive Student Assessment



United Nations
Educational, Scientific and
Cultural Organization



International Bureau
of Education

The information provided in the material and case studies does not necessarily represent the views of the IBE-UNESCO. The designations employed and the presentation of the material, including maps, do not imply the expression of any opinion whatsoever on the part of the IBE-UNESCO concerning the legal status of any country, territory, city or area, or of its authorities, or concerning the delimitation of its frontiers or boundaries.

Users are granted permission to reproduce or translate short extracts, provided that the source is acknowledged and print out hard copies for personal use, provided that it is not for the purpose of private or commercial gain. The IBE-UNESCO cannot undertake responsibility for any subsequent use of the material, or the accuracy of the information provided. The IBE-UNESCO strongly recommends establishing hyperlinks to the present material instead of using and posting it on any other website for any purpose, and does not encourage the posting of PDFs on other websites in the absence of a previous agreement.

Hyperlinks to other websites are provided for the user's convenience. The IBE-UNESCO does not control or guarantee the accuracy, relevance, timeliness or completeness of this outside information. The inclusion of hyperlinks to other websites is not intended to reflect their importance, nor is it intended to endorse any views expressed, or products or services offered, on these sites, or the organizations maintaining the sites. Further, the brief description of the information and services provided on outside sites is not exhaustive and does not represent an evaluation of these sites on the part of the IBE-UNESCO.



Editors: Renato Opertti and Lili Ji

Published in Geneva, June 2017, by:

IBE-UNESCO

C.P. 199

1211 Geneva 20

Switzerland

Tel.: +41.22.917.78.00

Fax: +41.22.917.78.01

Email: ibe.training@unesco.org

<http://www.ibe.unesco.org>

Ref: IBE/2017/OP/CD/03

Training Tools for Curriculum Development: Inclusive Student Assessment

© Cover Photo: IBE-UNESCO

Table of contents

Conceptual framework	5
Rationale	5
Purpose	5
Module Structure	6
Introduction and Assessment Concepts.....	7
Training manual	21
ACTIVITY 1. How “inclusive” are our current student assessment systems?	21
Task 1. “Friendliness” of current student assessment	22
Task 2. Criterion-referenced assessment.....	23
Task 3. Student assessment and its effects on learning	23
Task 4. Object of assessment.....	24
Task 5. Recent efforts in changing assessment.....	25
Task 6. Curriculum and student assessment.....	27
ACTIVITY 2. How can we adjust student assessment so that it contributes to high quality education and inclusiveness?	28
Task 1. Characteristics of an inclusive assessment	29
Task 2. Attributes and strategies of formative assessment	30
Task 3. Student assessment reform	33
Task 4. Summative assessment	35
Task 5. Features of student assessment and a sample assessment plan	37
ACTIVITY 3. What are the basic conditions and competencies needed to make student assessment more inclusive?	40
Task 1. Classroom-level assessment	43
Task 2. Basic conditions for formative assessment.....	45
Task 3. Formative assessment strategies	47
ACTIVITY 4. Relationships between international, sub-regional, national, and school-level student assessments, and their connection to the curriculum.....	50
Task 1. Different levels of assessment	51
Task 2. Assessments and examinations in your context.....	52
Task 3. Cultures of evaluation	53
ACTIVITY 5. Adopting a holistic approach to student assessment	54
Task 1. Supporting factors for assessment reform	55
Task 2. A comprehensive approach to student assessment	56
Task 3. Effectiveness of student assessment systems	58

List of Tables

Table 1.1. Comparison of Assessment for Learning and Assessment of Learning 12

Table 1.2. Assessment Modes and Strategies for Authentic Assessment Practices 13

Table 1.3. Traditional vs. Authentic Assessment Methods..... 14

Table 3.1. Features of Classroom Assessment Competency 41

Conceptual Framework

Rationale

The *Inclusive Student Assessment* Module is intended to support the training of teachers, principals, inspectors, teacher trainers, educators and curriculum development specialists by strengthening capacities in the field of student assessment under a broadened concept of inclusive education and inclusive curriculum. Through integrating perspectives on assessment from around the world, it also draws upon an international comparative perspective on student assessment that practitioners can use to further reflect upon education in their own contexts.

The module aims to equip professionals with the foundational understanding and tools required to effectively develop student assessment mechanisms by providing an expansive sequence of exercises that:

1. Introduce users to essential student assessment concepts and principles;
2. Highlight the need to revisit student assessment in light of the broadened concept of Inclusive Education;¹
3. Introduce different approaches and activities to promote inclusivity in student assessment;
4. Offer opportunities to gain hands-on experience in forging inclusive student assessment strategies and tools;
5. Explore key competencies required for educators to effectively analyse and implement renewed student assessment practices;
6. Provide insight into how teachers can transform knowledge-based student assessment into student-centred formative assessment that contributes to developing relevant competencies, including attitudes and behaviours;
7. Explore the connections between student assessment at different levels (international, sub-regional, national and school-based); and
8. Relate student assessment issues to other aspects of education (e.g. curriculum, teacher capacity development) and society in a holistic manner.

Purpose

As the UNESCO institute specialized in the field of curriculum, the International Bureau of Education (IBE-UNESCO) emphasizes the key role of high-quality and pertinent curriculum development as a way to improve the excellence, relevance and equity of education, and in particular to enhance and democratize learning opportunities. This module is designed to be completed in conjunction with, or subsequent to the study of the general guidance on curriculum design and development found in the IBE-UNESCO's *Training Tools for Curriculum Development – A Resource Pack (2013)*, **Module 8: Student Assessment and Curriculum Evaluation**.

¹ The UNESCO IBE website (<http://www.ibe.unesco.org>) hosts the Conclusions and Recommendations of the 48th session of the International Conference on Education (Geneva, 25-28 November 2008), "Inclusive Education: The Way of the Future", as well as a vast array of publications, information documents and working documents on this topic.

This Module is not a manual which prescribes a sequence of steps to be taken as no unique solution can be applied in the same way in all contexts. Rather, it concretely orients existing work while providing valuable inspiration to curriculum developers and other stakeholders for future efforts related to student assessment.

Module Structure

This module first provides a **Conceptual Framework**, which combines concepts and theories, including the evolving purposes and different types of student assessment. The Conceptual Framework is intended to ground practitioners' understanding of student assessment in light of inclusive education in order to also promote inclusivity in assessment.

The **Training Activities** section of the module which follows the Conceptual Framework is structured around five key areas of student assessment:

- 1. How “inclusive” are our current student assessment systems?** This activity takes readers through a series of tasks linking student assessment to the broadened concepts of inclusive education and inclusive curriculum.
- 2. How can we make student assessment contribute to better quality and more inclusive education?** Introduces interesting practices and models that promote greater inclusivity in education, including the development of formative assessment, the use of summative assessment for inclusive purposes, and the design and use of quality rubrics for student assessment.
- 3. What basic competencies are needed to make student assessment more inclusive?** Explores the basic competencies (knowledge, skills, values, attitudes, etc.) that are required to assess student learning in light of inclusive education.
- 4. What is the relationship between international, sub-regional, national, and school level student assessments and how do they relate to curriculum?** Explores the relationship between student assessment at different levels – international (PISA, etc.), regional, national and school levels, as well as various countries' approaches to these assessments.
- 5. Adopting a holistic approach to student assessment.** Introduces a series of tools and models for harmonising student assessment with curriculum, teacher training and other sectors of the education system and society.

Each of the five areas has several training tasks attached to it, each of which requires the completion of a number of sub-tasks.

The purpose of the training activities is to place you, the reader or participant, in the position of a curriculum developer, so that you can analyse or apply assessment practices in response to specific issues or challenges that you may encounter during the educational process. You may be required to do this individually, as a member of a country team, or as a member of a smaller group.

Some of the tasks will request that you read an external source or complete a worksheet. Where this is the case, the documents can be readily accessed through the [hyperlink](#) embedded in the text.

Introducing Assessment and Related Concepts

The international community's education agenda is based upon the fundamental human right to education, and is increasingly informed by discussions that support the attainment of the Education for All Goals (EFA Goals), the Millennium Development Goals (MDGs), as well as the Sustainable Development Goals (SDGs). Thus far, progress towards these goals has revealed that ensuring a holistic vision and approach to education and the right to lifelong learning is essential for achieving inclusive and equitable societies.

Quality Education

A high quality education is, among others, one that “satisfies basic learning needs, and enriches the lives of learners and their overall experience of living.”² It is also “an education that is inclusive”, which aims for “the full participation by all learners, teaches attitudes and behaviours of tolerance, and is therefore a vehicle for the construction of an inclusive and participative economy and society.”³ In quality education:

1. goals and content should address the needs of all learners, as well as the values and aspirations of the community and society;
2. goals and objectives should be assessed by looking at the effectiveness of information, knowledge, skills, values, attitudes and behavioural changes, etc.; and
3. resources should be adequately allocated and efficiently used, so as to improve learning outcomes.⁴

Inclusive Education

The concept of inclusive education is based on the fundamental right of all learners to a quality and equitable education that meets their basic learning needs, and considers the diversity of backgrounds and abilities a learning opportunity. Inclusive education is intertwined with the international vision and efforts towards achieving EFA.⁵

However, confusion still exists within the education field about what inclusive education encompasses, partly due to the fact that the idea of inclusive education can be defined in a variety of ways.⁶ It is also important to remember that there is no one perspective on inclusion within a single society or school.⁷ A recent analysis of international research⁸ suggests a typology of five ways of conceptualizing inclusion:

2 The World Education Forum. (2000). Dakar Framework for Action. Senegal.

3 The 47th session of International Conference on Education. (2004). Discussion Paper, Workshop 2 “Quality Education and Social Inclusion”. Geneva: IBE-UNESCO.

4 *Ibid.*

5 Ainscow, M.; Miles, S. (2008). Making Education for All inclusive: where next? *Prospects*, vol. 145, no. 1, pp. 15–34.

6 Ainscow, M.; Farrell, P.; Tweddle, D. (2000). Developing policies for inclusive education: a study of the role of local education authorities. *International journal of inclusive education*, 4(3), pp. 211-229.

7 Booth, T. (1996) A perspective on inclusion from England. *Cambridge review of education* 26(1), 87–99. Booth, T.; Ainscow, M. (eds.) (1998) *From them to us: an international study of inclusion in education*. London: Routledge.

8 Ainscow, M. et al. (2006). *Schools and special needs: issues of innovation and inclusion*. London: Paul Chapman.

8 Ainscow, M. et al. (2006) *Improving schools, developing inclusion*. London: Routledge.

1. inclusion concerned with disability and “special educational needs”;
2. inclusion as a response to disciplinary exclusions;
3. inclusion as relating to all groups vulnerable to exclusion;
4. inclusion as the promotion of a school for all; and
5. inclusion as Education for All.

Over the past decade, and particularly after the 48th session of the International Conference on Education (ICE) (25 to 28 November 2008, Geneva, Switzerland), UNESCO has become a leader in promoting inclusive education, both at the policy and implementation levels. UNESCO defines inclusion as “a process of addressing and responding to the diversity of needs of all learners through increasing participation in learning, cultures and communities, and reducing exclusion within and as a result of education. Inclusion involves modifying content, approaches, structures and strategies, towards a common vision that includes all children of the appropriate age range and a conviction that it is the responsibility of the regular system to educate all children.”⁹ The 153 Member States participating in the 2008 ICE declared that “a broadened concept of inclusive education can be viewed as a general guiding principle to strengthen education for sustainable development, lifelong learning for *all* and equal access of all levels of society to learning opportunities so as to implement the principles of inclusive education.”¹⁰

Inclusive Curriculum

Moving towards an inclusive education system and achieving the EFA goals requires developing, implementing, and achieving curricula that are inclusive. In fact, educational reforms throughout the world have become increasingly focused on sound curriculum processes and products.¹¹ As previously mentioned, the IBE defines the curriculum as a political and technical concept, which is deeply rooted in the paradigm of society, politics and education. Particularly, inclusive curricula address the learning needs of all and take into account the concerns of multiple stakeholders, both within and outside the education system. In other words, inclusive curricula:

- reflect the kind of society to which we aspire;
- are flexible and adaptable to diverse contexts and needs; and
- contribute to ensuring equity and quality as twin concepts in the development of human capital.

Teamwork and cooperation are essential to developing and implementing an inclusive curriculum. Local and national education systems, schools, teachers, communities and learners should be engaged as co-developers of inclusive curricula. The implementation of inclusive education at the school level – making sure that all children and young people are able to participate and learn – greatly relies on teachers. Teachers need to be supported by the whole education system in order to effectively own and apply inclusive pedagogic strategies in their classrooms. On the other hand, interaction, co-operation and interdisciplinary teamwork among practitioners have demonstrated the benefits of working with diversity.

9 United Nations Educational, Scientific and Cultural Organisation. (2005). Guidelines for inclusion: Ensuring access to Education for All. Paris: UNESCO.

10 United Nations Educational, Scientific and Cultural Organisation. (2008). Conclusions and recommendations of the 48th session of the ICE. Geneva: UNESCO-IBE.

11 Moreno, J. M. (2006). Chapter 11. The dynamics of curriculum design and development: scenarios for curriculum evolution. School knowledge in comparative and historical perspective. Changing curricula in Primary and Secondary Education (edited by Aaron Benavot and Cecilia Braslavsky), Comparative Education Research Centre, the University of Hong Kong, Springer, Hong Kong, China, 2006, p.195-209.

In order to respond to the diverse characteristics of learners, an inclusive curriculum has to be flexible while defining the content, processes and outcomes of teaching and learning. Curricular flexibility based on universal concepts (human rights, lifelong learning, etc.) should underpin curriculum development at the school and classroom levels. There is a pressing need for educational systems to cater for *differences* as an ordinary aspect of the human condition. An inclusive curriculum is also “glo-local” by being responsive to global, national and local contexts and considerations.

Assessing and monitoring the inclusiveness of a curriculum requires consensus on the definition of success in light of inclusion. In other words, there is a need to revisit what we value in terms of presence, participation and achievement in education. Meaningful participation goes beyond access to education; it involves active engagement with the purposes and methods of teaching, learning and assessment. Thus, an inclusive curriculum provides grounds for students and teachers to construct interactive and collaborative relationships. The curriculum guides teachers through complex planning for teaching and learning. Such processes should take into account individual learners’ characteristics, as well as each teacher’s background, capacities and professional development, while avoiding pre-conceived expectations or clichés.

Student Assessment

Student assessment is the process of documenting students’ acquisition and mastery of knowledge, skills and competencies in order to make informed decisions about the next steps in an educational process. This implies consideration of students’ aptitudes, attitudes, learning styles, progressions and outcomes. Evaluation methods may include written tests, responses to oral questioning, and computer-adaptive testing models, etc. Decisions premised on the results may vary from designing system-wide programs for improving teaching and learning in classrooms, to adjusting classroom instruction, to determining students’ admission to a university.¹²

During the last decades, there have been shifts towards more modern and holistic ways of teaching and learning, aimed at preparing children to become responsible citizens and better equipping them for their future lives. However, researchers and practitioners assert that actual assessment practices have not yet followed these shifts.

Assessment Systems

Assessment systems are made up of the policies, structures, practices and tools that generate and use information about student learning. Effective assessment systems are those that obtain quality information to satisfy decision-making needs, as well as to support and enhance student learning.¹³ A student assessment system can inform and improve learning and instruction, measure progress, determine achievement, and provide partial accountability information. All of these purposes and decisions are expected to ultimately lead to improved quality in education.

¹² Clarke, M. (2011). *Framework for building an effective Student Assessment System*. Washington, DC: The World Bank Group.

¹³ Ravela, P.; Arregui, P.; Valverde, G.; Wolfe, R.; Ferrer, G.; Rizo, F. M.; Aylwin, M.; and Wolff, L. (2009). *The Educational Assessments that Latin America Needs*. Washington, DC: PREAL.

Types of Assessment

Assessment can be labelled using various distinctions:

*Classroom assessments, examinations and large-scale assessments*¹⁴

Classroom assessments, also referred to as continuous assessments, are those carried out by teachers and students in the course of daily activity (Airasian and Russell, 2007). They encompass a variety of standardized and non-standardized instruments and procedures for collecting and interpreting written, oral, and other forms of evidence on student learning or achievement. Examples include oral questioning and feedback, homework assignments, student presentations, diagnostic tests, and end-of-unit quizzes. The main purpose of these assessments is to provide 'real time' information to support teaching and learning. They involve assessment *for* learning (engaging and supporting students' learning progressions) and assessment *as* learning (helping students to gain awareness of and reflect upon their own thinking and learning processes and outcomes).

Examinations, which are often qualified by the terms 'public', 'external', or 'end of cycle', provide information for high-stakes decision-making about individual students. For example, exams may indicate whether students are assigned to a particular type of school or program, graduate from high school, or can be admitted to university. Whether externally administered or school-based, their typically standardized nature is meant to ensure that all students are given an equal opportunity to demonstrate what they know and can do in relation to an official curriculum or other identified body of knowledge and skills (Madaus and Clarke, 2001).

For example, the exit certification exams at the end of compulsory education in many education systems are considered "high-stakes" because they have a significant impact on what is taught and learned (resulting in "teaching to the test" or even "teaching the test"). They may also influence the skills and knowledge profiles of graduates (West and Crighton, 1999). These tests may discourage students who do not score well, and effectively contribute to their exclusion from the education of their choice. Students from disadvantaged groups are the most vulnerable to such risks. Practices associated with examinations that create inequities among students include scoring, ranking, exam registration fees, private tutoring, and examination in a language with which students are not familiar, etc. The use of quota systems to mitigate differences in performance associated with location, ethnicity, or language-group membership also creates inequity among students (Greaney and Kellaghan, 1995).

Large-scale assessments are designed to provide information on system performance levels and related or contributing factors, typically in relation to an agreed-upon set of standards or learning goals, in order to inform educational policy and practice. Examples include international assessments of student achievement levels such as TIMSS (Trends in International Mathematics and Science Study), PIRLS (Progress in International Reading Literacy Study), and PISA (Program for International Student Assessment). Other examples include regional assessments such as PASEC (Programme d'analyse des systèmes éducatifs / *Program on the Analysis of Education Systems*) in Francophone Africa, SACMEQ (Southern and Eastern Africa Consortium for Monitoring Educational Quality) in Anglophone Africa, and LLECE (Latin American Laboratory for Assessment of the Quality of Education) in Latin America. Large-scale assessment may also exist at the national-level such as SIMCE (Sistema de Medición de Calidad de la Educación) in Chile and subnational level such as state-level testing in the United States or Canada. These assessments may vary in the grades or age levels tested, coverage of the target population (sample or census), subjects or skill areas covered, types of background data gathered, and the frequency with which they are administered. They also vary in how the results are reported and used. For example, while some stop at the reporting of results to policy makers or the general public, others use the results to hold specific groups accountable in the education system. The

¹⁴ Source: same as footnote 12.

application of a census-based approach to data collection and the use of background data can contribute to controlling for non-school factors that affect student achievement.

Several factors contribute to countries' decisions on how to design and implement assessment:

- the official vision and goals for the education system and the perceived role of assessment in achieving that vision;
- the historical legacy of assessment in a particular education system, which can lead towards a particular type of assessment activity (Madaus, Clarke, and O'Leary, 2003);
- the capacity of various stakeholders in the system to effectively carry out different types of assessment activities (Greaney and Kellaghan, 2008); and
- the perceived or real costs of assessment activities (Wolff, 2007).

Formative and summative

Assessment is often divided into formative and summative categories, depending on the purposes of the assessment practices. *Summative assessment* is generally employed at the end of a course or project. In an educational setting, summative assessments are typically used to assign a course grade to students. Summative assessments are evaluative. *Formative assessment* is generally employed throughout a course or project. Formative assessment, also referred to as "educative assessment", is used to aid learning. In an educational setting, formative assessment might be a teacher (or peer) or the learner providing feedback on a student's work, and would not necessarily be used for grading purposes. Formative assessment can be diagnostic.

Summative and formative assessments are often referred to in a learning context as the "assessment of learning" and "assessment for learning," respectively. "Assessment of learning" is generally summative in nature, intended to measure learning outcomes and report those outcomes to stakeholders (students, parents, and administrators, etc.). "Assessment of learning" generally occurs at the conclusion of a class, course, semester, or academic year. On the other hand, "assessment for learning" is generally formative in nature, and is mostly used by teachers to consider strategies for individual learners and whole classes.¹⁵

One way to differentiate among the three aforementioned types of assessment activities – classroom assessments, examinations and large-scale assessments – lies in that classroom assessment is mainly about assessment as learning or for learning (and hence is primarily formative in nature) while examinations and surveys are mainly about assessment of learning (and hence are primarily summative in nature). These distinctions do not always appear clearly in practice and hybrid approaches are becoming more common.¹⁶

¹⁵ Earl, L. (2003). *Assessment as Learning: Using Classroom Assessment to Maximise Student Learning*. Thousand Oaks: Corwin Press.

¹⁶ Source: same as footnote 12.

Table 1.1. Comparison of Assessment for Learning and Assessment of Learning

Parameters	Assessment for Learning	Assessment of Learning
Purpose	<ul style="list-style-type: none"> • Supports learning 	<ul style="list-style-type: none"> • Assesses accountability (linked to predetermined standards)
Goals	<ul style="list-style-type: none"> • Informs teaching and learning • Promotes further advances in learning • Focuses on improving learning • Develops pupils' reflection skills 	<ul style="list-style-type: none"> • Collects information about what has been achieved (a record of marks or grades) • Compares to targets that have been pre-established • Focuses on achievement
Actors	<ul style="list-style-type: none"> • Teachers • Pupils • Parents • Peers • Other school professionals 	<ul style="list-style-type: none"> • Teachers • External practitioners
When	<ul style="list-style-type: none"> • Ongoing 	<ul style="list-style-type: none"> • At fixed- and pre-determined times
Tools	<ul style="list-style-type: none"> • Discussion, observation, self-assessment, peer-assessment, teacher debate, comment-only, dialogue, questioning, feedback, no-grading, portfolio, individual education plan, etc. 	<ul style="list-style-type: none"> • Tests, drilling, grading, marking, questioning, observation, etc.

Source: Based on Harlen, W. (2007a). *Assessment of Learning*. London: Sage.

Diagnostic assessment is a regular method of formative assessment. It provides an indication of a learner's aptitude and preparedness for a course, unit or program of study and identifies possible learning challenges.

Self-assessment is a form of diagnostic assessment which entails students assessing themselves. It is a process during which students reflect on the quality of their work, judge the degree to which it reflects stated goals or criteria, and revise their work accordingly. The emphasis is on the word *formative* because self-assessment is done on ongoing work (e.g. drafts in progress) in order to inform revision and improvement. However, students do not determine their own grades. *Forward-looking assessment* asks students to consider themselves in hypothetical situations in the future.¹⁷ In contrast, *self-evaluation* refers to approaches that involve students in grading their work, perhaps as part of their final grade for an assessment or class¹⁸.

¹⁷ Reed, D. (2009). *Diagnostic Assessment in Language Teaching and Learning*. Center for Language Education and Research.

¹⁸ Andrade, H. and Valtcheva, A. (2008). "Promoting Learning and Achievement Through Self-Assessment", *Theory Into Practice*, 48: 1, pp. 12-19.

Authentic assessment, which aims to bring authenticity into assessment, is a concept that emerged in the mid-1990s to create an antidote to highly-standardized testing. Authentic assessment is any type of assessment that is aligned with the curriculum and requires students to demonstrate skills and competencies needed to address problems or situations that are likely to be encountered in daily life. Students are required to integrate knowledge, produce ideas and complete tasks that have real-world application. Such approaches require the assessor to use human judgement in the application of criterion-referenced standards (Archbald, 1991). The table below presents the main features of authentic assessment:

Table 1.2. Assessment Modes and Strategies for Authentic Assessment Practices

Type of assessment	Strategy	Recordkeeping
Observations	Group work, problem-solving, learning centres, excursions or field trips, developmental play, learning stories	Anecdotal notes, skills checklists, grading criteria, photography, video/audio/digital recordings
Work Samples	Individual work items: concept mapping, drawing, activity sheets, writing tasks, reflections, visual representations, surveys, position papers	Portfolios (digital and hard copy), student profiles, scrap books, files
Performances	Dramatic enactments, debates, interviews, operas, raps, poetry, songs, dance, panel discussions	Marking criteria, rubrics, peer- and self-assessment, descriptive feedback, oral or written
Products	Models, murals, collages, written projects, community projects, presentations, design briefs	Marking criteria, rubrics, peer- and self-assessment, descriptive feedback (oral or written)

Source: Moss, J., Godinho, S. From profiles to rich tasks: the situated nature of 'authenticity' in the context of reforming curriculum and assessment practices. Paper presented at the British Educational Research Association Annual Conference, University of Glamorgan, 14-17 September 2005.

In authentic assessment, students apply recalled information to produce an original product, participate in a performance, or complete a project. Students are assessed according to specific criteria that are known to them in advance. They do not simply recall information; rather, they apply what they know to new tasks. For example, there is a difference between asking students to identify all the metaphors in a story and asking them to discuss why the author used particular metaphors and the effect they had on the story. In the latter case, students must put their knowledge and skills to work just as they might naturally do outside of school (Valencia, 1997).

Authentic assessment (designed to demonstrate learners' mastery of the skills and competencies needed) is sometimes regarded as synonymous with formative assessment (designed to provide them with feedback on their progress in order to inform their development¹⁹), as both modes are integrated into the teaching and learning process and aim to enhance student learning. These are not competing concepts and they are not mutually exclusive. Rather, they are two ways of framing a related approach to assessment. The following table compares traditional assessment to authentic assessment:

¹⁹ Collins, R. (2013). "Authentic assessment: assessment for learning", Curriculum Leadership Journal, Vol. 11, No. 7. Education Services Australia. Available at: http://www.curriculum.edu.au/leader/authentic_assessment_assessment_for_learning_36251.html?issueID=12745.

Table 1.3. Traditional vs. Authentic Assessment Methods

Traditional Assessment	Authentic Assessment
Generally relies on forced-choice, written measures	Promotes integration of various written and performance measures
Relies on proxy measures of student learning to represent target skills	Relies on direct measures of target skills
Encourages memorization of correct answers	Encourages divergent thinking in generating possible answers
Goal is to measure acquisition of knowledge	Goal is to enhance development of meaningful skills
Emphasis on developing a body of knowledge	Emphasis on ensuring proficiency in real-world tasks
Promotes knowledge of “what”	Promotes knowledge about “how”
Provides a one-time snapshot of student understanding	Provides an examination of learning over time
Emphasizes competition	Emphasizes cooperation
Targets simplistic skills or tasks in a concrete, singular manner	Prepares students for ambiguities and exceptions that are found in realistic problem settings
Priority on summative outcomes or products	Priority on the learning sequence or process

Source: Lombardi, M. M. (2008). Making the Grade: The Role of Assessment in Authentic Learning, ELI Paper, vol. 1.

There is no specific form or method of assessment that can be called formative. Rather, there are different forms and methods that can be used for formative purposes: “the distinction between formative and summative applies not to the assessment itself, but to the use to which the information arising from the assessment is put.”²⁰ In fact, an important aspect of assessment for learning is the formative use of summative data.²¹

Performance-based assessment is similar to summative assessment, as it focuses on achievement. It is often aligned with the standards-based education reform and outcomes-based education movement. Ideally, performance-based tests are significantly different from traditional multiple choice tests; however, they are most commonly associated with standards-based assessment in which free-form responses to standard questions are scored by humans on a standards-based scale. In such systems, learners meet, fall below, or exceed a performance standard and they are not ranked against a curve. A well-defined task is identified and students are asked to create, produce or engage in an activity, often in settings that involve real-world application of knowledge and skills. Proficiency is demonstrated by providing an extended response to the activity. Presentation formats are further differentiated into products and performances. A task may result in a product, such as a painting, portfolio, paper or exhibition, or it may consist of a performance, such as a speech, athletic skill, musical recital or reading.

²⁰ William, D. The Meanings and Consequences of Educational Assessments *in* Critical Quarterly, 42(1), pp. 105-127 (2000).

²¹ Collins, R. (2013), “*Authentic assessment: assessment for learning*”, Curriculum Leadership Journal, Vol. 11, No. 7. Education Services Australia. Available at: http://www.curriculum.edu.au/leader/authentic_assessment_assessment_for_learning_36251.html?issueID=12745.

A *portfolio* is an authentic collection of a learner's work that is able to document his/her learning over a period of time. It can refer to an individual collection of daily drawings, writings and other materials that document a learner's strengths (Schipper and Rossi, 1997). It can also be "a purposeful collection of student work that exhibits learners' efforts, progress, and achievements in one or more areas" (Paulson, Paulson, and Meyer, 1991, pp. 60-61). Therefore, portfolio assessment may mean different things according to the reason for using a portfolio. It may improve communication about complex student learning targets and processes, especially as teachers are encouraged to ensure that:

- The portfolio includes criteria for content selection and criteria for judging merit;
- Students participate in the selection of contents;
- Students comment on why the contents are selected and what they represent;
- Students reflect on what they learned as a result and include evidence of such self-reflection in the portfolio; and
- The portfolio-making and experience-sharing is safe (especially in the way teachers treat students and the way they allow students to treat others).

Portfolios provide a direct method for students to reflect more deeply upon their learning and retaining the lessons longer. They are a medium through which students can exercise responsibility for their own learning and engage in the critical thinking required for self-assessment and self-reflection.²²

Objective and subjective

Assessment (either summative or formative) can also be categorized as either objective or subjective. *Objective assessment* is a form of questioning that has a single correct answer. *Subjective assessment* is a form of questioning which may have more than one way of expressing the correct answer. There are various types of objective and subjective questions. Objective question types include true/false statements, multiple choice, multiple-response and matching questions. Subjective questions include extended-response questions and essays. Objective assessment is well-suited to the increasingly popular computerized or online assessment format.

Some have argued that the distinction between objective and subjective assessments is neither useful nor accurate because, in reality, no assessment is fully "objective." In fact, all assessments are created with inherent biases built into decisions about learning, including cultural biases that relate to class, ethnicity, gender, etc.²³

Referencing

Criterion-referenced assessment, typically using a criterion-referenced test, occurs when candidates are measured against defined criteria. Criterion-referenced assessment is often, but not always, used to establish a person's competence. For example, a driving test is a criterion-referenced assessment, where student drivers are measured against a range of explicit criteria. In this case, a criterion would be "not endangering other road users or pedestrians".

²² Chappuis, J.; Stiggins, R. J.; Arter, A. (2012). *Classroom Assessment for Student Learning: Doing It Right - Using It Well*. Pearson.

²³ Joint Information Systems Committee (JISC). What Do We Mean by e-Assessment in Merriam-Webster Dictionary. (2005). Available at Dictionary.reference.com.

Norm-referenced assessment (colloquially known as "grading on the curve"), typically using a norm-referenced test, is not measured against defined criteria. This type of assessment is relative to the entire student body's assessment performance and is effectively a way of comparing students. Many entrance tests are norm-referenced, allowing a fixed proportion of students to be accepted into a school or university (rather than assessing an explicit level of ability). This implies that standards may vary from year to year, depending on the quality of the cohort. As for criterion-referenced assessments, they generally do not vary from year to year, unless the established criteria change.²⁴

Ipsative assessment is self-comparison either in the same domain over time, or compared to other domains within the same student.

Informal and formal

Assessment can be either formal or informal. *Formal assessment* usually involves a written document, such as a test, quiz, or paper. A formal assessment is given a numerical score or grade based on student performance, whereas an *informal assessment* such as a copy and pasted discussion question does not contribute to a student's final grade. An informal assessment usually occurs in a more casual manner and may include observation, inventories, checklists, performance, portfolio assessments, participation, peer and self-assessment, and discussion.²⁵

Internal and external

Internal assessment is established and marked by the school (i.e. teachers). Students receive the mark and feedback regarding the assessment. *External assessment* is set by a governing body, and is marked by non-biased personnel. Some external assessments offer limited feedback while others provide students with detailed feedback so that students are aware of their learning achievements and opportunities for improvement.

The Quality of Assessment

Reliability and validity are two overarching technical issues in assessment. Generally, high-quality assessments are considered those with a high level of reliability and validity. Approaches to reliability and validity may vary.

Reliability

Reliability relates to the consistency and accuracy of an assessment. It is an important consideration for high-stakes examinations and for monitoring trends over time. A reliable assessment is one that achieves consistent results with the same (or a similar) cohort of students. Various factors affect reliability, including ambiguous questions, too many options within a paper, vague marking instructions and poorly-trained markers. Traditionally, the reliability of an assessment is based on the following:

- *Temporal stability*: performance on a test is comparable on two or more separate occasions;
- *Form equivalence*: performance among examinees is equivalent on different forms of a test based on the same content;
- *Internal consistency*: responses on a test are consistent across questions. For example, in a survey that asks respondents to rate attitudes on technology, consistency would be measured across questions such as, "I feel very negative about computers in general" and "I enjoy using computers."²⁶

²⁴ Educational Technologies. Assessment Purposes. VirginiaTech DesignShop: Lessons in Effective Teaching. Available at Edtech.vt.edu.

²⁵ Valencia, S. W. (1997). What Are the Different Forms of Authentic Assessment? *in* Understanding Authentic Classroom-Based Literacy Assessment, available at Eduplace.com.

²⁶ Yu, C. (2005). Reliability and Validity *in* Educational Assessment. Available at Creative-wisdom.com.

Validity

A *valid assessment* is one which succeeds in measuring what it is intended to measure. For example, when the language of instruction differs from the language of assessment, it is difficult for a child to demonstrate what they know and can do. Likewise, it would not be valid to assess driving skills through a written test alone. A more valid way of assessing driving skills would be through a combination of tests, such as through both a written test of driving knowledge and a performance assessment of actual driving. When teachers complain that some examinations do not properly assess the syllabus upon which learning is based, they are effectively questioning the validity of the exam.

Validity of an assessment is generally gauged through an examination of evidence in the following categories:

- *Content*: Does the test's content measure given objectives?
- *Criterion*: Do the test scores correlate to an external reference? (e.g. Do high scores on a 4th grade reading test accurately predict reading skill in future grades?)
- *Construct*: Does the assessment correspond to other significant variables?²⁷ (e.g. Do students learning English as a Second Language consistently perform differently on a writing exam than native English speakers?); and
- *Face*: Does the item or theory make sense, and is it seemingly correct to an expert reader?²⁸

A good assessment has both validity and reliability, plus the other quality attributes noted above for a specific context and purpose. In practice, an assessment is rarely totally valid or reliable. A test that consists of entirely multiple-choice questions will have high reliability, whereas having essay and fill-in-the-blank questions will provide higher validity as to the actual mastery of a subject. We may generalize from this: reliability and validity can be negatively correlated. It is also important to note that there are at least thirteen sources of invalidity, which can be estimated for individual students in test situations. They never are. Perhaps this is because their social purpose demands the absence of any error, and validity errors are usually so high that they would destabilize the entire assessment system. The use of assessment results requires the careful consideration of possible consequences, including social, economic and other impacts on different groups of the population.

Evaluation Standards

These standards provide guidelines for designing, implementing, assessing and improving the identified forms of evaluation, in order to promote educational evaluations that are appropriate, useful, feasible, and accurate. For example, the student accuracy standards help to ensure that student evaluations will provide sound, accurate, and credible information about student learning and performance.

Learning Targets and Rubrics²⁹

All instructional activities should be designed to help students practice or demonstrate proficiency on specific *learning targets*. Assessment materials emphasise the same learning targets.

²⁷ Moskal, B. M.; Leydens, J. A. (2000). Scoring Rubric Development: Validity and Reliability *in* Practical Assessment, Research and Evaluation, 7(10).

²⁸ Vergis, A.; Hardy, K. (2010). Principles of Assessment: A Primer for Medical Educators in the Clinical Years *in* The Internet Journal of Medical Education, 1 (1). Available at: http://www.ispub.com/journal/the_internet_journal_of_medical_education/volume_1_number_1_74/article_printable/principles-of-assessment-a-primer-for-medical-educators-in-the-clinical-years-4.html.

²⁹ Arter, J. (2009). Creating and Recognising Quality Rubrics. Pearson Assessment Training Institute (U.S.). Flinders University (Australia). Using rubrics in assessment. Available at: <http://www.flinders.edu.au/teaching/teaching-strategies/assessment/grading/using-rubrics.cfm>.

A *rubric* presents written criteria that describe levels of quality in complex student reasoning, performances, or products. Rubrics can be used to help students identify and understand the learning targets through examples, teacher feedback, peer- and self-assessment (in assessment *for* learning). Later, the rubrics can also be used to evaluate the final performances and products (in assessment *of* learning).

Carefully developed rubrics provide more than just a mechanism for consistent grading. They can also be used to accomplish two broad aims: for teachers to *educate* students (track student mastery of important learning targets), and for students to develop their capacity to *judge* their own work, thereby sending the message that ownership of their own learning is respected and valued.

Additionally, a rubric can be used to foster a greater level of understanding and trust between the teacher and students. When a rubric reveals the grading or scoring rules that inform assessment judgements and decisions, grading is no longer a special secret activity.

Further, rubrics may also be a tool for more effective grading by helping assessors to make more efficient and consistent judgments and to provide basic feedback on key dimensions of performance and learning. However, rubrics are not appropriate for all assessment tasks and should not be relied upon as the only form of feedback for students.

Controversy

High-stakes testing and standardized tests have long been a key focus of concern when considering how best to apply assessment practices across public school systems. There are divergent perspectives on how to gauge student progress, teacher quality, and school-, district-, or state-wide educational success.

High-stakes testing

Assessments can incite the greatest controversy when those who failed to pass high-stakes examinations are denied further educational opportunities. For instance, in the U.S., high school graduation examinations are used to decide whether students will be granted a high school diploma. Opponents maintain that no student that has invested four years of seat time should be denied a diploma for repeatedly failing a test, or for not knowing the required material.³⁰

High-stakes testing can also undermine the pleasure of learning. Among students and teachers, illness and test anxiety have been attributed to *high-stakes tests*. In some cases, teachers have even chosen to narrow the curriculum to the content the teacher believes will be tested. Other critics have questioned the use of test materials that extend beyond the standard cognitive levels of students' age groups.³¹

High-stakes testing are also questioned for their reliability. For instance, in the U.S., the use of IQ tests has been banned in some states for educational purposes. Similarly, norm-referenced tests, which rank students from "best" to "worst", have been criticized for bias against minorities. Most educational officials support criterion-referenced tests where each student's score depends solely upon whether he or she has answered the questions correctly, regardless of how his or her peers performed. Furthermore, in comparison to portfolio assessments, simple multiple-choice tests are much less expensive, less prone to disagreement between scorers, and can be scored quickly enough to be returned before the end of the school year. For these reasons, standardized tests, where all students take the same test under the same conditions, often employ multiple-choice formats. However, some authors have argued that students may not be consistent in their performance, as they can have "good days" and "bad days".

³⁰ Dang, N. Reform Education, not Exit Exams *in* Daily Bruin (18 March 2003). Weinkopf, C. Blame the Test: LAUSD Denies Responsibility for Low Scores *in* Daily News (2 June 2002).

³¹ Bach, D.; Blanchard, J. WASL Worries Stress Kids, Schools *in* Seattle Post-Intelligencer (April 19, 2005).

There are also inconsistencies in the ways that assessments are marked. Most importantly, the choice of questions included in each test can already be subject to bias.³²

The European Agency for Development in Special Needs Education recommended that “educational standards [should] be evaluated, but ‘snapshot’ assessments [should] not be used as the basis for decisions about individual pupils, teachers, schools, policies, financing or resourcing.”³³

Evolving purposes of student assessment

In the past few decades, while summative assessment has been the prevalent form of student assessment, there has been a growing interest in other forms of assessment. The conceptualisation of student assessment and its purpose has evolved over time. In the past, many countries lacked the resources to provide all learners with advanced education. Therefore, assessment became increasingly selective at higher age and educational levels. Today, in light of the EFA Goals and their socio-economic aims, countries have become more sensitive to the allocation of greater resources and opportunities for education. As the number of secondary school and university students continues to grow, there have been calls for revisiting the purposes of assessment, and more specifically, a transition from “*selecting talents*” to “*developing talents*.”³⁴

21st Century assessment

With the development of internet technologies and social media, learning has become increasingly collaborative and knowledge is increasingly distributed across many members of a learning community. Traditional assessment practices which primarily focus upon the individual will be less likely to account for knowledge-building and learning in context. The cultural shifts arising from the emergence of a more participatory culture will require the education community to explore new methods for assessing learners.³⁵

The Introduction of Inclusive Student Assessment

In the process of exploring the concept of inclusive assessment, the 2008 *European Conference on Assessment in Inclusive Settings* is an important milestone. The Conference gathered representatives from 29 European countries in Cyprus and culminated in a Recommendation document which is a useful reference for the understanding and practice of inclusive assessment. According to the Cyprus Recommendations, inclusive education:

- “is an approach to assessment in all educational settings where policy and practice are designed to promote the learning of all pupils as far as possible;”
- “aims to prevent segregation by avoiding labelling and by focusing on learning and teaching practice that promotes inclusion;”
- “can only be realised within an appropriate policy framework and with the appropriate organisation of schools and support to teachers who have a positive attitude towards inclusion.”

The Conference also recommended that:

- All assessments be coherent, coordinated, and guided by the goal of informing and promoting learning;

³² William, D. The Meanings and Consequences of Educational Assessments *in* *Critical Quarterly*, 42(1), pp. 105-127 (2000).

³³ European Agency for Development in Special Needs Education (2007) *Assessment in inclusive Settings. Key Issues for Policy and Practice.*

³⁴ Cf. Guskey, T. R. Five Obstacles to Grading Reform *in* *Educational Leadership*, Vol. 69, No. 3, November 2011.

³⁵ Fadel, C.; Honey, M.; Pasnik, S. (2007). Assessment in the Age of Innovation *in* *Education Week* (May 18, 2007).

- Pupils be entitled to inform assessment procedures in which they are participants;
- All assessment be reliable, valid and meet the specific needs of individual pupils;
- The summative and formative approaches and uses of assessment be aligned and combined;
- Investment be made in training and support for formative assessment;
- Innovation in assessment be encouraged; and
- Research, policy and practice be bridged in the area of assessment.

In light of recent developments with regard to the understanding of Inclusive Assessment, an inclusive student assessment system should be one that serves the ideals of inclusive education. Such an assessment system would not only allow students to maximize access to learning opportunities, but also cater for learners' individual differences and contribute to improving the quality of education at the macro, meso and micro levels of a system. Strategies adopted and data collected through an assessment that is truly inclusive should not only be used for "traditional" purposes such as selection, certification, accountability or comparison, but also and more prominently for the improvement of performance and equity in education at all levels – systemic, institutional and individual –.

Keywords:

Quality Education; Inclusive Education; Inclusive Curriculum; Student Assessment; Types of Student Assessment; Quality of Assessment; Inclusive Assessment

Training manual

ACTIVITY 1

How “inclusive” are our current student assessment systems?

Objectives

1. Link student assessment to the broadened concepts of inclusive education and inclusive curriculum.

Background information and discussion

Despite the long-standing efforts and progress that countries around the world have made towards the Education for All Goals, there are still 61 million out-of-school children and among them, 26% are considered drop-outs.³⁶ Although such challenges are particularly relevant in the least developed countries, more developed countries have also faced setbacks in improving student learning achievement. In the European Union countries, on average, 14% of young people aged 18 to 24 have no education or training beyond lower secondary levels.³⁷ Further, educational communities have increasingly started to consider whether some features of current education systems may “push” learners out of the education system or distance them from the pleasure of learning.

In the last decades, there have been shifts towards more holistic and sustainable ways of teaching and learning, aimed at preparing students to become responsible citizens, and equipping them for the future. However, even in more developed countries, researchers and practitioners have argued that assessment practices have not kept up with these changes. Others have suggested that if two students are being assessed on something they have not had equal opportunities to learn, the assessment is unjust.³⁸ Others have voiced concern that some assessment practices are inadequate or discourage students.

As we progress in conceptualizing inclusive education and inclusive curriculum (cf. *Conceptual Framework*), “inclusive student assessment” should support the attainment and be an integral part of such education and curriculum.

Focus of the activity

Effects of current assessment practices; test culture vs. broad assessment culture; summative and formative assessment; reducing exclusion through assessment

³⁶ Figures as of 2010. Cf. UNESCO (2012). 2012 Global Monitoring Report on Education for All. Youth and Skills: Putting Education to Work. Paris: UNESCO.

³⁷ *Ibid.*

³⁸ Gee, J.P. (2003) ‘Opportunity to learn: a language-based perspective on assessment’, *Assessment in Education*, vol. 10, no. 1, pp. 27-46.

TASK 1 Individual reflection on the “friendliness” of current student assessment

Read the following text:

“Unfriendly” aspects of traditional student assessment³⁹

In contemporary education systems, some traditional forms of assessment can negatively impact the teaching-learning environment on several fronts:

- **Learning unfriendly:** traditional tests and examinations may reveal what students recall in light of what we think they should remember, but do not necessarily uncover purpose, meaning or utility of learning. Such testing basically represents a form of assessment that is confined (only assesses a very narrow range of academic qualities), “sudden death” (candidates obtain credit only for what they do during the brief examination period) and non-informative (the grade does not reflect what a candidate knows and can do).
 - **Curriculum unfriendly:** examinations emphasize recall of factual knowledge with a heavy reliance on memory and rote learning. Consequently, they can lead teachers to coach their students in a narrow range of test-taking skills, rather than encourage them to teach a broader range of higher-order competencies and understanding. As a result, students develop problem-solving strategies that help them pass examinations without reaching the intended learning goals.
 - **Teacher unfriendly:** among teachers, the knowledge that test scores are the sole means of describing and judging schools cultivates the determination to accomplish what is necessary but results in a ‘test-like’ teaching programme – to avoid being publicly embarrassed and humiliated. In fact, many teachers may even be willing to engage in questionable, educationally-indefensible practices to increase their students’ test scores.
 - **Student unfriendly:** students find themselves not so much described by examinations as constructed by them. An unwelcome byproduct of this ‘intrusive’ process is labelling, which can subconsciously establish limits on student achievement. When teachers have low expectations, apart from actively inhibiting student performance, such an outlook may also promote self-sabotage within students. Instead of motivating all students to work harder and continue learning, such environments may demotivate them and effectively force them out of the system.
-

Reflect upon the following:

- Do you agree with these criticisms of traditional student assessment? Which are the ones that you agree with, and which do you not agree with? Why?
- Which of the “unfriendly” aspects apply to student assessment in your context, and which do not apply? Conduct research to obtain support (evidence, official data and communications) for your position.

PRODUCT: Bulleted list of arguments, agreeing or disagreeing with the criticism of traditional student assessment, a summary of those “unfriendly” aspects that apply or do not apply to your context, as well as evidence to support your position.

³⁹ Extracted from Buhagiar, M. A. (2007). Classroom assessment within the alternative assessment paradigm: revisiting the territory, *Curriculum Journal*, vol. 18, no. 1, pp. 39-56.

TASK 2 Work in small groups on criterion-referenced assessment

In a small group, examine the image located on the following website:

<http://gjismyp.files.wordpress.com/2011/10/fair-assessment.jpg>

Reflect upon the following:

- To what extent does this image reflect criterion-referenced assessment? What weaknesses of such assessment does this image attempt to convey, if any?
- Is criterion-referenced student assessment employed in your context? If so, under what circumstances?
- In your context, what are the benefits of applying criterion-referenced assessment? What are the challenges?
- How can these deficiencies be overcome?

After shared reflection on the image, each group should present a summary of the discussion to the plenary.

PRODUCT: Group presentation on the use, strengths and deficiencies of criterion-referenced student assessment in group members' contexts.

TASK 3 Work in small groups on student assessment and its effects on learning

In small groups, preferably organized by country, complete the worksheet [Factors/Practices of current student assessment and their effects on learning](#) by responding to the following two questions:

- Within your country's current student assessment mechanisms, what factors or practices are conducive to fostering students' desire to learn and encourage positive learning experiences and achievements?
- Within your country's current student assessment mechanisms, what factors or practices undermine students' desire to learn or hinder their long-term learning experience and achievements?

After the exercise, each group should present their worksheet to the plenary.

Worksheet: Factors/Practices of current student assessment and their effects on learning

Conducive to learning		Hindering learning	
Factors/Practices	Effects	Factors/Practices	Effects

PRODUCT: Worksheet with group inputs.

TASK 4 Individual reading and reflection on the object of assessment

Read the comparative case study *What are we assessing: the student or the work?*

What are we assessing: the student or the work?⁴⁰

Excerpt (see extended version in Annex)

Teachers' assessment of their pupils' work on a daily basis is not limited to determining whether the work is 'right' or 'wrong'. Other factors are frequently taken into account: children's efforts, their progress, and neatness. These three elements featured in all the infant schools that were observed in England and in France. However, they were given different status, and this was reflected in the content of written evaluation.

In France, various scales are used for marking pupils' work, some verbal (very good, good, fair, NA), others numerical (a mark out of 10 or 20). Marks are meant as an objective measurement of attainment. They are a statement on the outcome of a task and are not related to the child who produced it, not to the child's efforts and past performance.

In the UK, the "good work check-up" has a formative dimension. It contains remarks on the quality of the work, on progress, on presentation and on efforts. The way of marking often reflects pupils' progression in learning, without an external standard of attainment. English teachers' judgment is relative to each child and they have greater sensitivity to issues of self-esteem; it is a personalized system. This is a major characteristic of the English primary education system.

The basis on which assessment rests – cognitive outcomes in France, the whole child in England – reflects a key difference between the two education systems. French educational practice relates to the republican ideals of equality. French teachers encourage their students not to consider a mark as being

⁴⁰ Extracted from Raveaud, M. (2004). Assessment in French and English infant schools: Assessing the work, the child or the culture?

a judgment passed on them as individuals. A 'distancing of self' in French education makes pupils less vulnerable to criticism. In England, the criteria used to evaluate children's works reveals a continuing concern to take into account the whole child, not simply academic achievement. Any education system needs to strike a balance between respecting children's needs and interests on the one hand, and pushing towards achievement on the other.

Reflect on the following:

- Have you identified any noteworthy points in the case study? Please make a short list of the salient points.
- Is the assessment culture and approach in your context similar to the French or the English system? If it is very different from either system, please explain how it is different.
- How can these two approaches shape your assessment culture and mechanisms to promote inclusiveness, especially in the development of learning and competencies for active citizenship, adult life and social cohesion? Why?

PRODUCT: A personal note containing the answers to the questions for reflection.

TASK 5 **Work in small groups on recent efforts in changing assessment**

Read the case studies *Shift from 'test culture' to 'broad assessment culture' in Belgium: challenges of translating theory into practice* and "*Striving towards a balance between summative and formative assessment in Hong Kong*."

Shift from 'test culture' to 'broad assessment culture' in Belgium: challenges of translating theory into practice⁴¹

Excerpt (see extended version in Annex)

In the Flemish Community of Belgium, the shift from the traditional 'test culture' towards a 'broad assessment culture' has recently gained importance. The broad assessment culture relates to the implementation of 'active learning methods'. Students are no longer perceived as passive subjects, but they are involved in the assessment of their performances and progress, and may even be involved in the assessment of their classmates.

The formative function of evaluation is strongly emphasized within the 'broad assessment culture': evaluation not merely serves the goal of judging students, but also (and more importantly) to guide students to gain knowledge, skills and attitudes through feedback mechanisms. It is more important to provide information on the progress of a student, than marking/grading what he/she achieved. This means that evaluation is conducted continuously, because both teachers and students need up-to-date information on individual performances in order to create effective learning environments. Evaluation as part of the learning process helps students in their development of holistic competences: next to assessment of cognitive performances, also dynamic-affective, relational and motor aspects are continuously assessed.

⁴¹ Extracted from Ministry of Education and Training Belgium and University of Antwerp Edubron Research Group (2010). Country Background Report for the Flemish Community of Belgium, OECD Review on Evaluation and Assessment Frameworks for Improving School Outcomes.

In the broad assessment culture, observation, portfolios, reflection sheets, peer- and self-assessment are only a few examples of the wide range of assessment possibilities. The teacher needs to consider which assessment method contributes most to guiding his students to the aspired objectives. Student assessment methods need to be realistic, adjusted to the student characteristics, fostering student activity and integrated in the learning processes. The norm in the broad assessment culture shifts from pre-determined or peer-group norms to an individual norm (without losing the pre-determined attainment targets and developmental objectives out of sight). The job profile and basic competences required from teachers emphasize that teachers should be competent in implementing the broad assessment culture.

There is, however, opposition towards this shift. Some teachers, parents and stakeholders claim that students' general knowledge is decreasing because schools are becoming less demanding as far as cognition is concerned and have been putting too much emphasis on the training of skills. A study carried out in 2002 in secondary schools pointed out that the traditional test culture still prevailed. The Inspectorate's report Education Mirror 2006 concluded that student assessment practices are the weakest link in the students' learning process towards the attainment targets and developmental objectives. Innovative methods like self-evaluations were hardly ever applied in the evaluation practices. Assessment focuses largely on knowledge reproduction. Furthermore the reliability and validity of tests and exams used in secondary education needed improvement. According to parent organizations, it is too often unclear which contents are evaluated and which assessment methods will be used.

Striving towards a balance between summative and formative assessment in Hong Kong⁴²

Excerpt (see extended version in Annex)

Before the expansion of compulsory schooling up to Junior Secondary level in 1979, the local assessment structure in Hong Kong had been criticized for its examination-oriented nature. The importance of changing assessment practices for improving teaching and learning has contributed to critical discussions in major reform policy documents. Curriculum change will not effect without making corresponding changes in assessment. These changes include more emphasis on "assessment for learning" than "assessment of learning".

Since 2000, the Government of Hong Kong has implemented a large scale educational reform in which a series of measures was employed to promote a new culture of assessment:

Instead of performing the major selective role as in the past, assessment "should provide the information necessary to help teachers and students improve teaching and learning". More importantly, it should possess "a positive rather than a negative effect on teaching and learning". Instead of the scoring, ranking, and grading that serve mainly summative assessment purposes, there is increasing demand on drawing a clear picture of students' learning progress among teachers, parents, and students.

However, the territory-wide Basic Competency Assessment (BCA) [designed to help teachers better understand the learning needs and problems of students, so as to provide timely assistance to enhance students' learning effectiveness] has created new problems. While the Government can get more information about schools from results of BCA, schools are put into a more vulnerable situation if their students have failed the system tests. Principals and teachers are worried about closing down of schools in consequence. The BCA first considered low-stake has been interpreted as new public examinations by some schools. In many schools, unnecessary rehearsals are being practiced.

Assessments, no matter they are "formative" or "summative", have their unique roles in the teaching and learning process. We see the effective assessment mode an integration of the two. Assessment will only be effective when "learning" is put in the integral part of the process.

⁴² Extracted from Yu, W. M. et al. (2006). Assessment Reform in Basic Education in Hong Kong: The Emergence of Assessment for Learning.

In small groups, discuss the following:

- In the case of the Flemish Community of Belgium, apart from theory and broad national programmes, what other measures and practices would be required to bring about the intended change in the assessment culture?
- In the case of Hong Kong, if there have been changes in assessment practices, why have some schools misinterpreted the Basic Competency Assessment system as high-stakes testing? What related perceptions and social phenomena do you think exist?

Each group reports on their discussions to the plenary and the facilitator provides an overview.

PRODUCT: Group report.

TASK 6 **Work in small groups and plenary on curriculum and student assessment**

Read the following statement:

In light of the evolving conceptualization of Inclusive Education, in order to preserve the “process of addressing and responding to the diversity of needs of all learners through increasing participation in learning, cultures and communities, and reducing exclusion within and from education,”⁴³ we need to reconsider and renew curriculum, including the content, approaches, structures and strategies used for learning.

1. In small groups, reflect on the following:
 - Is curriculum reform enough to reshape learning? What would the role of student assessment be in such reform?
 - What would the relationship between student assessment and curriculum be? Should curriculum include guidelines for student assessment?
2. Each group draws a diagram to describe the connections and dynamics that were discussed. Groups may add as many connecting actors and factors as they desire.
3. Each group should present their responses to the questions to the larger group, as well as the diagram produced, citing some examples from real experiences.
4. Following each presentation, the other groups should assess the groups that have presented, indicating an assessment result and the reason for that result, as well as the assessment approaches adopted.
5. The presenting groups will share their opinion on the assessment results provided by the other groups.

PRODUCT: Group responses, diagrams on curriculum and student assessment, and inputs from the plenary.

⁴³ United Nations Educational, Scientific and Cultural Organisation. (2005). Guidelines for inclusion: Ensuring access to Education for All. Paris: UNESCO.

ACTIVITY 2

How can we adjust student assessment so that it contributes to high-quality education and inclusiveness?

Objectives

1. Review interesting practices and models that can result in greater inclusivity in education, e.g. the development of formative assessment, the use of summative assessment for inclusive purposes, the design and use of quality rubrics for student assessment.

Background information and discussion

As we have mentioned, assessment has formative and summative purposes. There is much research on how formative assessment can contribute to student learning so that students perform better on final (“real”, summative and selective) exams.

Most of the time, formative assessment is directly integrated and inseparable from teaching and learning. It can, for example, include “classroom interactions, questioning, structured classroom activities, and feedback aimed at helping students to close learning gaps.”⁴⁴

According to Shavelson et al.,⁴⁵

“Formative assessment takes place on a continual basis. It is conducted by teachers with the intent of informing teachers and students about the gap between what students know and can do, and what they are expected to know and be able to do with immediate, informative feedback. Classroom formative assessment ranges on a continuum from informal to formal. Where a particular formative assessment practice falls on the continuum depends on the amount of planning involved, its formality, the nature and quality of the data sought, and the nature of the feedback offered to students by the teacher”.

Promoting wider, deeper and more sustained formative assessment practices would have policy implications,⁴⁶ including measures to:

1. Maintain the focus on teaching and learning;
2. Align summative and formative assessment approaches;
3. Ensure that data gathered at the classroom, school and system levels are linked, and are used formatively;
4. Invest in training and support for formative assessment;
5. Encourage innovation; and
6. Build stronger ties between research, policy and practice.

⁴⁴ Looney, J. W. (2011). Integrating Formative and Summative Assessment: Progress Toward a Seamless System?, OECD Education Working Papers, No. 58, OECD Publishing.

⁴⁵ Shavelson, R. J. et al. (2008). On the Impact of Curriculum-Embedded Formative Assessment on Learning: A Collaboration between Curriculum and Assessment Developers, *Applied Measurement in Education*, 21:4, pp. 295-314.

⁴⁶ OECD (2008). Assessment for Learning. Formative Assessment, OECD/CERI International Conference “Learning in the 21st Century: Research, Innovation and Policy”.

Summative assessment does not necessarily lead to exclusion. For instance, summative data can be used to analyze learning inequalities, thereby informing policies and practices that promote inclusion.

Focus of the activity

Policy implications of formative assessment; characteristics of inclusive assessment; attributes of formative assessment; student assessment reform with inclusive features; summative assessment for inclusion; student self-assessment; quality rubrics for student assessment

TASK 1 Work in small groups on the characteristics of an inclusive assessment

In small groups (preferably organized by country), record the following in the worksheet [Characteristics of an inclusive assessment](#):

1. Bulleted list on the characteristics you would expect an inclusive assessment to have;
2. How is student assessment organized in your country at different educational levels? What are the main features of the current student assessment mechanism?;
3. How can these features be modified to better reflect the characteristics of inclusive assessment that your group just identified; and
4. What needs to be achieved in order to implement these changes in your context?

Worksheet: Characteristics of an inclusive assessment		
<ul style="list-style-type: none"> • • • • 		
Current student assessment mechanism: main features	Possible changes towards greater inclusion	Conditions/ Actions needed

PRODUCT: Group worksheet.

TASK 2 **Work in small groups on the attributes and strategies of formative assessment**

1. Working in the same groups from Task 1, read the case study [Cyprus Recommendations on Inclusive Assessment](#) and the documents [The Attributes of Formative Assessment: A Literature Review](#); [Three Anchor Points of Formative Assessment Continuum](#).

Cyprus Recommendations on Inclusive Assessment⁴⁷

Excerpt (see extended version in Annex)

More than 150 participants representing 29 countries met during the Conference 'Assessment in Inclusive Settings' jointly organized by the Cypriot Ministry of Education and Culture and the European Agency for Development in Special Needs Education (Limassol, Cyprus, 23-24 October 2008). The Conference Representatives agree:

- Upon the concept of Inclusive Assessment: an approach to assessment in all educational settings where policy and practice are designed to promote the learning of all pupils as far as possible;
- That the overall goal of inclusive assessment is that all assessment policies and procedures should support and enhance the successful inclusion and participation - physical, social and academic - of all pupils including pupils who are vulnerable to exclusion and especially pupils with special educational needs;
- That the principles underpinning inclusive assessment are:
 - All assessment procedures should focus on informing and promoting learning;
 - Pupils should be entitled to inform assessment procedures in which they are participants;
 - All pupils should be entitled to be part of assessment procedures that are reliable, valid and accommodated to meet the specific needs of individual pupils;
 - All assessment procedures should be constructed using the principles of universal design so that they give all pupils the opportunity to demonstrate their learning achievements, skills and knowledge;
 - The needs of pupils with special educational needs should be considered and accounted for within all general as well as special education specific assessment policies;
 - All assessment procedures should be complementary and inform each other;
 - All assessment procedures should aim to take into full account and also celebrate diversity by identifying and valuing all pupils' individual learning progress and achievements;
 - All assessment procedures should be coherent, coordinated and guided by the goal of supporting learning and teaching;
 - Inclusive assessment explicitly aims to prevent segregation by avoiding – as far as possible – forms of labelling and by focusing on learning and teaching practice that promotes inclusion.

⁴⁷ Extracted from European Agency for Development in Special Needs Education (2008). Assessment in inclusive Settings. Cyprus Recommendations on Inclusive Assessment.

The Representatives recommend that:

- All pupils are involved in and have opportunities to influence their own assessment and the development, implementation and evaluation of their own learning targets and plans (individual education plan or similar tool);
- Parents are involved in and have opportunities to influence all assessment procedures involving their child;
- Teachers use assessment for learning as a means of improving learning opportunities for all pupils. This involves setting goals/targets for and with pupils (in relation to effective teaching strategies for a specific pupil) as well as for themselves. It also involves providing feedback on learning to pupils in a way that meets their needs and supports their learning;
- Schools implement an assessment plan that describes the purposes and use, roles and responsibilities for assessment as well as a clear statement on how assessment is used to support the diverse needs of all pupils;
- Multi-disciplinary assessment teams – no matter what their professional composition or team membership – work to support inclusion and teaching and learning processes for all pupils;
- Assessment policies and procedures support and enhance the successful inclusion and participation of all pupils including pupils who are vulnerable to exclusion, particularly pupils with special educational needs;
- Assessment legislation should promote the effective implementation of inclusive assessment at all times.

The Attributes of Formative Assessment: A Literature Review⁴⁸

Excerpt (see extended version in Annex)

According to research findings, the formative assessment process would focus on the following elements:

- Learning progressions. Learning progressions identify the enabling knowledge and skills students need to reach the learning goal as well as provide a map of future learning opportunities.
- Learning goals. Perhaps the first step of the whole assessment for learning process is the establishment of student learning goals (Wiggins and McTighe, 2000).
- Criteria for success. Success criteria are the indications that teachers and students use as checks on learning.
- Identifying the gap.
- Eliciting evidence of learning. The various strategies for evidence gathering can be categorized into the following three broad types: curriculum-embedded (or systematic), planned, and on-the-fly (or spontaneous) (Heritage, 2007).

⁴⁸ Extracted from Grant Wood Area Education Agency (Iowa). "Assessment for Learning (Formative Assessment)". Accessible at: http://www.gwaea.org/iowacorecurriculum/docs/AssessmentForLearning_LitReviewFinal.pdf.

- Teacher assessment. Accordingly, “a critical component of quality formative assessment is teachers’ use of the evidence obtained from students’ performance on assessment tasks to adjust instruction and to guide students in adjusting their learning strategies” (DeMeester and Jones, 2009).
- Teacher feedback. Appropriate descriptive feedback should provide students with some combination of verification (linked to the learning goal and criteria for success) with elaboration (e.g., written communication) (Shute, 2008).
- Student involvement through peer assessment (students serve as learning resources for one another), self-assessment (students monitor and regulate their learning), and a learning culture (where teachers, students and parents view assessment as a source of insight rather than a form or meting out rewards and punishments) (Shepard, 2000).

Three Anchor Points of Formative Assessment Continuum⁴⁹

Excerpt (see extended version in Annex)

Three anchor points could be described on the formative assessment continuum:

- On-the-fly formative assessment arises when a “teachable moment” unexpectedly occurs, for example, when a teacher circulating and listening to the conversation among students in small groups, recognizes students’ grasp of concepts, and challenges students with other learning materials.
- Planned-for-interaction formative assessment is deliberate. A teacher plans for and crafts ways to find the gap between what students know and what they need to know.
- Embedded-in-the-curriculum formative assessment comes “ready-to-use”; teachers or curriculum developers place formal assessments ahead of time in the ongoing curriculum to create goal-directed “teachable moments.” Embedded assessments inform the teacher about what students currently know, and what they still need to learn (i.e., “the gap”) so that teachers can provide timely feedback.

Often, a combination of different methods can be used together for formative assessment.

-
2. Review and refine your group’s worksheet from Task 1, drawing upon the information provided in the readings. Identify and highlight the most realistic changes for student assessment in your context.

<p>PRODUCT: Reviewed worksheet with the most realistic changes highlighted.</p>
--

⁴⁹ Extracted from Shavelson, R. J. et al. (2008). “On the Impact of Curriculum-Embedded Formative Assessment on Learning: A Collaboration between Curriculum and Assessment Developers”, *Applied Measurement in Education*, 21:4, pp. 295-314.

TASK 3 Individual reflection and group discussion on student assessment reform

1. Read the case studies [Student assessment in the Finnish basic education, Moving from assessment of learning to assessment for learning in Hong Kong](#) and [Student assessment reform with some inclusive features in Romania](#).

Student assessment in the Finnish basic education⁵⁰

Excerpt (see extended version in Annex)

In the Finnish basic education, teachers and students are not burdened with national testing. Once every 5 years, each school is included in a national sample assessment of learning outcomes in one subject. Teachers receive information on their own school's results, but those results are not published, and schools are not compared with each other.

During basic education, students are not streamed in any way, and they usually study in heterogeneous groups. The same goals are set for all students, but the Basic Education Act and the core curriculum require schools to consider students' individual needs and learning styles using diverse pedagogical means, and by adapting study contents to support all students' learning processes. Teachers are expected to identify all students' strengths and needs, in order to provide personal study plans. In recent years, great emphasis has been placed on creating diverse learning environments, as well as an encouraging and interactive atmosphere.

Early intervention in basic education means that teachers can tackle problems in learning and development, as soon as they detect them. Thus students rarely repeat grades, beyond the roughly 2% who repeat either the first or second grade; and the dropout rate is below 0.5%. A wide array of support measures are systematically used, helping everyone to complete basic education.

Moving from assessment of learning to assessment for learning in Hong Kong⁵¹

Excerpt (see extended version in Annex)

Traditionally, schools in Hong Kong have adopted tests and examinations as the major forms of assessment in the education system. These forms of assessment carry summative and high-stake characteristics that not only can generate pressure to the students but also jeopardize teaching and learning.

After 1997, when the sovereignty of Hong Kong was returned to the government of the People's Republic of China from the U.K., a large-scale education reform coupled with curriculum reform initiated and was expected to be implemented over a 10-year span. One of the major issues of the reforms relating to assessment was reducing public examinations and unnecessary excessive tests and examinations existing in both secondary and primary sectors. Assessment was defined as "the practice of collecting evidence of student learning in terms of knowledge, skills, values and attitudes through observation of student behaviour when carrying out tasks, tests, examinations, etc." (Curriculum Development Council, 2002, p.2) It has been highlighted that both the processes (e.g. inquiring, independent learning, use of generic skills, reflections) and the products of learning (e.g.

⁵⁰ Extracted from Halinen, I. and Järvinen, R. (2008). Towards inclusive education: the case of Finland, Prospects, vol. 38, p. 89 and 91.

⁵¹ Extracted from Chan, J. K. et al. (2006). Assessment Policy in Hong Kong: Implementation Issues for New Forms of Assessment, Paper presented at the 32nd Annual Conference of the International Association for Educational Assessment: Assessment in an Era of Rapid Change: Innovations and Best Practices, Singapore, 21-26 May 2006.

knowledge/concepts, problem-solving capabilities) should be assessed by the methods most suited to them (e.g. oral tests for oral communication, discussion for collaboration, presentation/performance for creativity, tests and examinations for knowledge).

For the government policy to be implemented in the school contexts, it is suggested that:

- teachers should be provided adequate knowledge of assessment, as well as opportunities to participate in discussions about how to implement the reform;
- adaptations could be allowed at both the system and school levels; and
- assessment of learning and assessment for learning could work in synergy.

Student assessment reform with some inclusive features in Romania⁵²

Excerpt (see extended version in Annex)

Since the collapse of the Ceaușescu regime, Romania has made significant progress in its broad, structural reform programme. The Constitution, endorsed in a national referendum of 1991, laid the foundations for a legal framework facilitating political pluralism and private initiatives.

At primary level, the Education Reform brought the introduction, in 1998/9, of a new assessment system for Grades I–IV. This replaced the traditional method of using ‘marks’ from 1–10 with criteria-related four general levels of performance (considered to be more informative and comprehensible): very good, good, sufficient, insufficient. These categories are used by each teacher, to make judgments about a student’s level of achievement and to communicate this to parents through the school’s reporting system. There is no formal examination at the end of primary education and transition to Grade V is usually automatic. There is, however, an increasing tendency for ‘good’ urban schools to organize selection tests, especially in foreign languages and ICT.

In the coming years, assessment at the secondary level is likely to change to reflect similar trends, including an increase in the weighting of internal assessment and less emphasis on external examinations; greater diversification in the assessment methods used; and stronger linkages between assessment standards and the competencies defined in National Standards. In particular, there is a move towards more comprehensive records of achievement for students. The challenge now is to help schools and teachers use such information to improve teaching and learning. In addition, educational policy-makers need to be convinced of the merits of using assessment-related data to guide strategic thinking.

-
2. Using the worksheet [Review of Case Studies](#), each participant should record the key lessons that can be extracted from the three case studies. Assuming the information is available, please comment on the following:
 - What were the initial challenges student assessment faced in each context?
 - What new visions, structures and practices of student assessment were adopted in each case?

⁵² Extracted from Bethell, G. and Mihail, R. (2005). ‘Profiles of educational assessment systems worldwide’.

- What stakeholders were considered for involvement in the process, and what were their roles?
- What new challenges emerged in each case?

Worksheet: Review of case studies

	Initial issues	New assessment visions	New assessment structures	New assessment practices	Roles of stakeholders	New challenges
Finland						
Hong Kong						
Romania						

3. In small groups, preferably organized by country, answer the following:

- In the three case studies, what were the new visions, structures and practices of student assessment that may be conducive to improving the inclusiveness and quality of education? Why are they advantageous?
- To what extent are these new aspects of student assessment relevant to your context? Why, or why not?

Each group presents their findings to the plenary, and the facilitator provides a summary of the entire discussion.

PRODUCT: Completed worksheets for each group member and group report.

TASK 4 Individual reading and work in small groups on summative assessment

1. Read the case study [Summative Assessment for Inclusion? The Chilean national learning outcome assessment system \(SIMCE\)](#).

Summative Assessment for Inclusion?

The Chilean national learning outcome assessment system (SIMCE)⁵³

Excerpt (see extended version in Annex)

The Chilean national learning outcome assessment system (Sistema de Medición de Calidad de la Educación, SIMCE) has carried out census-based assessments since 1988 and publishes the results at both the national and school levels. Its aim is to improve the quality and equity of education by providing data about learning outcomes at the national and school levels, in order to:

- Monitor and inform decision-making in educational policy.
- Provide feedback to schools in order to improve teaching practices.
- Foment accountability and parental and school community involvement.

For more than 20 years, individual data on students' performance has been protected to avoid exclusion or academic selection of pupils. In 2009, new laws were passed allowing individual reports of test results to be given to families, while protecting students' privacy. These laws also restrict the selection of primary students on the basis of academic potential. On the other hand, SIMCE data have provided and publicised evidence about the close relationship between socio-economic inequities and learning gaps. This information has been the basis of policies focused on equity. SIMCE data are also used by a merit-based teacher reward system.

SIMCE is acknowledged within Latin America as an outstanding assessment system, especially with respect to the dissemination and extensive use of data in designing, implementing and evaluating education policies. It is necessary to achieve a balance between the demands for accountability and the creation of capacities to face such demands. Achieving this balance remains a challenge for Chile.

-
2. In a small group, preferably organized by country, reflect upon the following:
 - What are the main features of the Chilean SIMCE system presented in the case study?
 - In your opinion, what opportunities for inclusion, or risks of exclusion would each of these features imply?
 - What are the current practices in your country with respect to each of these features?
 - What lessons can your country borrow from the Chilean case?
 3. Each group reports back to the plenary on the discussions.

PRODUCT: Group report on each point of discussion.

⁵³ Extracted from Meckes, L. and Carrasco, R. (2010), Two decades of SIMCE: an overview of the National Assessment System in Chile, *Assessment in Education: Principles, Policy and Practice*, vol. 17, no. 2, pp. 233-248.

TASK 5 Individual activity or work in small groups on the features of student assessment and a sample assessment plan

1. Read the case studies and documents [Assessment in the Sudbury model of democratic education schools](#); [Effective self-assessment: purposes, process and value](#); [Using rubrics for authentic assessment at Jurong Secondary School in Singapore](#) and [Developing and using quality rubrics: some hints](#).
-

Assessment in the Sudbury model of democratic education schools⁵⁴

Excerpt (see extended version in Annex)

The Sudbury model of democratic education schools in the U.S. do not perform and do not offer assessments, evaluations, transcripts, or recommendations, asserting that they do not rate people, and that school is not a judge. Comparing students to each other, or to some standard that has been set is considered a violation of the student's right to privacy and to self-determination. Students decide for themselves how to measure their progress as self-starting learners in a process of self-evaluation. The no-grading and no-rating policy helps to create an atmosphere free of competition among students or battles for adult approval, and encourages a positive cooperative environment amongst the student body⁵⁵. The final stage of a Sudbury education is the graduation thesis. Each student writes on the topic of how they have prepared themselves for adulthood and entering the community at large. This thesis is submitted to the Assembly, who reviews it. The final stage of the thesis process is an oral defence given by the student in which they open the floor for questions, challenges and comments from all Assembly members. At the end, the Assembly votes by secret ballot on whether or not to award a diploma.

Sudbury Valley School has published two studies of their alumni over the past forty years. They have learned, among other things, that about 80% of their students have graduated from college⁵⁶, and that they have gone on to become successful in many areas of life⁵⁷. Anecdotally, other Sudbury schools seem to have similar results.⁵⁸

Effective self-assessment: purposes, process and value⁵⁹

Excerpt (see extended version in Annex)

The primary purposes of engaging students in careful self-assessment are to boost learning and achievement, and to promote academic self-regulation, or the tendency to monitor and manage one's own learning (Pintrich, 2000; Zimmerman and Schunk, 2004). As a result of self-assessment, both self-regulation and achievement can increase (Schunk, 2003).

⁵⁴ Extracted from Wikipedia. Democratic education. Available at: http://en.wikipedia.org/wiki/Democratic_education.

⁵⁵ Greenberg, D. (1987). Evaluation in Free at Last – The Sudbury Valley School. Sudbury Valley School Press.

⁵⁶ Greenberg, D; Sadofsky, M (1992). Legacy of Trust. The Sudbury Valley School Press.

⁵⁷ Greenberg, D.; Sadofsky, M.; Lempka, J. (2005). The Pursuit of Happiness: The Lives of Sudbury Valley Alumni. United States: Sudbury Valley School Press.

⁵⁸ Greenberg Greenberg, D. (1996). Outcomes. The Sudbury Valley School Press Newsletter.

⁵⁹ Extracted from Andrade, H. and Valtcheva, A. (2008). Promoting Learning and Achievement Through Self-Assessment, Theory Into Practice, 48: 1, pp. 12-19.

In order for effective self-assessment to occur, students need (Goodrich, 1996):

- awareness of the value of self-assessment,
- access to clear criteria on which to base the assessment,
- a specific task or performance to assess,
- models of self-assessment,
- direct instruction and assistance in self-assessment,
- practice,
- cues regarding when it is appropriate to self-assess, and
- opportunities to revise and improve the task or performance.

Access to clear criteria on which to base self-assessment can be met by introducing a rubric. A rubric is usually a one- or two-page document that lists criteria and describes varying levels of quality, from excellent to poor, for a specific assignment (Andrade, 2000; Arter and Chappuis, 2007; Goodrich, 1997).

Among the conditions under which self-assessment is likely to meet with success, we recommend the following:

- provide sufficient time for revision after self-assessment, and
- do not turn self-assessment into self-evaluation by counting it toward a grade.

Using rubrics for authentic assessment at Jurong Secondary School in Singapore⁶⁰

Excerpt (see extended version in Annex)

At Jurong Secondary School [in Singapore], Project Work is an integral part of the school curriculum and is offered to all secondary 2 and 3 students. Students are presented with problems in real world contexts that affect them in their immediate environment and community. Through solving the project tasks, students raise their level of social awareness and learn important values.

Students are assessed in Project Work using a set of tailored scoring rubrics. The scoring rubrics give emphasis on the process of doing the project and not merely the product. The students are assessed on learning outcomes such as information processing, Habits of Mind, complex thinking, collaboration and team dynamics, declarative knowledge and effective communication.

Teachers concur that "Project Work is an invaluable learning experience through which students can integrate cross-disciplinary knowledge, learn to work in teams, and demonstrate creativity and imagination in coming up with the solutions". Students also reflected that "learning is fun" and coming up with solutions to the problem tasks gave them a sense of satisfaction.

⁶⁰ Extracted from Hoe, W. M. and Tong, I. (2006). Authentic Assessment through Innovative Project Work in Jurong Secondary School.

Developing and using quality rubrics: some hints⁶¹

Excerpt (see extended version in Annex)

In order to develop good rubrics, we need to:

- Have clear learning targets/criteria/expectations (e.g. knowledge, reasoning, performance skill, product, etc.);
- Use rubrics that reflect these targets;
- Match rubric design to the purpose of using the rubric.

Rubrics for instructional purposes are most effective when they:

- reflect the intended targets of learning;
- are organised in a way that facilitates understanding of the important dimensions of the learning target; and
- provide descriptive detail on various levels of performance in terms of quality of proficiency.

Revisions based on input from both students and instructor peers will result in higher quality rubrics. In assessment *for* learning, rubrics are used to provide feedback to improve performance. It is recommended that converting rubric scores to grades should be avoided, especially on individual pieces of work. In assessment *of* learning, rubrics are used to report level of proficiency or a grade. Converting rubric scores to grades requires developing a table that relates a grade to a logical range of rubric scores⁶², instead of merely calculating the percentage of points received.

-
2. Record the important features of the types of student assessment that are presented in the readings.
 3. What features are relevant to your context? Use these context-relevant features to draft a sample assessment plan *for* learning to guide teachers, specifying the:
 - Intended class/grade level(s) and learning area(s);
 - Objectives and expected outcomes;
 - Plan structure, including components, criteria and rubrics (if any), modalities and dynamics (e.g. teacher-student interactions), frequency, etc.;
 - How the plan relates to other types of assessment; and
 - How the plan can be used to inform and enhance student learning.

Each group or individual should present their activity outputs to the plenary for discussion of the findings.

PRODUCT: Group or individual report on the activity outputs.

⁶¹ Extracted from Arter, J. (2009). Creating and Recognising Quality Rubrics. Pearson Assessment Training Institute (U.S.). Help for Instructors: Rubrics. The University of Arizona (U.S.). Available at: <http://help.d2l.arizona.edu/instructors/rubrics>. iRubric Gallery. RCampus. Available at: <http://www.rcampus.com/rubricshellc.cfm?mode=gallery&sms=publicrub>.

⁶² See how this can be carried out in Arter, J. (2009). Creating and Recognising Quality Rubrics. Pearson Assessment Training Institute (U.S.). Available at: http://ati.pearson.com/downloads/car_studyguide_forprinting_final.pdf.

ACTIVITY 3

What are the basic conditions and competencies needed to make student assessment more inclusive?

Objectives

1. Explore the basic competencies (knowledge, skills, values, attitudes, etc.) that are required to assess student learning in light of inclusive education.

Background information and discussion

The learning needs of today's students are shifting from simply learning facts and basic skills to acquiring more complex skills in higher-order thinking and problem solving. Learners must be able to select, organize, interpret and use information they receive in relation to the knowledge they already possess. They must also be able to transfer that information to new contexts to solve complex problems, so that understanding is complete. This implies affective and metacognitive skills, the capacity to work collaboratively and across disciplines, as well as learning through tasks or situations that they may encounter in adult life.

Recent school improvement initiatives involving assessment require teachers to effectively perform classroom-level assessment. However, such assessment expertise may not have been offered through pre-service or in-service trainings. An assessment-literate educator⁶³ should be aware:

- that assessment can serve a variety of *users* and fulfill *purposes* of both supporting and verifying learning;
- that quality assessment arises from *clear achievement targets* and are designed and constructed to satisfy specific *assessment quality control* criteria;
- assessment results must be *delivered into the hands of the intended user* in a timely and comprehensible form;
- assessment can no longer be viewed solely as something adults do to students; and
- how to *engage students in productive self-assessments* that support their learning success.

The table below illustrates key shifts in thought and practice that are features of classroom assessment competency:⁶⁴

⁶³ Chappuis, J., Stiggins, R. J., Arter, A. (2012), *Classroom Assessment for Student Learning: Doing It Right - Using It Well*. Pearson.

⁶⁴ *Ibid.*

Table 3.1. Features of Classroom Assessment Competency

Classroom assessment: key shifts	
From	To
Classroom tests disconnected from the focus of instruction	Classroom tests that reflect the written and taught curriculum
Assessments using only select response formats	Assessment methods intentionally selected to reflect specific kinds of learning targets
“Mystery” assessments, where students don’t know in advance what they are accountable for learning	Transparency in assessments, where students know in advance what they will be held accountable for learning
All assessments and assignments, including practice, “count” towards the grade	Some assessments and assignments “count” towards the grade; others are for practice or other formative uses
Students are passive participants in the assessment process	Students are active users of assessment as a learning experience
Students are not aware of what they are good at and what they need to work on until after grades are distributed	Students are able to identify their strengths and areas for further study during learning

Source: Chappuis, J., Stiggins, R. J., Arter, A. (2012), *Classroom Assessment for Student Learning: Doing It Right - Using It Well*. Pearson.

There are five keys to quality classroom assessment,⁶⁵ each requiring a number of competencies from teachers:

1. *Clear purposes*: assessment instruments, processes and results serve *clear and appropriate purposes*. The teacher is expected to:
 - identify the key users of classroom assessment information and discern their informational needs; and
 - understand formative and summative assessment uses and know when to apply each.
2. *Clear targets*: assessments reflect clear student *learning targets*. The teacher is expected to:
 - know how to identify learning targets;
 - know how to transform broad statements of content standards into classroom-level learning targets;
 - begin instructional planning with clear learning targets; and
 - translate learning targets into a student-friendly language.

⁶⁵ *Ibid.*

3. *Sound design*: learning targets are translated into assessments that *accurately measure* student achievement. The teacher is expected to:
 - design assessments to serve intended formative and summative purposes;
 - select assessment methods to match intended learning targets;
 - understand and apply principles of sampling learning appropriately;
 - write and/or select assessment items, tasks, scoring guides, and rubrics that meet standards of quality; and
 - identify and avoid sources of bias that distort results.
4. *Effective communication*: assessment results contribute to increase student achievement through *effective management and communication*. The teacher is expected to:
 - use assessment information to plan instruction;
 - offer effective feedback to students during learning;
 - record formative and summative assessment information accurately; and
 - combine and summarize information appropriately to accurately reflect current levels of student learning.
5. *Student involvement*: students are *active participants* in the assessment process through self-assessment, goal setting, tracking, reflecting on and sharing their learning. The teacher is expected to:
 - identify students as important users of assessment information;
 - share learning targets and standards of quality with students;
 - design assessments so students can self-assess and set goals on the basis of results; and
 - involve students in tracking, reflection and sharing their own learning progress.

In other words, for assessment to have validity for a student's learning and progress, the student should arguably be at the centre of the process. For the student's judgments to become reliable and valid, they must have:⁶⁶

- well-understood success criteria;
- "other" actors (peers, adults) involved in supporting the process of making judgments; and
- belief in the value of being part of the assessment and a commitment to acting on the inferences drawn from assessments.

⁶⁶ New Zealand Educational Institute (NZEI Te Riu Roa) and Lester Flockton, (2009). *Key Competencies. The New Zealand Curriculum*.

Focus of the activity

Classroom assessment competency; five keys to quality classroom competency; approaches to “mistake”; differentiated / non-differentiated teaching; teacher training for assessment competencies; elements of formative assessment; authentic assessment; good feedback practices

TASK 1 **Work as individual or in groups on classroom-level assessment**

1. Read the comparative study [What does it mean to make a mistake?](#)

What does it mean to make a mistake?⁶⁷

Excerpt (see extended version in Annex)

French and English views of what constitutes a ‘mistake’ reflect different approaches to teaching and the underlying beliefs in terms of children’s needs. In the emergent writing approach [in England], making ‘mistakes’ is a central and constructive part of the learning process. Only gradually are children expected to adopt standard spelling systematically. In the French system, children are intended to spell words “correctly” as soon as they start writing them so that they do not memorize them with mistakes.

Differentiated tasks are a key characteristic of English teaching, while owing to egalitarian principles, French teachers preferred to provide the same tasks for all. Some children will inevitably face greater difficulties than others. Already at the age of four, some French pupils were incapable of doing any of the tasks required of them, and were threatened with “repeating” the year. “Teachers in France did not believe that knowing the grim truths about oneself hindered a child’s achievement” (Anderson-Levitt 2002). They repeated that it was normal to make mistakes, that it was a healthy part of the learning process.

Children’s self-esteem is taken into account in both countries but teachers intervene at different times, and in culturally shaped ways. The English approach can be said to be preventive, making the environment child-friendly so as to avoid accidents. The French approach on the other hand is curative. The child has to jump, and the teacher helps those who fall to pick themselves up and to clear the obstacle the next time they try.

-
2. Complete the worksheet [Comparing the characteristics of classroom-level assessment in English and French schools](#).

⁶⁷ Extracted from Raveaud, M. (2004). Assessment in French and English infant schools: Assessing the work, the child or the culture? *Assessment in Education*, Vol. 11, No. 2, pp. 193-211.

Worksheet: Comparing the characteristics of classroom-level assessment in English and French schools

	English school	French school
Approaches to the notion of “mistake”	• • • •	• • • •
Differentiated / non-differentiated teaching	• • • •	• • • •
Teacher feedback	• • • •	• • • •
Approaches to addressing difficulty in learning	• • • •	• • • •

3. In light of your own experiences, write a personal reflection (no more than one page) on the questions below and provide illustrative examples:
- What do you think are the strengths and weaknesses of the classroom-level assessment approaches in each case?
 - In each case, what approaches and practices are likely to be conducive to inclusive education? Why? What competencies do they require from teachers?

PRODUCT: Comparative table and personal reflection with examples.

TASK 2 **Work as individual on basic conditions for formative assessment**

1. Read the documents [Changing the practice of Oral English Assessment in Malaysia](#), [Integrating assessment competences into teacher training in Belgium](#) and [The Elements of Formative Assessment: OECD Case Study Findings](#).

Changing the practice of Oral English Assessment in Malaysia⁶⁸

Excerpt (see extended version in Annex)

In school-based Oral English Assessment (OEA), students are given feedback on their assessment. Then they are given the chance to redo the assessment until both the teacher and student are satisfied with the grade. Hence, both parties are fully involved in the teaching and learning process. Teachers' competency and commitment are key ingredients in the success of any assessment endeavours.

However, teachers who participated in the study differed in the interpretation of the objectives and how they can be achieved at the school level. Such differences could be due to insufficient guidelines provided by the education authorities. Teachers' difference in implementing the OEA could affect consistency in the grading of student performance, thus having a direct implication on the validity and reliability of the assessment. More standardisation was needed in this aspect.

Integrating assessment competences into teacher training in Belgium⁶⁹

Excerpt (see extended version in Annex)

In the Flemish Community of Belgium, there are no central agencies that evaluate the performances of all the students. Assessing whether a student meets the attainment targets established by the Flemish government, falls under the jurisdiction of individual schools. The freedom of education granted to schools implies that student assessment practices may differ considerably across schools.

The basic competences for both primary and secondary Flemish school teachers include the expectation that a recently graduated teacher “*is able to choose and draw up tasks and assignments using different evaluation methods, in respect with the objectives. In the light of ‘broad student assessment’, the teacher is able to use observation instruments in co-operation with colleagues and knows how to give concrete and personal feedback to parents. Based on the results of the evaluation prospective teachers can adjust their educational handling and implement methods for differentiation between students.*”

In spite of the efforts from the Institutes for Initial Teacher Education, the Inspectorate found in 2007 that starting teachers experienced the drafting of an adaptive or alternative assessment system as one of their most difficult tasks. Courses on assessment are an important part of the wide range of in-service training possibilities. The schools and teachers are free to determine the content of their professionalization courses. In many schools drafting tests and other tools for student assessment happens in consultation with colleagues, enabling the starting teacher to become gradually competent in assessing students' progress and performances.

⁶⁸ Extracted from Malakolunthu, S. and Hoon, S. K. (2010). Teacher perspective of school-based assessment in a secondary school in Kuala Lumpur, *Procedia Social and Behavioral Sciences*, vol. 9, pp. 1170-1176.

⁶⁹ Extracted from Ministry of Education and Training Belgium and University of Antwerp Edubron Research Group (2010). Country Background Report for the Flemish Community of Belgium, *OECD Review on Evaluation and Assessment Frameworks for Improving School Outcomes*.

The Elements of Formative Assessment: OECD Case Study Findings⁷⁰

Excerpt (see extended version in Annex)

Recent case studies in OECD countries have shown the following key elements of formative assessment:

1. Establishment of a classroom culture that encourages interaction and the use of assessment tools.
2. Establishment of learning goals, and tracking of individual student progress toward those goals.
3. Use of varied instruction methods to meet diverse student needs.
4. Use of varied approaches to assessing student understanding.
5. Feedback on student performance and adaptation of instruction to meet identified needs.
6. Active involvement of students in the learning process.

In all the case studies, while teachers may have placed different emphases on the various elements, they had incorporated each of the six elements into regular practice.

Summative results, when embedded in the wider teaching and learning environment, are more likely to be used formatively. They also help to lower the stress of tests, which can have a negative impact on the self-esteem of lower achieving students.

2. Write a short paper on the following:

- What new points did you learn from these readings?
- What are the basic conditions that must be present to effectively carry out formative assessment? Who are the key actors that should be involved? What competencies would be required from each actor?
- Are any of the six key elements of formative assessment also being implemented in your country or context? Which ones? Through which concrete practices? What are the effects? Are there any challenges? Please illustrate your answer with examples. If applicable, interview relevant practitioners around you.
- Are any of the six key elements insufficiently understood/implemented in your context? Which ones? In what concrete ways can they be further developed?

PRODUCT: Short paper.

⁷⁰ Extracted from OECD (2008). Assessment for Learning. Formative Assessment, OECD/CERI International Conference "Learning in the 21st Century: Research, Innovation and Policy".

TASK 3 Work as individual or in small groups on formative assessment strategies

1. Read the resources [Authentic assessment: assessment for learning](#) and [Good feedback practices](#). Among the assessment strategies and practices presented in the extracts, list those that are new to you.

Authentic assessment: assessment for learning⁷¹

Excerpt (see extended version in Annex)

The move away from behavioural to cognitive theory has shifted the assessment emphasis:

From	To
Responding	Constructing knowledge
Discrete isolated skills	Integrated and cross-disciplinary activities
Accumulation of isolated facts and skills	Application and use of knowledge
Paper-and-pencil tests	Authentic assessments on multiple-choice, contextualised problems that are relevant and meaningful, emphasize higher-order thinking, do not have a single correct answer, have public standards known in advance, and are not timed
End-of-term tests/exams	Samples of work over time which provide a basis for assessment by teachers, students and parents
Individual assessment tasks only	Collaborative group-process tasks which mirror real-life learning
Assessment focused on reproducing information	Authentic assessment which measures student understanding, as well as knowledge

Key characteristics of assessment for learning are:

- using effective questioning techniques;
- using marking and feedback strategies;
- sharing learning goals; and
- peer- and self-assessment.

⁷¹ Extracted from Collins, R. (2013). Authentic assessment: assessment for learning, Curriculum Leadership Journal, Vol. 11, No. 7. Education Services Australia.

According to Dietel, R.J., et al. (1991), excellence in assessment requires the curriculum and instruction to:

- identify core concepts;
- involve a comprehensive repertoire of learning and assessment strategies;
- be committed to collaborative teaching and learning;
- build new information on student strengths;
- use authentic tasks that are related to everyday problems;
- provide opportunities for students to engage in learning and assessment out of school with community members;
- include homework that is challenging enough to be interesting but not too difficult so as to cause failure;
- enable students to develop an understanding of the purposes and methods of assessment; and
- allow opportunities for students to decide performance criteria and method.

A 2005 OECD report states that the achievement gains associated with formative assessment are “among the largest ever reported for educational interventions”. The report suggests that teachers can still work toward standards, providing they identify the factors behind the variation in student achievements and adapt their teaching to meet individual needs.

Good Feedback Practices⁷²

Excerpt (see extended version in Annex)

Nicol and Macfarlane-Dick (2006) suggest seven principles of good feedback practice:

1. It helps clarify what good performance is (goals, criteria, expected standards);
2. It facilitates the development of self-assessment (reflection) in learning;
3. It delivers high quality information to students about their learning;
4. It encourages teacher and peer dialogue around learning;
5. It encourages positive motivational beliefs and self-esteem;
6. It provides opportunities to close the gap between current and desired performance;
7. It provides information to teachers that can be used to help shape teaching.

⁷² Extracted from Nicol, D. J. and Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: a model and seven principles of good feedback practice, *Studies in Higher Education*, vol. 31, no. 2, pp. 199-218.

Strategies that increase the quality of teacher feedback include:

- prioritising areas for improvement;
 - making the feedback timely, corrective, succinct and accessible online;
 - allowing students to rewrite selected pieces of work;
 - providing students with “action points”; etc.
-

2. In small groups, discuss the points you listed:

- Why are these strategies and practices interesting to you;
- Of these strategies and practices, which are the ones that can be adopted to your context; and
- What are the challenges in adapting some of them to your context?

3. While considering the readings and discussions, describe the assessment (features, methods, processes, etc.) that you deal with in your daily work and life. Provide a proposal on how to improve such assessment, indicating the competencies such improvements would require from teachers.

PRODUCT: List of new strategies and practices, as well as the analysis and proposal regarding assessment in your context.

ACTIVITY 4

Relationships between international, sub-regional, national, and school-level student assessments, and their connection to the curriculum

Objectives

1. Explore the relationship between student assessment at different levels – international (PISA, etc.), regional, national and school levels, as well as various countries' approaches to these different assessments.

Background information and discussion

Many assessment system reforms have occurred internationally, and at least two broad trends can be identified among them:

1. There has been an increase in national and international assessments (PISA, etc.) using summative methods to monitor educational progress; and
2. Many countries have sought to incorporate more formative assessment methods to improve student learning and/or equity in assessment.

As the OECD recently reported,⁷³ summative assessments remain the most prevalent in schools. They are used for several purposes, including measuring what students have learned at the end of a unit, promoting students, ensuring they have met the required standards towards graduation or entry to certain occupations, or as a method for selecting students for admission to higher education. Ministries or departments of education may use summative assessments and evaluations as a way to hold publicly-funded schools accountable for providing quality education. Increasingly, international summative assessments such as the OECD's Programme for International Student Assessment (PISA) have been critical for comparing the developments of national education systems.

In classrooms, formative assessment refers to frequent, interactive assessments of student progress and understanding to identify learning needs and adjust teaching appropriately. But there are major barriers to wider practice, including perceived tensions facing high-visibility summative tests, used to hold schools accountable for student achievement, as well as a lack of cohesion between systemic, school and classroom approaches to assessment and evaluation.

In environments where summative tests prevail, teachers often feel compelled to "teach to the test," and students are encouraged to meet performance goals at the expense of learning goals (that is, to understand and master new knowledge). Many teachers perceive these external assessments as conflicting with the practice of formative assessment. Poorly designed external tests, media league tables which use a narrow set of data to compare performance across schools, and lack of connection between tests and curriculum can also be detrimental to the quality of education.

In all cases, the use of data to inform the planning of classroom teaching and learning, or at the national level to inform and improve policies, might be considered secondary levels of formative assessment.

⁷³ OECD (2008). *Assessment for Learning. Formative Assessment*, OECD/CERI International Conference "Learning in the 21st Century: Research, Innovation and Policy".

Focus of the activity

National examinations; functions of examinations; international assessments, national assessments; regional assessments; participation of developing countries; culture of evaluation in OECD countries

TASK 1 Work as individual on different levels of assessment

1. Read the document [Measuring What Students Learn and Master](#).

Measuring What Students Learn and Master⁷⁴

Excerpt (see extended version in Annex)

Examinations can be used to distribute opportunities for further education and as a condition for accessing labour markets. In most countries secondary education school-leaving examinations are high-stakes and highly politicized. Examinations are a powerful tool for influencing and shaping the secondary school curriculum.

Policy makers and the public have become interested in international comparative studies of student achievement because of concern about the competitiveness of countries' current and future labour force in the global market and the extent to which citizens can adapt to rapidly changing societies. Most popular among these studies are the Programme on International Student Assessment (PISA) and the Trends in International Mathematics and Science Study (TIMSS). International assessments give policy makers a means of further examining their education systems.

National assessments assess students' ability, and mainly focus on the system level. In the 1990s, the use of national assessments as tools for school and teacher accountability became prominent, especially in countries such as Australia, Canada, the United Kingdom, and the United States. Public examinations measure individual students' ability and are normally used for selective purposes. Large-scale summative assessments are considered unsuitable for assessing some of the key competencies needed by labour markets. Classroom-based student achievement monitoring is considered important but still does not receive enough attention.

Further reinforcing the trend from norm-referenced to criterion-referenced assessment, there are now strong signs of movement toward the incorporation of such non-examination practices as portfolios, profiles, and school records in the final evaluation of secondary school leavers. These trends symbolize the way forward in building a new culture of educational evaluation and assessment that is both technically sophisticated and socially inclusive.

Prepare a 10-slide PowerPoint presentation on the different levels of assessment (national assessment, public examination, international assessment, classroom assessment, etc.) that are portrayed and discussed in the extract.

2. Deliver your presentation to the plenary and respond to questions from the floor.

PRODUCT: PowerPoint presentation on different levels of assessment.

⁷⁴ Extracted from The World Bank (2005). Responding to the Twin Challenges: Curriculum and Assessment. In Expanding Opportunities and Building Competencies for Young People: A New Agenda for Secondary Education, pp. 94-102.

TASK 2 **Work as individual on assessments and examinations in your context**

1. Read the case study [Participation in national, regional and international learning assessments: trends among developing countries, 1960-2009](#).

Participation in national, regional and international learning assessments: trends among developing countries, 1960-2009⁷⁵

Excerpt (see extended version in Annex)

The World Bank's support for the assessment, evaluation and comparison of student learning formalised on-going trends, beginning in the 1960s and growing significantly in the 1990s, of widening country participation in comparative studies of educational achievement as well as other forms of student testing and national assessments.

Concurrently, regional learning assessments also took root, including the Latin American Laboratory for the Assessment of the Quality of Education (LLECE), the Southern African Consortium for Monitoring Education Quality (SACMEQ) and the Francophone African countries' Programme d'analyse des systèmes éducatifs (PASEC).

National learning assessments – i.e. non-standardised, context sensitive and non-comparable learning assessments – have become well entrenched in the developing world as the preferred tool of education policy makers. But the continuation of this trend will also depend on the policies of international agencies, NGOs and regional associations that support and advise countries on assessment.

If high-stakes international assessment drives out national and regional assessment, the risk is developing countries in particular will be left with a 'horse race' mentality and crises of confidence. They will end up imitating the fate of educational reform in the developed world, where endless and pernicious cycles of educational reform emerge every few years, each one dedicated to finding the most effective means of increasing test scores (Baker and LeTendre 2005). On the other hand, national assessments are becoming increasingly similar in content and format. As globalisation intensifies, the international and regional effects of assessment may become commonplace.

-
2. Write a personal note (not to exceed two pages) by answering the following questions:
 - What are the national, sub-regional and international assessments or examinations that are currently applied to students in your context?
 - What are the purposes, contents and modalities of each assessment and examination? How are their results used?
 - Interview stakeholders around you, including policy-makers, schools leaders, teachers, parents, students, employers, and communities. What are their attitudes and opinions towards these assessments and examinations?
 3. In your view, what are the statuses, functions and impacts of each assessment and examination? How could they better co-exist or evolve?

⁷⁵ Extracted from Kamens, D. H., Benavot, A. (2011) National, regional and international learning assessments: trends among developing countries, 1960-2009. In *Globalisation, Societies and Education*, vol. 9, no. 2, pp. 285-300.

PRODUCT: Personal note on assessment and examinations in your context.

TASK 3 **Work as individual or small group on cultures of evaluation**

1. Read [Strengthening cultures of evaluation: findings from OECD countries](#).

Strengthening cultures of evaluation: findings from OECD countries⁷⁶

Excerpt (see extended version in Annex)

In a culture of evaluation, teachers and school leaders use information on students to generate new knowledge on what works and why, share their knowledge with colleagues, and build their ability to address a greater range of their students' learning needs. Several OECD countries support school-based evaluation as a key component, either as the primary or only form of school-level evaluation, or as a complement to external testing, inspections and programme evaluation. School-based evaluation involves school staff directly, incorporates local knowledge, and potentially, directly shapes school improvement.

Ideally, information gathered in assessments and evaluations is used to shape strategies for improvement at each level of the education system: at the classroom level, at the school level, and at the policy level. In this way, summative information is used formatively. Teachers, school leaders and policy officials are more likely to use assessment information when assessments are well coordinated, and it is clear why and how the information is relevant to their work.

-
2. In groups organized by country, discuss the following:
 - What are the main ideas you can extract from this study?
 - In your context, what existing assessments and evaluations can be used to support student learning, school improvement and systemic improvement?
 - In your view, what conditions are needed, and what mechanisms should be established for such assessments to result in the improvement of student learning, school improvement and systemic improvement? Record the conditions and mechanisms you suggest in bullet points.
 - Each group reports their findings to the plenary, before the facilitator recaps the discussions.

PRODUCT: Group report.

⁷⁶ Extracted from OECD (2008). Assessment for Learning. Formative Assessment, OECD/CERI International Conference "Learning in the 21st Century: Research, Innovation and Policy".

ACTIVITY 5

Adopting a holistic approach to student assessment

Objectives

1. Introduce interesting tools and models for harmonising student assessment with curriculum, teacher training and other sectors of the education system and society.

Background information and discussion

Assessment practices can be strongly influenced by traditions and stakeholders' personal and collective assumptions. When assessment practices evolve, the roles of the involved stakeholders also change. Further, assessment practices may also encounter the resistance or limitations of stakeholders representing different interest groups. For example, teachers' actions may be dependent on school leaders, while school leaders may be dependent on central or local education authorities, etc. Moreover, there can be competing demands with respect to assessment — while some stakeholders advocate for making assessments more inclusive, others believe that assessment will always have some selective purpose.

“Traditional” assessment practices may put teachers and students into unbalanced power relations. Teachers can exercise power through assessment. Such power may function as a barrier to changing assessment and pedagogical practices. It is never easy to relinquish power, especially as this power might be perceived as part of teacher identity and core business.

Teachers play a central role in making assessment more inclusive. It is important to include assessment-related issues in teacher training at both pre-service and in-service levels. Further, teachers should have enough time and material on hand, as well as the freedom to practice different forms of assessment. However, even when teachers favour alternative forms of assessment, they may encounter constraints when attempting to change their practices. For example, teachers in diverse cultural contexts have different perceptions and understanding of assessment. It is important for policymakers to take such differences into account.

Students are also important stakeholders and participants in their own assessment. Assessment is of key importance for students' learning and for their lives. When assessment reforms are planned, changes need to be clearly explained, not only to teachers, but also to students. Students have their understanding and values that should also be taken into account. In reality, exam designers often lack knowledge about student perceptions and expectations. For instance, they may not know student beliefs about assessment and its purpose, the kind of feedback that students wish to receive, as well as the type of assessment students find most important.

Adopting a holistic approach to student assessment may include several dimensions:

- Providing students with multifaceted assessment so they have fairer, more humane and comprehensive awareness of their progress. An assessor can collect and use evidence or information from the learner through multiple dimensions (e.g. naturally occurring evidence, experiential learning), in relation to his or her development over time, in order to assess the learner's overall performance holistically. The assessor should take into account holistic learning outcomes such as self-conception and creativity, in addition to content-based learning outcomes such as literacy and numeracy.

- During the assessment design and implementation processes, considering assessment's relation to other areas within education (i.e. policy, planning, curriculum, teaching-learning, etc.) and beyond education, particularly the world of work.
- In assessment exercises, it is not only important to acknowledge that there are different stakeholders (educational authorities, schools, teachers, students, parents, employers, communities, international stakeholders, etc.), but also to examine various stakeholders' roles, and the effects of their roles in the processes of designing, implementing, utilising and reforming assessment.

Focus of the activity

Holistic approach to student assessment; stakeholder involvement in student assessment; comprehensive approach to student assessment; framework for effective student assessment

TASK 1 Work in small groups on supporting factors for assessment reform

1. Read the case studies [Stakeholder involvement in the Assessment is for Learning \(AifL\) Programme in Scotland](#) and [Portfolio Assessment of Cooperative Learning Groups in Small Classes: Experiences from Hong Kong](#).

Stakeholder involvement in the Assessment is for Learning (AifL) Programme in Scotland⁷⁷

Excerpt (see extended version in Annex)

In Scotland curriculum and assessment have been inextricably linked since the late 1970s. There has been a long tradition of teacher engagement in curriculum and assessment. A debate on assessment in the Scottish Parliament in 2001 set the context for the 'Assessment is for Learning' (AifL) programme: assessment should be integral to learning and teaching and be concerned to improve learning and achievement; the approach to change should be evolutionary; teachers' informed professional judgment should be at the heart of the system; and information for monitoring purposes should be streamlined into a single, integrated system.

The AifL Programme team worked with schools across Scotland. A wide range of stakeholder groups was established to build supportive, informed communities. There is evaluation evidence that this ambitious and complex project structure has led to quite significant changes in assessment practice. There are very significant numbers of teachers who are now committed to formative assessment practices in their classrooms.

The AifL programme emphasizes the idea that all participants are learners: teachers, children, young people, researchers, and policymakers. This approach is consonant with the wider drive towards enhanced local autonomy, a crucial element of which is self-evaluation as part of planning for improvement. An AifL school is described as one which actively and coherently promotes Assessment for Learning, Assessment as Learning, and Assessment of Learning.

⁷⁷ Extracted from Hayward, E. L. (2007). Curriculum, pedagogies and assessment in Scotland: the quest for social justice. *Assessment in Education: Principles, Policy and Practice*, vol. 14, no. 2, pp. 251-268.

Portfolio Assessment of Cooperative Learning Groups in Small Classes:

Experiences from Hong Kong⁷⁸

Excerpt (see extended version in Annex)

There is no unanimous agreement on the definition of portfolio assessment. It can refer to an individual collection of daily drawings, writings and other materials that provide documentation of a child's strengths (Schipper and Rossi, 1997). It can also be "a purposeful collection of student work that exhibits the learners' efforts, progress, and achievements in one or more areas. The collection must include student participation in selecting contents, the criteria for selection, the criteria for judging merit, and evidence of student self-reflection" (Paulson, Paulson, and Meyer, 1991, pp. 60-61). All kinds of portfolio assessment are characterized by their authenticity and ability to document a learner's learning over a period of time.

Lessons that can be learnt from the successful implementation of portfolio assessment in Hong Kong include:

- it is of paramount importance that adequate professional training has to be provided for teachers to master the concepts and the use of various kinds of portfolios for different purposes;
- a small class size can facilitate the implementation of portfolio assessment by allowing time for teachers to review portfolios together with individual pupils;
- cooperative learning can enhance students' ability of conducting self-assessment which enables them to record learning progression in their portfolios.

2. In small groups, draw a graph that illustrates the factors and actors that were important for the introduction of a new assessment programme in the cases of Scotland and Hong Kong, as well as the relationships and interactions between these factors and actors.

3. Each group presents the graph to the plenary.

PRODUCT: Graph demonstrating factors, actors and their interactions.

TASK 2 Work as individual or small group on a comprehensive approach to student assessment

1. Read the document [Accountability, student assessment, and the need for a comprehensive approach](#).

⁷⁸ Extracted from Chan, K. W. (2006). Portfolio Assessment of Cooperative Learning Groups in Small Classes. Paper presented at the International Association for Educational Assessment 32nd Annual Conference.

Accountability, student assessment, and the need for a comprehensive approach⁷⁹

Excerpt (see extended version in Annex)

Holding individual teachers, schools, and districts accountable for student performance continues to be a central feature of educational reform throughout the Western world. Assessment of learning (i.e., for accountability purposes) has far outpaced the desire to promote assessment for learning (i.e., for instructional purposes) (Stiggins, 2002). Standardized tests often sample a restricted range of student knowledge and skills: a number of important curriculum areas are often excluded from these measures.

A comprehensive approach to student assessment is one that: (1) expands the assessment repertoire to include sound classroom-based assessment data; (2) supplements classroom assessment data with appropriate standardized measures; (3) sets realistic targets and standards for students, teachers, and schools; (4) downplays cross-school comparisons in favor of improvement for individual students; (5) utilizes a value-added criterion to interpret student performance in context; (6) provides teachers and administrators with professional development aimed at enhancing assessment literacy; (7) monitors and reviews the assessment system regularly. These seven components were frequently noted within the broader assessment, accountability, and educational reform literature.

In many respects, a comprehensive approach to student assessment is improvement-oriented rather than results-oriented. It also takes into account the perspectives of a variety of stakeholders, including parents, teachers, administrators, policy-makers, community members, and elected officials.

2. The author proposed seven components for a comprehensive approach to student assessment. He also stated that although frequently noted within the broader literature on assessment accountability and education reform, “the decision to incorporate a component was ultimately guided by the desire to develop a robust framework that promotes assessment for accountability as well as instructional purposes.” In small groups, organized by country, discuss the following:
 - Does the notion of “comprehensive approach to student assessment” exist in your context? If yes, how is it being understood and applied?
 - Among the seven components listed in the document, when, how often, and to what extent is each component implemented in your context? What are the consequences on students, teachers, schools and the system?
 - What are the existing gaps regarding the implementation of each component?
 - Are there any plans to strengthen, or introduce the implementation of any of the seven components?
3. Each group reports their findings to the plenary, before the facilitator recaps the discussions.

PRODUCT: Group report.

⁷⁹ Extracted from Volante, L. (2005). Accountability, student assessment, and the need for a comprehensive approach, *International Electronic Journal for Leadership in Learning*, vol. 9, no. 6.

TASK 3 Work as individual on the effectiveness of student assessment systems

1. Read the document [Framework for Building an Effective Student Assessment System](#).

Framework for Building an Effective Student Assessment System⁸⁰

Excerpt (see extended version in Annex)

There are three main drivers of information quality in an assessment system (AERA, APA, and NCME, 1999; Darling-Hammond and Wentworth, 2010):

- enabling context (broader legislative or policy framework for assessment; institutional and organizational structures for assessment; availability of funding; and presence of competent assessment practitioners)
- system alignment (connection between assessment and the rest of the education system: learning goals, standards, curriculum, pre- and in-service teacher training, etc.), and
- assessment quality (psychometric quality of the instruments, processes, and procedures used for assessment [AERA, APA, and NCME, 1999]).

According to Clark (2011), assessment systems that are at the *Emerging* stage tend to have enabling contexts, as well as levels of system alignment and assessment quality, that are just taking shape. Assessment systems that are at the *Established* stage tend to have enabling contexts, as well as levels of system alignment and assessment quality, that are stable, assured, or consolidated in nature. Assessment systems that are at the *Cutting Edge* stage tend to have enabling contexts, as well as levels of system alignment and assessment quality that are highly developed in nature.

-
2. Respond to the following questions through a personal note:

- To what extent do the guidelines and principles provided in this reading apply to your context?
- Among the referenced “stages of student assessment system development” that were proposed, where would you position the different types of assessment that are practiced in your context? Why? Provide evidence.
- What is your personal view on the effectiveness of student assessment systems? Do you see any issues/ points that can be added to the guidelines the author presented?

PRODUCT: Personal note.

⁸⁰ Extracted from Clarke, M. (2011). Framework for building an effective Student Assessment System. Washington, DC: The World Bank Group.

Training Tools for Curriculum Development

Inclusive Student Assessment

“Student assessment is the process of documenting students’ acquisition and mastery of knowledge, skills and competencies in order to make informed decisions about the next steps in an educational process. This implies consideration of students’ aptitudes, attitudes, learning styles, progressions and outcomes.”

The IBE series of Training Tools for Curriculum Development is designed to support Member States with regard to strengthening leadership for education and curriculum reforms and development processes. Specifically, “Training Tools for Curriculum Development: Inclusive Student Assessment” aims to strengthen capacities in the field of student assessment under a broadened conceptualization of inclusive education and inclusive curriculum. It draws upon international comparative perspectives on student assessment that practitioners can use to further reflect upon education in their own contexts.

This Thematic Module provides comprehensive guidance to teachers, principals, inspectors, teacher trainers, educators, and curriculum developers. It consists of:

A conceptual framework – This intends to ground practitioners’ understanding of student assessment in light of inclusive education in order to promote inclusivity in assessment.

Training activities – These are structured around five key areas of student assessment. Each of the five areas has several training tasks attached to it, which place the readers in the position of a curriculum developer. Each activity requires the completion of a number of sub-tasks, related to decision-making or formulation of responses. The central challenge for the readers is to transfer and apply the generic content of the module to their country or education system circumstances.

IBE-UNESCO
C.P. 199
1211 Geneva 20
Switzerland

Tel.: +41.22.917.78.00
Fax: +41.22.917.78.01
<http://www.ibe.unesco.org>



United Nations
Educational, Scientific and
Cultural Organization



International Bureau
of Education