Assessment: Ensuring a Balanced and Principled Approach

Brian North (Eurocentres Foundation)

1. The shift in the 1990s
2. Assessment for learning / Assessment of level
3. "Constructive alignment" & "Backward design"
4. The CEFR linking planning, teaching, assessment
5. How CEFR descriptors help:
6. The place in assessment of the new descriptors

COUNCIL OF EUROPE
CONSEIL DE L'EUROPE

- **Frustration with assessment as standardised tests, pass or fail…Summative assessment is just one function**

- The Assessment Reform Group (ARG) (1989 – 2010): study of the implications of assessment policy and practice

- *A Practical Guide to Alternative Assessment* (Herman, Aeschbacher & Winters, 1992)

- *Assessment for learning: Beyond the Black Box* (ARG, 1999)

- *Beyond Testing* (Gipps 1994)

- *Inside the black box: Raising standards through classroom assessment* (Black & William, 1998)

- *Aligning teaching and assessment to curriculum objectives: Imaginative curriculum* (Biggs, 2003)

- **Growing awareness of need for teacher assessment of tasks/assignments related to learning objectives**

1. The shift in the 1990s

   Assessment for learning / Assessment of level

3. "Constructive alignment" & "Backward design"

4. The CEFR linking planning, teaching, assessment

5. How CEFR descriptors help:

6. The place in assessment of the new descriptors

Assessment for learning = formative assessment:

The process of seeking evidence and interpreting it in order to decide:

- where the learners are in their learning

- where they need to go

- and how best to get there

- Diagnostic assessment
- Continuous assessment
- Dynamic assessment

Assessment of level is proficiency assessment.

The natural **criterion** to assess proficiency is real-world, practical use of whatever is learnt.

Criterion-referenced assessment (CR) situates a learner on the continuum of ability in the subject or for the trait concerned. (Glaser, 1963, 1994a, 1994b)

CR is in contrast to more traditional norm-referenced, summative assessment (in relation to peers, previous years etc.)

1. The shift in the 1990s
2. Assessment for learning / Assessment of level
   "Constructive alignment" & "Backward design"
4. The CEFR linking planning, teaching, assessment
5. How CEFR descriptors help:
6. The place in assessment of the new descriptors

- **Constructive alignment:** The process of ensuring coherence between curriculum, teaching and assessment, usually by planning both the curriculum and assessment in relation to desired outcomes through **backward design**.

- **Backward design:** Identifying learning needs and necessary outcomes, usually in terms of 'real world' competences, and then planning the necessary enabling objectives, learning activities and other course content 'backwards' from those outcomes

Curriculum as a dynamic system: (Graves 2008)
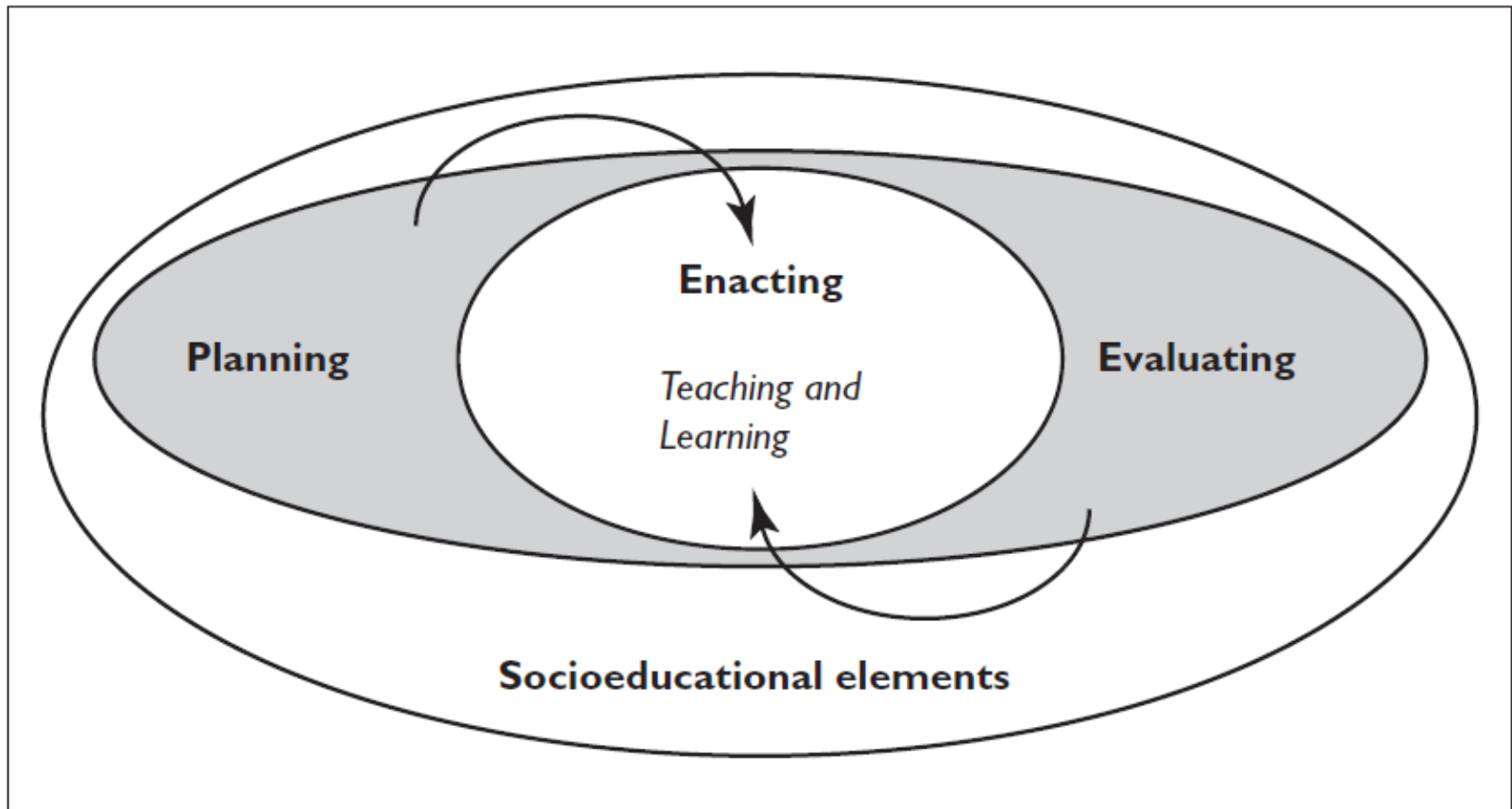
"Planning, implementation and evaluation decisions should be consistent and interdependent rather than undertaken in a lockstep or piecemeal approach." (Graves', 2008: 148)

1. The shift in the 1990s

2. Assessment for learning / Assessment of level

3. "Constructive alignment" & "Backward design"

   The CEFR linking planning, teaching, assessment

5. How CEFR descriptors help:

6. The place in assessment of the new descriptors

"One of the most important curriculum publications in the last decade is the Council of Europe's Common European Framework of Reference (Council of Europe 2001) that seeks to provide descriptions that apply to competence in all languages across all levels." (Graves, 2008: 148)

"Perhaps the most widespread example of backward design using standards in current use is the Common European Framework for Reference for Languages (CEFR) (Council of Europe, 2001)…" (Richards, 2013: 26)

A transparent, coherent curriculum

Planning:

Sign-posting with 'can do' descriptors

Teaching:

Action-oriented approach: language and activity

Assessment:

Tasks: descriptors as criteria

1. The shift in the 1990s
2. Assessment for learning / Assessment of level
3. "Constructive alignment" & "Backward design"
4. The CEFR linking planning, teaching, assessment

   How CEFR descriptors help:
6. The place in assessment of the new descriptors

- Descriptors can help to provide transparent and coherent alignment between the curriculum aims and
  a. the detailed objectives teachers use to implement them
  b. the assessment of achievement in relation to them

In this process, descriptors can be used for assessment in two complementary ways:
  - continuous assessment checklist (ELP) approach
  - module: task observation approach

- In either case, the assessment rubric can be
  - a checklist of descriptors
  - a grid with categories and 'levels' (e.g. A2+ **B1**, B1+) defined by descriptors for each category

"Continuous assessment" Checklist/ELP approach	"Module / task approach" observation approach
• simple and transparent • clear link between aims and assessment • good for reporting	• useful form of professional training in itself • focuses on specific descriptors, and learner performances • judgements are evidence-based and considered. • teachers and students enjoy the approach.

Both approaches can be used for teacher assessment, self-assessment – and in principle – peer assessment

Distinguish between:

- Descriptors of communicative activities: Ch. 4
- Descriptors of aspects of proficiency related to particular competences: Chapter 5
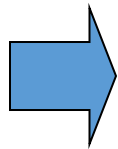
Distinguish between:

- Descriptors of communicative activities: Ch. 4
- Descriptors of aspects of proficiency related to particular competences: Chapter 5

"The former are very suitable for teacher- or self-assessment with regard to real-world tasks. Such teacher- or self-assessments are made on the basis of a detailed picture of the learner's language ability built up during the course concerned. They are attractive because they can help to focus both learners and teachers on an action-oriented approach."
(CEFR p 180)

Distinguish between:

- Descriptors of communicative activities: Ch. 4

➡ Descriptors of aspects of proficiency related to particular competences: Chapter 5

"The former are very suitable for teacher- or self-assessment with regard to real-world tasks. Such teacher- or self-assessments are made on the basis of a detailed picture of the learner's language ability built up during the course concerned. They are attractive because they can help to focus both learners and teachers on an action-oriented approach."
(CEFR p 180)

	RANGE	ACCURACY	FLUENCY	INTERACTION	COHERENCE
B2	Has a sufficient range of language to be able to give clear descriptions, express viewpoints on most general topics, without much conspicuous searching for words, using some complex sentence forms to do so.	Shows a relatively high degree of grammatical control. Does not make errors which cause misunderstanding, and can correct most of his/her mistakes.	Can produce stretches of language with a fairly even tempo; although he/she can be hesitant as he or she searches for patterns and expressions, there are few noticeably long pauses.	Can initiate discourse, take his/her turn when appropriate and end conversation when he/she needs to, though he/she may not always do this elegantly. Can help the discussion along on familiar ground confirming comprehension, inviting others in, etc.	Can use a limited number of cohesive devices to link his/her utterances into clear, coherent discourse, though there may be some "jumpiness" in a long contribution.
B1+					
B1	Has enough language to get by, with sufficient vocabulary to express him/herself with some hesitation and circumlocutions on topics such as family, hobbies and interests, work, travel, and current events.	Uses reasonably accurately a repertoire of frequently used "routines" and patterns associated with more predictable situations.	Can keep going comprehensibly, even though pausing for grammatical and lexical planning and repair is very evident, especially in longer stretches of free production.	Can initiate, maintain and close simple face-to-face conversation on topics that are familiar or of personal interest. Can repeat back part of what someone has said to confirm mutual understanding.	Can link a series of shorter, discrete simple elements into a connected, linear sequence of points.
A2+					
A2	Uses basic sentence patterns with memorised phrases, groups of a few words and formulae in order to communicate limited information in simple everyday situations.	Uses some simple structures correctly, but still systematically makes basic mistakes.	Can make him/herself understood in very short utterances, even though pauses, false starts and reformulation are very evident.	Can answer questions and respond to simple statements. Can indicate when he/she is following but is rarely able to understand enough to keep conversation going of his/her own accord.	Can link groups of words with simple connectors like "and, "but" and "because".
A1	Has a very basic repertoire of words and simple phrases related to personal details and particular concrete situations.	Shows only limited control of a few simple grammatical structures and sentence patterns in a memorised repertoire.	Can manage very short, isolated, mainly pre-packaged utterances, with much pausing to search for expressions, to articulate less familiar words, and to repair communication.	Can ask and answer questions about personal details. Can interact in a simple way but communication is totally dependent on repetition, rephrasing and repair.	Can link words or groups of words with very basic linear connectors like "and" or "then".

COUNCIL OF EUROPE
CONSEIL DE L'EUROPE

	RANGE	ACCURACY	FLUENCY
A2+	▪ routine, everyday transactions ▪ familiar situations and topics BUT ▪ compromise the message ▪ search for words.	▪ use some simple structures correctly	▪ adapt memorised simple phrases to particular situations ▪ handle short routine exchanges without undue effort, BUT ▪ hesitation and false starts.
A2	▪ communicate limited information ▪ simple everyday situations. ▪ basic sentence patterns ▪ memorised phrases ▪ groups of a few words and formulae	▪ Uses some simple structures correctly, BUT ▪ systematically makes basic mistakes.	▪ makes him/herself understood ▪ very short utterances, BUT ▪ pauses, false starts and reformulation
A1	▪ words and simple phrases ▪ personal details ▪ particular concrete situations.	▪ memorised repertoire. ▪ limited control ▪ a few simple grammatical structures	▪ very short, isolated, utterances ▪ mainly pre-packaged BUT much pausing to ▪ search for expressions, ▪ articulate less familiar words, ▪ repair communication

1. The shift in the 1990s

2. Assessment for learning / Assessment of level

3. "Constructive alignment" & "Backward design"

4. The CEFR linking planning, teaching, assessment

5. How CEFR descriptors help:

   The place in assessment of the new descriptors

We are talking about:

- Mediating a text
- Mediating concepts
- Mediating communication
- Online interaction
- Plurilingual and pluricultural competences

Phonology descriptors are designed for both curriculum orientation and assessment: can help inform qualitative assessment criteria

Why do we need to assess everything?

What is not assessed is not learnt (washback)?

What areas is it feasible and desirable to assess?

What types of assessment are feasible and desirable for those areas?

Cross-linguistic mediation in examinations:

- Greek KPG exam:
  - Processing a text in writing

- German KMK exams:
  - Relaying specific information in speech / in writing
  - Processing a text in writing
  - Acting as an intermediary (but in a written dialogue!)

- Austrian Matura: Certificate of Plurilingualism:
  - Explaining data
  - Processing a text in speech
  - Acting as an intermediary

Place of new descriptors

Consultation survey result	Assessment		Goal-setting/self-assess	
	Individuals	Institut	Individuals	Institut
Online conversation and discussion	89.44%	87.5%	95.77%	95.8%
Goal-oriented online transaction/collab	84.51%	87.5%	90.85%	95.8%
Express a pers. response to creative txt	90.57%	82.6%	90.57%	91.3%
Analysis and criticism of creative text	80.19%	78.3%	85.85%	82.6%
Facilitating collab. interaction with peers	85.50%	87.0%	90.84%	91.3%
Collaborating to construct meaning	83.21%	91.3%	89.31%	95.7%
Managing interaction	79.09%	69.6%	91.82%	95.7%
Encouraging conceptual talk	72.48%	69.6%	81.65%	91.3%
Facilitating pluricultural space	77.27%	60.9%	89.09%	87.0%
Facilitating comm. in delicate situations	79.09%	69.6%	87.27%	91.3%
Plurilingual comprehension	72.64%	69.6%	86.79%	91.3%
Building on plurilingual repertoire	66.67%	60.9%	81.90%	91.3%
Building on pluricultural repertoire	66.67%	65.2%	80.00%	91.3%

COUNCIL OF EUROPE
CONSEIL DE L'EUROPE

PILOTING

Online interaction

Teacher assessment & self-assessment

Tertiary/YL teaching practice B2: The descriptors seem **appropriate to measure students' communicative skills** in the target language using technology means. ... They help students focused on the task ... They serve for both students and the professor alike.

Tertiary - Teacher Education B2: These descriptors **allow learners to make judgments about their own work. They can help students critique their own work and make judgments about its strengths and weakness,** which makes it extremely valuable in developing learner autonomy.

Tertiary: A2.1/B1/B2.2: We used the descriptors to **rate our students' answers to the Google form and their contribution to the WhatsApp conversation.** After the students had been classified into different levels through the new descriptors, we checked whether the newly assigned levels **matched the levels they had previously attained in their placement tests**. In general terms, **we found a positive correlation between them.**

COUNCIL OF EUROPE
CONSEIL DE L'EUROPE

PILOTING

Mediating a creative text (literature)

Teacher assessment & self-assessment

Upper sec B1 / B2: For the students : They were happy and very cooperative to give feedback and **appreciated I was interested in their self-assessment**. For their self-assessment they had to tick a scale of 1 – 5 (very well – not sufficiently) for each descriptor.

Tertiary B1/B2: (*Students reacted*) very enthusiastically. **They found self-and peer-assessment based on the scales very much effective**, especially if followed by teachers' comments and feedback on their performance.

Upper Sec B1 / B2: The descriptors helped me to be more concentrated on the assessment from the point of view of the *can-do* descriptors and **better evaluate the students' performances rather than their knowledge of content.**

Tertiary B2/C1: The descriptors are useful to **measure the literature competences** and the critical view on literary masterpieces of students in a clear way.

COUNCIL OF EUROPE
CONSEIL DE L'EUROPE

PILOTING

Mediating concepts

Teacher assessment & self-assessment

Tertiary A1/A2: I really like them (*the descriptors*). I think they make the process of language learning more transparent for beginning students. The team of tutors got together and we selected the descriptors we wanted to use and then after that we reviewed the table and reduced it further to be able to use it both as **self-assessment and teacher assessment criteria**. We also used it in the **oral exam just as teacher assessment**.

Tertiary B1/B2: We found these descriptors to be quite useful in framing tasks based on interaction and collaboration and **gave us the tools to assess this interaction.** We saw how the participants moved from needing to clarify and confirm mutual understanding to interacting more effortlessly by building upon each other's ideas and presenting one's own ideas to invoke discussion.

Teenage B2: The benefits gained from the activities were considerably beneficial to the students in organising collaborative work, exploiting a task, taking roles to perform it well and **self-evaluating it.**

COUNCIL OF EUROPE
CONSEIL DE L'EUROPE

PILOTING

Plurilingual / pluricultural

Goal-setting & self-assessment

***Primary A1*: I found the assessment criteria to be clear and coherent and, thus, facilitating the way to promote the plurilingual competence.** The activity served to help students see that although the struggle at times with English, they have the resources and tools available to them to make meaning of new language.

***Tertiary B1*:** The CEFR descriptors represent an extremely valuable tool to be integrated in these activities in terms of **learning aims and students' self assessment**.

***Tertiary B2/C1*:** The descriptors informed the design of 10 tasks (each of about 30 minutes) that I piloted first with my own students and were later implemented by 4 teachers. These are teachers teaching English for Academic Purposes in a university program. **The descriptors also informed the design of a 5-point Likert Scale survey that I gave the students before the EAP program (Time 1) and 3 months later, at the end of the EAP program (Time 2).**

Conclusions:

- All the new descriptors seem suitable for goal-setting and self-assessment

- Those for online interaction, mediating creative text (incl. literature) and mediating concepts in group interaction appear **also** suitable for teacher assessment

- Those for mediating informative text and for acting as an intermediary appear **also** suitable for examinations