

RELANG

Relating language examinations to the common European reference levels of language proficiency: *promoting quality assurance in education and facilitating mobility*

Standardization & Benchmarking

European Centre for Modern Languages and European Commission cooperation on Innovative Methodologies And Assessment In Language Learning

Linking Procedures in the Manual

- Familiarisation with the CEFR
- Linking on the basis of specification of examination content
- **Standardization and Benchmarking**
- Standard setting
- Validation: checking that exam results relate to CEFR levels as intended

Claims through Specification and Standardization

- If a claim of a link to the CEFR is based on **specification** only, we do not know what is the *score* at which candidates can claim that their ability is at the CEFR-level of the test.
- Claims can be further substantiated through **standardisation** of judgements: item-difficulties are judged in relation to CEFR levels.

Steps in the standardization phase

1. Assuring adequate familiarisation with the CEFR.
2. Training in rating productive skills using tables and scales in the CEFR or the Manual & scales or specific rating scales
3. Training in rating receptive tasks using tables in the Manual & specifications developed for the examinations or tests in question
4. Benchmarking performance samples
5. Standard setting of receptive tasks

Group Decisions

Both ***benchmarking*** and ***standard setting*** are procedures which require group decisions, which in turn have to be carefully prepared by appropriate training.

Benchmarking in Direct Tests

- In holistically rated tests, the judgment on the CEFR level is direct, and therefore it is important to assist raters in giving valid judgments. The main tool used for this special type of standard setting is called ***benchmarking***.
- Benchmarking involves providing one (or more) typical sample(s) to illustrate performance at a given level both for standardisation training and to serve as a point of reference in making future decisions about performances of candidates.