RELANG

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

Benchmarking & Standard-setting

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

Linking Procedures in the Manual

- Familiarisation
- Specification
- Standardization and Benchmarking
- Standard setting
- Validation

Claims through Specification and Standardization

- Claimed link to CEFR is based on *specification* only
  - no indication of *score* required for CEFR-level in question
- Claims to be substantiated through *standardisation*
  - item-difficulties judged in relation to CEFR levels.

Group Decisions

- **Benchmarking** and **standard setting** are procedures requiring group decisions
    - to be carefully prepared by appropriate training.

Benchmarking in Direct Tests

- In tests productive skills the judgment on the CEFR level tends to be direct:
  - Assistance needed for raters in giving valid judgments
  - Main tool used for this special type of standard setting: *benchmarking*
- Benchmarking: providing one (or more) typical sample(s) to illustrate performance at a given level

Standard-setting in Indirect Tests

- For tests with numerical score, performance standards to be set
  - Receptive skills (reading, listening)
  - Underlying competences (grammar, vocabulary)
- Performance standard
  - Boundary (cut-off score) between two levels on the scale
- Process to arrive at cut-off score: **standard-setting**

Basket Method

- Basic question: *At what CEFR level can a test taker already answer the following item correctly?*
  - What is the minimum CEFR level required to give a correct response to this item?
- Panelists to put each item in a "basket" corresponding to one of the relevant CEFR levels
- Candidates at higher levels to give correct response
- Correct response not required at lower levels

Angoff Method

- Basic concept: *minimally acceptable person/borderline person*

- *Borderline person at B1*
  - has the ability to be labelled B1
  - slightest decrease in ability: no longer B1

- Task for panel: keep *borderline person* in mind for all the judgments

- For each item: What is the **probability** that the *borderline person* gives the correct answer ?