RELANG

ΚΥΠΡΙΑΚΗ ΔΗΜΟΚΡΑΤΙΑ
ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ, ΠΟΛΙΤΙΣΜΟΥ,
ΑΘΛΗΤΙΣΜΟΥ ΚΑΙ ΝΕΟΛΑΙΑΣ

EUROPEAN UNION

EUROPEAN CENTRE FOR MODERN LANGUAGES
ECML CELV
CENTRE EUROPEEN POUR LES LANGUES VIVANTES

COUNCIL OF EUROPE
CONSEIL DE L'EUROPE

Constructing valid language tests and examinations on the CEFR and its Companion Volume

Nicosia, 10-12 November 2021

Validity and the CEFR

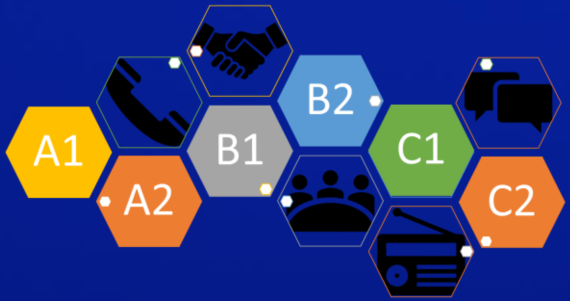
- A test is valid if it measures what we intend it to measure.
- Demonstrating that a learner reported to be at B1 actually *is* at B1 according to the evidence
- Consequential validity refers to the positive or negative social consequences of a particular test (see fairness)
- If focus is on *use:* validity evidence relates to language used for communicative purposes

Reliability in Testing

- **Consistency**
  - Same or similar results on repeated use
  - High reliability does not mean the test is valid
- **Minimising likely sources of error**
- **Using statistics to estimate reliability of test scores**
- **Reliability figures depend on task type and way of marking**

Introduction to the CEFR Companion Volume (CV) with New Descriptors

- Highlighting CEFR areas for which no descriptor scales had yet been provided, especially mediation and plurilingual / pluricultural competence.
- Extended definition of ‘plus levels’ and a new ‘Pre-A1’ level.
- More elaborate description of listening and reading in existing scales.
- Enriching the description at A1, and at the C levels, particularly C2.



Focus on the CV project

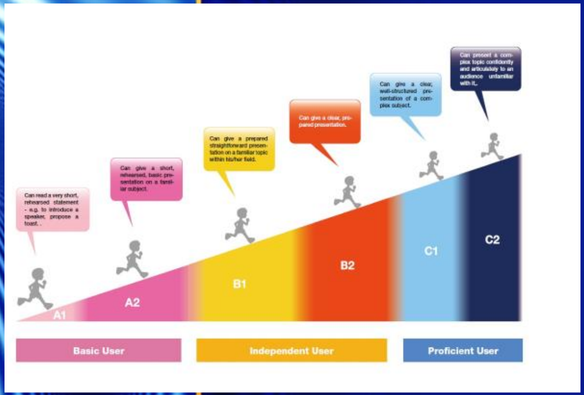
- **Pre-A1 level descriptors** - relate to simple, general tasks, which were scaled below Level A1 (Pre-A1), but can constitute useful objectives for beginners.

- **Mediation -** Treatment of mediation in the CEFR is not limited to cross-linguistic mediation, passing on information in another language:

  Mediating a text

  Mediating concepts

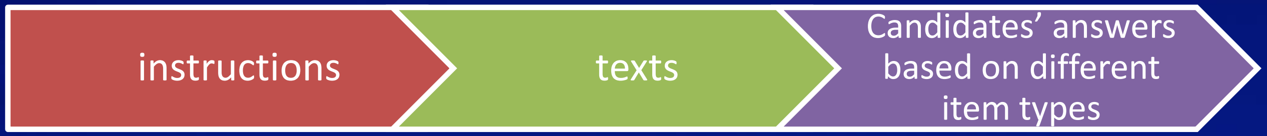
  Mediating communication


Can read a very short, rehearsed statement - e.g. to introduce a speaker, propose a toast. .
Can give a short, rehearsed, basic presentation on a familiar subject.
Can give a prepared straightforward presentation on a familiar topic within his/her field.
Can give a clear, prepared presentation.
Can give a clear, well-structured presentation of a complex subject.
Can present a complex topic confidently and articulately to an audience unfamiliar with it..
A1
A2
B1
B2
C1
C2
Basic User
Independent User
Proficient User

Testing Receptive Skills

Tasks and items

- Certain number of tasks
- Tasks composed of

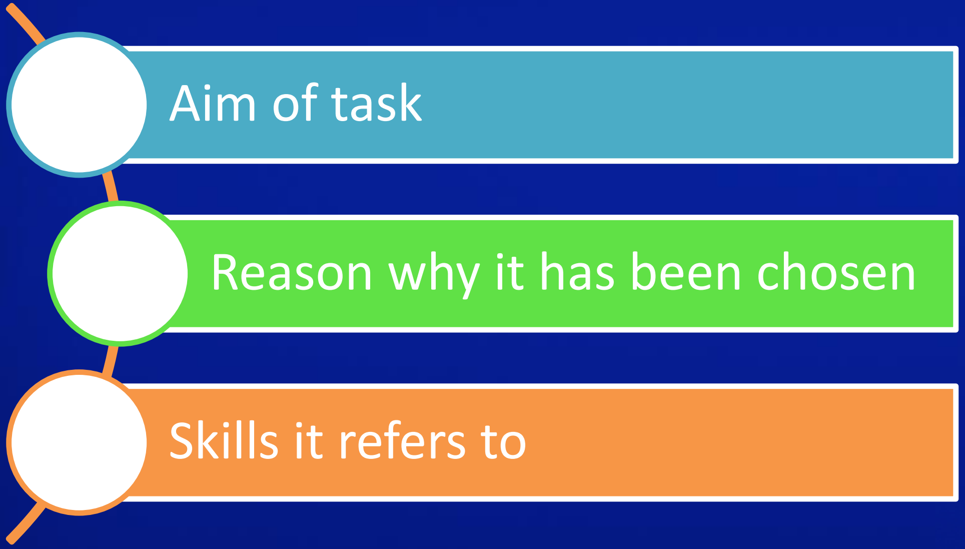


instructions

texts

Candidates’ answers based on different item types

- Tasks assessed using making grid/rating scale
- Test developers to have clear idea about:



Aim of task

Reason why it has been chosen

Skills it refers to

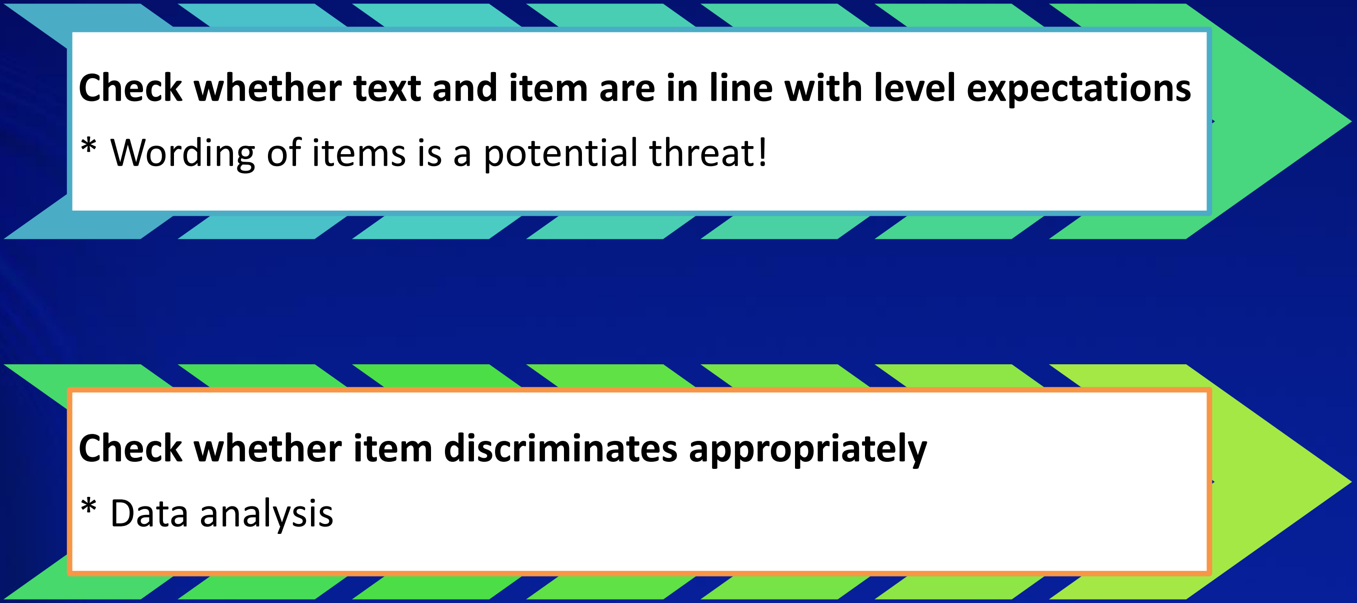
Quality criteria: Relevance

Check whether item focuses on targeted construct

* Authenticity

Check whether item does **not** focus on elements outside of targeted construct

Quality criteria: Level



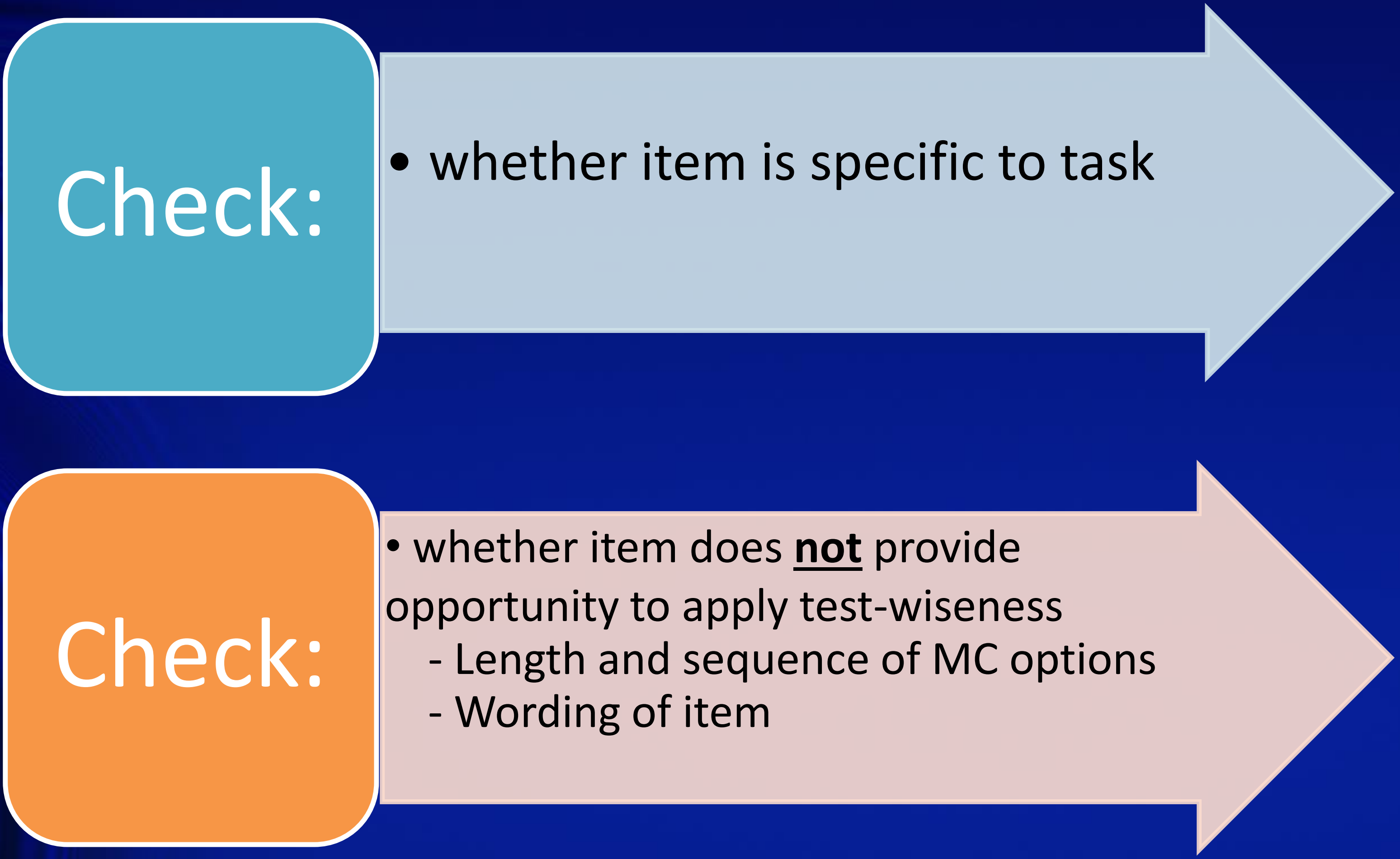
Check whether text and item are in line with level expectations

* Wording of items is a potential threat!

Check whether item discriminates appropriately

* Data analysis

Quality criteria: Specificity



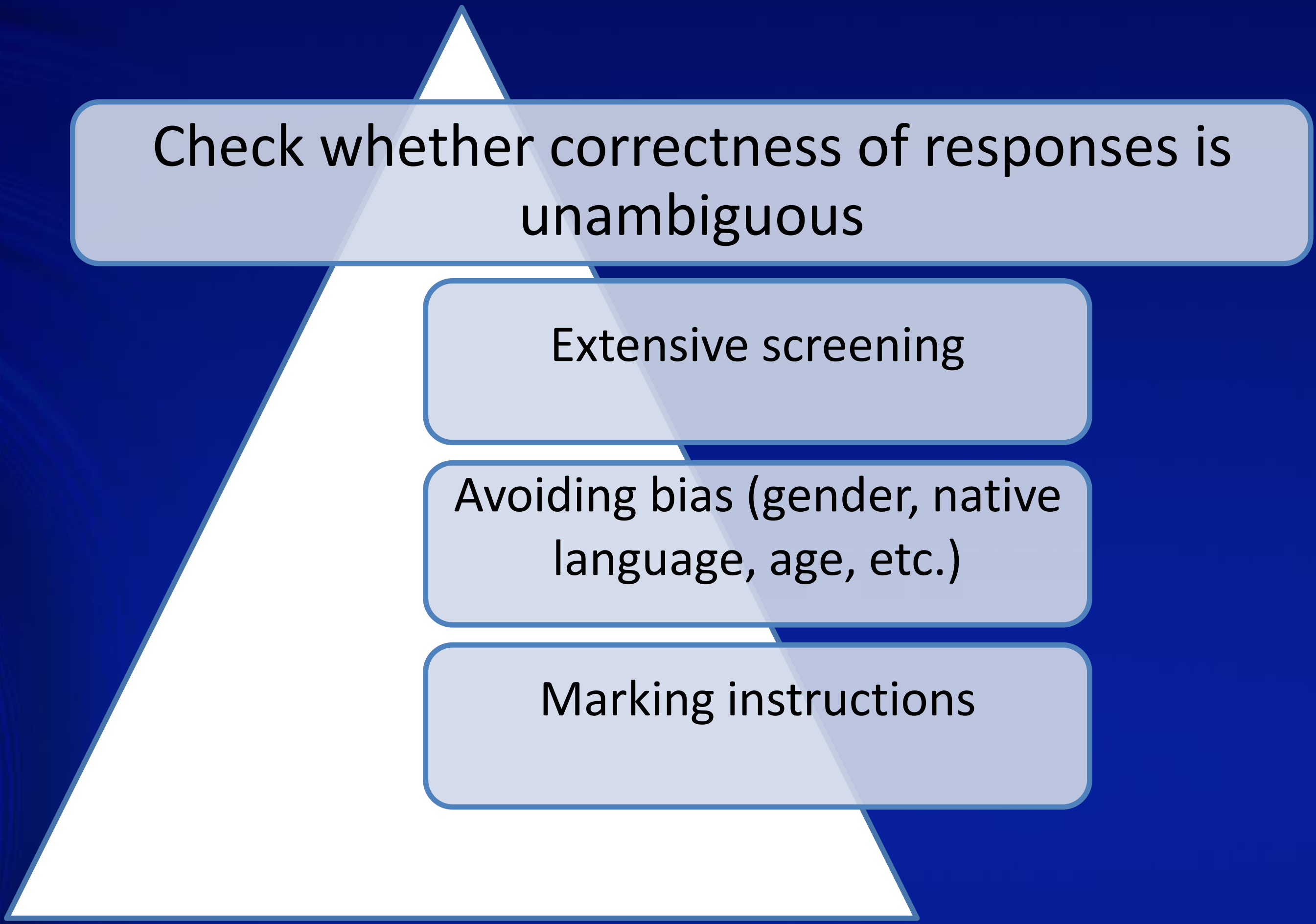
Check:

- whether item is specific to task

Check:

- whether item does **not** provide opportunity to apply test-wiseness
  - Length and sequence of MC options
  - Wording of item

Quality criteria: Objectivity



Check whether correctness of responses is unambiguous

Extensive screening

Avoiding bias (gender, native language, age, etc.)

Marking instructions

Quality criteria: Acceptability

Check whether instructions are clear

Check whether item does **not** deliberately mislead candidates (Instructions for item writers)

Quality criteria: Transparency

Check whether response specifics are clearly indicated

Check whether item is in line with candidate expectations

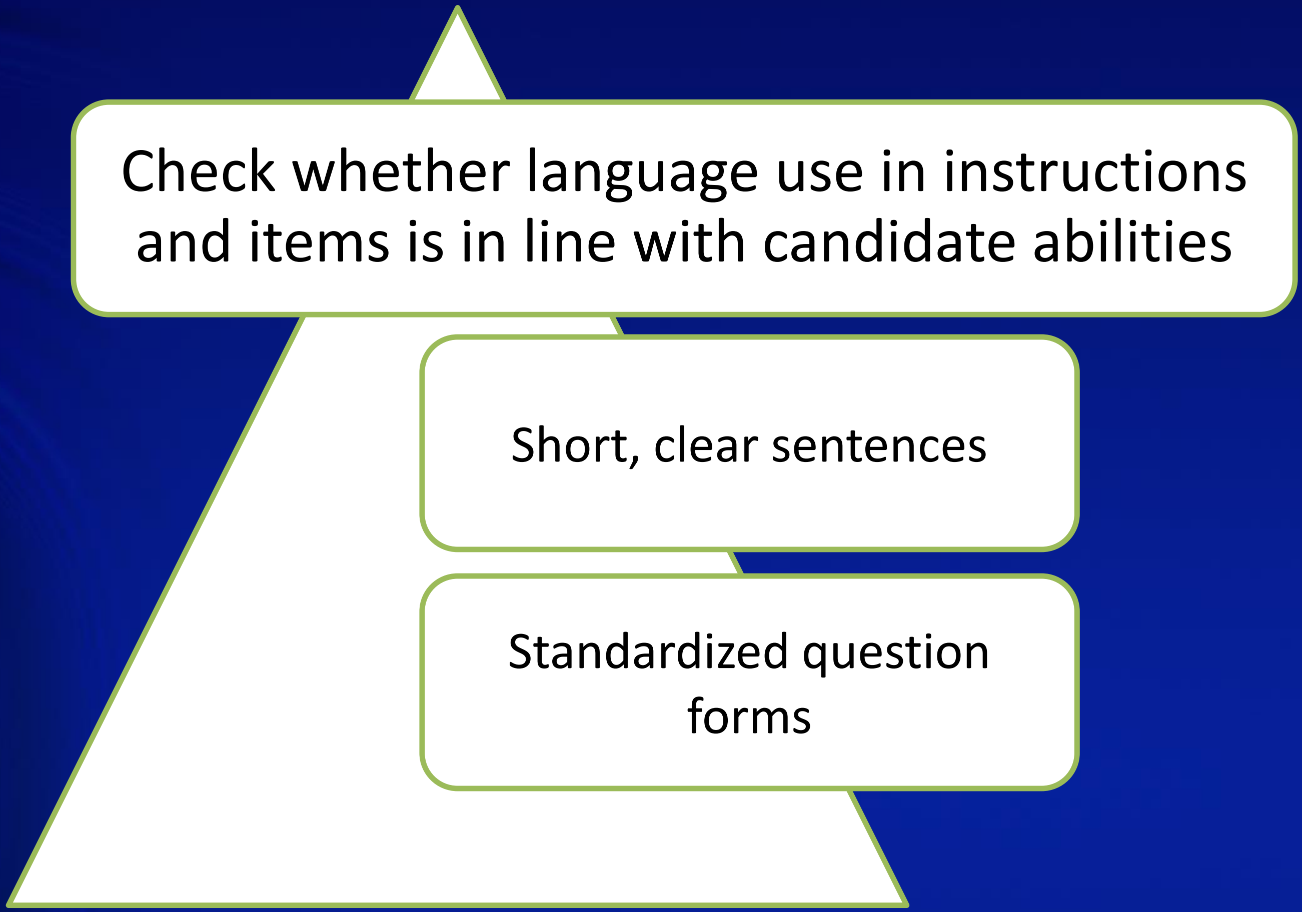
Check whether item type is familiar to candidates - specifications

Quality criteria: Efficiency

Check whether information is presented in the most efficient way to candidates

- Time
- Length/complexity of responses

Quality criteria: Language use

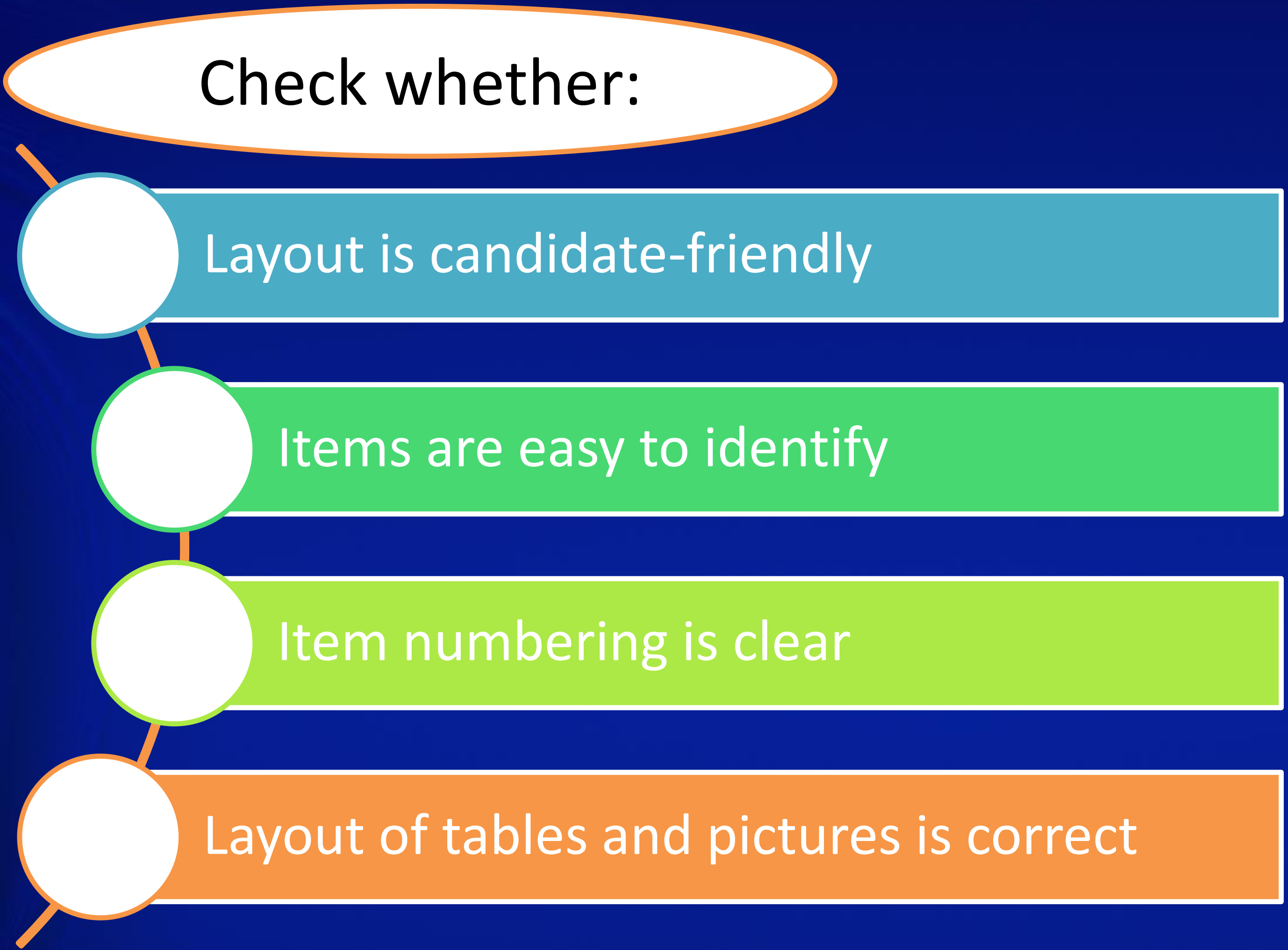


Check whether language use in instructions and items is in line with candidate abilities

Short, clear sentences

Standardized question forms

Quality criteria: Layout



Check whether:

- Layout is candidate-friendly
- Items are easy to identify
- Item numbering is clear
- Layout of tables and pictures is correct

Testing Productive Skills speaking and writing

Speaking & Writing in the CEFR

- **Actions performed by persons**

- **Illustrative scales for:**
  - Spoken/Written production
  - Spoken/Written interaction

- **Various contexts**

- **Communicative competences**
  - Linguistic competence
  - Sociolinguistic competence
  - Pragmatic competence

Production Activities (CV)

ORAL PRODUCTION

- Overall oral production
- Sustained monologue: giving information
- Sustained monologue: describing experience
- Sustained monologue: putting a case
- Public announcements
- Addressing audiences

WRITTEN PRODUCTION

- Overall written production
- Creative writing
- Reports and essays

Production Strategies

- **Planning**

rehearsing, locating resources, considering audience, task adjustment, message adjustment

- **Compensating**

building on previous knowledge, trying out

- **Monitoring and repair**

monitoring success and self-correction

Interaction Activities (CV)

Oral interaction

- Overall oral interaction
- Understanding an interlocutor
- Conversation
- Informal discussion
- Formal discussion
- Goal-oriented co-operation
- Obtaining goods and services
- Information exchange
- Interviewing and being interviewed
- Using telecommunications

Written interaction

- Overall written interaction
- Correspondence
- Notes, messages and forms

Interaction Strategies (CV)

- Turntaking
- Co-operating
- Asking for clarification

Task Characteristics

Stimulus as short and clear as possible

- Providing an adequate framework for candidates’ speaking/writing performance

Pictures/visual material can be used as stimulus

- Difficulties of interpretation

Candidates to know who they are speaking/writing to and how this will affect what they speak/write about

Checklist for speaking/writing tasks

- □ Is topic of text to be produced accessible to candidates?
- □ Is context realistic?
- □ Is language in rubrics accessible to candidates at target level?
- □ Is purpose of task clearly indicated?
- □ Does task provide opportunity for candidates to show their range?
- □ Is marking scheme provided?

Standardization & Benchmarking

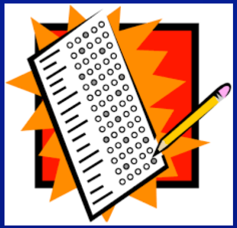
Steps in the standardization phase

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

Benchmarking in Direct Tests

- In tests of productive skills the judgment on the CEFR level is 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

** ***Benchmarking*** and ***standard setting*** are procedures requiring group decisions



Standard Setting

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**

How to arrive at standards?

- Group decisions (panel)
- Group is familiar with CEFR
- Test content specified in terms of the CEFR
- Standard setting procedures formalized
- Careful selection and training of panel members

ECML Teacher Trainers:

José Noijons
Evelyne Bérard

PPT prepared by:

Andrea Chrysostomidou

EFL Advisor