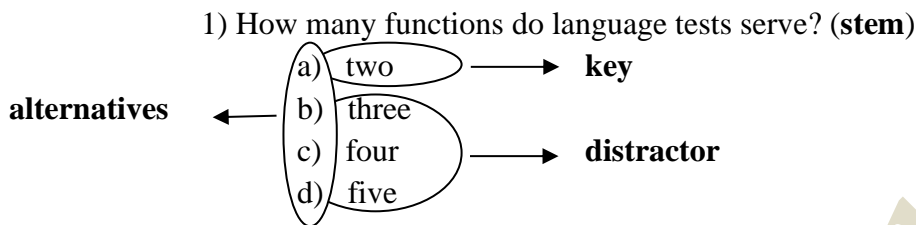


Chapter Three: Forms of Language Test

Physical appearance of a test

1. STRUCTURE OF AN ITEM

An item, the smallest unit of a test, consists of two parts: the *stem* and the *response*.



2. CLASSIFICATION OF ITEM FORMS

2.1. Subjective vs. objective items

Subjective → the scorer must make an opinionated judgment

Objective → the correctness of the test taker's response is determined by objective criteria

☛ **Note:** objectivity and subjectivity refers to the way a test item is scored.

The most beautiful season is

1) spring 2) summer 3) fall 4) winter

There are seasons in a year.

1) four 2) three 3) two 4) five

2.2. Essay-type vs. multiple-choice items

Essay-type items → examinee *produces* language elements

Multiple-choice items → examinee *selects* the correct response from among given alternatives

2.3. Suppletion vs. recognition items

Suppletion items → testee supplies the missing part(s)

Recognition items → testee selects an answer from a list of possibilities

3. TYPES OF ITEMS

3.1. Receptive response items

Advantages of Multiple-Choice (MC) items:

- ☺ Highly structured → test writer can get directly at the specific skills → diagnostic function
- ☺ Practicality → scoring can be done quickly → include a large number
- ☺ No judgments as to degrees of correctness → reliability

Disadvantages of Multiple-Choice (MC) items:

- ☹ Test only recognition knowledge but not language communication → passive

- ⊗ Harmful washback,
- ⊗ Expose students to errors,
- ⊗ De-contextualized,
- ⊗ Time-consuming to construct,
- ⊗ Simpler to answer than subjective tests,
- ⊗ Encourage guessing

To compensate for students' guessing on tests:

$$\text{Guessing Correction Formula} = \text{Right} - \frac{\text{Wrong}}{n - 1}$$

n refers to the number of options

Example: In a test which consisted of 80 items with five options, a student answered 44 items correctly. After applying guessing correction formula his score would be -----.

3.2. Personal response items

An attempt to develop non-test assessment options

Against the notion that all people could be measured by traditional tests → data triangulation

3.2.1. Self-assessment

Any items wherein students are asked to rate their own knowledge, skills, or performances

Provides the teacher with some idea of how the students view their own language abilities

- ☺ Speed
- ☺ Direct involvement of students → increased motivation
- ☺ Encouragement of autonomy
- ⊗ Subjectivity

🌟* **Note:** *Peer-assessment* is a variation on this theme that requires students to rate each other

Categories of self- and peer-assessment:

- Direct assessment of a specific performance
- Indirect assessment of general competence
- Student-generated tests

3.2.2. Journal

An interaction between the teacher and the student through dialogues

- ☺ Teachers become acquainted with their students' learning progress and their affective states → meet students' individual needs → offer various kinds of feedback
- ⊗ Too free to be assessed accurately
- ⊗ Critics have expressed ethical concerns

3.2.3. Conference

Any assessment procedure that involves students visiting the teacher’s office alone or in groups for brief meetings.

☺ Teachers are able to direct feedback toward a student’s specific needs

Teachers assume the role of a facilitator and guide, rather than a master controller and deliverer of final grades.

3.2.4. Portfolio

A purposeful collection of students’ work that demonstrates to students and others their efforts, progress, and achievements in given areas

CRADLE → Collecting, Reflecting, Assessing, Documenting, Linking, Evaluating

4. ALTERNATIVE vs. TRADITIONAL ASSESSMENT

Traditional Assessment	Alternative Assessment
One-shot, standardized exams Timed, multiple-choice format Decontextualized test items Scores suffice for feedback Norm-referenced scores Focus on the right answer Summative Oriented to product Non-interactive performance Fosters extrinsic motivation	Continuous long-term assessment Untimed, free-response format Contextualized communicative tasks Individualized feedback and washback Criterion-referenced scores Open-ended, creative answers Formative Oriented to process Interactive performance Fosters intrinsic motivation