



CHARACTERISTICS OF A GOOD TEST

RELIABILITY : THE STABILITY OF TEST SCORES

Types of estimating reliability:

1. Test retest

one test is administered to the same students at least two different times

2. Parallel forms

two different version of test at the same time are administered to the same students

3. Two scorers

one test is administered once to the students and the scores are separated into two parts

4. Split half procedures

one test is administered to the students and the scores are separated into two parts, 1st and 2nd part

Test retest

Administered the test twice to a group of students

Use the **Pearson Product Moment Formula (PPMF)** to find the correlation coefficient of the two sets of scores

$$r_{xy} = \frac{N \sum XY - (\sum X)(\sum Y)}{\sqrt{\{(N \sum X^2 - (\sum X)^2)\}\{(N \sum Y^2 - (\sum Y)^2)\}}}$$

X= first score

Y= second score

N= number of students

Parallel form

- Administer two version of a test to a group of students
- Find the correlation coefficient of the two sets of scores by using PPMF

Parallel:

1. Version 1
2. Version 2

Two raters

- Administer once to a group of students then it scored by two scorers/raters

usually in a form of free test (writing/speaking)

- Find the correlation coefficient of the two sets of scores by using PPMF

Split half procedure

- Administer the test once to a group of students
- Divide the test items into two parts
- Count each students scores in the two parts
- Find the correlation coefficient of the two sets of scores by using PPMF
- Find the reliability coefficient of the formula of the whole test by using **Spearman Brown Formula**

$$r_{ii} = \frac{2 r_{xy}}{(1 + r_{xy})}$$

VALIDITY : THE DEGREE TO WHICH A TEST MEASURE WHAT IT SHOULD MEASURE

Types of validity:

1. Content validity

the content of the tests represent the content of syllabus

TABLE OF SPECIFICATION **Vocabulary Teaching Material**

No	Topics	Noun	Verb	Adjective	Adverb
1					
2					

TABLE OF SPECIFICATION of Vocabulary test

No	Topics	Noun	Verb	Adjective	Adverb
1					
2					

2. Empirical validity

the test correlate with another independent criterion

▣ Concurrent validity

- If the test correlates with another criterion that exist at about the same time

▣ Predictive validity

- If the test can predict what will happen at the end of the semester

3. Face validity

how the test “look” to the examine, examiners, test administrator, etc

factors:

- ❑ the language
- ❑ the direction (instruction must be clear)
- ❑ complete options
- ❑ no missing words
- ❑ good printing quality

PRACTICALITY

1. Ease of Economy

The more persons who must be involved in giving the score of a test, the more costly the test becomes

2. Ease of administration

Consider: clear directions
good printing quality

3. Ease of scoring

decide to use subjective or objective scoring
what kind of answer sheet is used