VALIDITY & RELIABILITY

Raja C. Bandaranayake
QUALITIES OF MEASUREMENT DEVICES

- **Validity**
  Does it measure what it is supposed to measure?

- **Reliability**
  How representative is the measurement?

- **Objectivity**
  Do independent scorers agree?

- **Practicality**
  Is it easy to construct, administer, score and interpret?
PURPOSES OF ASSESSING STUDENTS

- Selection
- Certification
- Comparison
  - of individuals (ranking)
  - of student groups
- Promotion
- Diagnosis
- Evaluating programs and teachers
VALIDITY

- Content related to objectives

- Construct related to other qualities e.g. test scores and anxiety measures

- Concurrent related to other measures of performance e.g. written scores & clerkship ratings

- Predictive related to future performance
VALIDITY

- **Content**
  related to objectives

- **Construct**
  related to other qualities
    
    *e.g. test scores and anxiety measures*
VALIDITY

- **Content**
  related to objectives

- **Construct**
  related to other qualities
    
  *e.g. test scores and anxiety measures*

- **Concurrent**
  related to other measures of performance
    
  *e.g. written scores & clerkship ratings*
VALIDITY

- **Content**
  related to objectives

- **Construct**
  related to other qualities
  *e.g. test scores and anxiety measures*

- **Concurrent**
  related to other measures of performance
  *e.g. written scores & clerkship ratings*

- **Predictive**
  related to future performance
VALIDITY

- **Content**
  - related to objectives

- **Construct**
  - related to other qualities
    - *e.g. test scores and anxiety measures*

- **Concurrent**
  - related to other measures of performance
    - *e.g. written scores & clerkship ratings*

- **Predictive**
  - related to future performance
QUALITIES OF MEASUREMENT DEVICES

- **Validity**
  Does it measure what it is supposed to measure?

- **Reliability**
  How representative is the measurement?

- **Objectivity**
  Do independent scorers agree?

- **Practicality**
  Is it easy to construct, administer, score and interpret
RELIABILITY

An estimate of the amount of error in the measurement.
RELIABILITY

An estimate of the amount of error in the measurement.

Thus can only be determined after the test is administered.
RELIABILITY

An estimate of the amount of error in the measurement.

Thus can only be determined after the test is administered.

Estimate based on group rather than individual scores.
SOURCES OF ERROR

- Examinee
- Examiner
- Examination
RELIABILITY

**Individual score**

\[
\text{Observed score} = \text{True score} \pm \text{Error score}
\]

\[
X_t = X_\infty \pm X_e
\]

**Group scores**

\[
\text{Observed variance} = \text{True variance} \pm \text{Error variance}
\]

\[
\sigma_0^2 = \sigma_\infty^2 \pm \sigma_e^2
\]
RELIABILITY

**Individual score**

Observed score = True score ± Error score

\[ X_t = X_\infty \pm X_e \]

**Group scores**

Observed variance = True variance ± Error variance

\[ \sigma_0^2 = \sigma_\infty^2 \pm \sigma_e^2 \]

Reliability of a set of measurements is the proportion of observed variance that is true variance

\[ r_{tt} = \frac{\sigma_\infty^2}{\sigma_0^2} \]

Standard error of measurement

\[ SE_{\text{meas}} = \sigma \sqrt{1 - r_{tt}} \]
RELATIONSHIP BETWEEN VALIDITY & RELIABILITY

Validity takes precedence over reliability

A test cannot be deemed valid unless the measurements resulting from it are reliable
RELATIONSHIP BETWEEN VALIDITY & RELIABILITY

Validity takes precedence over reliability

A test cannot be deemed valid unless the measurements resulting from it are reliable
QUALITIES OF MEASUREMENT DEVICES

- **Validity**: Does it measure what it is supposed to measure?
- **Reliability**: How representative is the measurement?
- **Objectivity**: Do independent scorers agree?
- **Practicality**: Is it easy to construct, administer, score and interpret