ImplicitifyAI

MULTIMODAL PSYCHOMETRIC ASSESSMENT

TEACHING RESOURCES

Reliability and Validity of Measures

Topic: Psychometrics
Slides: 4

How we evaluate measurement quality: reliability coefficients, validity evidence, and item-level statistics.

Reliability

- Reliability is the consistency of measurement across items, time, or raters.
- Internal consistency (Cronbach's alpha), test–retest, and inter-rater are common forms.
- Reliability sets a ceiling on validity: an unreliable measure cannot be valid.

Validity evidence

- Content validity: items adequately sample the construct's domain.
- Criterion validity: scores relate to an external criterion (concurrent or predictive).
- Construct validity: convergent and discriminant evidence support the intended meaning.

Item statistics

- Item difficulty (p) is the proportion endorsing or answering correctly.
- Item discrimination indexes how well an item separates high and low scorers.
- Item-total correlations flag items that do not cohere with the scale.

Norms and standardization

- Norm-referenced scores compare a person to a reference sample.
- Percentiles, z-scores, and T-scores are common standardized metrics.
- Norms are only valid for populations resembling the standardization sample.

Answer Key — Practice Items

Item-bank questions for the topic “Psychometrics”.

1. Cronbach's alpha is an index of:

A. Criterion validity
B. Internal-consistency reliability
C. Test difficulty
D. External validity

Answer: Internal-consistency reliability

Rationale: Alpha summarizes how consistently a set of items measure the same thing.

2. Reliability and validity are related such that:

A. A valid test must be reliable
B. A reliable test must be valid
C. They are unrelated
D. Validity caps reliability

Answer: A valid test must be reliable

Rationale: Reliability sets a ceiling on validity; a measure cannot be valid for a purpose if it is not reliable.

3. Convergent and discriminant evidence are forms of:

A. Content validity
B. Construct validity
C. Inter-rater reliability
D. Face validity

Answer: Construct validity

Rationale: Convergent (relates to similar constructs) and discriminant (unrelated to different constructs) evidence support construct validity.

4. An item-total correlation near zero suggests the item:

A. Is too difficult
B. Does not cohere with the rest of the scale
C. Has high discrimination
D. Is perfectly reliable

Answer: Does not cohere with the rest of the scale

Rationale: Low item-total correlation indicates the item is not measuring the same construct as the scale.

5. A T-score has a mean and standard deviation of:

A. 0 and 1
B. 50 and 10
C. 100 and 15
D. 5 and 2

Answer: 50 and 10

Rationale: T-scores are standardized to M = 50, SD = 10.