The Concept of Validity

Dr Wan Nor Arifin

Unit of Biostatistics and Research Methodology,
Universiti Sains Malaysia.
wnarifin@usm.my / wnarifin.github.io

Outlines

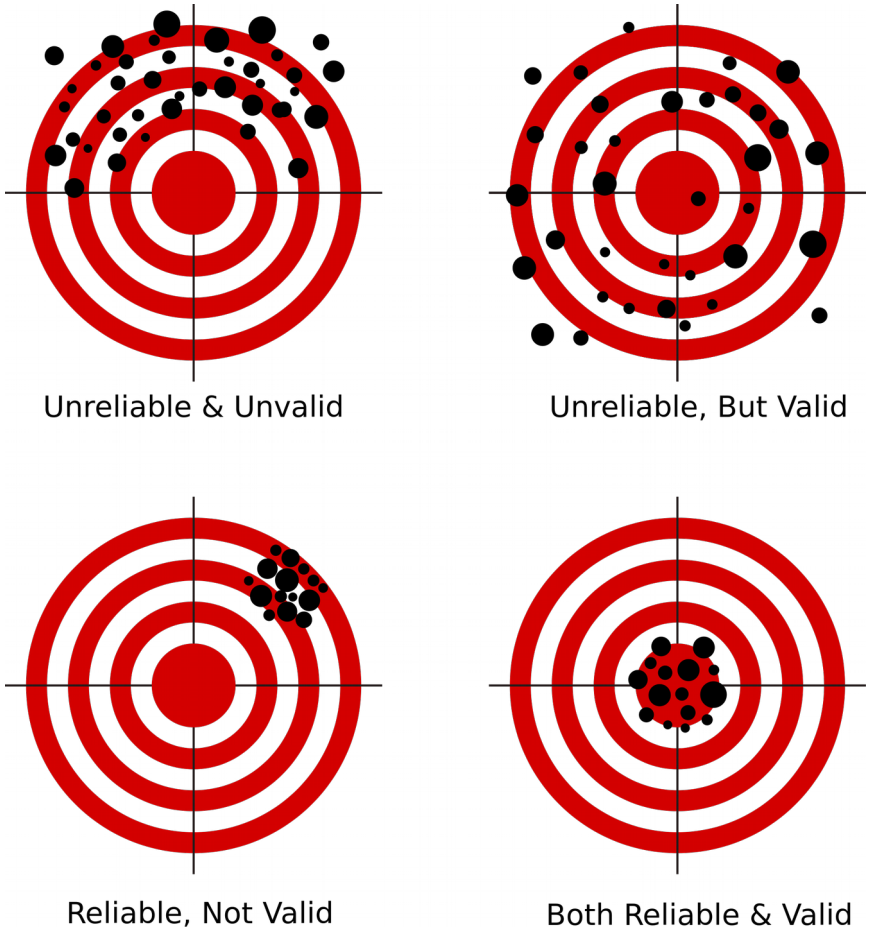
1. **Measurement validity and reliability**
2. **The classical view of measurement validity**
3. **The Validity**

Measurement validity and reliability

Measurement validity and reliability

- **Measurement → Process of observing & recording.**
- **Measurement validity → Accuracy.**
- **Measurement reliability → Precision, consistency, repeatability, reproducibility.**

Measurement validity and reliability



The classical view of measurement validity

The classical view of measurement validity

- **3Cs (Fletcher, Fletcher and Wagner, 1996; Streiner and Norman, 2008):**

1. Content
   - Content of a questionnaire.
2. Criterion
   - Concurrent.
   - Predictive
3. Construct
   - Convergent
   - Discriminant

The Validity

The Validity

- **Unitary concept. Also called "construct validity".**
- **Degree of evidence → Purpose & Intended use of a tool.**
- **Evidence from 5 sources (AERA, APA & NCME, 1999):**

1. Content.
2. Internal structure.
3. Relations to other variables
4. Response process.
5. Consequences.

Content

- **How well a measure includes all the facets of an idea or concept, which a researcher intends to measure**

  (Fletcher, Fletcher and Wagner, 1996).

- **Judged on three aspects (Streiner and Norman, 2008):**

  1. Relevance: How relevant and related the items to the concept.
  2. Coverage: Adequate number of items to cover the concept.
  3. Representativeness: Number of items covering the item is proportionate to the importance of the concept.

Internal Structure

- **The degree of the relationships among items and constructs as proposed or hypothesized (AERA, APA & NCME, 1999).**
- **Proven on the basis of analyses that can prove the correlatedness (i.e. correlations coefficients, factor loadings) and dimensionality (number of factors) (Cook, Thomas & Beckman, 2006):**

  1. Factor analysis (exploratory and confirmatory).
  2. Reliability.

- **The analyses are based on variables available internal to the test itself (i.e. the questions, items), hence the name internal evidence.**
- **D2 workshop.**

Relations to other variables

- **Prove the relationship of the measurement tool scores to other external variables, which may include other measurement tools/questionnaires, and other observable variables or criteria.**
- **Can be done by:**
  - Convergent and discriminant evidence
  - Test-criterion relationship

Relations to other variables

- **Convergent and discriminant evidence**
  - Convergent: vs Qs measuring same concept.
  - Discriminant: vs Qs measuring something else.
- **Test-criterion relationship**
  - Concurrent: vs criterion/gold standard available NOW.
  - Predictive: vs criterion/gold standard available LATER.

Response process

- **It is concerned with the process of responding to the questions.**
- **May be done in cognitive debriefing (next lecture) by probing the respondent as to how he comes up with a response per question.**
- **For interviewer rated, may observe how the interviewer/rater comes up with a rating.**

Consequences

- **It is concerned with the evidence regarding the intended and unintended consequences of the result from a measurement tool.**
- **For example, if a person is rated as depressed, what would be the consequence of that? Referral to psychiatric clinic (intended)? Losing job (unintended)? Etc.**
- **As an additional source of evidence to support the rest of evidence.**

References

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (1999). Standards for educational and psychological testing. Washington DC: American Educational Research Association.

Cook, D. A., & Beckman, T. J. (2006). Current concepts in validity and reliability for psychometric instruments: theory and application. The American journal of medicine, 119, 166.e7-166.e16.

Fletcher, R. H., Fletcher, S. W., & Wagner, E. H. (1996). Clinical epidemiology: the essentials (3rd ed.). Maryland: Williams & Wilkins.

Streiner, D. L. & Norman, G. R. (2008). Health measurement scales: a practical guide to their development and use. New York: Oxford University Press.