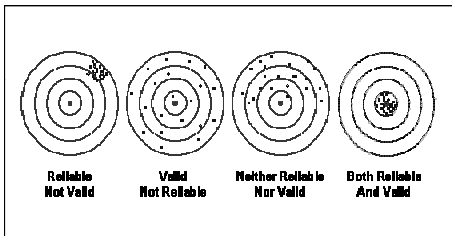


Measurement Reliability

- Measurements consist of two components. Both must be inferred because they can't be directly observed.
 - True score
 - The real score of the variable that exists hypothetically and is considered a constant.
 - Measurement error
 - The portion of the measurement result that varies from measurement to measurement
- The higher the correlation (r near 1) between two measurements, the more reliable the measurement.

34

Measurement Analogy



35

Assessing the Reliability of Psychological Measures

- Test-Retest Reliability
 - Take the measurement at two different times and correlate the results ($r = .8$ to be considered reliable).
- Internal Consistency Reliability
 - Split test in half and correlate the scores on the first half with the scores on the second half.
 - Correlate every item score with every other item score (Cronbach's alpha)
 - Correlate every item score with the total score.
- Interrater Reliability
 - Correlate rater scores with each other.

36

Measurement Validity

- If a measurement tool is valid, it is really measuring the construct it purports to measure rather than some other characteristic.
 - A valid measurement tool is said to have "*Construct Validity*".

37

Methods to Assess Construct Validity

- Face Validity (does it appear valid)
 - An IQ test that only asked questions about a person's income, address and clothing preferences would have little face validity.
- Criterion-Oriented Validity (does it allow accurate predictions)
 - A safe-driving test that passed people who continued to have driving accidents would lack criterion-oriented validity.

38

Forms of Criterion-Oriented Validity

- Predictive Validity- *ability to predict something it should theoretically be able to predict.*
 - A measure of math ability accurately predicts who will be a successful engineer
- Concurrent Validity- *ability to distinguish between groups that it should theoretically be able to distinguish between*
 - A measure of anxiety distinguishes individuals with an anxiety disorder from those with major depression.

39

Forms of Criterion-Oriented Validity

- **Convergent Validity**- *results from the measure are similar to other measures of the same construct*
 - The block assembly IQ test yields high IQ scores with the same people who got high IQ scores on a spatial problem solving IQ test.
- **Discriminant Validity**- *results from the measure are NOT similar to measures of theoretically different constructs.*
 - An abstract-reasoning test yields different results than a rote-memory test

40

Measurement Reactivity

- **Reactivity in measurement** is when the very act of measuring a variable *changes* that variable.
 - Example: recording the amount of food eaten each a day causes the person to eat less
- Using *unobtrusive* measures reduces the probability of measurement reactivity.

41

Measurement Scales

- **Nominal**
 - No numerical properties, only named categories
 - Examples: male or female, married or single
- **Ordinal**
 - Has numerical properties but only for the purpose of ordering the categories from first to last
 - Examples: Academic grades, places in a race, ranking schools or sports teams

42

Measurement Scales

■ Interval

- Has numerical properties with equal distance between each number on the scale but no zero point (representing a total absence of the thing being measured).
- Example: Rating beauty on a scale of 1-10

■ Ratio

- Same as the interval scale (equal intervals between numbers) except there is a zero point.
- Examples: time measures, response rate measures, other physical measurements.
