

Validity and Reliability



Edgar Degas: Portraits in a New Orleans Cotton Office, 1873



Validity and Reliability

A more complete explanation of validity and reliability is provided at the accompanying web page for this topic.

See: [Validity and Reliability](#)



Validity and Reliability

1. Validity

- The extent to which a “test” measures what it is supposed to measure.
 - Non-empirical evaluations.
 - Empirical evaluations.

2. Reliability

- The extent to which a “test” provides consistent measurement.
 - Negatively affected by measurement error.



Validity and Reliability

3. Importance

- Without validity and reliability one cannot test an hypothesis.
- Without hypothesis testing, one cannot support a theory.
- Without a supported theory one cannot explain why events occur.
- Without adequate explanation one cannot develop effective material and non-material technologies, including programs designed for positive social change.



Validity

1. Non-Empirical Validity

- “Does the concept seem to measure what it is supposed to measure?”
- This evaluation depends entirely upon the *intersubjective* opinion of the community of scholars.
- It is not entirely subjective because it represents the consensus opinion of many subjective individual opinions.



Validity

2. Empirical

- Data analysis evaluation of validity.
- “Does the concept predict another concept that it is supposed to predict?”
- This evaluation depends upon empirical evaluation, using either quantitative or qualitative data.
- Empirical validity does not imply content validity.



Validity

3. Outcomes of Content Validity Assessment

- If the community of scholars approves the content validity of a measure, then it is accepted for further use.
- If the community of scholars rejects the content validity of a measure, then it must be redesigned by the researcher(s).



Validity

4. Outcomes of Empirical Validity Assessment

- Assessments of construct validity imply three types of hypotheses:
 1. Do the items used to measure the construct have a statistically significant correlation with the construct? [Construct Validity]
 2. Does the construct have a statistically significant correlation with related constructs? [Concurrent Validity]
 3. Does the construct statistically predict another construct? [Predictive Validity]



Validity

5. Outcomes of Empirical Validity Assessment

- If the null hypotheses are rejected, then the empirical results lend support for the empirical validity of the measure.
- If one or more null hypotheses are not rejected, then:
 1. The measure of the construct might be invalid.
 2. The measure of the predicted construct might be invalid.
 3. The theory linking the two constructs might be flawed.



Reliability

1. Measurement Error

- Measurement errors are random events that distort the true measurement of a concept.
- The average of all measurement errors is equal to zero.
- The distribution of measurement error is incorporated within the standard error.
- This distribution has the form of a bell-shaped curve, identical to the distribution of the standard deviation about a mean.



Reliability

2. Measurement Error and Hypothesis Testing

- With a small standard error, even a weak relationship between the independent and dependent variables can be judged as statistically significant.
- With a large standard error, even a strong relationship between the independent and dependent variables can be judged as not statistically significant.
- Therefore, it is important to reduce measurement error as much as possible.



Reliability

3. Achieving Reliability

- Concepts must be defined as precisely as possible.
- Researchers must develop good indicators of these concepts. This task sometimes requires much work and time to create a valid and reliable measuring instrument.
- Data must be collected with as much care as possible.



Reliability

4. Assessing Reliability

- Test-Retest Procedure.
 - Advantages: Intuitive appeal.
 - Disadvantages: History, Maturation, Cueing.
- Alternative Forms Procedure.
 - Advantages: Little cueing.
 - Disadvantages: History, Maturation.



Reliability

4. Assessing Reliability (Continued)

- Split-Halves Procedure.
 - Advantages: No History, Maturation, or Cueing.
 - Disadvantages: Different splits yield different assessments of reliability.
- Internal Consistency Procedure.
 - Advantages: No History, Maturation, or Cueing.
 - Disadvantages: Number of items can affect assessment of reliability.



Errors that Affect Validity and Reliability

1. Measurement Error

- Random, but can be reduced with care.

2. Sampling Error

- Errors in selecting a sample for study.

3. Random Error

- Things that are unknown, unanticipated, uncontrollable.

4. Data Analysis Errors

- Type-I: False assumption of causality.
- Type-II: False assumption of no causality.

Questions?