VALIDITY AND RELIABILITY OF QUESTIONNAIRES

Dr. Arlene S. Opina
Research Instruments

1. Questionnaires
2. Tests
3. Observation schedule
4. Checklist
5. Rating scales
6. Interviews and etc.

Must meet the certain requirements before using them to gather data.
Three Requirements

1. Usability
2. Validity
3. Reliability
1. Usability

• Practical

• Can be used by researchers without spending much time, money, and effort
1. Usability

• Ease of administration
• Ease of scoring
• Ease of interpretation
• Low cost
• Proper mechanical make-up
  • Font size
Fill in the blank.

• **Usability** is the most ____________ quality of a good instrument.
Fill in the blank.

_________________ is the degree to which the research instrument gives proper mechanical make-up up to the subjects of the study.
Validity VS Reliability

• Validity – The results must satisfy the objectives of the test.

• Reliability - The results must be consistent.
The concept of validity

• Validity is the ability of an instrument to measure what it is intended to measure.

• Degree to which the researcher has measured what he has set out to measure (Smith, 1991)

• Are we measuring what we think we are measuring? (Kerlinger, 1973)

• Extent to which an empirical measure adequately reflects the real meaning of the concept under consideration (Babbie, 1989)
Three general categories of instrument validity

1. Content-Related Evidence (also known as Face Validity)
2. Criterion-Related Evidence
   a) Predictive Validity
   b) Concurrent Validity
3. Construct-Related Evidence
Types of validity

Validity

- Content validity
  - Face validity
- Criterion related
  - Concurrent
  - Predictive
- Construct validity
Content Validity

• The items in the questionnaire truly measure the intended purpose.
• Covers a representative sample of the behavior domain to be measured.
• All major aspects are covered by the test items in correct proportion.
• All items are relevant to all types of criteria.
• Commonly used in evaluating achievement test.
CONTENT VALIDITY

Experts: Enhancement of content of questionnaire (three or more experts)
CONTENT OR FACE VALIDITY

• Evaluate in terms of:

A research survey cannot be assumed to be reliable even if it has high face validity.
Criterion-Related Evidence by comparing the instrument with some future or current criteria

– Predictive Validity
– Concurrent Validity
Construct Validity

Example:

• Intelligence test
• Language proficiency
• Psychometric tests
  • are designed to measure candidates' suitability for a role based on the required personality characteristics and aptitude (or cognitive abilities).
Fill in the blank.

_____________________ is the degree to which the research instrument measures what it purports to measure.
Fill in the blank.
The two forms of criterion-related validity and their difference
_____________________ and
_____________________.
Fill in the blank.

My measure has __________________ validity if it allows me to successfully predict future behavioral outcomes.
Fill in the blank.

How do you obtain content-related evidence of validity? What is the most common way to do this?
RELIABILITY

• It is the ability of an instrument to create reproducible results.

• Each time it is used, similar scores should be obtained.
RELIABILITY

• A questionnaire is said to be reliable if we get same/similar answers repeatedly.

• Though it cannot be calculated exactly, it can be measured by estimating correlation coefficients.
Reliability is measured in aspects of:

**STABILITY**

- Done to ensure that same results are obtained when used consecutively for two or more times
- Test-retest method is used

**INTERNAL CONSISTENCY**

- To ensure all subparts of an instrument measure the same characteristic (Homogeneity)
  - Split-half method
  - Cronbach’s Alpha

**EQUIVALENCE**

- Used when two observers study a single phenomenon simultaneously
  - Inter-rater reliability
Cronbach’s alpha

• Developed by Lee Cronbach in 1951
• The most widely used objective measure of reliability.
• It is generally used as a measure of internal consistency of an instrument with items in rating scale form ex. Likert scale; rating scale: 1 = poor, 5 = excellent).
Cronbach’s alpha

• Generally, alpha coefficient ranges in value from 0 to 1
• Cronbach's alpha generally increases when the correlations between the items increase. For this reason the coefficient is also called the internal consistency reliability of the test.
• Values of at least 0.6 to 0.7 or higher indicate internal consistency. **PLEASE eCOPY EXCEL FORMULA from me.**
• Some professionals insist on a reliability score of 0.70 or higher in order to use a psychometric instrument. This rule should be applied with caution when $\alpha$ has been computed from items that are not correlated.
Fill in the blanks.

- ___________ is concerned with consistency of responses from moment to moment even when the subjects take the same instrument twice.
Steps in questionnaire validation
1. Establish Face Validity
2. Pilot test
3. Clean Dataset
4. Principal Components Analysis
5. Cronbach’s Alpha
6. Revise (if needed)
Step 1: Establish Face Validity

• Have your survey reviewed by two different parties.
  • The first is a group familiar with your topic who can evaluate if your questions successfully capture your topic.
  • The second review should come from someone who is an expert on question construction, ensuring that your survey does not contain common errors such as leading, confusing or double-barreled questions.
Step 2: Run a Pilot Test

• Select a subset of your intended survey participants and run a pilot test of the survey. Suggested sample sizes vary, although about 10 percent of your total population is a solid number of participants. The more participants you can round up, the better, although even a smaller sample can help you weed out irrelevant or weak questions.
**Step 3: Clean Collected Data**

- Enter your collected responses into a spreadsheet to clean the data. Having one person read the values aloud and another entering them into the spreadsheet greatly reduces the risk of error. Once data is entered, your next step is to reverse code negatively phrased questions.

- If respondents have responded carefully, their answers to questions that are phrased negatively should be consistent with their answers to similar questions that are phrased positively. If that is not the case, you may want to think about eliminating that respondent from the survey.

- Also double-check minimum and maximum values for your overall dataset. If you’ve used a five-point scale and you see a response indicating the number six, you may have an error with data entry.
Step 4: Check Internal Consistency

• You can review the internal consistency with a standard test known as Cronbach’s Alpha (CA). Test values range from 0 to 1.0, and values should generally be at least 0.6 to 0.7 or higher to indicate internal consistency. If you have a value lower than 0.6, some CA programs let you delete a question from the test to see if it improves consistency. If it does, you may want to consider deleting the question from the survey.
Step 5: Revise Your Survey

• The final stage of the validation process is to revise your survey based on the information you gathered from your principal components analysis and Cronbach’s Alpha.

• If major changes were made, especially if you removed a substantial amount of questions, another pilot test and round of PCA and CA is probably in order.