Validity and Reliability

Chapter 8

Validity and Reliability

- Validity is an important consideration in the choice of an instrument to be used in a research investigation
  - It should measure what it is supposed to measure
  - Researchers want instruments that will allow them to make warranted conclusions about the characteristics of the subjects they study
- Reliability is another important consideration, since researchers want consistent results from instrumentation
  - Consistency gives researchers confidence that the results actually represent the achievement of the individuals involved

Reliability

- Test-retest reliability
- Inter-rater reliability
- Parallel forms reliability
- Internal consistency (a.K.A. Cronbach's alpha)

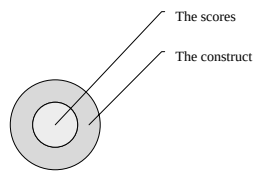
Validity

- Face
  - Does it appear to measure what it purports to measure?
- Content
  - Do the items cover the domain?
- Construct
  - Does it measure the unobservable attribute that it purports to measure?

Validity

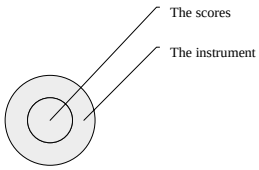
- Criterion
  - Predictive
  - Concurrent
- Consequential

Types of validity (cont.)

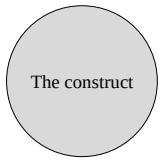


Here the instrument samples some and only of the construct

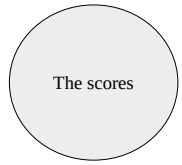
Types of validity

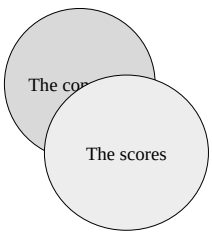


Here the instrument samples all and more of the construct



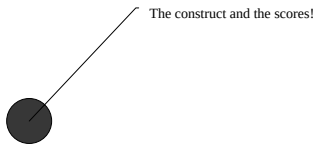
Here the instrument fails to sample ANY of the construct





Here the instrument samples some but not all of the construct

Perfection!



Reliability and Validity

The bulls-eye in each target represents the information that is desired. Each dot represents a separate score obtained with the instrument. A dot in the bulls-eye indicates that the information obtained (the score) is the information the researcher desires.

Experimental Research Designs

Inferring Causality

Sir Bradford Hill

- Strength of association
- Consistency
- Specificity
- Temporal order
- Dose-Response (biological gradient)
- Plausibility
- Experimental evidence
- Analogy

Fundamentals of Experimental and Quasi-Experimental Research

- Random selection and random assignment :
  - » Distinguish between "selection" and "assignment"
  - » Random selection helps to assure population validity
  - » If you incorporate *random assignment*

    → Experimental research
  - » If you do not use *random assignment*

    → Quasi-experimental research

Fundamentals of Experimental and Quasi-Experimental Research (cont'd.)

- When to use experimental research design :
  - » If you strongly suspect a cause-and-effect relationship exists between two conditions, ***and***
  - » The independent variable can be introduced to participants and can be manipulated, ***and***
  - » The resulting dependent variable can be measured for all participants

Internal and External Validity

- "Validity of research" refers to the degree to which the conclusions are accurate and generalizable
- Both experimental and quasi-experimental research are subject to threats to validity
- If threats are not controlled for, they may introduce error into the study, which will lead to misleading conclusions

Internal and External Validity

- "Validity of research" refers to the degree to which the conclusions are accurate and generalizable
- Both experimental and quasi-experimental research are subject to threats to validity
- If threats are not controlled for, they may introduce error into the study, which will lead to misleading conclusions

Threats to External Validity

- ***External validity***—extent to which the results can be generalized to other groups or settings
  - » *Population validity*—degree of similarity among sample used, population from which it came, and target population
  - » *Ecological validity*—physical or emotional situation or setting that may have been unique to the experiment
    - » If the treatment effects can be obtained only under a limited set of conditions or only by the original researcher the findings have low ecological validity.

Threats to Internal Validity

- ***Internal validity***—extent to which differences on the dependent variable are a direct result of the manipulation of the independent variable
  - » *History*—when factors other than treatment can exert influence over the results; problematic over time
  - » *Maturation*—when changes occur in dependent variable that may be due to natural developmental changes; problematic over time
  - » *Testing*—also known as "pretest sensitization"; pretest may give clues to treatment or posttest and may result in improved posttest scores
  - » *Instrumentation* – Nature of outcome measure has changed.

Threats to Internal Validity (cont'd.)

- » *Regression* – Tendency of extreme scores to be nearer to the mean at retest
- » Implementation-A group treated in an unintentional differential manner.
- » Attitude-Hawthorne effect, compensatory rivalry.
- » *Differential selection of participants*—participants are not selected/assigned randomly
- » *Attrition (mortality)*—loss of participants
- » *Experimental treatment diffusion* – Control conditions receive experimental treatment.

Experimental and Quasi-Experimental Research Designs

- Commonly used experimental design notation :
  - » X_1 = treatment group
  - » X_2 = control/comparison group
  - » O = observation (pretest, posttest, etc.)
  - » R = random assignment

Common Experimental Designs

- *Single-group pretest-treatment-posttest design*:

  O X O

  - » Technically, a *pre-experimental design* (only one group; therefore, no random assignment exists)
  - » Overall, a weak design
    - »Why?

Common Experimental Designs (cont'd.)

- *Two-group treatment-posttest-only design*:

  R X_1 O
  R X_2 O

  - » Here, we have random assignment to experimental, control groups
  - » A better design, but still weak—cannot be sure that groups were equivalent to begin with

Common Experimental Designs (cont'd.)

- *Two-group pretest-treatment-posttest design*:

  R O X_1 O
  R O X_2 O

  - » A substantially improved design—previously identified errors have been reduced

Common Experimental Designs (cont'd.)

- *Solomon four-group design*:

R	O	X_1	O
R	O	X_2	O
R		X_1	O
R		X_2	O

  - » A much improved design—how??
  - » One serious drawback—requires twice as many participants

Common Experimental Designs (cont'd.)

- *Factorial designs*:

R	O	X_1	γ_1	O
R	O	X_2	γ_1	O
R	O	X_1	γ_2	O
R	O	X_2	γ_2	O

  - » Incorporates two or more *factors*
  - » Enables researcher to detect *differential differences* (effects apparent only on certain combinations of levels of independent variables)

Common Experimental Designs (cont'd.)

- *Single-participant measurement-treatment-measurement designs*:

O O O | X O X O | O O O

  - » Purpose is to monitor effects on one subject
  - » Results can be generalized only with great caution

Common Quasi-Experimental Designs

- *Posttest-only design with nonequivalent groups*:

X_1	O
X_2	O

  - » Uses two groups from same population
  - » Questions must be addressed regarding equivalency of groups prior to introduction of treatment

Common Quasi-Experimental Designs (cont'd.)

- *Pretest-posttest design with nonequivalent groups*:

O	X_1	O
O	X_2	O

  - » A stronger design—pretest may be used to establish group equivalency

Similarities Between Experimental and Quasi-Experimental Research

- Cause-and-effect relationship is hypothesized
- Participants are randomly assigned (experimental) or nonrandomly assigned (quasi-experimental)
- Application of an experimental treatment by researcher
- Following the treatment, all participants are measured on the dependent variable
- Data are usually quantitative and analyzed by looking for significant differences on the dependent variable