Validity and Quantitative Research

RCS 6740
6/16/04

What is Validity?

• Valid Definition (Dictionary.com):
  – Well grounded; just: a valid objection.
  – Producing the desired results; efficacious: valid methods.
  – Having legal force; effective or binding: a valid title.

• Logic:
  – Containing premises from which the conclusion may logically be derived: a valid argument.
  – Correctly inferred or deduced from a premise: a valid conclusion.

What is Validity Cont.

• Validity Definition (Dictionary.com):
  • [Va*lid"i*ty], n. [Cf. F. validité, L. validitas strength.]
    – 1. The quality or state of being valid; strength; force; especially, power to convince; justness; soundness; as, the validity of an argument or proof; the validity of an objection.
    – 2. (Law) Legal strength, force, or authority; that quality of a thing which renders it supportable in law, or equity; as, the validity of a will; the validity of a contract, claim, or title.
What is Validity Cont.

- Additional Definitions and Quotes:
- **Validity is the best approximation to the truth or falsity of propositions** (Cook & Campbell, 1979).
- Validity is at best approximate or tentative “since one can never know what is true. At best, one can know what has not yet been ruled out as false” (Cook & Campbell, 1979, p. 37).
- And, as we have seen in examining the logic of hypothesis testing, statistical power, and the validity of outcome measures, we don’t really prove that something is false. In other words, we never really prove a null hypothesis; we only fail to reject it.

Why is Validity Important?

- “Varying degrees of ‘confirmation’ are conferred upon a theory through the number of plausible rival hypotheses available to account for the data. The fewer such plausible rival hypotheses remaining, the greater the degree of ‘confirmation’” (Campbell & Stanley, 1963, p. 36).
- Thus, research is a field of varying degrees of certainty. [Analogy: Monet’s garden.]
- And, “continuous, multiple experimentation is more typical of science than once-and-for all definitive experiments” (Campbell & Stanley, 1963, p. 3). [Analogy: Brush strokes in the painting.]

Why is Validity Important Cont?

- Threats to validity are plausible rival hypotheses (i.e., other ways of explaining the results rather than the author(s) hypothesis).
- Research design helps us to eliminate some threats to validity in individual studies.
- And, multiple studies with different participants, investigators, and conditions increase the degree of confirmation that can be accorded to a particular theory.
4 Types of Validity

- **Internal Validity:**
  - Validity with which statements can be made about whether there is a relationship between the variables in the form in which the variables were manipulated or measured.

- **Statistical Conclusion Validity:**
  - Certainty of inferences about presumed covariation of variables at specified alpha level and variances – or, in other words, relative probability that the results of the statistical tests are representative of actual relationships in the data.

- **Construct Validity:**
  - Approximate validity with which we can make generalizations about higher-order constructs from research operations.

- **External Validity:**
  - Certainty of generalizability across populations, persons, settings, times, etc.

4 Types of Validity Cont.

- There is much debate about
  - (a) which types of validity apply to which types of research,
  - (b) the relative priority of types of validity, and
  - (c) the interrelation of the types (Pedhazur & Schmelkin, 1991).
  - To some degree, the types of validity that one considers applicable depend on the definitions used for each type of validity and the type of research attempted.

4 Types of Validity Cont.

- In the broad sense, all of the types of validity may be considered applicable to all group quantitative designs.
- To a more limited degree, the four types of validity can be considered to be applicable to single subject designs.
Internal Validity

• Validity with which statements can be made about whether there is a relationship between the variables in the form in which the variables were manipulated or measured (Cook & Campbell, 1979).

Internal Validity Cont.

• Internal validity may or may not relate to whether or not a causal relationship can be established.
• Internal validity refers to our relative certainty that our outcomes resulted from what we did or what we tested (Tuckman, 1988).

Threats to Internal Validity

O (DV) X (IV) O (DV)=One group, Pretest/Posttest Design
• Study Example: This class, and this class only, takes a Validity knowledge pretest today. Today and Monday after break, this class attends lectures about Validity. The following Wednesday, this class takes a posttest to gauge Validity knowledge.
• Although this design includes a comparison (pretest to posttest), it does not account for several alternative explanations as to why the students will score the way they will on the posttest (DV)
  – History
  – Maturation
  – Testing
  – Instrument Decay
  – Regression Towards the Mean
  – Mortality
  – Selection
Threat to Internal Validity Cont.

History
- History refers to any event that occurs between the first and second test that could act as a confounding variable
- Example: A day before the posttest, all of the students in this class are kidnapped by aliens. Despising research, the aliens erase all memory of this class from the student’s brains. Needless to say, they all scored very low on the posttest.
- Any other example of History?

Threat to Internal Validity Cont.

Maturation
- Maturation refers to natural changes that individuals make over time
- These changes may effect the DV
- Example: John, an 90 year old student, was the oldest in the class. During the course of the semester, John’s memory began to rapidly deteriorate. Although he paid attention during the Validity lecture, John scored lower on the posttest than on the pretest.
- Any other examples of Maturation?

Threat to Internal Validity Cont.

Testing
- Testing refers to the notion that people may change simply because of taking a pretest
- Example: Upon realizing that she bombed the Validity pretest, Mary-Jane contacted Dr. Swett to receive extra training on the subject. Needless to say, Mary-Jane aced the validity posttest.
- Can you positively attribute the classes’ 2 days of Validity training to the rise in Mary-Jane’s score? NO
Threat to Internal Validity Cont.

**Instrument Decay**
- Instrument Decay refers to the notion that the basic characteristics of the measure (or measurement technique) change over time.
- Example: Dan became so tired of grading Validity posttest scores that he decided to give everyone a 100%.
- Example: Smokers are asked to keep a journal logging every cigarette that they smoke. At first, they are so excited that they record every single one. After a week, the task becomes so burdensome that they only log one out of three smoked.

Threat to Internal Validity Cont.

**Statistical Regression**
Regression Towards the Mean is a phenomenon where scores (both high and low) become closer towards the mean (average) as additional measures are given.

Also called Statistical Regression
- Example: The Validity pretest mean for the class was 83%. The high for the class was a 95% and the low was a 65%. The mean for the Validity posttest was also an 83%. The high was an 88% and the low was a 72%.
- Remember that any measurement score contains the true score plus measurement error. Therefore, unless a test is 100% reliable, scores that are high and low will gravitate towards the mean with time.
- Sports Illustrated Cover Jinx Example!

Threat to Internal Validity Cont.

**Mortality**
- The loss of subjects during research due to death, absence, etc.
- This is a particular problem if the treatment causes mortality; the treatment groups’ posttest mean would be contaminated by mortality.
  - Smoking Example
- Example 2: Two students in this class went away to Jamaica for the summer break, fell in love with the lifestyle there, and decided to stay there for good.
Threat to Internal Validity Cont.

Selection
• Selection occurs when subjects are assigned to treatment and control groups on a nonrandom basis. This results in the groups being different on many variables.
• Example: Being lazy, I decided to only give the males in the class the Validity training. Females in the class were given an extended summer break. Both males and females took the posttest and the males scored much higher.

Reducing Threats to Internal Validity
• Random assignment of subjects to treatment or control groups
• Holding extraneous variables constant or restricting their range
• Including extraneous variables in the design to measure their effects
• Employing methods of statistical control
• Matching subjects in the treatment and control groups on contaminating, extraneous variables (Parker, 1990)
• Note that these methods of control are listed in their order of preference.

Statistical Conclusion Validity
• Certainty of inferences about presumed covariation of variables at specified alpha level and variances (Cook & Campbell, 1979) – or, in other words, relative probability that the results of the statistical tests are representative of actual relationships in the data.
Threats to Statistical Conclusion Validity

- Low statistical power (topic of earlier lecture)
- Violated assumptions of statistical tests
- Fishing and the error rate problem (e.g., multiple F-Tests)
- Mistaken acceptance of null hypothesis
- Reliability of measures
- Reliability of treatment implementation (how much subjects learned, degree of program implementation)

Reducing Threats to Statistical Conclusion Validity

- Do a pre-analysis statistical power estimation and consider obtaining more participants, raising alpha, or using a more powerful statistical test to achieve higher power. Power of .80 is desirable (Cohen, 1988).
- Use alpha reduction procedures when running multiple comparisons.
- Be sure that the instruments that you use are reliable.
- Avoid using gain scores, or use them with appropriate caution. Ferguson and Takane (1989, pp. 474-475) and Cook and Campbell (1979, pp. 182-185) address the problems of gain scores.
- Be sure that treatments are fully implemented.
- Understand the assumptions that accompany the statistical tests that you are using and the consequences of their violation under various circumstances.
- Remember, we don’t prove the null hypothesis; we only fail to reject it.

Construct Validity

"Approximate validity with which we can make generalizations about higher-order constructs from research operations" (Cook & Campbell, 1979, p. 38).

- Construct: An attribute of an individual or a phenomenon that is not directly observable, but which is theoretically based or is inferred from empirical evidence
Threats to Construct Validity

- Inadequate preoperational explication of constructs (e.g., the wrong bait)
- Mono-operational bias (e.g., just one kind of fish net)
- Evaluation apprehension (e.g., test anxiety)
- Experimenter expectancies (e.g., the "Rosenthal effect")
- Interaction of different treatments
- Restricted generalization across constructs

Reducing Threats to Construct Validity

- Provide a good operational definition for the construct. Recall from our discussions of variables (i.e., a construct of interest) the importance of operational definitions and defined measurement scales. In other words, be specific in defining the construct of interest.
- Be sure to choose a dependent variable or outcome measure that really measures your intervention. For example, does the Zung Depression Scale really measure depression in people with stroke?
- Isolate the other constructs that can confound or confuse the issue
- USE MULTIPLE MEASURES AND MANIPULATIONS WHEN POSSIBLE (Cook & Campbell, 1979). Analogy: More than one type of fishnet, more than one type of fishing procedure

External Validity

- The certainty of generalizability across populations, persons, settings, times, etc. (Cook & Campbell, 1979; Tuckman, 1988).
Threats to External Validity

- Interaction of Testing and Treatment – the pretest increases or decreases the respondents’ responsiveness or sensitivity to the treatment; as a consequence, the results are not generalizable to the nonpretested population from which the treatment group was selected.
- Interaction of Selection and Treatment – research subjects are frequently volunteers or individuals who are prone to seek out research participation. Such persons may have traits that tend to enhance or diminish the effects of the treatment. Thus the results are not generalizable to the population of interest, which includes nonvolunteers.
- Reactive Arrangements – the treatment employed in a study, particularly if administered in an artificial, laboratory setting, may not be identical to the treatment utilized in applied settings. Therefore, the results of the research may not be generalizable to the field.

Threats to External Validity Cont.

- The following threats to external validity are explained in Pedhazur and Schmelkin (1991).
- Treatment-Attributes Interactions – The interaction between the attributes of the participant and the treatment.
- Treatment-Setting Interactions – The interaction between the setting and the treatment.
- Posttest Sensitization – There usually are no “posttests” in the real world.

Reducing Threats to External Validity

- Random selection
- Deliberate sampling for heterogeneity
- Using a “Hawthorne” control group
- Examining potential interactions
- Note: “The process of doing an experiment – that is, exercising some control over the environment – contributes to internal validity while producing some limitation in external validity.”
- “External validity is of little value without some reasonable degree of internal validity” (Tuckman, 1988, p. 6).