Validity 1

The attached checklist can help when one is evaluating the threats to validity of a study.

VALIDITY CHECKLIST

Recall that these types are only illustrative. There are many more.

INTERNAL VALIDITY

__ history
__ maturation
__ testing
__ instrumentation
__ statistical regression
__ mortality
__ selection

STATISTICAL CONCLUSION VALIDITY

__ low statistical power
__ violated assumptions of statistical tests
__ fishing and error rate (e.g., multiple F-tests)
__ mistaken acceptance of null hypothesis
__ reliability of measures
__ reliability of treatment implementation

CONSTRUCT VALIDITY

__ inadequate preoperational explication of constructs
__ mono-operational bias
__ evaluation apprehension
__ experimenter expectancies
__ interaction of different treatments
__ restricted generalization across constructs

EXTERNAL VALIDITY

__ interaction of testing and treatment
__ interaction of selection of selection and treatment
__ reactive arrangements
__ treatment-attributes interactions
__ treatment-setting interactions
__ posttest sensitization
Evaluating Articles for Threats to Validity

The following lists are not exhaustive. Rather they illustrate some ways in which designs address threats to validity and some types of threats to validity that can remain.

**Internal Validity**
Internal validity is a focus if there is an intervention.

The research design can **address** threats to validity through
- random assignment of subjects to groups (experimental or control)
- holding extraneous variables constant or restricting their range (for example, focusing only on young adults)
- including extraneous variables in the design to measure their effects (e.g., including pre-test measures to see how pre-test levels influence effectiveness of the treatment)
- employing methods of statistical control (e.g., Analysis of Covariance [ANCOVA])
- matching subjects in the treatment and control groups on contaminating, extraneous variables

Threats to validity **remain** when any of the following happen any of the following situations, described in the lecture notes, occur:
- history, maturation, testing, instrumentation, statistical regression, mortality, and selection

**Statistical Conclusion Validity**
Statistical conclusion validity is an issue whenever statistical tests are used to test hypotheses.

The research design can **address** threats to validity through
- considerations of statistical power
- alpha reduction procedures (e.g., Bonferroni technique) when multiple tests are used
- use of reliable instruments
- assurance of treatment implementation
- testing and abiding by the assumptions for valid use of statistical procedures
- never accept the null hypothesis (rather, we fail to reject it).

Threats to validity **remain** when
- there is low statistical power (this is only a threat if no significant differences are found)
- assumptions for statistical tests are violated
- multiple tests (more than 3) are conducted without alpha reduction
- a null hypothesis is falsely accepted
- measures with low reliability are used or the reliability for the sample is not reported
- reliability of treatment implementation is not assured

**Construct Validity**
Construct validity deals with what we are measuring and how we are measuring it.

The research design can **address** threats to construct validity through
- providing clear operational definitions of variables
- using multiple measures (e.g., multiple instruments) to evaluate the same construct
- using multiple ways of measuring a construct (e.g., behavioral observation and a standardized measure);
- using a measure that is free from confounding constructs and exactly measures the construct of interest (e.g., the right net)
- removing confounding constructs, like participant test anxiety and experimenter expectancies

Threats to validity **remain** when
- only one measure is used
- the measure is not clearly defined
- the construct of interest is only measured in one way
- confounding constructs are present (e.g., test anxiety, experimenter expectancies)

**External Validity**
External validity deals with the extent to which you can generalize results beyond the study sample.

The research design **addresses** external validity when
- the sample is randomly selected from a broad population
- the sample is deliberately selected for heterogeneity (e.g., multi-stage cluster sampling)
- a Hawthorne control group is used
- interactions of selection and treatment or testing and treatment are examined.

Threats to validity **remain** when
- volunteers or a purposive or convenience sample are used
- a pre-test interacts with the treatment
- the setting of the intervention is considerably different from the real world
- attribute treatment interactions are not addressed (e.g., people of different levels on the pre-test benefit differently from the intervention)
Validity 4

**RELATIONSHIP OF VALIDITY TO RESEARCH DESIGN IN QUANTITATIVE RESEARCH**

**Research design** is “the plan and structure of investigation so conceived as to obtain answers to research questions” (Kerlinger, 1986, p. 279). Its purposes are:
- to provide answers to research questions and
- to control variance (Kerlinger, 1986).

And, according to Cohen (1988), “Experimental [research] design is an area of inquiry wholly devoted to the removal of irrelevant sources of variability for the increase of precision and therefore the increase of the statistical power of tests of null hypotheses” (p. 8).

In other words, *research design is a method of reducing the alternative explanations (i.e., rival hypotheses) related to a study.*

**Recall that 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.

**Control**

Control is a word that we encounter often in research articles and research texts. Like many words that we encounter in research, it has a variety of different meanings, which depend on the purposes of the researcher.

**Control is the major tool used by research design to eliminate threats to validity (i.e., rival hypotheses).**

Control has two purposes. They are:
1. ruling out valid threats to inference
2. adding precision, that is increasing the ability (i.e., statistical power) to detect small observed effects (Cook & Campbell, 1979)

   Analogy: Decreasing the static on the phone line.

Pedhazur and Schmelkin (1991) identified four types of control: (a) manipulation, (b) elimination or inclusion, (c) statistical, and (d) randomization. Another way to categorize methods of control is:

- control of the situation, that is keeping out extraneous forces (e.g., testing students in the same room)
• control over the independent variable (i.e., the treatment) (for example, control of assignment of persons to different treatment groups, assurance of control over the nature and extent of the implementation of the treatment)
• controlling for an identified threat to validity through research design (for example, one can control for gender by only testing only one gender). Another example might involve measuring the effect of a particular program on changing participant’s attitudes towards persons with disabilities. Pre-program attitudes could be controlled for by pre-testing and controlling for statistically controlling for pretest scores (e.g., through ANCOVA) when comparing the experimental and control groups on the post-test.

Control, in whatever way it is categorized or implemented, has the major purpose of removing threats to validity. Such threats can also be conceptualized as plausible rival hypotheses.

**Putting it Together in Research**

Parker (1990) correctly observed that “the perfectly designed study exists only in textbooks; in reality there is no such thing as flawless research” (p. 620).

Before you get discouraged with all the uncertainty, think of Monet. Each research study is like a brush stroke in the painting of a field of inquiry. One study doesn’t determine the whole picture; it is but a brush stroke. Nonetheless, each brush stroke should be carefully planned.

Parker (1990) suggests a twofold approach for persons designing research.

First, avoid errors by selecting appropriate research design and statistical analyses at the planning stage.

Second, be aware of and report threats to validity.