RESEARCH VALIDITY

Winfred Arthur, Jr.
Department of Psychological and Brain Sciences
and Department of Management
Research Methods

• means of discovering truth
Research Methods

• means of discovering truth
• what is truth?
Research Methods

• means of discovering truth
• what is truth?
  – Riveda Sandhyavandanam ⇒ "There is only one truth [but] people often see it in different ways"
  – see also Ludwig Wittenstein (1889-1952) ⇒ *The metaphysics of space and motion and the wave structure of matter (WSM) gives absolute truth and meaning to language*
Research Methods

• means of discovering truth

• what is truth?
  – VALIDITY?
  – a conclusion based on a piece of research is valid when it corresponds to the actual or true state of the world
Validity

- a key (maybe THE key) criterion in the evaluation of any piece of research or test (measure)
- the appropriateness of **inferences** drawn from **data**
  - data = results of research study ➞ research validity
  - data = test scores ➞ test and measurement validity
Research Validity

• two philosophies, world views, or approaches to truth discovery
  – Donald Campbell and Donald Rubin
    – Shadish (2010; Campbell and Rubin: A primer and comparison of their approaches to causal inference in field settings. *Psychological Methods, 15*, 3-17)

• Campbell's causal model $\Rightarrow$ methodologically and design driven

• Rubin's causal model $\Rightarrow$ quantitatively and statistically driven
Research Validity ⇒ re Campbell’s causal model [CCM]

- facets of research validity
  - internal
  - external
  - statistical-conclusion
  - construct

- threats to research validity

- controls
• **internal validity** $\Rightarrow$ the validity of inferences about whether observed covariation b/n $A$ (presumed treatment) and $B$ (presumed outcome) reflects a causal relationship from $A$ to $B$, as those variables were manipulated or measured

| Threats          | Controls?
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>history</td>
<td>attrition or mortality</td>
</tr>
<tr>
<td>maturation</td>
<td>selection</td>
</tr>
<tr>
<td>testing</td>
<td>regression</td>
</tr>
</tbody>
</table>
• history effects (events outside the lab)
  ▪ observed effect between IV and DV might be due to events occurring between the pretest and posttest when these events are not the treatment of research interest

• maturation effects
  ▪ source of error in a study related to the amount of time between measurements
• testing effects
  ▪ effects due to the number of times particular responses are measured
  ▪ resulting from familiarity with the measurement instrument

• attrition or mortality effects
  ▪ the dropping out of some participants before a study is completed, causing a threat to validity
• selection effects
  ▪ result from biases associated with the selection of, and assignment of research participants into groups

• regression effects
  ▪ tendency of participants with extreme scores on first measure to score closer to the mean on a second testing
  ▪ a statistical threat
• **internal validity**

  ▪ extent to which we can infer that a relationship between two variables is causal or that absence of a relationship implies absence of cause
  
  ▪ is the observed effect real or artifactual?
  
  ▪ threats

• correction \(\Rightarrow\) randomization
• **external validity** ⇒ the validity of inferences about whether the cause-effect relationship holds over variations in persons, settings, time, treatment variables, and measurement variables

**Threats**

- population validity
- ecological validity
- temporal validity

**Controls?**
• other participants
  ▪ interaction of selection and treatment
  ▪ population validity

• other settings
  ▪ interaction of setting and treatment
  ▪ ecological validity

• other times
  ▪ interaction of history and treatment
  ▪ temporal validity
• external validity
  ▪ generalizability

• enhanced or increased \(\Rightarrow\) random sampling for representativeness

• trade-offs between internal and external validity?
• **statistical conclusion validity** ⇒ the validity of inferences about the correlation (covariation) b/n treatment and outcome

**Threats**
- low statistical power
- violation of test assumptions
- poor test reliability

**Controls?**
• low statistical power
  ▪ power analysis
    ▪ sample size (n)
    ▪ effect size (magnitude of effect)
    ▪ power (.80)
    ▪ alpha (p-value, .05)

• violations of statistical test assumptions

• poor reliability of measures

• address threats ⇒ adequate power, meet test assumptions, and use reliable measures
• **construct validity** ⇒ the validity with which inferences can be made from the operations in a study to the theoretical constructs those operations are intended to represent

<table>
<thead>
<tr>
<th>Threats</th>
<th>Controls?</th>
</tr>
</thead>
<tbody>
<tr>
<td>• loose connection between theory and study</td>
<td></td>
</tr>
<tr>
<td>• evaluation apprehension</td>
<td></td>
</tr>
<tr>
<td>• experimenter expectancies (&quot;good-subject&quot; response)</td>
<td></td>
</tr>
</tbody>
</table>
• **construct validity**
  - use of appropriate theories, theoretical labels, or models to describe and explain phenomenon being studied

Fatal driving crash involvement: Locus of control vs attribution theory

<table>
<thead>
<tr>
<th>Fatal crashes drivers</th>
<th>Locus of control measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fatal-crashes drivers = externals</td>
<td></td>
</tr>
<tr>
<td>No-crashes drivers = internals</td>
<td></td>
</tr>
</tbody>
</table>

TIME

confounds locus of control and attribution theory
• **construct validity**
  - use of appropriate theories, theoretical labels, or models to describe and explain phenomenon being studied

**Fatal driving crash involvement: Locus of control vs attribution theory**

- **Fatal-crashes drivers**
  - Locus of control measure
  - Fatal-crashes drivers = externals

- **No-crashes drivers**
  - No-crashes drivers = internals

- **All drivers**
  - Locus of control measure
  - Crashes → assess relationship b/n LoC and crashes.
  - Do externals have more fatal crashes than internals?

**TIME**

confounds locus of control and attribution theory

eliminates post-event attributions as a confound
- **threats**
  - loose connection between theory and study
  - changes in participants' behavior as a result of being studied
    - Hawthorne effect
    - "good subject" response
    - social desirability responding
    - evaluation apprehension
    - responses to experimenter expectancies, and experimenter effects
  - control or minimize threats
    - double-blind procedures
    - single-blind procedures
    - deception
• interrelatedness of different facets or dimensions of research validity

  – statistical conclusion
  – internal
  – external
  – construct
• good research design or study
  – free from threats
  – no alternative explanations
  – permits robust conclusions about relationships between study’s variables

• if research is conceptualized as the "discovery of truth", it then becomes blatantly obvious that **no single [primary] study will be sufficient to permit a claim of truth discovery**
  – at best, it can only be claimed to be another datum in the cumulative body of knowledge that slowly moves us closer to the discovery of truth
Philosophies of Causal Inference ⇒
Conditions of Causality

• contiguity
• temporal precedence
• constant conjunction

• summary
  – cause (X) must precede consequence (Y) in time; that is, X is manipulated and then Y is measured [temporal precedence and constant conjunction]
  – statistical covariation b/n X and Y must be present [contiguity]
  – alternative causes of Y must be controlled via random assignment to groups [experimental dsgn]
The Empirical Research Cycle

Research process $\Rightarrow$ summarized as 5-step sequence

1. Statement of the problem
2. Design of research study
3. Measurement of variables
4. Analysis of data
5. Conclusions from research
The Empirical Research Cycle

Research process $\Rightarrow$ summarized as 5-step sequence

1. Statement of the problem
2. Design of research study
3. Measurement of variables
4. Analysis of data
5. Conclusions from research
Research Design

• naturalness of the research setting
  – lab or field
  – "naturalness" or "artificiality" of the setting
  – field research typically employs a real-life setting, while lab research is more contrived and artificial
Research Design

• degree of control
  – confounding and extraneous variables
  – manipulation $\Rightarrow$ this is reflective of a high degree of control
  – research designs that permit manipulation are technically referred to as "experiments"
A diagram shows a matrix with two dimensions: **SETTING** and **CONTROL**.

- **SETTING** has two levels: **Artificial** and **Natural**.
- **CONTROL** has two levels: **High** and **Low**.

- **Artificial Control** High:
  - **Lab experiment**
  - Observational (expt)

- **Artificial Control** Low:
  - Quasi-experiment
  - Correlational design
  - Survey research
  - Observational (non-expt)

- **Natural Control** High:
  - Field experiment
  - Observational (expt)

This diagram illustrates the relationship between different research methodologies based on their settings and control levels.
Experimental Designs

• **experiment**
  - a research method in which the investigator manipulates a variable under carefully controlled conditions and observes whether changes occur in a second variable
  - experiments are used in order to detect *cause-and-effect* relationships

• **conditions that make a true experiment**
  - *manipulation* of independent variables
  - *random assignment* into experimental conditions (experimental & control conditions)

resultant effect ⇒ high levels of control over the who, what, when, where, and how, etc.
Experimental Designs

- random assignment
- manipulation
- high levels of control over the who, what, when, where, and how, etc.
Experimental Designs

• experimental and control groups
  – logic of the experimental approach
    • if the two groups are identical in all respects [both known and unknown factors] except for the variation created by the manipulation of independent variable, then any differences between the experimental and control groups on the dependent variable must be due to the manipulation of the independent variable
Effect of sensory-integrated building micro-adjustments on workplace productivity

Sample

Experimental
wearable sensor

Control
stnd wall-mounted

Measure DV
workplace productivity

random sample (ideally)

random assignment

manipulation
- Quasi-Experimental designs
- Correlational Designs
- Survey Research
Additional Design and Methods Issues

• time ⇒ longitudinal and cross-sectional designs
• data collection protocols ⇒ observational designs?
• secondary research designs ⇒ meta-analysis
Additional Design and Methods Issues

- control techniques and strategies
  - random assignment to groups
  - matching
  - within-subject designs
    - order and carry-over effects
    - counter-balancing
  - manipulation checks
  - instrumentation of response
    - common method variance
  - building nuisance variables into the study [moderators]
  - statistical control
  - experimenter effects and bias reduction
    - double-blind procedures
    - automation
QUESTIONS
Quasi-Experimental designs

- self-selection groups
- pre-existing groups

- nonequivalent control group designs
  - delayed control group
  - mixed factorials

- designs without control groups
  - interrupted time series
  - repeated treatment
Correlational Designs

- measurement of two or more variables and assessing relationship/association between them
  - predictive
  - concurrent
  - postdictive
Survey Research

- measurement and assessment of opinions, attitudes, and such, usually by means of questionnaires and sampling methods

**STRESS IN AMERICA™: THE STATE OF OUR NATION**

A SHARED VIEW ACROSS GENERATIONS

No matter their age, more than half of Americans believe this is the lowest point in our nation’s history that they can remember.

- **Ages 72+: OLDER ADULTS**
  - Pearl Harbor
  - World War II
  - 56%

- **Ages 33-71: BABY BOOMERS**
  - Vietnam War
  - JFK and MLK assassinations
  - 57%

- **Ages 39-52: GEN XERS**
  - Gulf War
  - Oklahoma City bombing
  - 61%

- **Ages 18-38: MILLENNIALS**
  - September 11
  - High profile mass shootings
  - 59%