Lecture 12: Structure & Logic

Dewayne E Perry
ENS 623
perry@ece.utexas.edu

Problems in Experimental Design

True Experimental Design

- Goal: *uncover causal mechanisms*
- Primary characteristic: random assignment to sampling units
- If not random, then only *Quasi Experimental*
- Without randomization, cannot rule out some systematic biases
- Types of designs
  - *Between subject designs*: sampling units are subjected to one treatment each
  - *Within subjects designs*: sampling units receive two or more treatments

Variables

- Independent and dependent
  - Dependent: effect in which the researcher is interested
  - Independent: cause of the effect
  - Any event or condition can be conceptualized as either an independent of dependent variable
- Concerned about the effects of *X* on *Y*
  - Ie, the causal effects of one on the other
  - Both in the labs and in the field

Independent Variables

- No single or standard way of classifying variables
- Useful categorization (not mutually exclusive):
  - Biological
    - Eg, affects of gender in mentoring developers
  - Environmental
    - Eg, schedule pressure and fault insertion
  - Hereditary
    - Eg, IQ effects on complexity
  - Previous training and experience
    - Eg, effects of first programming languages
  - Maturity
    - Eg, age and elegance of program structures

Dependent Variables

- Many possible flavors, literally thousands
- Eg, learning new design techniques
  - Direction of observed change
  - Amount of change
  - The ease with which change effected
  - Persistence of changes over time
- 2 general classes
  - Diffusion – fan out
    - Eg, technology insertion and adoption
  - Hierarchical variations – changes in ranking
    - Eg, changing roles in organizational structures

Hume's Classical Rules

- C&E must be contiguous in space and time
- C must be prior to E
- A constant union between C and E
- The same C always produces the same E and the same E never arises but from the same C
- Like Es imply like Cs
- Like Cs produce like Es
- Cs may have multiple components, ie subCs
- Some Cs are not complete in themselves

Distillation

- *Co-variation Rule* : cause is positively correlated with effect
- *Temporal Precedence Rule* : causes must precede effects
- *Internal Validity Rule* :all plausible alternative explanations must be ruled out
- *Reality* : must settle for best available evidence

Control

- Constancy of conditions
  - Maintaining extraneous conditions that might affect variables
  - Calibrating various elements in an experiment
- Control series
  - Expectation control
  - Behavior control
- Control condition
  - Of primary concern to us
  - Control + treatment

Mill's Method

- Method of agreement
  - If *X* then *Y*
    - If several instances of this and only *X* present
    - Then *X* is a sufficient condition for *Y*
- Method of Difference
  - If ~*X* then ~*Y*
    - If *Y* does not occur when *X* is absent
    - Then *X* is a necessary condition of *Y*
- Joint Method
  - Should lead to better, more highly justified conclusions than either method separately
- Method of Concomitant Variation
  - Relates changes in the amount or degree of change
  - *Y = f(x)*
    - *Y* is functionally related to variations in *X*
  - Eg, stronger treatments show larger effects

Practical Application

- Practical problems:
  - Rare to find perfect covariance, $r = 1.0$
  - Rare to rule out every plausible alternative hypothesis
  - Experience needed to determine adequate control conditions
- Two group design
  - Treatment: if X then Y
  - Control: if $\sim X$ then $\sim Y$

Internal Validity

- Never completely satisfied
- Campbell, Stanley and Cook
  - Standardize choice of control groups
  - Try to isolate potential invalidity sources
- Confounding effects
  - Treatment effect and some other effect cannot be separated
- Confounding sources of internal invalidity
  - H: *History*
    - takes place between pre and post test
    - May contaminate post test results
  - M: *Maturation*
    - older/wiser/better between pre/post
  - I: *Instrumentation*
    - change due to test instrument
  - S: *Selection*
    - nature of participants
    - Control over assignment may have effects

Practical Application

- Distinguish two kinds of control groups
  - No treatment
    - Ok for physical effects
    - Problems where belief may confound
  - Placebo
    - Rule out belief effects
- Practical decision is not easy which to use
  - Question of greatest interest
  - Experience or knowledge of the general area
  - Easy to make mistakes in a new area

Basic Designs

- Design 1
  - One shot case study: X O
  - H- M- I(NR) S-
    - Deficient in terms of any reasonable controls
      - History may be alternative explanation
      - Maturation not controlled for
      - May be changes in instruments or judges
      - Unknown state of participants
    - Instrumentation not a factor: no pre-measurement
- Design 2
  - One group pretest: O X O
  - Slight improvement, but no comparison

Basic Designs

- Design 3 – Solomon Design
  - True experimental, 4 group
    - I R O X O
    - II R X O
    - III R O O
    - IV R O
  - H+ M+ I+ S+
    - All well controlled for

Basic Designs

- Solomon provides elegant illustration of logic of control
  - Pretest performance scores in I & III to estimate pretest scores in II & IV
    - Requires a leap of faith even if scores the same
    - Even if differ greatly, II & IV could be equal to the mean of I & III
  - Use estimated pre-test scores to enrich factorial analysis of variance of post test scores
  - Tells us if any confounding of pre-test and treatment

Basic Designs

- Pre/post test effects:
  - I+ II- III+ IV-
- Experimental treatment effects:
  - I+ II+ III- IV-
- Pretest & X sensitization:
  - I+ II- III- IV-
  - Pretest sensitization = $\left(\overline{Y}_I - \overline{Y}_{III}\right) - \left(\overline{Y}_{II} - \overline{Y}_{IV}\right)$
- Extraneous effects:
  - I+ II+ III+ IV+

Basic Designs

- External Validity
  - All 3 suffer from the possible confounding of selection and treatment
  - Design 3 - Solomon:
    - controls for confounding of treatment and pre-test sensitization
  - Design 4:
    - I R O X O III R O O
    - IV: H+ M+ I+ S+
    - Deficient in pre-test sensitization – eg, problem in attitude change or learning experiments
  - Design 5
    - II R X O IV R O
    - IV: H+ M+ I+ S+
    - Avoids pretest sensitization issues

Basic Designs

- Within subjects designs
  - Each subject receives all treatments in turn
  - Useful in SWE/CS – repeated measures design
  - Advantages:
    - Same number of subjects used more effectively
    - Each sampling unit serves as its own control
    - Can examine relationships longitudinally
  - Difficulties:
    - Sensitization problems
      - ✓ Learning etc
    - Order of treatments may produce differences in successive measures
  - Another threat to IV in longitudinal studies
    - Regression towards mean
      - ✓ When linear relationship is imperfect
      - ✓ Eg, overweight people appear to lose weight, low IQs appear to become brighter
      - ✓ Observed when variables consists of the same measure taken at two points in time and the correlation $r < 1$

Basic Designs

- Solve threat by standard *Z* score
  - A raw score from which the sample mean has been subtracted and the difference then divided by the standard deviation
  - Regression equation: $Z_Y = r_{XY} Z_X$
    - The estimated score of *Y* is predicted from the *XY* correlation r times the standard score of *X*
    - If there is a perfect correlation, the *Z* scores will be equivalent; otherwise not if $r < 1$