Empirical Methods: Experimental Design

Research Methods - Barbara Russo

SwSE - Software and Systems Engineering

unibz
Freie Universität Bozen
Libera Università di Bolzano
Università Liedia de Bulsan

The Experiment

"Refers to that portion of research in which variables are manipulated and their effects upon other variables observed"

Campbell & Stanley 1963

Purpose of the Experiment

Exploring cause - effect relationships

The design

- The design is a schema that models the experiment
- It sets the elements in an experiments
  - The variables,
  - the treatment that determines the cause
  - the sequence of activities to perform,
  - the resource to allocate to the experiment,
  - the individuals involved in the experiment
- It predicts the type of output

Types of Experimental Designs

- Pre-Experimental
- True Experimental
- Quasi-Experimental

For further details see Neutens, pp. 76-86

Design Definitions

- A group is a set of individuals, objects, entities, … on which an experiment is run
- A group on which the treatment is supplied is called experimental group a group without any treatment is called control group

Design Notation

- **X** an independent variable that is manipulated through the treatment (X potential "cause")
- **O** a specific observation (i.e. measurement of the effect)
- **R** subjects are randomly assigned to groups
- When X's and O's appear on the same line, this means that the apply to the same subject(s)
- **(X)** treatment not relevant
- **---** dashed lines between groups indicate no random assignment of subjects to groups

Design Issues

- Internal Validity Issues
  - The basic minimum without which any experiment is non interpretable
- External Validity
  - Asks the question of generalizability:
    - To what populations, settings, treatment variables, and measurement variables can this effect be generalized?
- Construct Validity
- Content Validity

From: Campbell & Stanley, 1963

Distinction between types of validity

- External Validity
- To whom and under what circumstances can the results be generalised?

- Internal Validity
- Is the experiment, in the particular case, responsible for the observed change(s)?

Threats To Internal Validity

- Depending on the design of an experiment, there may be explanations for changes other than the treatment

An example

- Some subjects develop a piece of software that controls a NASA shuttle
- A pre-test post-test experiment has been administrated
  - The subjects are tested on the quality of their production before and after the experiment
- A treatment X has been administrated; X="use of eXtreme Programming to develop the software"
- Suppose to observe an average gain of 9 points from pretest to posttest:
  - They improved the quality of the product of 9 points

Threat to internal validity: Maturation

- Another possibility is the internal threat called Maturation:
  - Perhaps, subjects matured during the period between the pretest and posttest: they became more adult and responsible
  - Increase may be due to maturation and not to treatment

Pre-Experimental Designs

- Poor internal validity
- Useful designs to detect the rival variables to internal validity

Types of Experimental Designs

- Pre-Experimental Designs
  - One-shot case study
  - One-group pretest-posttest
  - Static-group comparison

Pre-Experimental Designs

- If X then Y
  - If the program is given,
    then the outcome occurs
- 3 types

Problems encountered when using Pre-Experimental Designs

- One Shot Case Study
  - Nothing to compare result to
- One-Group Pretest-Posttest
  - No control group – no way to tell if extraneous variables are really causing a change
- Static Group Comparison
  - Since there was no random assignment of subjects, we cannot assume groups were equally heterogeneous from the beginning

One Shot Case Study

X O

One-Group Pretest-Posttest Design

O_1 X O_2

Static-Group Comparison Design

X O

O

True Experimental Designs

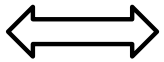
- Involve random assignment to treatments
- Strong in terms of their internal validity
- Three types

True Experimental Designs

- If X, then Y

and

- If not X, then not Y



- Also said the *if only if* design
- Example: If the treatment is given, then the outcome occurs and if the treatment is not given, then the outcome does not occur

True Experimental Designs

- True Experimental Designs
  - Posttest-only control group design
  - Pretest-Posttest control group design
  - Solomon four-group design

True Experimental Designs

Pretest-Posttest Control Group Design

R	O_1	X	O_2
R	O_3		O_4

Posttest-Only Control Group Design

R	X	O_1
R		O_2

Solomon Four-Group Design

R	O_1	X	O_2
R	O_3		O_4
R		X	O_5
R			O_6

Advantage / Disadvantages of pretest

- Can determine how much each group has gained
  - Not just whether they are different at the end of the experiment
- It can sensitise subjects to the experimental treatment
  - In effect, become part of the treatment
- Example. Experimental course
  - Subjects can gain an overview of what will be covered, in how much depth the material will be covered, etc.

Advantage of Solomon randomised four-group design

- Can compare the first two groups
  - Determine how much gain is made
- Can compare the last two groups
  - Determine whether the treatment is more effective than the control condition in the absence of a pretest (without pretest sensitisation)
  - Understand whether pretest is influencing the conclusions

Potential drawback of Solomon randomized four-group design

- Must have a reasonably large pool of subjects to begin with
  - So that when they are divided into four groups, each of the groups will have a sufficient number to yield reliable results
- Probably would be unwise to use the four-group design instead of a two-group design if you have a total pool of only 48 subjects

Quasi-experimental design

- It looks a bit like an experimental design but in general lacks the key ingredient -- random assignment.
- They give the experimental purists a queasy feeling
- With respect to internal validity, they often appear to be inferior to randomized experiments

Types of Experimental Design

- Quasi-experimental Designs
  - **Time series experiment**
  - **Non equivalent Control Group Design**
  - **Separate Sample Pretest-Posttest Design**
  - Separate Sample Pretest-Posttest Control Group Design
  - Multiple Time-Series design

Time Series Design

- Includes repeated measures on only one group before and after the treatment, represented by following diagram
- No control group and randomization

Time Series Design

O_1 O_2 O_3 O_4 X O_5 O_6 O_7 O_8

Non-equivalent Control Group Design

- This design is identical to the Pretest-Posttest Control Group design, with the exception of randomization
- Comparison groups are naturally-occurring but chosen to be as similar as possible in all experimentally-relevant aspects.

Non Equivalent Control Group Design

O_1 X O_2

O_3 O_4

Separate-samples pre test post test design

- Randomization of both samples
- One group pre test only and a (X) treatment not relevant
  - X is not relevant because no observation afterwards
  - Alone makes little sense
- and one group post test only and a treatment

Separate-samples pretest post-test Design

R O_1 (X)
R X O_4

Exercise

- Define an real case experiment that foresees each of the quasi experimental design