Experiments

ESP178 Research Methods
Dillon Fitch
1/26/16

Adapted from lecture by Professor Susan Handy
Recap – Causal Validity

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Association</td>
<td></td>
</tr>
<tr>
<td>Non-spurious</td>
<td></td>
</tr>
<tr>
<td>Time order</td>
<td></td>
</tr>
<tr>
<td>Causal Mechanism</td>
<td></td>
</tr>
<tr>
<td>Context</td>
<td></td>
</tr>
</tbody>
</table>

Experimental Designs
based on true events
THE STANFORD PRISON EXPERIMENT
IN THEATERS JULY 17
Experimental Design Terminology

• Depends on discipline

• Depends on sub-discipline

• Confusing

• Almost as bad as terminology in statistics....we’ll get to that later as well.
Experimental Design Concepts

• Treatment (Condition)- unique IV of concern that can change
  • Levels- discrete measures of the treatment

• Experimental unit – what gets “treated”

• Manipulation- change in the treatment level
  • Experimenter controlled (Intervention experiment)
  • Non-experimenter controlled (Natural experiment)

• Assignment- treatments given to experimental units
  • Random
  • Non-Random (e.g. matching, counterbalanced)

• Control
  • Experimental (manipulation, assignment, exclusion)
  • Statistical (inclusion) – “Control is out of control” Richard McElreath
Experimental Design Classification

- Quasi vs. “True”
- Between-Subject (parallel groups) vs. Within-Subject (repeated measures)
- By name (e.g. factorial, randomized block, etc.)
Experimental Design

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Aspect of Experimental Design</th>
</tr>
</thead>
<tbody>
<tr>
<td>Association</td>
<td>Treatment group versus control group</td>
</tr>
</tbody>
</table>

http://0ptimus.com/what-do-we-mean-when-we-say-experiment/?r=1&l=ri&fst=0
### Experimental Design

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Aspect of Experimental Design</th>
</tr>
</thead>
<tbody>
<tr>
<td>Association</td>
<td>Treatment group versus control group</td>
</tr>
</tbody>
</table>

*Diagram showing a control group versus a group labeled 'out of control group.'*
## Experimental Design

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Aspect of Experimental Design</th>
</tr>
</thead>
<tbody>
<tr>
<td>Association</td>
<td>Treatment group versus control group</td>
</tr>
<tr>
<td>Nonspuriousness</td>
<td>Randomly assign participants to treatment and control</td>
</tr>
</tbody>
</table>

![Diagram of experimental design](image)
Experimental Design

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Aspect of Experimental Design</th>
</tr>
</thead>
<tbody>
<tr>
<td>Association</td>
<td>Treatment group versus control group</td>
</tr>
<tr>
<td>Nonspurousness</td>
<td>Randomly assign participants to treatment and control</td>
</tr>
</tbody>
</table>
## Experimental Design

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Aspect of Experimental Design</th>
</tr>
</thead>
<tbody>
<tr>
<td>Association</td>
<td>Treatment group versus control group</td>
</tr>
<tr>
<td>Nonspuriousness</td>
<td>Randomly assign participants to treatment and control Match control group members to treatment group members</td>
</tr>
</tbody>
</table>

![Match](GROUP_1.png) ![Equivalent groups](GROUP_2.png)
## Experimental Design

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Aspect of Experimental Design</th>
</tr>
</thead>
<tbody>
<tr>
<td>Association</td>
<td>Treatment group versus control group</td>
</tr>
<tr>
<td>Nonspuriousness</td>
<td>Randomly assign participants to treatment and control</td>
</tr>
<tr>
<td>Time Order</td>
<td>Measure before and after treatment</td>
</tr>
</tbody>
</table>
Three Basic Experimental Designs

Notation

R/N = randomize/non-randomize assignment
O = observation
X = treatment
E = effect size
Between-Subject Post-Test Experiment

\[ E = O_1 - O_2 \]

Treatment Group

Control Group
Visualizing Effect Size

Treatment effect = 10

Treatment Group
Control Group

Treatment effect = 10
Between-Subject Post-Test Experiment

\[ R \xleftarrow{X} O_1 \quad O_2 \quad E = O_1 - O_2 \]

- Pros:
  - Simple
  - Handles internal validity well:
    - Maturation, Testing, Regression to the mean, Selection, pre-post test interaction

- Cons:
  - Need large sample (for Randomization to work)
  - Time order not fully accounted for
  - Mortality may still be an issue
  - Individual differences must be expected to washout (Intra subject variability)
Between-Subject Pre and Post-Test Experiment

\[ E = (O_2 - O_1) - (O_4 - O_3) \]
Visualizing Effect Size

Treatment effect = 15

Dependent (Response) Variable

Time
Between-Subject Pre and Post-Test Experiment

\[ E = (O_2 - O_1) - (O_4 - O_3) \]

- **Pros:**
  - Simple
  - Handles internal validity well:
    - Maturation, Testing, Regression to the mean, Selection
  - Time order
- **Cons:**
  - Need large sample (for Randomization to work)
  - Pre-post test interaction
  - Mortality may still be an issue
  - Individual differences must be expected to washout (Intra subject variability)
Within-Subject Experiment
(AKA Repeated Measures, AKA Crossover)

\[ E(X_1) = \frac{1}{n_A} \sum_{i=1}^{n_A} (O_{2i} - O_{1i}) + \frac{1}{n_B} \sum_{i=1}^{n_B} (O_{6i} - O_{4i}) \]

\[
\frac{1}{n_A} \left( \sum_{i=1}^{n_A} O_{2i} - \sum_{i=1}^{n_A} O_{1i} \right) + \frac{1}{n_B} \left( \sum_{i=1}^{n_B} O_{6i} - \sum_{i=1}^{n_B} O_{4i} \right)
\]

A Group

B Group
Visualizing Effect Size

**Group A:**
- X1 effect = 2
- X2 effect = 1

**Group B:**
- X1 effect = -1
- X2 effect = 4.5

**E:**
- X1 effect = 0.5
- X2 effect = 2.75
Within-Subject Experiment
(AKA Repeated Measures, AKA Crossover)

O₁ X₁ O₂ X₂ O₃
R/N
O₄ X₂ O₅ X₁ O₆

A Group
B Group

• Pros:
  • Small sample sizes OK (individuals act as own control)
  • Cost efficient
  • Handles mortality well
  • Confounding covariates reduced (Individual difference are OK)

• Cons:
  • Time based effects:
    • Order effects
    • Carry-over effects
    • Learning effects (testing)
  • Pre-post test interaction (for all treatments)
  • Regression to the mean
  • Maturity
My crossover study

What is the relationship between the road environment and bicyclist stress?
Stress and Heart Rate Variability (HRV)

- HRV: variability in beat-to-beat time intervals, and shown to be related to stress
My crossover study

Sample: 20 Female Undergrads

Dependent Variable:
Heart Rate Variability (SDNN)

Independent Variables:
Experimental Condition
Exertion
Personal Characteristics
Weather
etc....
My crossover study

Treatments: B St.  Russell Blvd.

Counterbalanced Treatments
Simulated Effect Size 😊
Can we be sure that the experiment has the effect we think it has?
I.e. Are our inferences correct?

<table>
<thead>
<tr>
<th>Judgment of Null Hypothesis ((H_0))</th>
<th>Table of error types</th>
<th>Null hypothesis ((H_0)) is Valid/True</th>
<th>Null hypothesis ((H_0)) is Invalid/False</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reject</td>
<td></td>
<td>Type I error (False Positive)</td>
<td>Correct inference (True Positive)</td>
</tr>
<tr>
<td>Fail to reject</td>
<td></td>
<td>Correct inference (True Negative)</td>
<td>Type II error (False Negative)</td>
</tr>
</tbody>
</table>
I.e. Are our inferences correct?

**Type S (Sign) error:**
Are we inferring the correct direction of treatment effect?

**Type M (Magnitude) error:**
Are we inflating (possibly deflating) the real magnitude of the treatment effect?

Source: Gelman and Carlin (2014)
Susan Handy’s “quasi-longitudinal” Natural Experiment

Movers

Do your children play outside more or less than they did before you moved?

Non-movers

Do your children play outside more or less than they did a year ago?
# Threats to Internal Validity for Experiments

<table>
<thead>
<tr>
<th>Threat</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Selection bias</td>
<td>when characteristics of treatment and control groups differ</td>
</tr>
</tbody>
</table>
Threats to Internal Validity for Experiments

<table>
<thead>
<tr>
<th>Threat</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Selection bias</td>
<td>when characteristics of treatment and control groups differ</td>
</tr>
<tr>
<td>Endogenous change</td>
<td>when subjects develop or change independent of treatment – testing, maturation</td>
</tr>
</tbody>
</table>

Testing
## Threats to Internal Validity for Experiments

<table>
<thead>
<tr>
<th>Threat</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Selection bias</td>
<td>when characteristics of treatment and control groups differ</td>
</tr>
<tr>
<td>Endogenous change</td>
<td>when subjects develop or change independent of treatment – testing, maturation</td>
</tr>
</tbody>
</table>

### Maturation

![Maturation Diagram](image)
Threats to Internal Validity for Experiments

<table>
<thead>
<tr>
<th>Threat</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Selection bias</td>
<td>when characteristics of treatment and control groups differ</td>
</tr>
<tr>
<td>Endogenous change</td>
<td>when subjects develop or change independent of treatment – testing, maturation</td>
</tr>
<tr>
<td>History effects</td>
<td>when something occurs during the experiment that influences outcomes</td>
</tr>
</tbody>
</table>
## Threats to Internal Validity for Experiments

<table>
<thead>
<tr>
<th>Threat</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Selection bias</td>
<td>when characteristics of treatment and control groups differ</td>
</tr>
<tr>
<td>Endogenous change</td>
<td>when subjects develop or change independent of treatment – testing, maturation</td>
</tr>
<tr>
<td>History effects</td>
<td>when something occurs during the experiment that influences outcomes</td>
</tr>
<tr>
<td>Contamination</td>
<td>when one group is aware of the other group and is influenced</td>
</tr>
</tbody>
</table>
## Threats to Internal Validity for Experiments

<table>
<thead>
<tr>
<th>Threat</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Selection bias</strong></td>
<td>when characteristics of treatment and control groups differ</td>
</tr>
<tr>
<td><strong>Endogenous change</strong></td>
<td>when subjects develop or change independent of treatment – testing, maturation</td>
</tr>
<tr>
<td><strong>History effects</strong></td>
<td>when something occurs during the experiment that influences outcomes</td>
</tr>
<tr>
<td><strong>Contamination</strong></td>
<td>when one group is aware of the other group and is influenced</td>
</tr>
<tr>
<td><strong>Treatment misidentification</strong></td>
<td>a process the researcher is not aware of is responsible for apparent effect of treatment – expectancy effect (nocebo vs. placebo), observer effect</td>
</tr>
</tbody>
</table>
# Threats to Internal Validity for Experiments

<table>
<thead>
<tr>
<th>Threat</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Selection bias</td>
<td>when characteristics of treatment and control groups differ</td>
</tr>
<tr>
<td>Endogenous change</td>
<td>when subjects develop or change independent of treatment – testing, maturation</td>
</tr>
<tr>
<td>History effects</td>
<td>when something occurs during the experiment that influences outcomes</td>
</tr>
<tr>
<td>Contamination</td>
<td>when one group is aware of the other group and is influenced</td>
</tr>
<tr>
<td>Treatment misidentification</td>
<td>a process the researcher is not aware of is responsible for apparent effect of treatment – expectancy effect (nocebo vs. placebo), observer effect</td>
</tr>
</tbody>
</table>
Threats to Internal Validity for Experiments

<table>
<thead>
<tr>
<th>Threat</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Selection bias</td>
<td>when characteristics of treatment and control groups differ</td>
</tr>
<tr>
<td>Endogenous change</td>
<td>when subjects develop or change independent of treatment – testing, maturation</td>
</tr>
<tr>
<td>History effects</td>
<td>when something occurs during the experiment that influences outcomes</td>
</tr>
<tr>
<td>Contamination</td>
<td>when one group is aware of the other group and is influenced</td>
</tr>
<tr>
<td>Treatment misidentification</td>
<td>a process the researcher is not aware of is responsible for apparent effect of treatment – expectancy effect (nocebo vs. placebo) observer effect</td>
</tr>
</tbody>
</table>

Control group gets a fake treatment

![Control Group vs. Treatment Group](image.png)
# Threats to Internal Validity for Experiments

<table>
<thead>
<tr>
<th>Threat</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Selection bias</td>
<td>when characteristics of treatment and control groups differ</td>
</tr>
<tr>
<td>Endogenous change</td>
<td>when subjects develop or change independent of treatment – testing, maturation</td>
</tr>
<tr>
<td>History effects</td>
<td>when something occurs during the experiment that influences outcomes</td>
</tr>
<tr>
<td>Contamination</td>
<td>when one group is aware of the other group and is influenced</td>
</tr>
<tr>
<td>Treatment misidentification</td>
<td>a process the researcher is not aware of is responsible for apparent effect of treatment – expectancy effect (nocebo vs. placebo), observer effect</td>
</tr>
</tbody>
</table>

Feeling expected to...
# Threats to Internal Validity for Experiments

## Definition

<table>
<thead>
<tr>
<th>Threat</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Selection bias</td>
<td>when characteristics of treatment and control groups differ</td>
</tr>
<tr>
<td>Endogenous change</td>
<td>when subjects develop or change independent of treatment – testing, maturation</td>
</tr>
<tr>
<td>History effects</td>
<td>when something occurs during the experiment that influences outcomes</td>
</tr>
<tr>
<td>Contamination</td>
<td>when one group is aware of the other group and is influenced</td>
</tr>
<tr>
<td>Treatment misidentification</td>
<td>a process the researcher is not aware of is responsible for apparent effect of treatment – expectancy effect (nocebo vs. placebo), observer effect</td>
</tr>
<tr>
<td>Mortality</td>
<td>Participants drop out at differential rates by group for systematic reason</td>
</tr>
</tbody>
</table>
Careful design and execution of experiments is critical!
Replication is important!
Can we be sure that the effect is the same beyond the experiment?
Generalizability – more next week
Interaction of testing and treatment
But we need them!
“careful before-and-after studies of policy interventions to promote more compact, mixed-used development to help determine what works and what does not”

“Natural experiments”

“Intervention studies”

“Policy evaluation”
Intervention Studies = Researcher designs and implements the treatment

Natural Experiments = Researcher does not control the treatment
Davis Natural Experiments

**Target Store opening:**
Shopping VMT before and after

**Fifth Street Road Diet:**
Mode split to downtown and bike/ped safety before and after

**Cannery Project:**
VMT and active travel before-and-after moving into net-zero energy community
Natural experiments for programs

- Educational programs
- Promotional programs
- Bike give-away programs
Natural Experiments for Infrastructure

Green Lane Project
Portland State University
5 cities

Expo Line Opening
UC Irvine, USC
1 line
To Test Housing Program, Some Are Denied Aid


“Half of the test subjects — people who are behind on rent and in danger of being evicted — are being denied assistance from the program for two years, with researchers tracking them to see if they end up homeless.”

Moving to Opportunity

“Moving to Opportunity for Fair Housing (MTO) is a 10-year research demonstration that combines tenant-based rental assistance with housing counseling to help very low-income families move from poverty-stricken urban areas to low-poverty neighborhoods.”

**Treatment group:** Randomly selected households with children receive housing counseling and vouchers that must be used in areas with less than 10 percent poverty

**Control groups:** One already receiving vouchers, one just coming into voucher program

Study progression

Cross-sectional studies
Establish associations
Basis for designing interventions

Intervention studies
Before-and-after measures
Establish causal relationship
To do

• Work on assignment 2!!!
• Office hours Friday
• Experiments exercise on Tuesday
• Read articles for Friday section – see website
• Don’t forget lecture notes on website!
Refining your conceptual model

• Your stated question should match your model
• Keep your question general, even if it is motivated by a specific place
• Make sure your variables are consistent with your level of analysis
• Do you have a “policy lever”? 
• Simplify where possible, but include the important control variables!