Practical Methods for Designing Research Studies and Analyzing Data in Educational Settings

Whitney Zimmerman
Assistant Teaching Professor, Dept. of Statistics, Penn State
Measuring Intervention Impact

Randomized Experiments

Pre-Test / Post-Test

Quasi-Experiments

Propensity Scoring
Randomized Experiments

Experiment + Randomization

A research study where the researchers randomly assign participants to different treatment conditions
Randomized Experiments

**Pros**

Causal conclusions

**Cons**

Still need to worry about sampling issues and threats to validity
Randomized Experiments

Pros
- Causal conclusions

Cons
- Still need to worry about sampling issues and threats to validity

Population → Sampling → Sample → Inference → Population
Randomized Experiments

Pros
Causal conclusions

Cons
Still need to worry about sampling issues and threats to validity
Randomized Experiments

Pros
Causal conclusions

Cons
Still need to worry about sampling issues and threats to validity

Threats to Validity
- Diffusion of Treatment
- Hawthorne Effect
- Novelty Effect
- Experimental Mortality
Randomized Experiments

Pros
Causal conclusions

Cons
Still need to worry about sampling issues and threats to validity

Threats to Validity
Diffusion of Treatment
Randomized Experiments

**Pros**
- Causal conclusions

**Cons**
- Still need to worry about sampling issues and threats to validity

**Threats to Validity**
- Hawthorne Effect

---

This Photo by Unknown Author is licensed under CC BY-SA
Randomized Experiments

Pros
Causal conclusions

Cons
Still need to worry about sampling issues and threats to validity

Threats to Validity
Novelty Effect
Randomized Experiments

**Pros**

Causal conclusions

**Cons**

Still need to worry about sampling issues and threats to validity

**Threats to Validity**

Experimental Mortality
Randomized Experiments

**Pros**

Causal conclusions  
Works well with shorter studies

**Cons**

Still need to worry about sampling issues and threats to validity  
More challenging with longer studies and those not in laboratory-type settings
Measuring Intervention Impact

Randomized Experiments

Pre-Test / Post-Test

Quasi-Experiments

Propensity Scoring
Pre-Test / Post-Test

Repeated Measures

 Pros

Greater statistical power
Smaller sample size
Logistically straightforward

 Cons

No control group
History
Maturation
Testing effect
## Pre-Test / Post-Test

### Solomon Four Group Design

<table>
<thead>
<tr>
<th>Group</th>
<th>Pre-Test</th>
<th>Treatment</th>
<th>Post-Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group #1</td>
<td>Pre-Test</td>
<td>Treatment</td>
<td>Post-Test</td>
</tr>
<tr>
<td>Group #2</td>
<td>Pre-Test</td>
<td>Post-Test</td>
<td></td>
</tr>
<tr>
<td>Group #3</td>
<td>Treatment</td>
<td>Post-Test</td>
<td></td>
</tr>
<tr>
<td>Group #4</td>
<td></td>
<td></td>
<td>Post-Test</td>
</tr>
</tbody>
</table>

### Pros
- Control groups allow for comparisons
- Can determine if there is a testing effect

### Cons
- More complicated
- Still need to worry about sampling issues and threats to validity
- Randomization required for a causal conclusion
Measuring Intervention Impact

Randomized Experiments

Pre-Test / Post-Test

Quasi-Experiments

Propensity Scoring
Quasi-Experimental Design

Lacks Randomization

§ 001

§ 002

Old Curriculum

New Curriculum

Helping you deliver on your online promise
qualitymatters.org
Quasi-Experimental Design

**Pros**

- Logistically easy to implement
- Consistent with real-life: “Ecological validity”
- Can avoid potential ethical issues

**Cons**

- Still need to worry about sampling issues and threats to validity
- Cannot make casual conclusions
Measuring Intervention Impact

Randomized Experiments

Pre-Test / Post-Test

Quasi-Experiments

Propensity Scoring
**Propensity Score Analysis**

Research question: Does using our tutoring service improve final exam scores?

In my class, the better performing students are the ones who use the tutoring services.

Traditional-aged students probably have more time than adult learners to use the tutoring services and they usually do better.

Used Tutoring

\[ \bar{x} = 92 \]

Declined

\[ \bar{x} = 88 \]
# Propensity Score Analysis

**Propensity score:** Probability of receiving a treatment given baseline characteristics

<table>
<thead>
<tr>
<th>Covariates</th>
<th>Used Tutoring</th>
<th>Declined</th>
</tr>
</thead>
<tbody>
<tr>
<td>GPA</td>
<td>3.5 GPA</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>24 years old</td>
<td></td>
</tr>
<tr>
<td>Employment status</td>
<td>Work part-time</td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td>Woman</td>
<td></td>
</tr>
<tr>
<td>Credits completed</td>
<td>32 credits completed</td>
<td></td>
</tr>
<tr>
<td>Credits currently enrolled</td>
<td>Enrolled in 12 credits</td>
<td></td>
</tr>
<tr>
<td>Online courses completed</td>
<td>4 online courses completed</td>
<td></td>
</tr>
</tbody>
</table>

Propensity score = 0.75
Propensity Score Matching

Used Tutoring

Declined
Propensity Score Analysis

**Pros**

Observational data can be analyzed in a way that “mimics” a randomized experiment (Austin, 2011)

Stronger conclusions

---

**Cons**

Requires data from more variables

Some baseline similarities are necessary

Does not guarantee comparable groups

More complicated

Practical Methods for Designing Research Studies and Analyzing Data in Educational Settings

Whitney Zimmerman
waz107@psu.edu
https://www.linkedin.com/in/waz107/