PSC 200
Hypotheses and Research Design

We want to formalize how to evaluate claims, statements, and research; and how we should conduct our own research so that it is less susceptible to criticism.

• How to formulate hypotheses (last lecture)

• Research design
  o Control groups
  o Randomization

• Threats to validity
  o Threats to internal validity
  o Threats to external validity
Want to make inferences concerning the “effect” of some variable/test/treatment/intervention on some other variable.

Examples: What is the effect of...

- Lower speed limit on traffic deaths
- Gun control on violent crime rates
- Democracy on war proneness
- New drug on arthritis pain
- Smaller class size on student learning/performance
- Prayer on world peace
- Presidential policy on the economy’s performance
- Shampoo Z on the luster of one’s hair
Two important aspects of political science research:

- Characteristics of people, institutions, and political phenomena:

- Relationships among people, institutions, and political phenomena

As previously mentioned, hypotheses can be formulated and tested for either.

Research design is a set of procedures for determining the “effects” of independent variables on dependent variables.

- Set of procedures used to test an hypothesis about a relationship

  H: A state with mandatory background checks will have fewer handgun-related homicides than if it does not implement background checks.
What is an “effect”?

Suppose we wanted to determine the effect of cutting off a person’s legs on their height. Consider a single person

\[ Y_{\text{before}} \quad X \quad Y_{\text{after}} \]

Ht before \quad cut off legs \quad Ht after

The effect of cutting of his legs would be the difference in ht = \( Y_{\text{after}} - Y_{\text{before}} \)
Now suppose we wanted to determine the average effect over a large number of people:

\[
\begin{array}{ccc}
Y_{\text{before},1} & X & Y_{\text{after},1} \\
Y_{\text{before},2} & X & Y_{\text{after},2} \\
Y_{\text{before},3} & X & Y_{\text{after},3} \\
\vdots & \vdots & \vdots \\
Y_{\text{before},N} & X & Y_{\text{after},N} \\
\end{array}
\]

Ht before \quad \text{cut off legs} \quad \text{Ht after}

Average effect of cutting of their legs would be the average difference in height

\[
= \frac{1}{N} \sum_{i=1}^{N} (Y_{\text{after},i} - Y_{\text{before},i})
\]
Depending on how we obtain our data and conduct our measurements, we might have to define “effect” slightly differently.

Example:

- Suppose instead that we measure the height of N people and calculated the average height.
- Then chopped off their legs.
- Then measured their height again and calculated the average height.
- But we didn’t keep track of each individual person’s measurements – just the before and after averages.

Effect of chopping off legs on height

\[
\frac{1}{N} \sum_{i=1}^{N} Y_{after,i} - \frac{1}{N} \sum_{i=1}^{N} Y_{before,i}
\]

When same number of people in both groups. This is same as previous. Commonly find different # in groups. “effect” here is the difference in means.
To make inferences concerning effect, we need at least two samples of data:

Before/After or Pre/Post

Examples:
- Effect of speed limits on auto fatalities in a given state
- Effect of presidential policy on economy
With/Without  a.k.a.  Control Group Comparison

Examples:
- Effect of a drug
- Effect of democracy on war involvement
Three broad types of “designs”

- **True Experiments**
  - Full control over sample selection, control groups (exposure), measurement, and analysis

- **Quasi-experiments**
  - No control over sample selection or control groups, but can determine measurement and analysis

- **Non-experimental**
  - Analyze existing data

More experimental control is usually better, but does not guarantee validity.
Threats to validity

Studies or claims can be criticized based on analytical methods.

However, here, we will focus on the research design

Main problems:

• Non-representative sample
  
  o Does the sample represent the population?
  o How generalizable are the results?

• Something else is responsible for the effect – rival explanations

Main fixes

• Randomization – individuals for a sample are randomly drawn and with equal probability
• Control group(s)
Threats to validity

**Internal validity**
Are our claims concerning the effects of the treatments valid in this particular experiment?

**External validity**
Can our results from this experiment be generalized to other populations and settings?

Let’s look at some threats to internal validity…
Example: The RIAA claims that illegal downloading of music on the internet has led to a precipitous drop in CD sales.

What is the implied hypothesis here?

What is the basic research design?

What are possible problems with their claim?

If you were omnipotent, what would be a better research design?
The Big Picture

DVDs and Videogames Make Big Gains

CD Prices Rise

Number of New CD Titles Falls

Sources: CDs & cassettes, RIAA; DVDs, Ernst & Young; videogames, Screen Digest in conjunction with the Entertainment & Leisure Software Publishers Association

These graphs show that more than just downloading and disc pirating may be contributing to the recent slump in CD sales. For instance, the graph above shows that, worldwide, people are spending more of their entertainment budget on DVDs, which cost slightly more than CDs, and videogames, which cost considerably more.
Threats to Internal Validity (Campbell & Stanley)

**History**
Other contemporaneous event(s) cause change in Y, not treatment.

Ex: Rise in DVD and game purchases.
Example:

At the age of 5 you are given an algebra exam. You fail it miserably. Your embarrassed parents think you don’t work hard enough. To supplement your normal education, they make you do 100 pushups every day until you are 15. At that point, you take the algebra exam again and pass. Your parents claim it was because they made you work harder every day.

What is the dependent variable?

What is the independent variable?

What is the underlying research design?

What are possible problems with their claim?
Threats to Internal Validity (Campbell & Stanley)

History  Other contemporaneous event(s) cause change in Y, not treatment. Ex: Rise in DVD and game purchases.

Maturation  Natural aging/process causes change in Y. Ex: Getting smarter or more experienced over time.

Testing  Change in Y due to testing or pretest reactivity
• Pretest may bias subsequent responses

Instrumentation  Something happens to the observers or measurement device that causes a change in Y Ex: fatigue, change in observers (different people)
Example:

A recent Kaplan advertisement features an excited testimonial by one of their students who scored 16 points higher on the LSATs after taking the Kaplan program.

What are the dependent and independent variables?

What is the implied research design?

What are possible problems with it?

What would be a better research design?
Threats to Internal Validity (Campbell & Stanley)

Statistical Regression
Regression to the mean. Groups selected based on extreme scores will tend to display lower scores next time.

EX: “There’s no reason to quit smoking. My uncle smoked from the time he was 13 and he lived to be 82.”
September 5, 2003

- Gunmen Wound Three at Baghdad Mosque | Iraq Trip | More
- Democrats Rip President Bush During Official Party Debate
- Dad Kills 14-Year-Old Son in Ambush, Then Kills Himself
- FBI Warns Al Qaeda May Try to Poison Food or Water
- Head of U.S. Bishops Rejects Plea to Allow Married Priests
- Bryant’s Lawyers Ask More Hospitals for Accuser’s Records
- Old Pros Trade Business Suits for Tennis Whites at U.S. Open
- Football Fever, From Halftime to Instant Replay | ’Skins | Video
- Former Child Star Macaulay Culkin Returns In Gritty New Film
- ‘Who You Gonna Call?’ Ghostbusters Sent to Ky. Police Station

- Disco Lights Luring Turtles to Their Deaths in Greece
- This Year’s Top Vacation Spot Is Closer to Home Than You Think
- Quote of the Day: Price of a Child’s Life
September 18, 2003

- U.S. Troops in Central Iraq Take Heavy Fire in Ambush | More
- EXCLUSIVE: Inside Group’s Attack on U.S. Troops in Iraq | Video
- Bush: Arafat ‘Failed as Palestinians’ Leader’ | Death Vow | More
- Cops: Tenn. School Gunman Left Note Saying He ‘Wanted to Die’
- NASA Readies Atlantis for Next Mission Under Intense Scrutiny
- NYSE Expects Changes After Chief Resigns | Grasso | Markets
- Why Does Pres Bush Want to Give You $3,000? | Jobless Rate
- Calif. Charges Former Health Chief With Exposing Lover to HIV
- Air Force Charges Cadet With Running Porn Site | Misconduct
- School Hands Out Tough Punishment for Alleged Team Hazing
- Family Tragedy Fuels Wrestler’s Fight to Knock Down Disease

- Overworked Monkeys Say: ‘Take This Task and Shove It!’
- Busy Beyonce Fighting Temptations of ‘Bootylicious’ Babble
- Quote of the Day: ‘Not a Hurricane They’ve Experienced Before’
Threats to Internal Validity (Campbell & Stanley)

**Statistical Regression**
Regression to the mean. Groups selected based on extreme scores will tend to display lower scores next time.

**Sample Selection**
Difference in comparison groups is not due to treatment, but to the fact that the groups were different from the start.

**Experimental Mortality**
Effect is due to a differential loss in comparison groups.
One Cure for Threats to Internal Validity

Randomization ensures 1b and 2b are equivalent

Control group 2 controls for temporal threats
Threats to External Validity

In general, anything that makes the effects found for the experiment groups unrepresentative of effects on a larger population or in a different setting.

Reactivity/Interaction of Testing with X

Pretest increases or decreases the sensitivity to the treatment. Results from the pre-tested group are therefore not representative of the larger population.

Ex: Persuasive effect of movies. Pre-movie questionnaire decreased effect of movie relative to group without pre-test.

Selection

Groups not representative of the larger population.
Ex: Political Science professors

Experimental Arrangements

Experimental setting is too artificial
Ex: Simulated vs real combat.
Multiple Treatment Interference
Treatment “memory.” Can’t erase effect of prior treatments on subject. Prior treatments may have a tendency to affect later treatments.
Ex: Drug tolerance

Extrapolating to Different Values of X
Setting is not within bounds of experimental values of treatment.
Ex: Suppose drug dosage ranges from 0 to 100 and that a study only administers dosages from 0 to 20. Suppose “general” claims are made about the effect of the drug. Claims outside 0 to 20 must be considered an extrapolation.