Week 11 Lecture - Experimental Confounds

Undergraduate Research Methods in Psychology

Quinton Quagliano, M.S., C.S.P

Department of Psychology

Table of Contents

1 Learning Objectives

1.1 Textbook Objectives

- Interrogate a study and decide whether it rules out twelve potential threats to internal validity.
- Describe how researchers can design studies to prevent internal validity threats.
- Interrogate an experiment with a null result to decide whether the study design obscured an effect or whether there is truly no effect to find.
- Describe how researchers can design studies to minimize possible obscuring factors.

1.2 Professor's Objectives

- Describe the importance of comparison groups in ruling out threats to validity
- Explain an example for each type of validity threat

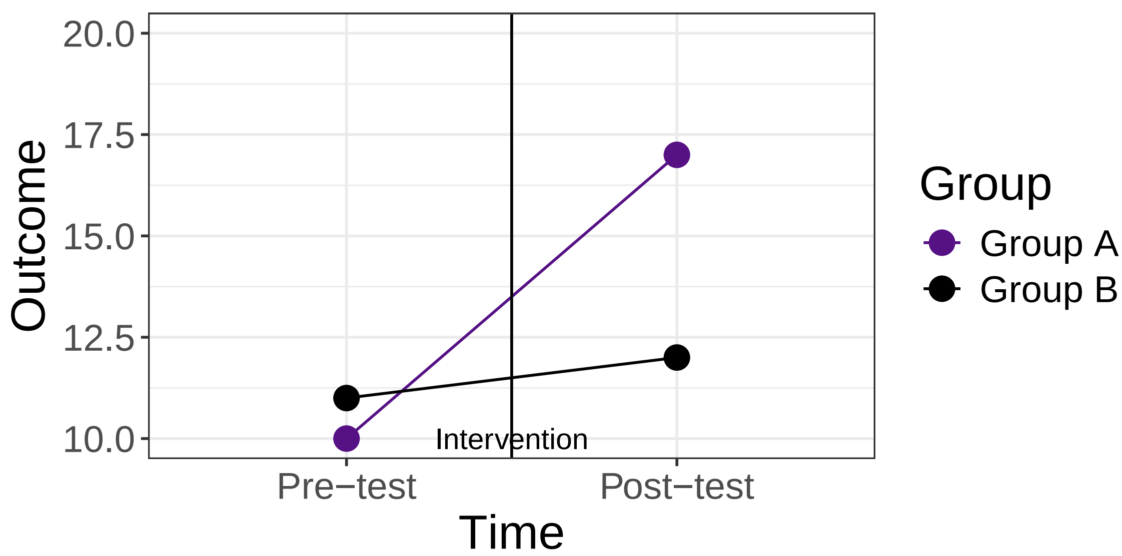
2 Chapter Overview

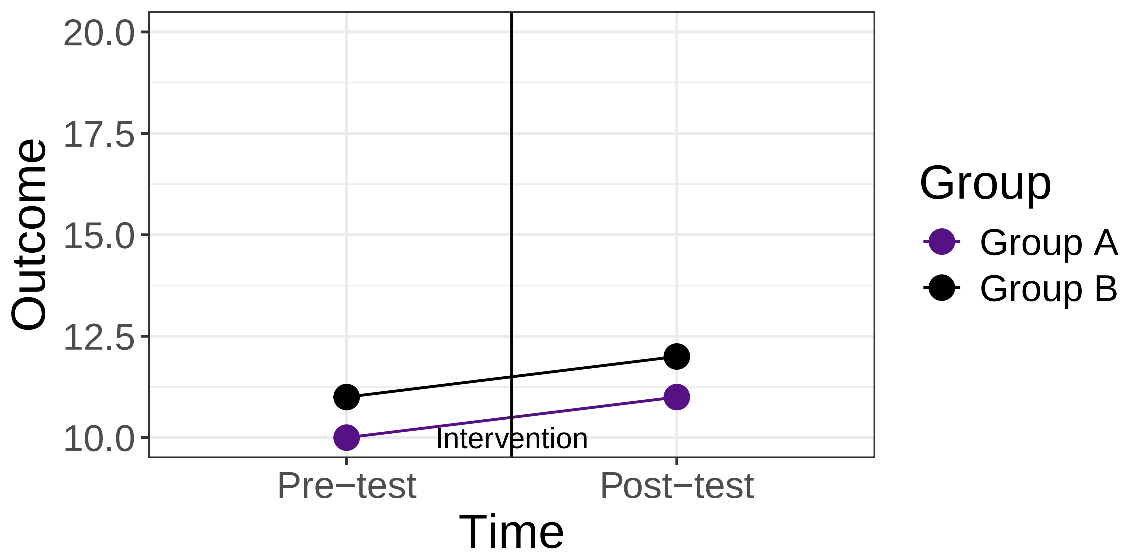
2.1 Chapter Overview

- Because of the ______ required to investigate causal claims, we must be cautious of a number of internal validity threats
- We may have to be especially mindful in our ______ in order to ensure our finding are "real". Design flaws may manifest as third variables.
- Failure to design around these threats may lead to confounds in our ______ validity, and therefore, may hurt our claims

3 More Internal Validity Threats

3.1 Working Example





3.2 Overview

- We have already discussed ________________ effects, design confounds, and order effects - as well as how they impact the ________________ -subjects and within-subjects designs differently
- These all result from a lack of random ________________ not being implemented

- However, there are many other threats to also be concerned with…

3.3 For One-group & Pretest/Posttest Designs

- Some threats are a particular concern in our between-subjects designs:

Maturation

- Certain behaviors may simply ________________ by themselves - we may describe this change as being "spontaneous" and unexpected
- This is change that is ________________ explainable by our intervention or lack-thereof
- *Prevention:*
  - Our ________________ group! Perhaps there is a maturation effect, but because we have a comparison group, we can see if there is a difference despite the maturation.
  - However, this can be most accurately demonstrated with a pretest/posttest design, to look at the trend of change over ________________.

History

- An effect may occur due to some ________________ that creates an unexpected change in our groups.
- However, to be a true history threat this must be ________________, i.e., happening in a outsized manner on only one of the groups.
- *Prevention:*
  - Once again, our comparison group saves the day! If both groups deal with the same event, then the effect is ________________ and not of major concern for the differences between the groups.
  - However, if the event somehow has a biased impact on one group, the experiment may have to be re-done under "normal" conditions.

Regression

- This threat revolves around regression to the ________________, where, over time, extreme scores tend to naturally converge towards the central tendency of the data

- Thus, extremely high scores naturally trend downward as time passes, and extremely low scores naturally trend ________________.
- *Prevention:*
  - Guess what saves us here? ________________ groups! (Notice a trend?)
  - If we see groups equal at the start, and a difference in trend, we know that one group did indeed have an effect above and beyond the effects of ________________.

Differences Between Maturation, History, & Regression

- *Maturation* threats deal with *spontaneous and unexplainable* change in behavior
- *History* threats are due to a ________________ event or know outside environmental influence
- *Regression* threats come form naturally ________________ scores converging on the mean or center of a scale.

Attrition

- Attrition occurs whenever we have some systematic ________________ from our sample (may also be known as "mortality")
- If attrition happens among all groups and ________________ of people, then concern is low.
- But, If there is high attrition among one of the groups, we have a differential effect that could confound
- *Prevention:*
  - It is often impossible to fully prevent people from not completing the study (because research is ________________!)
  - Instead, we may opt to ________________ delete, or completely remove the data corresponding to those who dropped out.

Testing

- As we learned with ________________ effects, participants may grow more skilled if they take an assessment more than once, not due to intervention, but just due to natural growth.
- *Prevention:*

  - Don't use a ________ - ironically, this makes investigating the other threats more difficult
  - Use two different, but equivalent forms of the same test! But how do we know they measure the same thing? $\rightarrow$ ________

Instrumentation

- This occurs when something about the ________ instrument changes over time.
- Put another way, our measurement is not behaving ________!
- *Prevention:*
  - We may use only a post-test design. The ________ can't change over time if it is only used once!
  - We may use construct validity statistics (e.g., ________ α) to assess validity and reliability of the measures at both time points
  - We may counterbalance order of two ________ forms

Combined Threats

- Our selection effects may ________ with any of the threats described here, or even multiple threats may be present.
- As a general rule, we are looking for sources of systematic ________ that may affect one comparison group, but *not* the other.

3.4 Any Study

- Even with good and proper use of comparison groups, we are not entirely out of the woods of threats.

Observer Bias

- When we described observational measures, we broached the problem of potential bias in how our research observers ________ and record measurements.
- This is still a major concern in experiments, even despite the rigorous procedures - effectively, a part of construct validity impacts the internal validity
- **Prevention:**

- Remember to ____________ your observers, i.e., make them unaware of the experiments groups and purpose prior to the recording.

Demand Characteristics

- *Review:* a demand characteristic is when a participant changes their behavior unnaturally due to understanding then nature of the experiment
- *Prevention:*
  - Ideally, both participants and observers should be blinded to study goals and hypotheses, to prevent possible ____________ from arising.
  - Little detail about the conditions should be shared until ____________

Placebo Effect

- Placebo effects occur when the mere ____________ in a treatment produces a pronounced positive effect - this is *extremely common*
- *Prevention:*
  - We can use double-blinding and measure whether the ____________ group sees a more pronounced effect than the placebo group

3.5 Validity in the Face of Many Threats

- Despite the numerous challenges discussed above, experiments are still ____________ and rigorous designs
- A comparison group already does wonders for preventing many of the threats discussed, and further issues can usually be addressed with specific attention paid to the ____________ methods above.

4 Null Effects

4.1 Overview

- Sometimes, we ____________ find a difference where we expect, as evidenced by a p > 0.05 or a 95% CI that contains 0. What gives?!
- This is actually very common in research, and not necessarily indicative of something having been done ____________.

- However, we should investigate whether changes to the design would have changed the statistical ________ of the results

4.2 Not Enough Differences Between Groups

- One root cause of having null effects is having insufficient evidence of difference between the independent variable ________.

Weak Manipulation

- It is possible that our manipulation (i.e., the difference between the two conditions) is simply not impactful enough to create a difference in ________.
- Put another way, the way we ________ our construct of interest was insufficient

Insensitive Measures

- We may have the reverse problem: where our manipulation is impactful, but our ________ variable measure is not sensitive enough to detect it.
- Here, we may need to consider using a ________ measure that is detailed enough to capture change in the outcome.

Ceiling and Floor Effect

- A **ceiling effect** happens when most participants have a score close to the ________ end of possible scores on the outcome measure.
- Inversely, a **floor effect** is when most scores are clustered towards the ________ of possible scores.
- These are symptoms of a ________ measure, and will often result in means that are close together, regardless of condition.

Value in Manipulation Checks

- We alluded to **manipulation checks** before, which offer some way to assess whether a manipulation truly caused any sort of change. This is often some additional variable we ________ alongside the outcome.

Reverse Design Confound

- A design confound may confuse the results of an experiment in the ________________ direction of the effect of the intervention, thus causing the appearance of null results.

4.3 Too Much Within-group Variation

- It may be that we do have a difference between our groups, but the variance within each condition is so large (i.e., ________________ or error) that it becomes difficult to determine.
- This often manifests as large standard errors / wide ________________ intervals.

Measurement Error

- Measurement error naturally occurs in ________________ instrument, but our goal is to limit this as much as possible.
- A Formula:

$$Observed = True + Error$$

- *The solution:*
  - Ideally: Use ________________ tools! e.g., recall chapter 5 and our various assessments of measurement reliability and validity
  - Less ideal: Just measure ________________ people! The inconsistency will balance out naturally (probably)

Individual Differences

- As previously stated, research is ________________, in other words, we capture general trends and mean differences. However, salient-enough individual differences can confuse between groups effects.
- *The solution:*
  - Bigger ________________! This dilutes the effects of individuals with odd or unusual characteristics within each group
  - Use a within-subject design, then each person's trait controls for themselves!

Situation Noise

- This is a much more real type of noise, in which the literal ______________ of the experiment may be playing some un-intended role in the outcome variable.
- This is why many experiments are done in barren, ______________ lab settings.

4.4 Statistical Power

- **Power**, in research, refers to the likelihood that a significant effect is found, when that difference is real.
- Ideally, we employ methods that lend themselves to a high amount of power such as:
  - ______________-groups designs
  - Larger ______________
  - Sensitive measures
  - Low "noise" and high control
- In addition to helping us get significant results, this also aids in the ______________ of the study.

4.5 Transparency About Null Effects

- *A misconception:* Non-significant effects are ______________ worth reporting.
- Null findings are ______________ to the process of self-correcting science