



Week 11 Lecture - Experimental Confounds

Undergraduate Research Methods in Psychology

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

Department of Psychology

Table of Contents

1	Chapter Overview	2
1.1	Learning Objectives	2
1.2	Chapter Overview	2
2	More Internal Validity Threats	3
2.1	Working Example	3
2.2	Overview	3
2.3	For One-group & Pretest/Posttest Designs	4
2.3.1	Maturation	4
2.3.2	History	5
2.3.3	Regression	5
2.3.4	Differences Between Maturation, History, & Regression	6
2.3.5	Attrition	6
2.3.6	Testing	6
2.3.7	Instrumentation	7
2.3.8	Combined Threats	7
2.4	Any Study	8
2.4.1	Observer Bias	8
2.4.2	Demand Characteristics	9
2.4.3	Placebo Effect	9
2.5	Validity in the Face of Many Threats	9
3	Null Effects	10
3.1	Overview	10
3.2	Not Enough Differences Between Groups	10
3.2.1	Weak Manipulation	10
3.2.2	Insensitive Measures	11
3.2.3	Ceiling and Floor Effect	11
3.2.4	Value in Manipulation Checks	11
3.2.5	Reverse Design Confound	11
3.3	Too Much Within-group Variation	12
3.3.1	Measurement Error	12
3.3.2	Individual Differences	12
3.3.3	Situation Noise	13
3.4	Statistical Power	13
3.5	Transparency About Null Effects	13

1 Chapter Overview

1.1 Learning Objectives

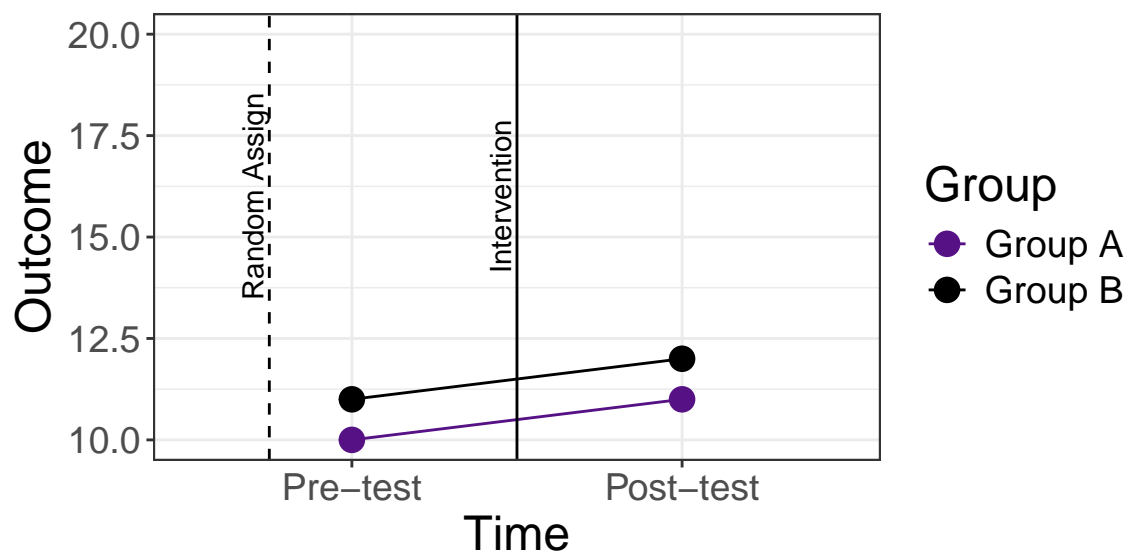
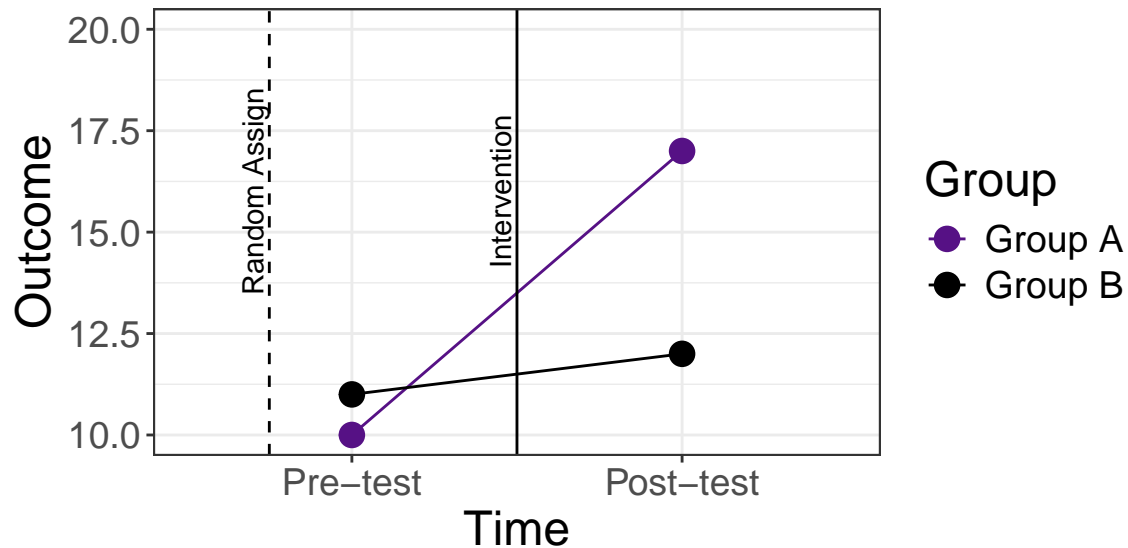
- Interrogate a study and decide whether it rules out twelve potential threats to internal validity.
- Describe how researchers can design studies to prevent internal validity threats.
- Interrogate an experiment with a null result to decide whether the study design obscured an effect or whether there is truly no effect to find.
- Describe how researchers can design studies to minimize possible obscuring factors.
- Describe the importance of comparison groups in ruling out threats to validity
- Explain an example for each type of validity threat

1.2 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 findings 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
-


2 More Internal Validity Threats

2.1 Working Example



2.2 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

 Discuss: We used two different names for experiments based on whether the intervention was done on separate groups or the same groups - what were the names of those types?


- However, there are many other _____ to also be concerned with...

2.3 For One-group & Pretest/Posttest Designs

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

2.3.1 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

 Discuss: What is an example of a behavior that we would expect goes away naturally over time, not necessarily due to a specific intervention?

- *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 _____.
-

2.3.2 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.

? Which of the following seems to be the best example of a history threat?

- A) All members of one condition are older than another condition
- B) The order of our conditions effects how participants respond to an intervention
- C) Some participants in both orders change their behaviors as they discover the hypothesis
- D) Some folks in one condition become extremely ill due to living all together

Explanation:

- *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 _____ impact on one group, the experiment may have to be re-done under “normal” conditions.

2.3.3 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 _____.
- *Example:* Over time, extreme IQ scores tends to converge to the mean as children develop more
- *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 _____.

2.3.4 Differences Between Maturation, History, & Regression

- *Maturation* threats deal with _____ 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.

2.3.5 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 _____

 Discuss: Other than death, what are some other reasons why a participant may leave the sample?

- *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.


2.3.6 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? → _____

2.3.7 Instrumentation

- This occurs when something about the _____ instrument changes over time.
- Put another way, our measurement is not behaving _____ !

 Discuss: Try describing the difference between a testing and an instrumentation threat

- *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

2.3.8 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.
-

? I have found that members of all 3 conditions in my experiment vary greatly in their personality types, regardless of the group they are in - what type of variance is this and should I be concerned?

- A) Unsystematic, and should be concerned
- B) Systematic, and should be concerned
- C) Unsystematic, and should not be concerned
- D) Systematic, and should not be concerned

Explanation:

2.4 Any Study

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

2.4.1 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 _____ concern in experiments, even despite the rigorous procedures - effectively, a part of construct validity impacts the internal validity

🔊 Discuss: Try explaining how the construct validity issues presented by observer bias also contribute to problems in internal validity

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

2.4.2 Demand Characteristics

- *Review:* a demand characteristic is when a participant changes their behavior unnaturally due to _____ 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 _____

2.4.3 Placebo Effect

- Placebo effects occur when the mere _____ in a treatment produces a pronounced positive effect - this is *extremely common*

? I give one condition a sugar pill with no medication and another condition with a novel medication, but I do not tell them which people got which. Which are the types of comparison groups?

- A) Sugar pill - placebo; medication - treatment
- B) Sugar pill - control; medication - control
- C) Sugar pill - treatment; medication - treatment
- D) Sugar pill - treatment; medication - placebo

Explanation:

- *Prevention:*
 - We can use double-blinding and measure whether the _____ group sees a more pronounced effect than the placebo group

2.5 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.
-

3 Null Effects

3.1 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

3.2 Not Enough Differences Between Groups

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

3.2.1 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

? Previously, we used a vocab term to describe a way to ensure our manipulation had the intended effect - what was that term?

- A) Manipulation check
- B) Manipulation insurance
- C) Pilot block
- D) Pre-test


Explanation:

3.2.2 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.

3.2.3 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.

 Discuss: Try to come up of an example of when you might run into a ceiling effect

- These are symptoms of a _____ measure, and will often result in means that are close together, regardless of condition.

3.2.4 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.

3.2.5 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.

3.3 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.

? What validity are standard errors and confidence intervals associated with?

- A) External validity
- B) Internal validity
- C) Construct validity
- D) Statistical validity

Explanation:

3.3.1 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)

3.3.2 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!

3.3.3 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.

3.4 Statistical Power

- **Power**, in research, refers to the _____ 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.

3.5 Transparency About Null Effects

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

 Discuss: Explain what purpose null results serve in the Theory-Hypothesis-Data framework