



---

# **Week 10 Lecture - Introduction to Experiments**

Undergraduate Research Methods in Psychology

---

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

Department of Psychology

## Table of Contents

<b>1</b>	<b>Week Overview</b>	<b>2</b>
1.1	Learning Objective(s) . . . . .	2
1.2	Week Overview . . . . .	2
<b>2</b>	<b>Experimental Variables</b>	<b>2</b>
2.1	Overview . . . . .	2
2.2	Independent & Dependent . . . . .	3
2.3	Control . . . . .	4
<b>3</b>	<b>Why are Experiments Causal?</b>	<b>4</b>
3.1	Overview . . . . .	4
3.2	Covariance . . . . .	5
3.3	Temporal Precedence . . . . .	6
3.4	Internal Validity . . . . .	6
3.4.1	Variability . . . . .	6
3.4.2	Selection Effects . . . . .	7
<b>4</b>	<b>Independent-groups Design</b>	<b>7</b>
4.1	Overview . . . . .	7
4.2	Posttest-only . . . . .	8
4.3	Pretest/Posttest . . . . .	8
4.4	Comparison . . . . .	9
<b>5</b>	<b>Within-groups Design</b>	<b>9</b>
5.1	Overview . . . . .	9
5.2	Repeated Measures . . . . .	9
5.3	Concurrent Measures . . . . .	10
5.4	Causal Criteria in Within-Groups . . . . .	10
5.4.1	Internal Validity . . . . .	10
5.5	Advantages . . . . .	11
5.6	Disadvantages . . . . .	12
5.7	Pretest-posttest vs Repeated Measures . . . . .	12
<b>6</b>	<b>Four Validities Critique</b>	<b>13</b>
6.1	Construct Validity . . . . .	13
6.2	External Validity . . . . .	14
6.3	Statistical Validity . . . . .	14
6.4	Internal Validity . . . . .	14

# 1 Week Overview

## 1.1 Learning Objective(s)

- Apply the three criteria for establishing causation to experiments and explain why experiments can support causal claims.
- Identify an experiment's independent, dependent, and control variables.
- Classify experiments as independent-groups and within-groups designs and explain why researchers might conduct each type of study.
- Evaluate three potential threats to internal validity in an experiment— design confounds, selection effects, and order effects—and explain how experimenters usually avoid them.
- Interrogate an experimental design using the four validities.


## 1.2 Week Overview

- Experiments are \_\_\_\_\_ investigations that are supportive of \_\_\_\_\_ claims.
- They tell us about \_\_\_\_\_ influence that one variable has upon another
- They also tend to be the most \_\_\_\_\_ and difficult studies, with lots of required resources

# 2 Experimental Variables

## 2.1 Overview


- Experiments will have at least one measured variable and at least one \_\_\_\_\_ variable
  - *Example (from the first test):* Changing pillow type (manipulated variable) results in change in sleep quality (measured variable)
-

 Discuss: What word in the prior example should really cue you to this being a casual claim?

- *Review:* a manipulated variable is one that we can directly influence and \_\_\_\_\_, to have a hypothesized impact on our measured variables
- *Review:* measured variables will be things like self-reports or \_\_\_\_\_ measures

## 2.2 Independent & Dependent

- An **independent variable** is one that we can change “independent” of other factors, i.e., it does not \_\_\_\_\_ on something else
  - The manipulated variable
  - It levels (i.e., possible values) are called \_\_\_\_\_
- A **dependent variable** is one that is somehow hypothesized to be connected to or \_\_\_\_\_ upon the condition or state of the independent variable
  - The measured variable
  - It may also be referred to as an outcome variable, similar to a multiple regression design
- Often, the \_\_\_\_\_ variable will be categorical/grouped and the dependent variable will be continuous - but this is not a hard and fast rule


 I am looking at whether depression symptoms are improved more by psychotherapy or medication treatments - what construct is being treated as dependent?

- A) Depression symptoms
- B) Medication
- C) psychotherapy
- D) None of these


Explanation:

## 2.3 Control

- **Control variables** are those that are \_\_\_\_\_ held at the same level across participants, to prevent it from having an impact on the outcome

 Discuss: What do we call some value that does not change values in our data?

- The purpose of keeping these still is to effectively \_\_\_\_\_ the effect of the independent variable(s), so that it does not confound our causal link.
- *Example:* I am looking at whether more exercise leads to an increase in energy. However, I suspect that caffeine level could confound this relationship, so I instruct participants to not take any caffeine. Caffeine level across participants is consistently 0, therefore it is a control variable and constant.

 What do we call a variable that interferes with internal validity, like in the prior example?

- A) A sample
- B) A covariate
- C) A third variable
- D) A criterion

Explanation:

## 3 Why are Experiments Causal?

### 3.1 Overview

- *Review:* Recall our three causal criteria
  - \_\_\_\_\_
  - Temporal precedence

- Internal validity
- Finally, with experiments we have a \_\_\_\_\_ by which to definitively establish all of these

### 3.2 Covariance

- If our separate \_\_\_\_\_ see differences in the outcome variable (i.e., mean differences on a continuous outcome), we would say there is covariance between the independent and dependent variables
- We have to distinguish what conditions we may assign or manipulate participants to:
  - **Comparison groups:** We have at least \_\_\_\_\_ conditions or levels in our independent variable, so that we can compare them

? What source of information did we explicitly say lacks comparison groups, early in the semester?

- A) Research
- B) Personal experience
- C) Authority
- D) Intuition

Explanation:


- We may use specific labels for types of comparison groups:
  - **Control group:** One that receives no or \_\_\_\_\_ intervention
  - **Treatment group:** One that receives the intervention or change of interest
  - **Placebo group:** A type of control, that is exposed to an inert or \_\_\_\_\_ treatment
- Covariance is \_\_\_\_\_ as important as the other causal criteria → no difference between conditions = no causation
- *Example:* I try two different teaching styles between classes to deliver a particular lesson. I take a test at the end of the class on the content. If there is a difference in test scores between the teaching styles, there is covariance.

### 3.3 Temporal Precedence

- By \_\_\_\_\_ the independent variable prior to measuring the dependent variable, we establish that it came prior and that we have temporal precedence.
- We will discuss how exactly to implement time in our discussions on the design types shortly

### 3.4 Internal Validity


- This is the most \_\_\_\_\_ and core feature that sets experiments apart from other designs, and satisfies the last criteria for causality
- Put simply, we must rule out other reasonable and theoretically-salient \_\_\_\_\_ for the effects that we see in the outcome variable
- **Confounding variables/confounds:** Those that offer alternative \_\_\_\_\_ for the effects seen in an outcome.
  - Often we will try to \_\_\_\_\_ possible confounding variables
  - When a confound results from poor planning in the research design → **design confound** → poor internal validity
- Experiments can implement \_\_\_\_\_ controls that prevent confounds.

 Discuss: Previously, we discussed a design that gives a statistical control, in contrast to the design control we discuss now - what design was that, and in what context did that control work?

#### 3.4.1 Variability

- \_\_\_\_\_ **Variability** is when a confound changes in a \_\_\_\_\_ a predictable manner *with* the independent variable
  - **Unsystematic Variability** is when a confound changes in some way \_\_\_\_\_ or indifferent to the independent variable. It may occur in a seemingly random fashion.
-

- The result is that there may be variability in the outcome, but if it happens \_\_\_\_\_ across all the conditions, it is not as much a concern.

 Discuss: Based on our discussions in this lecture so far, which of these seems worse for us, and why?

### 3.4.2 Selection Effects

- **Selection Effects** occur when there is systematic variance of some factor across the conditions of the independent variable
  - This ends up being a confound → is the change in outcome due to this variance of a third variable or because of the condition? We \_\_\_\_\_ know
  - *Example:* In a treatment study on memory treatments, we assign some people to treatment A and some to treatment B. However, the group with treatment A comprises of mostly older individuals, whereas group B is mostly younger. Age may play a role in resulting memory impairment.
- This can be largely prevented by using **random assignment**, where members of the \_\_\_\_\_ are sort at random into conditions of the IV
- Using **matched groups** is another method to reduce the \_\_\_\_\_ of selection effects
  - This is where participants are measured on some third variable that may confound and ranked. Then each \_\_\_\_\_ is randomly placed in an alternating group
  - A major difficulty of this method is you must accurately identify what \_\_\_\_\_ should be matched on.

## 4 Independent-groups Design

### 4.1 Overview

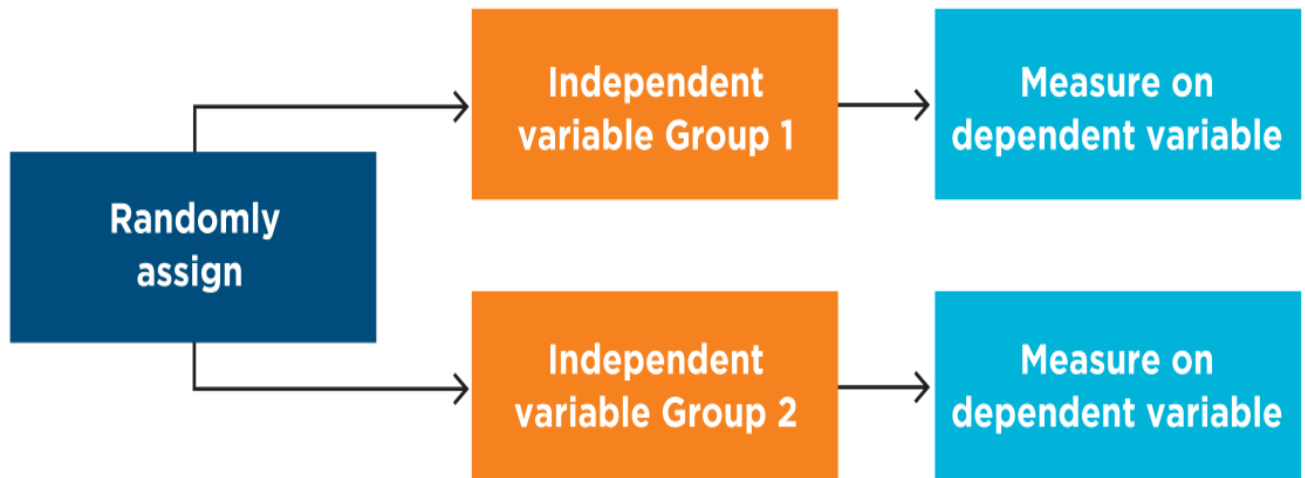
- **Independent-groups designs** are those in which the participants of each condition are entirely \_\_\_\_\_ from one another



- A.K.A. as between-subject or between-groups

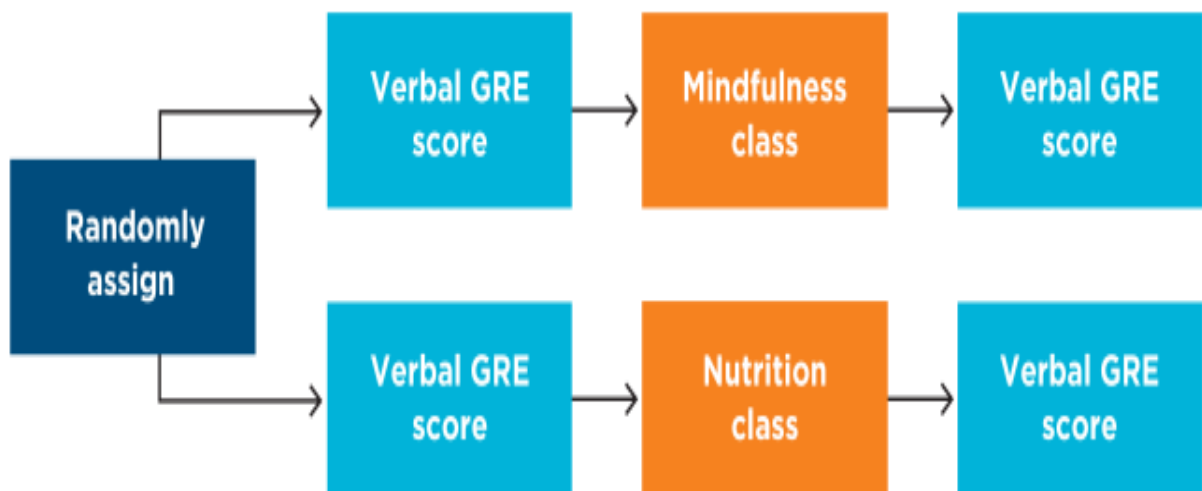
## 4.2 Posttest-only

- **Posttest-only design:** Separate groups of participants, assigned to different conditions, are measured only after the assignment and \_\_\_\_\_ has occurred.



## 4.3 Pretest/Posttest

- **Pretest/Posttest design:** Separate groups are tested on the outcome \_\_\_\_\_ the random assignment, but before *and* after the intervention



## 4.4 Comparison

- Both the posttest-only and the pretest/posttest designs are perfectly \_\_\_\_\_ experimental designs to investigate causal claims
- The primary difference is:
  - Posttest-only \_\_\_\_\_ that random assignment produces equivalent groups prior to intervention
  - Pretest/Posttest does not make this assumption and gives measures to validate
- Ideally, pretest/posttest is the “better” design as it provides more \_\_\_\_\_ and data - but we may use posttest-only for practicality

? I measure a sample of folks on their extraversion and happiness using two continuous measures and then measure them again in a year - what design is this?

- A) Pretest/Posttest
- B) Bivariate correlational
- C) Posttest-only
- D) Multivariate correlational

Explanation:

## 5 Within-groups Design

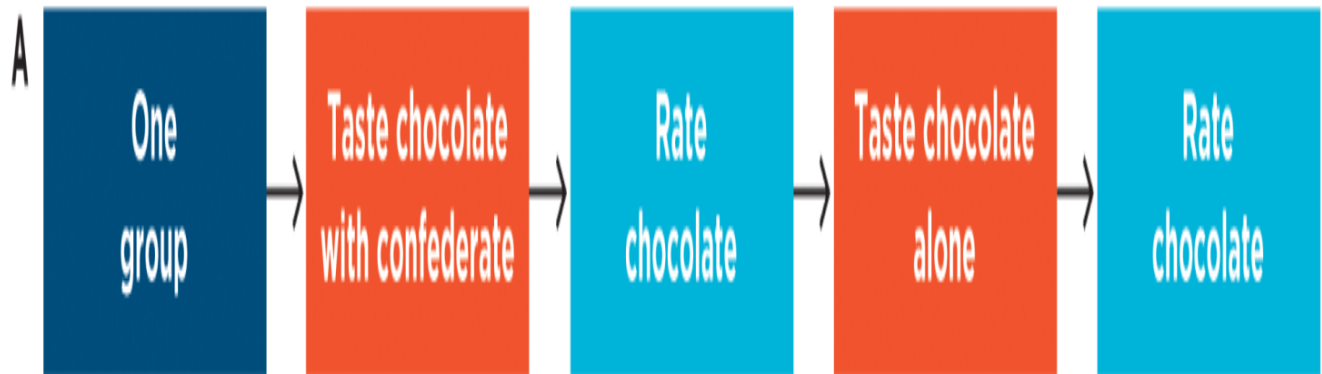
### 5.1 Overview

- **Within-groups designs** are when the \_\_\_\_\_ group of individuals is exposed to each condition
  - This is in contrast to the separate groups of the prior designs

### 5.2 Repeated Measures

- As the name would suggest, **repeated-measures** is when measurement on the outcome(s) of interest is taken after exposure to each condition, which happens

\_\_\_\_\_



### 5.3 Concurrent Measures

- **Concurrent measures** is when both “conditions” are displayed to the participant at the \_\_\_\_\_ time and the measure of interest is some function of the behavior or affect towards the conditions
  - Review: Remember Harlow’s \_\_\_\_\_ ?
- Depending on the conditions, this may not be a viable strategy


### 5.4 Causal Criteria in Within-Groups

- Do we have covariance?
  - Manipulate and measure two separate variables to see their \_\_\_\_\_
- Do we have temporal precedence?
  - Each measure comes \_\_\_\_\_ the intervention or manipulation

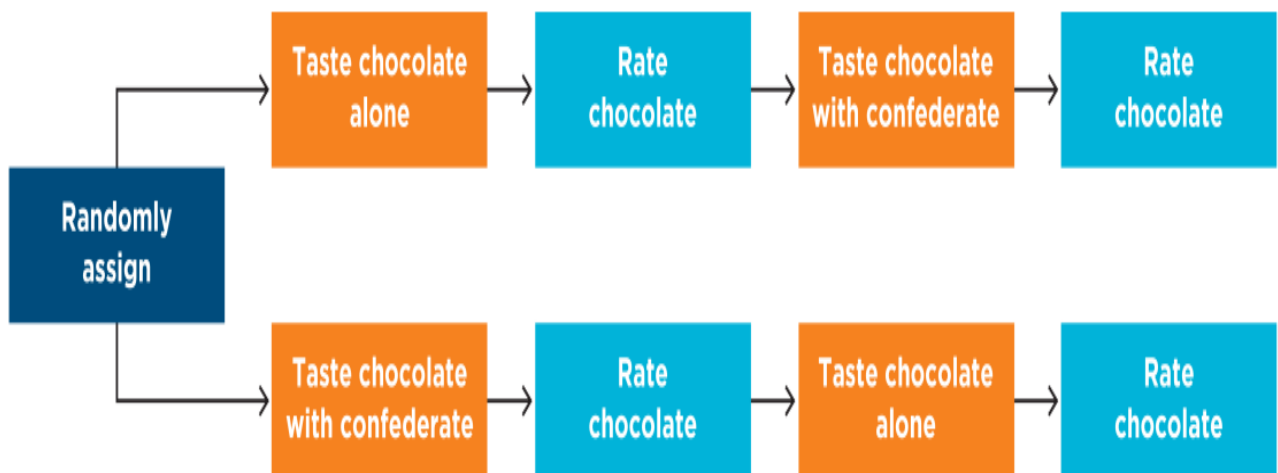
#### 5.4.1 Internal Validity

- Participants \_\_\_\_\_ for themselves, i.e., random assignment isn’t necessary
  - Because the individual is the same across conditions, we don’t need to ensure they are \_\_\_\_\_ on some variable like with between groups
- However, we do need to be concerned with **order effects**, where exposure to one condition can have an \_\_\_\_\_ on the next reaction to a different condition
  - **Practice effects:** When someone becomes more \_\_\_\_\_ at a task due to practice on the prior conditions

- **Fatigue effects:** When someone becomes \_\_\_\_\_ skilled at a task due to tiring from repeated testing
- **Carryover effects:** When the effect of the previous condition is still \_\_\_\_\_ at the start of the next condition

 Discuss: Give an example of a situation where you think practice effects could occur on a certain measure

- Order effects are best prevented by **counterbalancing**, that is assigning some individuals to one \_\_\_\_\_ of conditions, and assigning the others to a separate order
  - **Full counterbalancing:** when \_\_\_\_\_ possible orders are done
  - \_\_\_\_\_ **counterbalancing:** when only some possible orders are represented



## 5.5 Advantages

- The “groups” are naturally equivalent, as the conditions are tested upon the same individual!
  - I.e., no concern for \_\_\_\_\_ effects, like in between-groups
  - We are able to express a strong \_\_\_\_\_ over all the various personal characteristics that could vary unsystematically

- Also requires \_\_\_\_\_ sample sizes for adequate power

## 5.6 Disadvantages

- As previously discussed, *counterbalancing* is necessary in order to \_\_\_\_\_ order effects from occurring
- We must attempt to “return” to baseline after each condition - but some interventions may have a \_\_\_\_\_ or prolonged effect on the participants, confounding the next condition
- We have to contend with the possibility that a participant \_\_\_\_\_ the hypothesis of a study, by virtue of seeing all the conditions, and changes their behavior in relation to this knowledge.
  - This is called a **demand characteristic**
  - *Review:* This is almost similar to “\_\_\_\_\_”, as we discussed with self-report measures

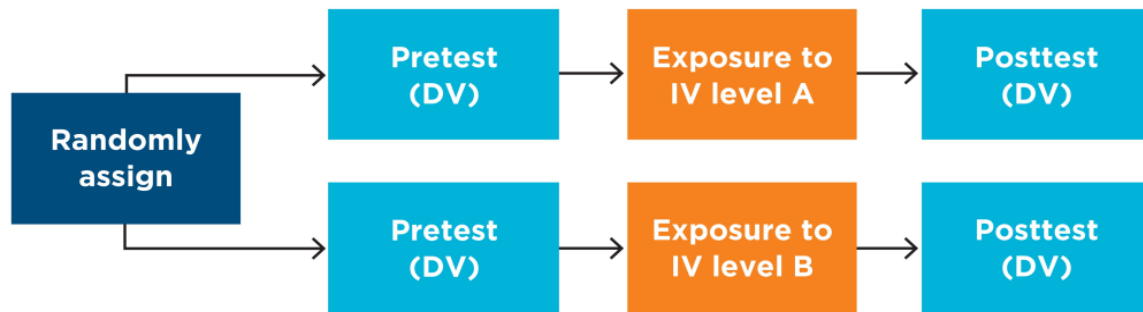
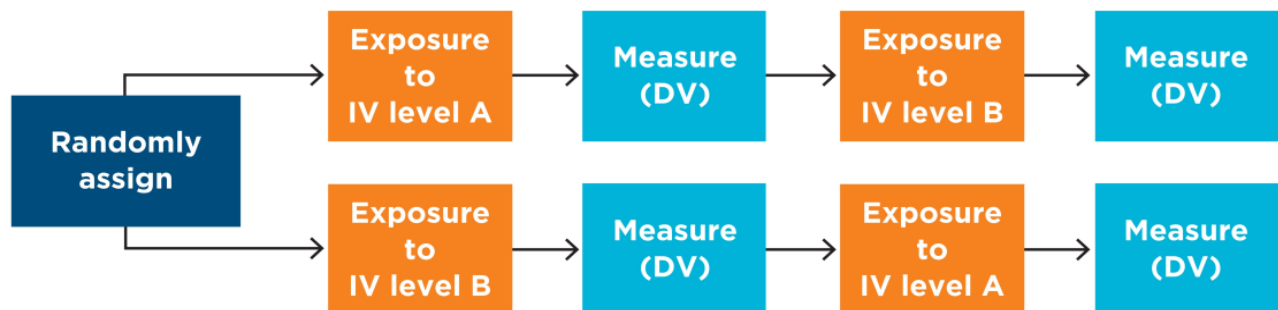
? What validity is a demand characteristic a threat to?

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

Explanation:

## 5.7 Pretest-posttest vs Repeated Measures


- The important difference is that whether conditions are \_\_\_\_\_ to different groups or the same people!

**Pretest/posttest design****Repeated-measures design**

## 6 Four Validities Critique

### 6.1 Construct Validity

- Dependent variables are likely to be \_\_\_\_\_ the same way we have previously discussed self-report and observational measures.


 Discuss: What were the 3 primary threats to observational design? Briefly describe them

- Independent variables are often based, at least initially, on whether they are valid.
  - We may use a **manipulation check**, or an additional measure meant to ensure that the intervention had the \_\_\_\_\_ effect

- This can be a good use for a **pilot study**
- All choices of variables to measure and manipulate should be guided by an understanding of the existing \_\_\_\_\_ in the area → we must assess whether the variables represent the constructs of the theory well

## 6.2 External Validity

- As previously discussed, sample \_\_\_\_\_ is the primary factor to consider in whether a sample is representative a specified population, and thus, whether results are generalizable.
- Extrapolating results to other, similar situations, requires \_\_\_\_\_ re-search
- Typically, experimental research tends to have worse external validity due to tight controls - the results aren't applicable to "normal" life - but they may be \_\_\_\_\_ with strong correlational designs (i.e., longitudinal or multiple regression) to help with this

 Discuss: What about sample sizes? What role will they play in external validity or the other validities?

## 6.3 Statistical Validity

- Just like previous statistical validity discussion, we are concerned with 4 aspects:
  - Effect \_\_\_\_\_ : magnitude of effect
  - Precision: What confidence do we have that this effect falls in a narrow range of values
  - Replication: When a study is replicated, are statistics \_\_\_\_\_.
  - Significance: Were these results just a fluke of \_\_\_\_\_ ?

## 6.4 Internal Validity

- Covered in the previous sections, but this is the key question:
-

- “Are the effects found here some \_\_\_\_\_ error, a third variable, or genuinely the intervention?”